

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

H04N 7/26 (2006.01)

H04N 7/50 (2006.01)

H04N 13/00 (2006.01)



# [12] 发明专利申请公布说明书

[21] 申请号 200880005395.7

[43] 公开日 2010年1月13日

[11] 公开号 CN 101627635A

[22] 申请日 2008.2.20

[21] 申请号 200880005395.7

[30] 优先权

[32] 2007.2.21 [33] US [31] 11/677,335

[86] 国际申请 PCT/US2008/054447 2008.2.20

[87] 国际公布 WO2008/103748 英 2008.8.28

[85] 进入国家阶段日期 2009.8.18

[71] 申请人 高通股份有限公司

地址 美国加利福尼亚州

[72] 发明人 杨凯希 王浩宏

哈立德·希勒米·厄勒-马列

沙拉什·曼朱纳什

[74] 专利代理机构 北京律盟知识产权代理有限公司  
代理人 刘国伟

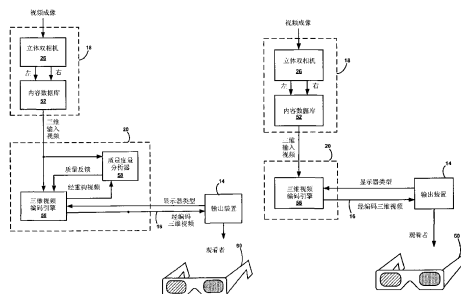
权利要求书4页 说明书23页 附图8页

## [54] 发明名称

三维视频编码

## [57] 摘要

立体三维视频帧包括经组合以产生立体图像的左分量及右分量。对于给定量的失真，所述左分量及右分量可由于人眼的失真响应的不对称性而对所述立体图像的感知视觉质量具有不同影响。三维视频编码器基于帧级位预算及三维视频的左分量与右分量之间的加权来调整所述三维视频的所述左分量与所述右分量之间的编码位分配。所述视频编码器可在  $\rho(\rho)$  域中产生所述位分配。可基于指示由所述左分量及右分量产生的总质量的质量度量来导出经加权的位分配。所述经加权的位分配补偿所述不对称的失真响应以减小所述立体图像中的总感知失真且进而增强或维持视觉质量。



1. 一种方法，其包含：  
获得界定可用于三维视频帧的编码位的数目的帧预算；及  
基于所述帧预算及所述三维视频帧的左分量与右分量之间的加权而针对所述三维视频帧的所述左分量及右分量内的块产生位分配值。
2. 根据权利要求 1 所述的方法，其中所述位分配值包括  $\rho$  域值，所述方法进一步包含：  
将所述  $\rho$  域值映射到对应量化参数（QP）值；及  
基于所述映射而将一定数目的编码位分配给所述左分量及右分量中的所述块中的每一者。
3. 根据权利要求 2 所述的方法，其进一步包含使用所述所分配的编码位来编码所述左分量及右分量中的所述块。
4. 根据权利要求 2 所述的方法，其进一步包含基于所述加权而将不同量的所述编码位分配给所述左分量及右分量。
5. 根据权利要求 1 所述的方法，其进一步包含基于用以观看所述左分量及右分量的显示装置类型而调整所述加权。
6. 根据权利要求 5 所述的方法，其进一步包含从所述三维视频帧所传输到的输出装置接收对所述显示装置类型的指示。
7. 根据权利要求 1 所述的方法，其中所述加权是基于人类观看者对所述左分量及右分量中的失真的敏感性。
8. 一种包含视频编码器的装置，所述视频编码器经配置以获得界定可用于三维视频帧的编码位的数目的帧预算，且基于所述帧预算及所述三维视频帧的左分量与右分量之间的加权而针对所述三维视频帧的所述左分量及右分量内的块产生位分配值。
9. 根据权利要求 8 所述的装置，其中所述位分配值包括  $\rho$  域值，且其中所述视频编码

器经配置以将所述  $\rho$  域值映射到对应量化参数 (QP) 值, 且基于所述映射而将一定数目的编码位分配给所述左分量及右分量中的所述块中的每一者。

10. 根据权利要求 9 所述的装置, 其中所述视频编码器经配置以使用所述所分配的编码位来编码所述左分量及右分量中的所述块。
11. 根据权利要求 9 所述的装置, 其中所述视频编码器经配置以基于所述加权而将不同量的所述编码位分配给所述左分量及右分量。
12. 根据权利要求 8 所述的装置, 其中所述视频编码器经配置以基于用以观看所述左分量及右分量的显示装置类型而调整所述加权。
13. 根据权利要求 12 所述的装置, 其中所述视频编码器经配置以从所述三维视频帧所传输到的输出装置接收对所述显示装置类型的指示。
14. 根据权利要求 8 所述的装置, 其中所述加权是基于人类观看者对所述左分量及右分量中的失真的敏感性。
15. 一种装置, 其包含:
  - 用于获得界定可用于三维视频帧的编码位的的数目的帧预算的装置; 及
  - 用于基于所述帧预算及所述三维视频帧的左分量与右分量之间的加权而针对所述三维视频帧的所述左分量及右分量内的块产生位分配值的装置。
16. 根据权利要求 15 所述的装置, 其中所述位分配值包括  $\rho$  域值, 所述装置进一步包含:
  - 用于将所述  $\rho$  域值映射到对应量化参数 (QP) 值的装置; 及
  - 用于基于所述映射而将一定数目的编码位分配给所述左分量及右分量中的所述块中的每一者的装置。
17. 根据权利要求 16 所述的装置, 其进一步包含用于使用所述所分配的编码位来编码所述左分量及右分量中的所述块的装置。

18. 根据权利要求 16 所述的装置，其进一步包含用于基于所述加权而将不同量的所述编码位分配给所述左分量及右分量的装置。
19. 根据权利要求 15 所述的装置，其进一步包含用于基于用以观看所述左分量及右分量的显示装置类型而调整所述加权的装置。
20. 根据权利要求 19 所述的装置，其进一步包含用于从所述三维视频帧所传输到的输出装置接收对所述显示装置类型的指示的装置。
21. 根据权利要求 15 所述的装置，其中所述加权是基于人类观看者对所述左分量及右分量中的失真的敏感性。
22. 一种包含计算机可读媒体的计算机程序产品，所述计算机可读媒体包含用以致使计算机执行以下操作的代码：
  - 获得界定可用于三维视频帧的编码位的的数目的帧预算；及
  - 基于所述帧预算及所述三维视频帧的左分量与右分量之间的加权而针对所述三维视频帧的所述左分量及右分量内的块产生位分配值。
23. 根据权利要求 22 所述的计算机程序产品，其中所述位分配值包括  $\rho$  域值，且其中所述计算机可读媒体进一步包含用以致使所述计算机将所述  $\rho$  域值映射到对应量化参数（QP）值且基于所述映射而将一定数目的编码位分配给所述左分量及右分量中的所述块中的每一者的代码。
24. 根据权利要求 23 所述的计算机程序产品，其中所述计算机可读媒体进一步包含用以致使所述计算机使用所述所分配的编码位来编码所述左分量及右分量中的所述块的代码。
25. 根据权利要求 22 所述的计算机程序产品，其中所述计算机可读媒体进一步包含用以致使所述计算机基于用以观看所述左分量及右分量的显示装置类型而调整所述加权的代码。
26. 根据权利要求 25 所述的计算机程序产品，其中所述计算机可读媒体进一步包含用

---

以致使所述计算机从所述三维视频帧所传输到的输出装置接收对所述显示装置类型的指示的代码。

27. 根据权利要求 22 所述的计算机程序产品，其中所述加权是基于人类观看者对所述左分量及右分量中的失真的敏感性。

## 三维视频编码

### 技术领域

本发明涉及数字视频编码，且更明确地说，涉及用于编码三维视频内容的技术。

### 背景技术

数字视频能力可并入到广泛范围的装置中，包括数字电视、数字直播系统、无线通信装置、个人数字助理（PDA）、膝上型计算机、桌上型计算机、数字音乐及视频播放器、手持式游戏装置、视频游戏控制台、数码相机、数字记录装置、蜂窝式或卫星无线电电话等。数字视频装置可在以增加的带宽效率来处理及传输视频序列中提供优于常规模拟视频系统的显著改进。

已建立用于编码数字视频序列的不同视频编码标准。举例来说，运动图像专家小组（MPEG）已开发出了包括 MPEG-1、MPEG-2 及 MPEG-4 的多个标准。其它实例包括国际电信联盟（ITU）-T H.263 标准以及 ITU-T H.264 标准及其对应物 ISO/IEC MPEG-4 第 10 部分（即，高级视频编码（AVC））。这些视频编码标准通过以压缩方式编码数据而支持视频序列的经改进传输效率。将视频编码标准应用于三维视频对于支持用户经由具有有限带宽的有线或无线网络进行三维观看来说为合乎需要的。

视频内容可以常规的二维（2D）格式或以三维（3D）格式来记录。三维视频内容可以不同方式来获得及呈现。作为实例，三维视频内容可由立体双相机获得且由左眼分量及右眼分量呈现。或者，常规的单相机可结合深度相机使用以获得镶嵌式视图及深度数据两者。作为另一替代方案，单相机可用于产生镶嵌式视图，其经处理以产生镶嵌式视图及深度数据两者。不同的三维视频源可能需要不同类型的视频输出装置，例如常规的二维显示器、由特殊眼镜辅助的立体显示器或自动立体显示器。

### 发明内容

本发明针对于用于三维视频编码的技术。立体三维视频帧包括经组合以产生立体图像的左分量及右分量。对于给定量的失真，左分量及右分量可由于人眼的失真响应的不对称性而对立体图像的感知视觉质量具有不同影响。三维视频编码器基于帧级位预算及三维视频的左分量与右分量之间的加权来调整三维视频的左分量与右分量之间的编码

位分配。

所述加权可基于人类观看者对左分量及右分量中的失真的敏感性。视频编码器可使用所述加权而在  $\rho$  域中产生位分配。可基于指示由左分量及右分量的作用产生的总质量的质量度量而导出经加权的位分配。所述经加权的位分配补偿不对称失真响应以减小立体图像中的总感知失真且进而增强或维持视觉质量。

在一个方面中，本发明提供一种方法，所述方法包含：获得界定可用于三维视频帧的编码位的的数目的帧预算；及基于帧预算及三维视频帧的左分量与右分量之间的加权来针对三维视频帧的左分量及右分量内的块产生位分配值。

在另一方面中，本发明提供一种装置，所述装置包含视频编码器，其经配置以获得界定可用于三维视频帧的编码位的的数目的帧预算，且基于帧预算及三维视频帧的左分量与右分量之间的加权来针对三维视频帧的左分量及右分量内的块产生位分配值。

本文中所描述的技术可以硬件、软件、固件或其任何组合来实施。如果以软件来实施，则所述技术可部分地由包含计算机可读媒体的计算机程序产品来实现，计算机可读媒体包含致使计算机执行本文中所描述的方法中的一者或一者以上的程序代码。

在附图及以下描述中陈述一个或一个以上实施例的细节。从描述及图式且从权利要求书将容易明白其它特征、目标及优点。

## 附图说明

图 1 为说明三维视频编码及解码系统的框图。

图 2 为说明用于三维视频产生、编码及呈现的不同技术的框图。

图 3A 为说明使用双遍技术由立体双相机进行三维视频产生以用于经由立体显示器来呈现的框图。

图 3B 为说明使用优化技术由立体双相机进行三维视频产生以用于经由立体显示器来呈现的框图。

图 4 为说明用于使用优化技术来产生经编码三维视频的视频编码引擎的框图。

图 5 为说明三维视频质量度量的产生及基于三维视频质量度量的编码位分配的流程图。

图 6 为说明基于三维视频质量度量而将编码位分配给三维视频分量的另一流程图。

图 7 为说明在  $\rho$  域中将编码位分配给三维视频分量的流程图。

## 具体实施方式

本发明针对于用于三维视频编码的技术。立体三维视频帧包括经组合以产生立体图像的左分量及右分量。对于给定量的失真，左分量及右分量可由于人眼的失真响应的不对称性而对立体图像的感知视觉质量具有不同影响。三维视频编码器基于帧级位预算及三维视频的左分量与右分量之间的加权来调整三维视频的左分量与右分量之间的编码位分配。所述加权可基于人类观看者对左分量及右分量中的失真的敏感性。视频编码器可在  $\rho$  域中产生位分配。可基于指示由左分量及右分量的作用产生的总质量的质量度量而导出经加权的位分配。所述经加权的位分配补偿不对称失真响应以减小总感知失真以便减小立体图像中的总感知失真且进而增强或维持视频质量并进而增强或维持立体图像中的视觉质量。

图 1 为说明三维视频编码及解码系统 10 的框图。如图 1 所示，系统 10 包括三维视频源装置 12 及三维视频输出装置 14。装置 12、14 经由传输信道 16 而通信。传输信道 16 可为有线或无线通信媒体。在图 1 的实例中，三维视频源装置 12 包括三维视频俘获装置或三维视频保存器 18 及三维视频编码器 20。三维视频输出装置 14 包括三维视频解码器 22 及三维视频显示器 24。三维视频源装置 12 产生用于传输到三维视频输出装置 14 的视频。然而，在一些方面中，装置 12、14 可以大致对称方式来操作。举例来说，装置 12、14 中的每一者可包括视频编码及解码组件。因此，系统 10 可支持视频装置 12、14 之间的单向或双向视频传输，例如用于视频串流、视频广播或视频电话。尽管为了易于说明起见而未图示，但为了支持有线及/或无线通信，装置中的一者或两者可包括适当的调制解调器、滤波器、频率转换及放大组件。

三维视频俘获装置/保存器 18 可包括多种三维视频俘获装置中的任一者，例如立体双相机、与深度相机组合的单相机或与二维到三维转换器组合的单相机。或者，三维视频俘获装置/保存器 18 可为存储先前所俘获的三维视频的保存器装置。作为其它替代方案，三维视频可从来自内容提供者的视频馈入获得或从计算机产生的图形数据再现。在任一情况下，所俘获、所预俘获或所再现的三维视频可由三维视频编码器 20 编码以用于经由传输信道 16 从三维视频源装置 12 传输到三维视频输出装置 14。作为又一实例，三维视频源装置 12 可包含广播装置，其广播一个或一个以上信道的三维视频数据，以及在一些情况下广播若干信道的二维视频数据。

系统 10 可根据会话起始协议 (SIP)、ITU-T H.323 标准、ITU-T H.324 标准或其它标准来支持视频电话或视频串流。而且，系统 10 可支持三维视频摄录像机应用。三维视频编码器 20 可根据例如 MPEG-2、MPEG-4、ITU-T H.263 或 ITU-T H.264 及其对应物 ISO/IEC MPEG-4 第 10 部分 (即，高级视频编码) 的视频压缩标准来产生经编码视频数



据。尽管图 1 中未展示，但视频编码器 20 及视频解码器 22 可分别与音频编码器及解码器集成，且可包括适当的多路复用器-多路分用器 (MUX-DEMUX) 单元或其它硬件及软件，以处理共同数据流或单独数据流中的音频与视频两者的编码。如果适用的话，则多路复用器-多路分用器单元可遵照 ITU H.223 多路复用器协议，或例如用户数据报协议 (UDP) 等其它协议。

在一些方面中，对于视频广播，本发明预期应用于增强型 H.264 视频编码以用于使用待公开为技术标准 TIA-1099 的仅前向链路 (FLO) 空中接口规范“用于陆地移动多媒体多播的仅前向链路空中接口规范 (Forward Link Only Air Interface Specification for Terrestrial Mobile Multimedia Multicast)” (“FLO 规范”) 来在陆地移动多媒体多播 (TM3) 系统中传递实时视频服务。FLO 规范包括界定流语法及语义以及适合于经由 FLO 空中接口传递服务的解码过程的实例。或者，可根据例如 DVB-H (数字视频广播-手持式)、ISDB-T (集成服务数字广播-陆地) 或 DMB (数字媒体广播) 等其它标准来广播视频。因此，源装置 12 可为移动无线终端、视频串流服务器或视频广播服务器。然而，本发明中所描述的技术不限于任何特定类型的广播、多播或点到点系统。如果用于广播系统中，则信道中的一者或一者以上可为三维信道且其它信道可为二维信道。

视频编码器 20 及视频解码器 22 各自可被实施为一个或一个以上处理器、数字信号处理器、专用集成电路 (ASIC)、现场可编程门阵列 (FPGA)、离散逻辑、软件、硬件、固件或其任何组合。视频编码器 20 及视频解码器 22 中的每一者可包括于一个或一个以上编码器或解码器中，所述编码器或解码器中的任一者可被集成为相应订户装置、广播装置、服务器等中的组合式编码器/解码器 (编解码器) 的部分。另外，在适用时，三维视频源装置 12 及三维视频输出装置各自可包括用于经编码视频的传输及接收的适当调制、解调、频率转换、滤波及放大器组件，其在无线通信的情况下包括射频 (RF) 无线组件及天线。然而，为了易于说明起见，图 1 中未展示此类组件。

为了支持三维视频，视频编码器 20 编码三维视频的左分量及右分量。左分量对应于在视觉上由人类观看者的左眼所感知的分量。右分量对应于在视觉上由人类观看者的右眼所感知的分量。共同地，左分量及右分量由人类观看者在视觉上整合以产生立体图像的总三维视图。左分量大体上可指代三维视频帧的经配置以由左眼观看的至少一部分。类似地，右分量大体上可指代三维视频帧的经配置以由右眼观看的至少一部分。每一分量包括界定视频帧的至少一部分的数据，使得与左分量及右分量相关联的视频帧的组合产生总三维立体视频帧。另外，每一分量可包括一视频块或多个视频块 (例如，形成左帧或右帧的宏块 (MB))。在一些情况下，对于具有滤色眼镜的立体显示器，左分

量可对应于红色数据且右分量可对应于蓝色及绿色数据（即，以补色立体图形式）。因此，左分量可携带来自帧的红色数据且右分量可携带来自帧的蓝色及绿色数据。

视频编码器 20 使用（例如）MPEG 4 或 H.264 编码来单独地编码左分量及右分量，且将经编码分量传输到三维视频输出装置 14 以供由三维视频解码器 22 解码。经解码的左三维视频分量及右三维视频分量用于驱动三维视频显示器 24 以直接产生或结合例如立体眼镜或取景器等额外显示附件来产生三维视频。视频显示器 24 可包括任何类型的显示器，例如液晶显示器（LCD）、有机发光二极管（OLED）或等离子体显示器，以及相关关联的显示驱动器硬件、固件及软件。三维视频的视觉质量可不仅取决于三维源视频，而且取决于应用于三维源视频的左分量及右分量的编码过程及显示装置的特性。

根据本发明，视频编码器 20 可经配置以增强或维持由三维视频帧的左分量及右分量产生的立体三维视频的总视觉质量。在一些方面中，视频编码器 20 可经配置以应用经加权位分配方案以减小由左分量及右分量产生的立体视频中的总感知失真。经加权位分配方案可利用加权因子来预配置，所述加权因子是（例如）针对给定显示装置而从左分量及右分量中的失真对立体图像中的总感知失真的相对作用的分析导出的。加权可基于人类观看者对左分量及右分量中的失真的敏感性。给定帧级位预算及加权因子，视频编码器 20 应用位分配方案以将编码位分配给左分量及右分量以减小经编码三维视频中的总失真。被分配给相应左分量及右分量的位的数目支配此类分量的有效编码速率。位分配方案可分配例如  $\rho$  域值或量化参数（QP）值等位分配值以将编码位分配给左分量及右分量。被分配给给定视频块或帧的编码位的数目的增加会增加所述视频块或帧的有效编码速率。举例来说，用于一个分量（例如，右）的位的数目的减少连同用于另一分量（例如，左）的位的数目的同量增加在一些情况下可增加总质量度量。视频编码器 20 以可调整的位分配来支持立体视频编码以增强或维持三维视觉质量。

位分配方案可在量化参数（QP）域、 $\rho$  域或某其它位分配域中操作，以分配位分配值且进而基于质量度量来调整左三维视频分量及右三维视频分量的编码速率。举例来说，在  $\rho$  域位分配的情况下， $\rho$  参数表示视频块（例如，MB）中的非零经量化 AC 变换系数的数目。宏块是形成帧的一部分的视频块。MB 的大小可为 16 乘 16 个像素。然而，其它 MB 大小为可能的，以及例如 16×8、8×16、8×8、4×8、8×4、4×4 等的子分区大小。本文中出于说明目的而描述宏块，其中应理解，宏块或其它视频块可具有多种不同大小。 $\rho$  域中的速率控制倾向于比 QP 域中的速率控制更准确。然而，在一些情况下可使用 QP 域位分配。 $\rho$  域或 QP 域位分配基于质量度量可用于实现分别被分配给左三维视频分量及右三维视频分量的编码位的有效控制，同时遵守经指定用于适用视频帧的总帧级

位预算。在本发明中，大体上将出于实例目的而描述  $\rho$  域位分配，而不仅限于用于在三维图像的左分量与右分量之间分配编码位的特定位分配域  $\rho$ 、QP 或其它。

三维视频输出装置 14 可被实施为多种有线或无线装置中的任一者，例如经装备用于视频串流、视频广播及/或视频电话的无线移动终端或有线终端。无线移动终端的实例包括移动无线电电话（例如，所谓的相机电话或视频电话）、移动个人数字助理（PDA）、移动计算机、移动游戏装置、移动电视或装备有无线通信能力及视频编码及/或解码能力的其它移动装置。有线终端的实例包括桌上型计算机、视频电话、网络器具、机顶盒、游戏控制台、交互式电视等。类似于三维视频输出装置 14，三维视频源装置 12 可为视频串流服务器、视频广播服务器，或有线或无线终端。任一装置 12、14 可经配置以发送视频信息、接收视频信息或发送及接收视频信息。

在视频串流、视频广播或视频电话中，且尤其在通过无线通信的移动视频串流、广播或电话中，由于通常需要极低位速率，所以带宽是重要的关注问题。明确地说，通信信道 16 可能具有有限带宽，从而使得质量视频序列经由信道 16 的有效实时传输非常有挑战性。举例来说，通信信道 16 可为由于信道 16 中的物理约束或可能由于服务质量（QoS）限制、延迟约束、缓冲器状态或由通信信道 16 的提供者强加的带宽分配约束而具有有限带宽的无线通信链路。另外，信道条件可由于位置或信道业务而改变，从而导致带宽改变。

信道条件、QoS 限制、载波要求、延迟约束、缓冲器状态及其它约束可产生帧级位预算，其支配可被分配给给定三维视频帧的位的最大数目。给定特定帧级位预算及人类观看者对左分量及右分量的不对称失真响应，三维视频的在左分量与右分量之间的选择性差异编码位分配可增强或维持三维视频的视觉质量。举例来说，将较多编码位选择性分配给左通道或右通道可产生视觉质量改进。换句话说，并非将位预算均匀地分裂于左分量与右分量之间，向一个分量分配比另一分量多的位可由于人类认知色彩内插机制而增强视觉质量。此外，失真对左分量对比右分量的视觉效应可由于在左眼与右眼之间响应于失真的不对称优势而为大致不同的。

图 2 为说明用于三维视频产生、编码及呈现的不同技术的框图。一般来说，三维视频内容的源可分成三个类别：（1）立体双相机 26；（2）具有深度相机的常规单相机 28；及（3）具有二维到三维转换 32 的常规单相机 30。立体双相机 26 实时地用两个相机同时俘获立体视频，且倾向于在最终观看者当中产生最佳类三维感知。单相机及深度相机 28 的使用为较不昂贵的解决方案。然而，由单相机及深度相机 28 执行的视频俘获过程通常需要较多时间以估计每一对象的深度，且可能会不能够实时地记录视频。具有二维

到三维转换 32 的单相机 30 的使用提供另一便宜的解决方案，且可与先前俘获二维视频序列向后兼容。然而，二维到三维转换可由于不准确的深度估计而产生不合需要的假象。

如图 2 所示，立体双相机 26 分别产生左眼视图 34 及右眼视图 36（即，左分量及右分量）。单相机及深度相机 28 的组合分别产生视频场景的镶嵌式视图 38 及深度数据 40。单相机 30 产生镶嵌式视图 42，其由二维到三维转换 32 转换为镶嵌式视图 38 及深度数据 40。视频编码器 20 编码由立体双相机 26 产生的左分量及右分量 34、36、由单+深度相机 28 产生的镶嵌式视图 38 及深度数据 40 或由单相机 30 及二维到三维转换 32 产生的镶嵌式视图 38 及深度数据 40，且产生一个或一个以上类型的显示数据以驱动显示装置以用于三维视频内容的视觉呈现。

举例来说，视频编码器 20 可产生二维显示数据 44、立体显示数据 46 及/或自动立体显示数据 48。立体显示数据 46 模仿人类认知色彩内插机制且通过用例如三维眼镜或取景器等滤色装置展示不同色彩通道（即，以补色立体图形式）而形成三维视频。立体显示数据 46 的使用相对较便宜且产生大体上可接受的视觉质量。自动立体显示器倾向于提供较自然且较好质量的三维体验，但相对较昂贵。二维显示数据 44 的使用与现有的二维显示装置一起良好地工作，但可能会牺牲视觉质量。

三维视频编码器 20 可经配置以将由立体双相机 26、单及深度相机 28 或单相机 30 及二维到三维转换模块 22 产生的三维视频数据变换（即，转换）为图 2 所示的显示数据类型 44、46、48 中的任一者（其它显示数据类型）。对于经加权位分配方案，如本发明中所描述，三维视频应被转换为包含经组合以产生立体视频的左视频分量及右视频分量的立体视频。在一些实施方案中，三维视频编码器 20 可接收显示器类型指示，其指示与三维视频输出装置 14 中的三维视频显示器 24 相关联的显示器类型。举例来说，如果所指示的显示器类型为立体的，则三维视频编码器 20 可在必要时变换输入三维数据，以匹配立体显示器类型。举例来说，如果所指示的显示器类型为立体的，则三维视频编码器 20 可将镶嵌式视图数据 38 及深度数据 40 转换为包括左眼视图数据及右眼视图数据的立体显示数据 46。作为另一实例，如果所指示的显示器类型为二维的，则三维视频编码器 20 可将左眼视图数据 34 及右眼视图数据 36 转换为二维显示数据 44。

可以多种方式来获得显示器类型指示。举例来说，三维视频输出装置 14 可（例如）使用带内信令或带外信令而将显示器类型指示发送到三维视频源装置 12。在此情况下，三维视频编码器 20 直接或间接从三维视频输出装置 14 接收显示器类型指示。或者，显示器类型指示可由制造商、运营商或用户提供，或基于例如源装置的类型、目标装置的类型、源装置被使用所针对的应用及/或源装置被使用所在的区域等因素来假定。显示器

类型指示不仅可用于选择三维内容到适当立体输出格式的任何必要变换，而且可用于选择用于产生质量度量及/或应用经加权位分配方案的加权因子  $\alpha$ ，如本发明中所描述。尽管三维视频编码器 20 可产生例如图 2 中所描绘的数据 44、46 及 48 等不同类型的三维数据，但出于说明目的，本发明将大体上描述包括左分量及右分量的立体显示数据 46 的编码。

在编码及立体图像重构阶段中，可能出现若干不同类型的视觉降级。三维视频内容需要比常规二维视频多的信息以构造立体图像。举例来说，三维视频内容进一步需要深度信息或双视频源。因此，对于给定带宽，三维视频通常将需要较高的压缩比以满足适用的帧级位预算。较高的压缩比可能引入较严重的压缩假象。另外，此类假象的影响可根据用于呈现三维视频的显示技术的类型而改变。举例来说，视觉假象可能在二维显示器中比在立体显示器或自动立体显示器中更显著。另外，由立体图像构造过程引入的若干新假象（例如，梯形畸变或纸板效应）可使所感知的视觉质量降级。质量度量的产生可支持视觉质量的客观评估及编码位的智能分配以减少假象。

图 3A 为说明使用双遍技术由立体双相机 26 进行三维视频产生以用于经由立体显示器呈现的框图。在图 3A 的实例中，视频编码器 20 将三维源（输入）视频与从经编码视频重构的视频进行比较以确定经编码视频的质量，且产生用于以可接受质量编码视频的质量反馈。换句话说，在第一遍中，视频编码器 20 以现有编码参数集合产生经编码视频帧，重构所述经编码帧，且将其与源视频进行比较。在第二遍中，视频编码器 20 基于从所述比较导出的质量反馈来重新编码视频帧。举例来说，如果质量反馈指示需要改进，则视频编码器 20 可调整左分量与右分量之间的位分配。视频编码器 20 可在逐帧的基础上应用此双遍编码过程以调整位分配。

尽管三维视频可由其它类型的三维视频俘获装置产生或由计算机产生，但将出于说明目的而描述立体双相机 26 的使用。在图 3 的实例中，立体双相机 26 产生左三维视频分量及右三维视频分量，且将其存储于内容数据库 52 中。将左及右视频对记录于数据库 52 中准许视频产生者在不考虑不同显示器及压缩问题的效应的情况下制造高质量视频。视频编码器 20 可经配置以编码包括左分量及右分量的立体视频帧，且可处理任何必要变换以产生左分量及右分量。然而，三维输入视频通常将由内容数据库 52 以包含左分量及右分量的立体视频的形式提供。如果三维视频由立体双相机获得，则左分量及右分量为直接可用的。如果三维视频由具有深度信息的单相机获得，则转换为产生左分量及右分量所必要的。同样地，如果单相机获得三维视频，则转换为产生左分量及右分量所必要的。三维视频编码器 20 包括视频编码引擎 56，其使用例如 MPEG4 或 H.264

的编码技术来编码来自内容数据库 52 的三维视频内容，且应用经设计以增强由左分量及右分量产生的立体图像的感知视觉质量的经加权位分配方案。

在图 3A 的实例中，质量度量分析器 58 分析由三维视频编码引擎 56 产生的经编码视频的视觉质量，且产生视觉质量度量作为到三维视频编码引擎 56 的质量反馈。明确地说，质量度量分析器 58 可相对于原始源（输入）三维视频的左分量及右分量而在重构后分析经编码三维视频的左分量及右分量的失真。如将描述，三维视频编码引擎 56 使用质量度量反馈以根据帧级位预算及基于质量度量的位分配方案而在左分量与右分量之间分配编码位，且进而增强或维持人类观看者的感知视觉质量。明确地说，视频编码引擎 56 可基于质量度量来调整位分配的加权。因此，视频编码引擎 56 可分配编码位以基于质量度量来调整左三维视频分量及右三维视频分量的编码速率，且应用经调整的编码速率以编码左三维视频分量及右三维视频分量。视频编码器 20 将经编码的左三维视频分量及右三维视频分量经由信道 16 传输到输出装置 14，以准许用户观看具有组合式立体格式的左三维视频分量及右三维视频分量。

输出装置 14 经由信道 16 接收及解码经编码的三维视频且向人类观看者呈现经重构的三维视频。在图 3 的实例中，输出装置 14 可包括由立体眼镜 60、立体取景器或适合于呈现立体三维视图的任何其它设备辅助的立体视频显示器。立体眼镜 60 可采取多种形式，例如双色补色立体图（红色-蓝色/青色）眼镜、液晶快门眼镜、偏光式透镜眼镜或头戴式显示器。在其它方面中，输出装置 14 可包括二维显示器或自动立体显示器。在每一情况下，输出装置 14 可将特定显示器类型、一般显示器类别或其它有用信息传送到三维视频编码引擎 56 以用于调整位分配方案，以将其定制用于与输出装置 14 相关联的显示器类型。

在图 3A 的实例中，眼镜 60 为具有右（蓝色）分量透镜及左（红色）分量透镜的滤色眼镜。滤色眼镜 60 呈现合理的成本及质量。如此项技术中众所周知的，眼镜 60 的基本操作原理在于，所述眼镜将色彩分离到独立通道中且向左眼及右眼分别呈现红色及蓝色。所得立体图像由分离式色彩通道及人类固有视觉机制形成。为了构造用于具有滤色眼镜 60 的立体显示器的适当图像文件，从左分量图像提取红色且从右分量图像提取蓝色及绿色，其被称为补色立体图。人类观看者在视觉上整合左分量图像及右分量图像以在智力上产生最终立体三维彩色图像。

显示器类型（包括显示器及例如眼镜的任何外部硬件）可对观看者可见的假象的程度及特性具有影响，且因此可用于影响位分配方案以支持较好的视觉质量。如将描述，在一些方面中，三维视频编码引擎 56 可基于显示器类型来选择加权因子以控制三维视

频的左分量与右分量之间的位分配加权。总三维视频中所呈现的假象可由左视图或右视图更占优势地产生。因此，朝向对总视觉质量具有较大效应的视图的经加权编码位分配可能为合乎需要的。加权可以位分配、质量度量或两者的组合来表达。经加权位分配协同考虑到由左分量与右分量两者作用的失真的总三维视觉质量度量可对增强或维持视觉质量为有效的。

贯穿本发明，将不同特征描述为模块、块或组件希望突出视频编码器 20 的不同功能方面，且未必暗示此类模块必须由单独硬件及/或软件组件来实现。而是，与例如三维视频编码引擎 56 及质量度量分析器 58 等一个或一个以上模块相关联的功能性可集成在共同或单独硬件及/或软件组件内。在一些情况下，此类特征可由包括经配置以致使计算机执行归因于此类特征的功能的代码的共同或单独软件或软件模块来实现。

图 3B 为说明使用优化技术由立体双相机进行三维视频产生以用于经由立体显示器呈现的框图。图 3B 大致与图 3A 一致，但说明基于优化的位分配技术。三维视频编码器 20 接收及编码包括左分量及右分量的立体视频帧，且可处理任何必要变换以产生左分量及右分量。然而，如在图 3A 的实例中，三维输入视频通常可由内容数据库 52 以包含第一分量及第二分量的立体视频的形式提供。在图 3B 的实例中，视频编码引擎 56 使用经设计以增强由左分量及右分量产生的立体图像的感知视觉质量的基于优化的经加权位分配方案来编码来自内容数据库 52 的三维视频内容。

视频编码引擎 56 接收来自内容数据库 52 的三维视频输入视频及（例如）来自输出装置 14 的显示器类型指示。通过使用显示器类型指示，三维视频编码引擎 56 选择加权因子且产生三维输入视频的左分量与右分量之间的经加权编码位分配。通过使用经加权位分配，三维视频编码引擎 56 编码三维输入视频且将经编码视频（例如）经由信道 16 传输到输出装置 14。给定三维视频的左分量及右分量的失真作用，基于优化的位分配方案可经配置以最小化或至少减小经编码三维视频的总感知视觉失真。明确地说，如果适用的话，则加权因子可经选择以补偿由于人类左眼及右眼的不对称失真响应引起的左分量及右分量的不同相对作用。

如在图 3A 的实例中，输出装置 14 经由信道 16 接收及解码经编码的三维视频且向人类观看者呈现经重构的三维视频。此外，如在图 3A 中，图 3B 的输出装置 14 可包括由立体眼镜 60、立体取景器或适合于呈现立体三维视图的任何其它设备辅助的立体视频显示器。如果发现显示器类型（包括显示器及例如眼镜的任何外部硬件）对观看者可见的假象的程度及特性具有显著影响，则三维视频编码引擎 56 可选择适合于特定显示装置或显示装置特性的加权因子。在任何情况下，在选择适当加权因子后，视频编码引

引擎 56 便提供朝向对总视觉质量具有较大效应的视图（左或右）的经加权编码位分配，其可能对于增强或维持视觉质量为合乎需要的。

图 4 为说明用于使用基于优化的位分配方案而在三维视频编码器 20 内产生经编码三维视频的视频编码引擎 56 的框图。如图 4 所示，视频编码引擎 56 可包括左/右分量权数产生器 62、三维  $\rho$  域位分配模块 64、帧级速率控制器 66、 $\rho$  到 QP 映射器 68 及视频编码单元 70。视频编码单元 70 连同左/右分量权数产生器 62、三维  $\rho$  域位分配模块 64、帧级速率控制器 66、 $\rho$  到 QP 映射器 68 形成总视频编码器 20 的部分，且负责根据例如 MPEG-4 或 H.264 的编码技术来应用编码位以产生经编码视频。如图 4 所示，视频编码单元 70 接收三维输入视频且编码所述视频以产生经编码三维视频。左-右分量权数产生器 62 接收显示器类型指示且产生用于由三维  $\rho$  域位分配模块 64 及编码单元 70 使用的加权因子  $\alpha$ 。位分配模块 64 基于加权因子而将位分配给三维视频帧的左分量及右分量。

此外，将不同特征描述为模块、块或组件希望突出不同功能方面且未必暗示此类模块必须由单独硬件及/或软件组件来实现。在一些情况下，视频编码引擎 56 的各种方面可被组合、合作或由共享硬件及/或软件来实现。因此，图 4 中所描绘的各种组件可以包含归于每一模块的功能性的多种方式来形成。在任何情况下，视频编码器 20 的各种组件可以硬件、软件、固件或其组合来实现。举例来说，此类组件可作为在一个或一个以上微处理器或数字信号处理器（DSP）、一个或一个以上专用集成电路（ASIC）、一个或一个以上现场可编程门阵列（FPGA）或其它等效集成或离散逻辑电路上执行的软件过程来操作。

在图 4 的实例中，左/右分量权数产生器 62 接收指示与输出装置 14 或输出装置群组相关联的显示器类型的输入。如先前所提及，显示器类型输入可为通过带内或带外信令从输出装置 14 传送的实际显示器类型输入。或者，显示器类型输入可由运营商、制造商或用户指定，或基于例如源装置的类型、输出装置的类型、源装置被使用所针对的应用及/或源装置或输出装置被使用所在的区域等因素来假定。基于显示器类型，左/右分量权数产生器 62 产生加权因子  $\alpha$ ，其出于失真分析且最终位分配的目的而指定朝向左分量或右分量的加权偏好。举例来说，左/右分量权数产生器 62 可从表中选择加权因子或计算加权因子。

视频编码单元 70 接收三维输入视频（即，原始源三维视频）的左分量及右分量。所述左分量及右分量可作为视频帧而从三维视频序列的左分量及右分量接收。通过使用加权因子  $\alpha$  以及帧级位预算  $R_{\text{BUDGET}}$ ，三维  $\rho$  域位分配模块 64 针对三维视频帧的左分量及右分量内的视频块（例如，宏块）产生位分配值。位分配值可为  $\rho$  域值。



加权因子  $\alpha$  可为优势因子，其表达左分量及右分量对三维视频的总视觉质量的相对影响。在此实例中，如果红色绿色蓝色（RGB）图像是从补色立体图图像构造且每一色彩通道是从不同相机获得，则左分量图像及右分量图像中的质量等级对观看者的总视觉质量具有不同作用。举例来说，给定人眼的感知质量，即使相同失真存在于左分量图像与右分量图像两者中，所述图像中的一者仍可能由于一个通道优于另一通道的不对称优势而对视觉质量具有显著不同的影响。因此，如果第一分量（左或右）中的失真与第二分量中的失真相比对总视觉质量具有更显著影响，则三维  $\rho$  域分配模块 64 可经配置以通过分配  $\rho$  域值以使得将额外位分配给第一分量以增强视觉质量来解决差异。

加权因子  $\alpha$  量化特定分量（左或右）影响总视觉质量的程度。如果左分量上的失真较强烈地影响视觉质量，则  $\alpha$  将较高。另一方面，如果右分量上的失真较强烈地影响视觉质量，则  $\alpha$  将较低。加权因子  $\alpha$  可基于准许识别左分量及右分量对立体图像的总感知视觉质量的相对作用的质量度量来确定。至少部分地基于加权因子  $\alpha$ ，三维  $\rho$  域位分配模块 64 偏置正由视频编码器 20 编码的视频帧的左分量与右分量之间的  $\rho$  域位分配。分配可针对视频帧内的个别视频块（例如，宏块（MB））或针对帧内的所有视频块来指定。分配可在逐块的基础上或在逐帧的基础上更新。

$\rho$  域位分配模块 64 接收来自帧级速率控制器 66 的帧级速率预算  $R_{\text{BUDGET}}$  及来自视频编码单元 70 的用于经编码 MB 的标准偏差  $\sigma$ 。标准偏差  $\sigma$  可为在运动估计之后所获得的实际残余的标准偏差，且可为来自先前帧的所存储残余统计。标准偏差可由左分量及右分量的标准偏差分量表示。举例来说，标准偏差可被表达为左分量标准偏差  $\sigma_{li}$  及右分量标准偏差  $\sigma_{ri}$ 。通过使用加权因子  $\alpha$  及标准偏差  $\sigma_{li}$  及  $\sigma_{ri}$ ，出于左分量与右分量之间的优先  $\rho$  域位分配的目的， $\rho$  域位分配模块 64 区分左分量 MB 与右分量 MB。 $\rho$  域位分配模块 64 针对在左分量及右分量的每一帧中的每一 MB 产生  $\rho$  (rho) 参数。 $\rho$  参数表示 MB 中的非零经量化 AC 变换系数的数目。 $\rho$  域中的速率控制倾向于比 QP 域中的速率控制更准确。

帧级速率控制器 66 产生对视频序列内的个别帧的位分配。明确地说，帧级速率控制器 54 产生指示可用于编码当前帧内的所有 MB 的位的数目的值  $R_{\text{BUDGET}}$ ，包括帧的左分量及右分量。帧级速率控制器 66 可为一般速率控制器，其基于由信道条件强加的带宽要求、载波要求、服务质量（QoS）限制、延迟约束、缓冲器状态或者其它要求或条件而将位分配给帧。通过使用帧级位预算  $R_{\text{BUDGET}}$ 、标准偏差  $\sigma$  及加权因子  $\alpha$ ， $\rho$  域位分配模块 64 针对左分量及右分量中的块产生位分配值。明确地说，在此实例中， $\rho$  域位分配模块 64 产生  $\rho$  域输出，其将用于每一左分量或右分量 MB 的  $\rho$  预算分别指示为  $\rho_{li}$  ( $\rho$

左图像)及 $\rho_{ri}$ ( $\rho$ 右图像)。 $\rho$ 域输出应用于 $\rho$ 到QP映射器68,其将 $\rho_{li}$ 及 $\rho_{ri}$ 值映射到用于每一MB的对应QP值 $QP_{li}$ 及 $QP_{ri}$ 。通过分别针对帧的左分量及右分量内的MB使用QP值 $QP_{li}$ 及 $QP_{ri}$ ,视频编码单元70编码三维输入视频以产生经编码的三维视频。

在与编码相关联的压缩及立体图像重构阶段中,可引入由于不同原因引起的若干降级。因为三维视频内容通常需要比常规二维信息(例如,图像及深度信息)或来自双视频源的右信息及左信息多的信息,所以可应用较高的压缩比以满足适用的位预算。因此,可针对给定带宽而由三维编码引入较严重的压缩假象。另外,其它假象可由立体图像构造过程引起,其还使所感知的质量降级。 $\rho$ 域位分配模块64准许使由伴随三维视频编码的压缩产生的假象减少。以下更详细地描述由 $\rho$ 域位分配模块64用于产生经加权位分配的实例等式。

出于产生质量度量的目的,可通过针对左分量将原始输入视频 $f_l$ 与经重构视频 $\tilde{f}_l$ 进行比较且针对右分量将原始输入视频 $f_r$ 与经重构视频 $\tilde{f}_r$ 进行比较来分析失真。由于经压缩立体图像的不对称特性,通过均匀地平均来自左图像与右图像两者的峰值信噪比(PSNR)来估计质量可为误导的。因此,可修改立体图像的PSNR,以便更准确地预测总立体图像质量。明确地说,加权因子 $\alpha$ 可用于表示由左分量及右分量产生的不同影响,且又可用于影响 $\rho$ 域位分配模块64所作的位分配。

以下陈述示范性优化技术的论述,所述技术可由视频编码引擎56的示范性组件且尤其是 $\rho$ 域位分配模块64执行,以便以增强或维持感知视觉质量的方式将编码位分配给三维图像的左分量及右分量。在以下论述中,值 $D_L$ 及 $D_R$ 分别表示立体图像的左分量及右分量的正规化每像素失真。加权因子 $\alpha$ 指示左眼对比右眼的感知重要性因子,且可被表示为在0与1之间的实数。

加权因子 $\alpha$ 的值可基于显示器类型或显示器类型特性来选择,且更明确地说,基于针对特定显示器类型或显示器类型特性的对左分量及右分量的相对失真作用的分析来选择。值得注意的是,如果显示装置为仅可支持二维显示数据(例如,类似图2的二维显示数据44)的二维显示器,则可将加权因子 $\alpha$ 设定为一或零,使得左分量及右分量中的仅一者接收所有编码位分配。换句话说,如果显示装置不能支持三维立体视频,则位分配可恢复到仅一个通道接收所有编码位的情形,即,通过将加权因子 $\alpha$ 设定为零或一。否则,如果输出装置14具有支持立体视频的显示器,则加权因子 $\alpha$ 的值将为在零与一之间的值,如由左分量及右分量对立体图像的总感知视觉质量的相对失真作用的分析所确定。

如果所提及的变量之间的关系被简化为线性函数，则质量度量可被表达为视频序列的总失真，其可由下式表示：

$$D = \alpha D_L(f_l, \tilde{f}_l) + (1 - \alpha) D_R(f_r, \tilde{f}_r) \quad (1)$$

其中  $f_l$  及  $\tilde{f}_l$  为原始及经重构左视图，且  $f_r$  及  $\tilde{f}_r$  为原始及经重构右视图。经重构视图是从经编码视频分量重构以用于与原始源视频进行比较。质量度量可通过以下方式来形成：用  $\alpha$  加权左三维视频分量的所估计质量  $D_L$  且用  $1 - \alpha$  加权右三维视频分量的所估计质量  $D_R$ ，或反之亦然，且基于经加权质量来形成质量度量。因子  $\alpha$  可为通常在 0 与 1 之间的数字。

等式 (1) 中的线性关系对于人类视觉感知观点是否有效的问题可由实验来验证。另外， $\alpha$  的适当值可由回归测试来确定，使得其针对特定输出显示装置或特定类别的显示装置特性而与大多数主观人类视觉感知紧密地对准。 $\alpha$  的适当值对于不同输出显示装置、不同种类的显示装置、不同观看应用或不同用户可为不同的。因此，左/右分量权数产生器 62 可基于用以观看左及右三维视频分量的显示装置类型来调整加权。举例来说，左/右分量权数产生器 62 可针对不同显示器类型指示而计算或选择  $\alpha$  的不同值。等式 (1) 可在立体视频编码及位分配中起重要作用，因为左/右经加权质量度量可作为输入而有助于导出支持总立体三维视频的所需主观感知质量的最佳编码参数选择及位分配方案。

考虑到立体视觉质量中的左分量及右分量的不同作用的质量度量（例如，如在以上等式 (1) 中所表达）可在通常评估视觉质量中为有用的。然而，此质量度量可尤其在视频编码过程中为有用的，且更明确地说，可对于三维  $\rho$  域位分配模块 64 的位分配为有用的。如果由帧级速率控制器 66 产生的  $R_{budget}$  表示给定双左分量及右分量帧的总位预算，且  $R$  为用于编码帧的位速率，则由位分配模块 64 所解决的位分配问题可由下式表示：

$$\begin{aligned} & \text{最小化 } \alpha D_L(f_l, \tilde{f}_l) + (1 - \alpha) D_R(f_r, \tilde{f}_r), \\ & \text{使得 } R \leq R_{budget}. \end{aligned} \quad (2)$$

此优化问题 (2) 可由拉格朗日松弛 (Lagrangian relaxation) 及动态编程来解决，例如，如 H·王 (H. Wang)、G·M·休斯特 (G. M. Schuster)、A·K·卡特珊格洛斯 (A. K.

Katsaggelos)的“用于基于对象的视频编码的速率失真最佳的位分配方案(Rate-distortion optimal bit allocation scheme for object-based video coding)”(IEEE 视频半导体电路系统技术(IEEE Trans. Circuits System for Video Technology),第15卷,第9号,第1113到1123页,2005年9月)中所描述。然而,此方法的计算复杂性通常将高于现有实时系统可承受的复杂性。因此,低复杂性的近最佳解决方案为优选的。

在立体视频编码中,使  $N$  表示帧中的宏块的数目,  $\{\rho_{li}\}$ 、 $\{\sigma_{li}\}$ 、 $\{R_{li}\}$  及  $\{D_{li}\}$  表示左视图中的第  $i$  个宏块的  $\rho$  集合、标准偏差集合、速率集合及失真集合(误差平方和),且  $\{\rho_{ri}\}$ 、 $\{\sigma_{ri}\}$ 、 $\{R_{ri}\}$  及  $\{D_{ri}\}$  表示右视图中的第  $i$  个宏块的  $\rho$  集合、标准偏差集合、速率集合及失真集合(误差平方和)。失真  $D_{li}$  表示经编码左三维视频分量的所估计质量,且失真  $D_{ri}$  表示经编码右三维视频分量的所估计质量。

基于左失真估计及右失真估计以及加权因子  $\alpha$ , 可形成由左分量及右分量形成的立体视频帧的总质量度量。举例来说, 立体视频帧的经加权失真  $D$  可由下式表示:

$$D = \sum_{i=1}^N (D_{li} + D_{ri}) = [\alpha D_L(f_l, \tilde{f}_l) + (1 - \alpha) D_R(f_r, \tilde{f}_r)] * 225^2 * 384 \quad (3)$$

因此, 给定等式(3), 为了支持  $\rho$  域位分配模块 64 所作的位分配, 问题(2)可被重写为:

$$\text{最小化 } D, \text{ 使得 } R \leq R_{budget} \quad (4)$$

问题(4)可使用基于建模的位分配方法来解决。如伯杰(Berger)的速率失真理论(Rate Distortion Theory)(新泽西州英格伍德克里夫: 学徒店堂(Englewood Cliffs, NJ:Prentice Hall), 1984年)中所描述, 自然图像中的 AC 系数的分布可最佳由拉普拉斯分布(Laplacian distribution)  $P(x) = \frac{\eta}{2} e^{-\eta|x|}$  来近似。因此, 根据 Z·何(Z. He)及 S·K·麦卓(S. K. Mitra)的“用于 DCT 视频编码的线性源模型及统一速率控制算法(A linear source model and a unified rate control algorithm for DCT video coding)”(IEEE 视频半导体电路系统技术(IEEE Trans. Circuits System for Video Technology), 第12卷, 第11号, 第970到982页, 2002年11月), 第  $i$  个宏块的速率及失真可在以下等式(5)及(6)中作为  $\rho$  的函数来建模, 例如, 如下:

$$R_{li} = A\rho_{li} + B, \quad (5)$$

及

$$R_{ri} = A\rho_{ri} + B, \quad (6)$$

其中 A 及 B 为恒定建模参数，且 A 可被视为编码非零系数所需要的位的平均数目，且 B 可被视为归因于非纹理信息的位。在以上等式 (5) 及 (6) 中，可针对左视图及右视图使用 B 的不同值。然而，为了简化起见，可基于两个视图非常类似的假定而针对左视图与右视图两者使用 B 的相同值。然而，重要的是应强调，简单性不会影响导出的一般性，因为值 B 在等式中为恒定值。

左分量及右分量的所估计失真可分别由以下等式 (7) 及 (8) 表示：

$$D_{li} = 384\sigma_{li}^2 e^{-\theta\rho_{li}/384}, \quad (7)$$

及

$$D_{ri} = 384\sigma_{ri}^2 e^{-\theta\rho_{ri}/384}, \quad (8)$$

其中  $\theta$  为未知常数，但可根据历史统计来估计，例如，作为  $\rho$  位速率曲线的斜率，其中位速率  $=\theta*\rho$ 。在等式 (7) 及 (8) 中，因为假定存在可用于从任何选定  $\rho_i$  产生适当量化器的足够准确的  $\rho$ -QP 表，所以值  $\rho_{li}$  及  $\rho_{ri}$  经优化而非量化器。

如本发明中所描述，等式 (5) 到 (8) 可用于表示速率失真表示集合，其被表达为编码参数 ( $\rho$ ) 的函数。这些等式可用于导出用于立体视频编码的  $\rho$  域中的最佳位分配策略。出于说明目的而呈现本发明中所描述的  $\rho$  域位分配策略。然而， $\rho$  域或量化域 (QP) 中的其它速率失真函数也可用于使用给定帧位预算及左分量与右分量之间的加权来导出位分配策略。

一般来说，等式 (4) 可使用拉格朗日松弛来解决，其中受约束问题如下被转换为无约束问题：

$$\text{最小化}_{\rho_i} J_{\lambda} = \lambda R + D = \sum_{i=1}^N [\lambda(R_{li} + R_{ri}) + \alpha D_{li} + (1-\alpha)D_{ri}] \quad (9)$$

$$= \sum_{i=1}^N \left[ \lambda(A(\rho_{li} + \rho_{ri}) + 2B) + 384(\alpha\sigma_{li}^2 e^{-\theta\rho_{li}/384} + (1-\alpha)\sigma_{ri}^2 e^{-\theta\rho_{ri}/384}) \right]$$

其中  $\lambda^*$  (等式 (9) 中的经优化  $\lambda$ ) 为实现  $\sum_{i=1}^N R_{li} + R_{ri} = R_{budget}$  的解。通过在等式 (9) 中将偏导数设定为零, 可针对经优化  $\rho_i$  获得以下表达式:

$$\frac{\partial J_{\lambda}}{\partial \rho_{li}} = 0, \text{ 且} \quad (10)$$

$$\frac{\partial J_{\lambda}}{\partial \rho_{ri}} = 0 \quad (11)$$

其为

$$\lambda A - \alpha\theta\sigma_{li}^2 e^{-\theta\rho_{li}/384} = 0, \quad (12)$$

因此, 由此可见:

$$e^{-\theta\rho_{li}/384} = \frac{\lambda A}{\alpha\theta\sigma_{li}^2} \quad (13)$$

且:

$$\rho_{li} = \frac{384}{\theta} \left[ \ln(\theta\sigma_{li}^2) - \ln(\alpha\lambda A) \right], \quad (14)$$

且类似地:

$$\rho_{ri} = \frac{384}{\theta} \left[ \ln(\theta\sigma_{ri}^2) - \ln((1-\alpha)\lambda A) \right]. \quad (15)$$

另一方面:

$$\begin{aligned}
R_{budget} &= \sum_{i=1}^N (R_{li} + R_{ri}), \\
&= \frac{384A}{\theta} \sum_{i=1}^N [\ln(\theta\sigma_{li}^2) + \ln(\theta\sigma_{ri}^2) - \ln(\alpha\lambda A) - \ln((1-\alpha)\lambda A)] + 2NB
\end{aligned} \tag{16}$$

因此:

$$\ln(\lambda A) = \frac{1}{2N} \sum_{i=1}^N [\ln(\theta\sigma_{li}^2) + \ln(\theta\sigma_{ri}^2)] - \frac{\theta}{768NA} (R_{budget} - NB) - \frac{\ln\alpha + \ln(1-\alpha)}{2}. \tag{17}$$

根据等式 (14) 到 (17), 位分配模块 64 可经配置以如下获得左分量及右分量的位分配:

$$\begin{aligned}
\rho_{li} &= \frac{384}{\theta} \left[ \ln(\theta\sigma_{li}^2) - \frac{1}{2N} \sum_{i=1}^N \ln(\theta\sigma_{li}^2) + \frac{\theta}{768NA} (R_{budget} - NB) - \frac{\ln\alpha - \ln(1-\alpha)}{2} \right] \\
&= \frac{R_{budget} - NB}{2NA} + \frac{384}{\theta} \left[ \ln(\theta\sigma_{li}^2) - \frac{\sum_{i=1}^N \ln(\theta\sigma_{li}^2)}{2N} - \frac{\ln\alpha - \ln(1-\alpha)}{2} \right].
\end{aligned} \tag{18}$$

且类似地:

$$\rho_{ri} = \frac{R_{budget} - NB}{2NA} + \frac{384}{\theta} \left[ \ln(\theta\sigma_{ri}^2) - \frac{\sum_{i=1}^N \ln(\theta\sigma_{ri}^2)}{2N} + \frac{\ln\alpha + \ln(1-\alpha)}{2} \right]. \tag{19}$$

因此, 在此实例中, 位分配模块 64 可使用以上等式 (18) 及 (19) 来分别产生左分量及右分量的  $\rho$  值  $\rho_{li}$  及  $\rho_{ri}$ 。值得注意的是, 给定加权因子  $\alpha$  以及左分量及右分量中的不同失真程度, 被分配给左三维视频分量及右三维视频分量的编码位的量可为不同的。换句话说, 如果由位分配模块 64 所指定的  $\rho$  值为不同的, 则左分量及右分量的有效编码速率为不同的。位分配模块 64 可经配置以应用等式 (18) 及 (19) 以产生左分量及右分量  $\rho$  值  $\rho_{li}$  及  $\rho_{ri}$ 。一旦针对三维立体视频帧的左分量及右分量中的宏块 (MB) 确定

了  $\rho_{li}$  及  $\rho_{ri}$  的值, 则可相应地基于由  $\rho$  到 QP 映射器 68 提供的历史  $\rho$ -QP 表来决定对应量化参数 (QP)。明确地说,  $\rho$  到 QP 映射器 68 将  $\rho$  值  $\rho_{li}$  及  $\rho_{ri}$  分别映射到每一 MB 的对应 QP 值  $QP_{li}$  及  $QP_{ri}$ 。映射器 68 可如上所描述使用查找表, 以将  $\rho$  值转换为对应 QP 值以供视频编码单元 70 使用。或者, 映射器 68 可使用数学函数。

对使用双相机记录且使用红色-蓝色眼镜观看的众多立体图像及视频所进行的实验验证了本发明中所描述的质量度量及位分配方案的功效。所述实验包括三个主要组成部分: (1) 对等式 (1) 中所陈述的线性关系满足“所预期”正常人类感知的验证; (2) 由回归测试所作的对等式 (1) 中的适当  $\alpha$  加权因子的确定; 及 (3) 通过在不考虑来自两个视图的组合式感知质量测量的情况下将最佳位分配方案与一般图像及视频编码方法进行比较所作的对最佳位分配方案的评估。

对于对等式 (1) 中的线性关系的验证, 主观测试指示线性关系对应于普通人类感知。随着在 0 与 1 之间调整  $\alpha$  加权因子, 来自等式 (1) 的客观质量评分与视觉所感知质量之间的对应性显著地改变。评估由具有不同质量等级的经压缩左图像及右图像的组合形成的立体图像。所使用的压缩方法为 JPEG2000。质量等级控制压缩效率及相关联的输出质量, 其从 1 变化到 100。较高值指示较好质量及较低压缩比。将左分量图像及右分量图像的质量等级分别设定为 5 及 100, 且形成立体图像的不同质量组合。

将预期由左分量图像及右分量图像形成的立体图像的人类所感知质量为非常类似的, 因为相同质量损失被引入于具有几乎相同内容但在不同眼睛 (左及右) 中的两个图像中。然而, 主观观测展示了, 人眼对左图像与右图像之间的总质量降级具有非常不同的敏感性。明确地说, 当将左图像质量设定为 5 且将右图像质量设定为 100 时, 立体图像质量比当将左图像质量设定为 100 且将右图像质量设定为 5 时好得多。明确地说, 当将左图像质量设定为 100 且将右图像质量设定为 5 时, 在立体图像中更可观测到失真。因为人眼对响应于来自左眼及右眼的失真具有不对称优势, 所以应用于左图像及右图像上的相同失真对于最终观看者可造成显著不同的影响。因此, 可鉴于不对称优势而以适当加权因子  $\alpha$  来应用左分量与右分量之间的差异位分配, 以支持所得立体三维图像中的较好的总视觉质量。

因子  $\alpha$  可被认为是人类观看者对来自左图像及右图像的失真的敏感性, 其中较高  $\alpha$  意味着人类对失真具有较小容许度。因此, 加权是基于人类观看者对左分量及右分量中的失真的敏感性。已执行了对由具有不同失真组合的左图像及右图像形成的立体图像的主观实验。输出为平均意见评分 (MOS), 其在从 1 到 5 的范围内, 其中较高值意味着较好质量。为了确保测试的高可靠性, 针对左图像与右图像两者仅使用质量等级 5、15、



20、30、50 及 100 以避免使测试者疲劳。在此情况下，针对每一场景产生立体图像的  $6 \times 6 = 36$  个不同组合。可用感知失真模型使用参数拟合方法来针对特定三维视频源及输出装置确定“适当” $\alpha$ 值：

$$MOS = \alpha [r_l D_L(f_l, \tilde{f}_l) + \phi_l] + (1 - \alpha) [r_r D_R(f_r, \tilde{f}_r) + \phi_r],$$

其中  $r_l$ 、 $\phi_l$ 、 $r_r$  及  $\phi_r$  为用以正规化 1 到 5 内的失真输出以与 MOS 数据拟合的因子。用最小平方法通过回归测试来获得所述参数。此外，可根据不同输出装置来调整  $\alpha$  值。因此，如先前所描述，显示器类型指示可用于触发对特定  $\alpha$  值的选择或计算。举例来说， $\alpha$  值的预定义表可用于响应于显示器类型来选择适当加权因子 ( $\alpha$ )。

对于双相机立体视频俘获以及依赖于显示屏及红色-蓝色立体眼镜的输出装置，已发现近似 0.3 的  $\alpha$  值可对产生合乎需要的视觉结果为尤其有效的。此结果大体上与红色、绿色及蓝色分量对像素强度的作用（如以下等式中所展示）一致：

$$\text{亮度} = 0.257R + 0.504G + 0.098B + 16, \quad (20)$$

且为合理的，因为左透镜（红色透镜）将大部分红光引导到左眼中且右眼镜（蓝色透镜）将大部分绿光及蓝光引导到右眼中，而人眼对来自右眼的绿光较敏感且对来自左眼的光较不敏感。眼睛倾向于对来自右视图的失真比对来自左视图的失真更敏感。在所述意义上，对于来自左视图的失真应给予  $\alpha$  值选择较低权数，其对应于等式 (1) 的基本假定。换句话说，等式 (1) 中的  $\alpha$  加权因子通常应经选择以使得出于产生质量度量且执行左分量与右分量之间的差异位分配的目的，给予左分量中的失真比右分量中的失真小的权数。

针对如本文中所描述由感知质量度量引导的位分配方案对比不依赖于感知质量度量的位分配方案考虑多个结果。基于所预期主观重要性来编码左分量图像及右分量图像可通过花费相同量的编码位但在左分量与右分量之间分配不同量的位来大致改进质量。当与向左分量分配比右分量多的编码位相比时，向右分量分配比左分量多的编码位导致立体图像中的较高总视觉质量。依据千位的数目，左分量及右分量的帧大小为近似相等的，但主观视觉质量影响为完全可观测的。因此，如本发明中所描述的质量度量及位分配方案可对增强从左分量图像及右分量图像产生的立体图像的总视觉质量为有效的。

图 5 为说明三维视频质量度量的产生及基于三维视频质量度量的编码位分配的流程

图。如图 5 所示，视频编码器 20 获得三维视频数据 (72) (例如，包括左分量及右分量)，产生左分量与右分量之间的加权 (74)，基于所述加权来调整左分量图像与右分量图像之间的编码位分配 (76)，且使用所述位分配来编码三维视频数据 (78)。位分配可在逐帧或逐块的基础上更新。以上所描述的操作可由视频编码器 20 的各种组成组件或模块来执行。举例来说，位分配模块 64 可调整三维视频的左分量与右分量之间的位分配。视频编码单元 70 基于所述位分配来编码左分量及右分量以产生三维经编码视频。

图 6 为说明基于三维视频质量度量而将编码位分配给三维视频分量的另一流程图。如图 6 所示，视频编码器 20 获得三维视频数据 (80)。在此实例中，左/右分量权数产生器 62 获得左/右加权因子  $\alpha$  (82)，例如，作为默认值、用户指定值、基于显示器类型指示所确定的值。显示器类型指示可由输出装置 14 提供，由制造商、运营商或用户指示，或基于例如源装置的类型、目标装置的类型、源装置被使用所针对的应用及/或源装置或输出装置被使用所在的区域等因素来假定。位分配模块 64 从帧级速率控制器 66 获得帧级位预算 (84)。位分配模块 64 基于左/右加权因子  $\alpha$  及位预算来调整左分量与右分量之间的编码位分配 (86)。编码过程接着应用经更新的编码位分配。明确地说，编码单元 70 使用位分配以 (例如) 根据例如 MPEG4 或 H.264 编码的视频编码方法来编码三维视频数据 (88)。

图 7 为说明在  $\rho$  域中将编码位分配给三维视频分量的流程图。 $\rho$  域中的位分配出于准确性目的而为合乎需要的。然而，如果需要的话，可完全在 QP 域内执行位分配。对于  $\rho$  域位分配，视频编码器 20 获得三维视频数据 (92)，获得左/右加权因子  $\alpha$  (94)，且获得帧级位预算 (96)。此外，所述操作中的一些或所有操作可由视频编码器 20 的组成组件或模块来执行。在获得帧级位预算 (96) 后，位分配模块 64 便使用加权因子  $\alpha$  及位预算来产生三维视频的左分量与右分量之间的  $\rho$  域位分配 (98)。 $\rho$  到 QP 映射器 68 接着将  $\rho$  值映射到对应 QP 值 (100)，且视频编码单元 70 应用 QP 值以编码三维视频 (102)。编码过程根据对位分配的基于帧或基于块的调整而继续进行。

在一些方面中，视频编码器 20 可支持基于对输出显示器类型的指示来将三维视频数据变换到不同立体格式。举例来说，视频编码器 20 可获得三维视频数据且确定与输出装置 14 相关联的显示器类型。基于显示器类型，视频编码器 20 确定任何变换 (例如，从一种立体格式到另一立体格式) 是否为必要的。如果变换为必要的，则视频编码器 20 基于所指示的显示器类型来选择适当立体格式，且将三维视频数据改变为选定立体格式。

视频编码器 20 接着可编码新三维视频数据。视频可根据例如 MPEG4 或 H.264 编码

的视频编码方法来编码。视频变换可准许三维视频内容依据经指定以接收三维视频内容的输出装置 14 的特性而改变为各种立体格式。举例来说，如果三维视频内容作为单相机信息加上深度信息而记录于内容数据库 52 中，则视频变换可在发现输出装置 14 包括立体显示器的情况下将三维视频内容转换为适合于经由立体显示器而呈现的右分量及左分量。类似地，如果输出装置 14 包括二维显示器且三维视频内容作为立体视频内容而记录于数据库 52 中，则变换可将立体数据转换为适合于二维显示器的二维数据。一般来说，变换可依赖于对显示器类型的指示以确定变换是否为必要的，且如果为必要的，则确定应选择哪一立体格式用于变换。如本发明中所描述，对显示器类型的指示可由输出装置 14 提供，或由运营商、制造商、用户、应用类型、区域、装置类型等指示。

本发明描述用于三维视频内容压缩及再现的框架，其可适用于例如装备有三维视频能力的下一代移动电话等广泛范围的装置。通过使用考虑三维视频的左分量及右分量的不同作用的感知质量度量以及基于质量度量的位分配方案，可使用不同显示格式及显示装置而将类似立体视频内容提供给不同观看者。三维源视频可以不同方式来俘获或再现，但存储于共同数据库中，使得视频产生者可在未必解决不同显示器及与编码相关联的压缩的效应的情况下产生高质量视频。而是，可提供视频编码器以执行将源视频变换到与经指定以接收视频的输出装置相关联的适当立体格式的任何必要变换。

另外或替代地，视频编码器可经装备以产生质量度量且调整视频的左分量与右分量之间的编码位分配以解决由假象及显示特性所造成的失真。明确地说，如本发明中所描述的视频解码器可动态地调整压缩比（经由位分配）及所预期观看者的所感知视频质量以满足位预算，同时实现可接受的视觉感知质量。为了平衡压缩比与感知质量，可如本发明中所界定而使用用于立体图像的质量度量。在由度量来引导的情况下，可应用由位分配方案体现的速率控制机制以控制编码位及由压缩造成的质量失真。

本文中所描述的技术可以硬件、软件、固件或其任何组合来实施。如果以软件来实施，则所述技术可至少部分地由计算机可读媒体上的一个或一个以上所存储或所传输的指令或代码来实现。计算机可读媒体可包括计算机存储媒体、通信媒体或两者，且可包括有助于将计算机程序从一个位置传送到另一位置的任何媒体。存储媒体可为可由计算机存取的任何可用媒体。

以实例方式而非限制方式，此类计算机可读媒体可包含例如 RAM 的数据存储媒体，例如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）、ROM、电可擦除可编程只读存储器（EEPROM）、EEPROM、快闪存储器、CD-ROM 或其它光盘存储装置、磁盘存储装置或其它磁性存储装置，或可用于以

指令或数据结构的形式来携带或存储所需程序代码且可由计算机存取的任何其它计算机可读数据存储媒体。

而且，适当地将任何连接称作计算机可读媒体。举例来说，如果使用同轴电缆、光纤电缆、双绞线、数字订户线（DSL）或例如红外线、无线电及微波等无线技术而从网站、服务器或其它远程源传输软件，则同轴电缆、光纤电缆、双绞线、DSL 或例如红外线、无线电及微波等无线技术包括于媒体的定义中。如本文中所使用的磁盘及光盘包括压缩光盘（CD）、激光光盘、光盘、数字通用光盘（DVD）、软盘及蓝光光盘，其中磁盘通常以磁性方式再生数据，而光盘（例如）用激光以光学方式再生数据。以上各项的组合还应包括于计算机可读媒体的范围内。

与计算机程序产品的计算机可读媒体相关联的代码可由计算机（例如，由一个或一个以上处理器）执行，所述处理器例如为一个或一个以上数字信号处理器（DSP）、通用微处理器、专用集成电路（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。在一些方面中，本文中所描述的功能性可提供于经配置以用于编码及解码的专用软件模块或硬件模块内，或并入于组合式视频编码器-解码器（编解码器）中。

已描述了所揭示视频编码技术的各方面。这些及其它方面在所附权利要求书的范围内。

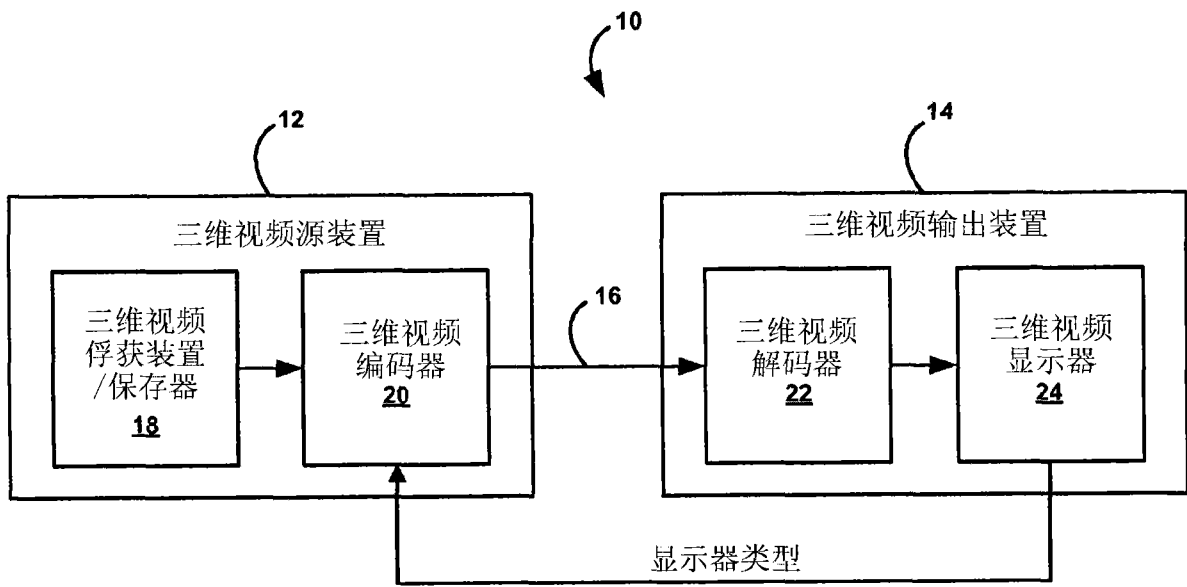


图1

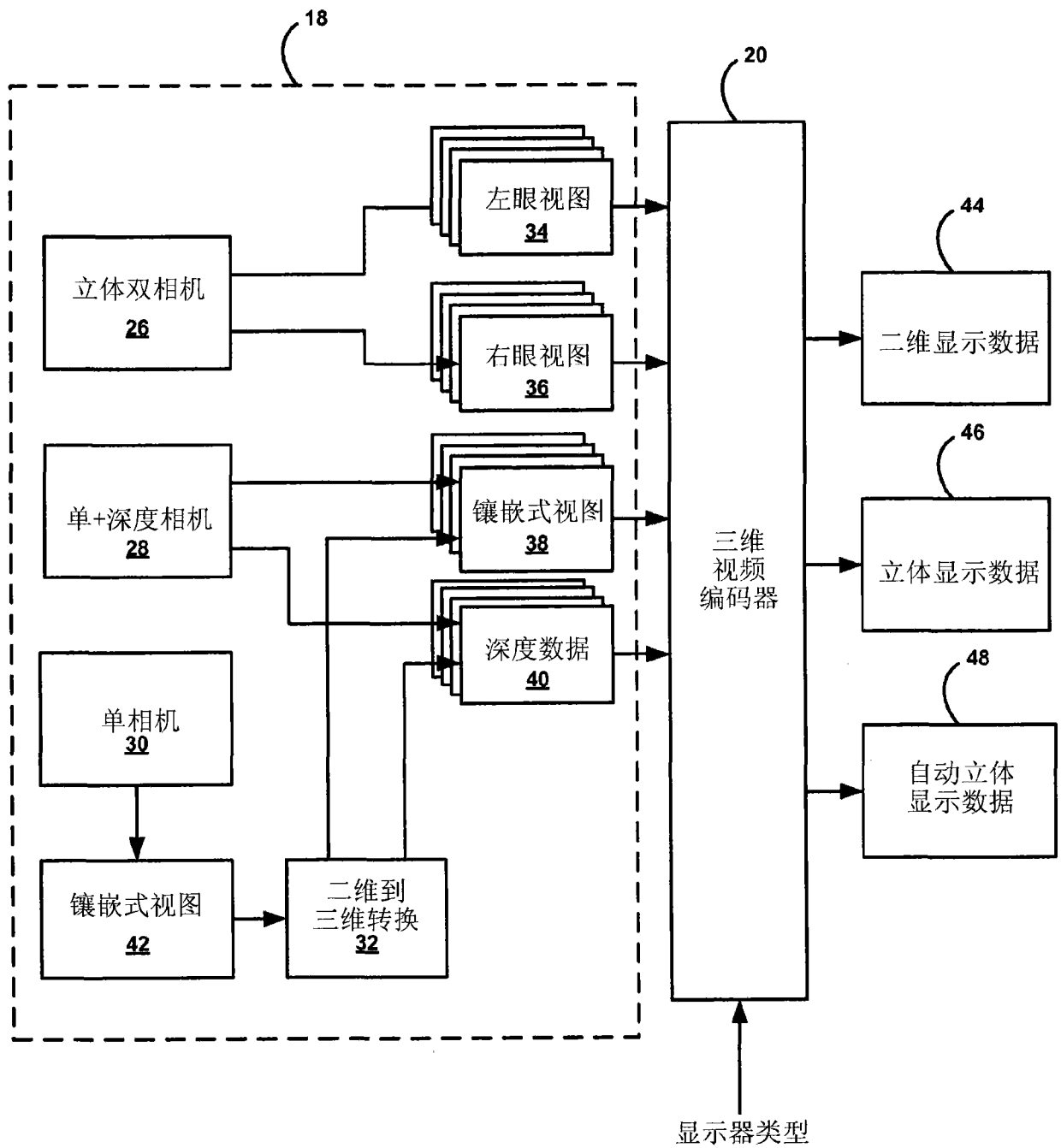


图2

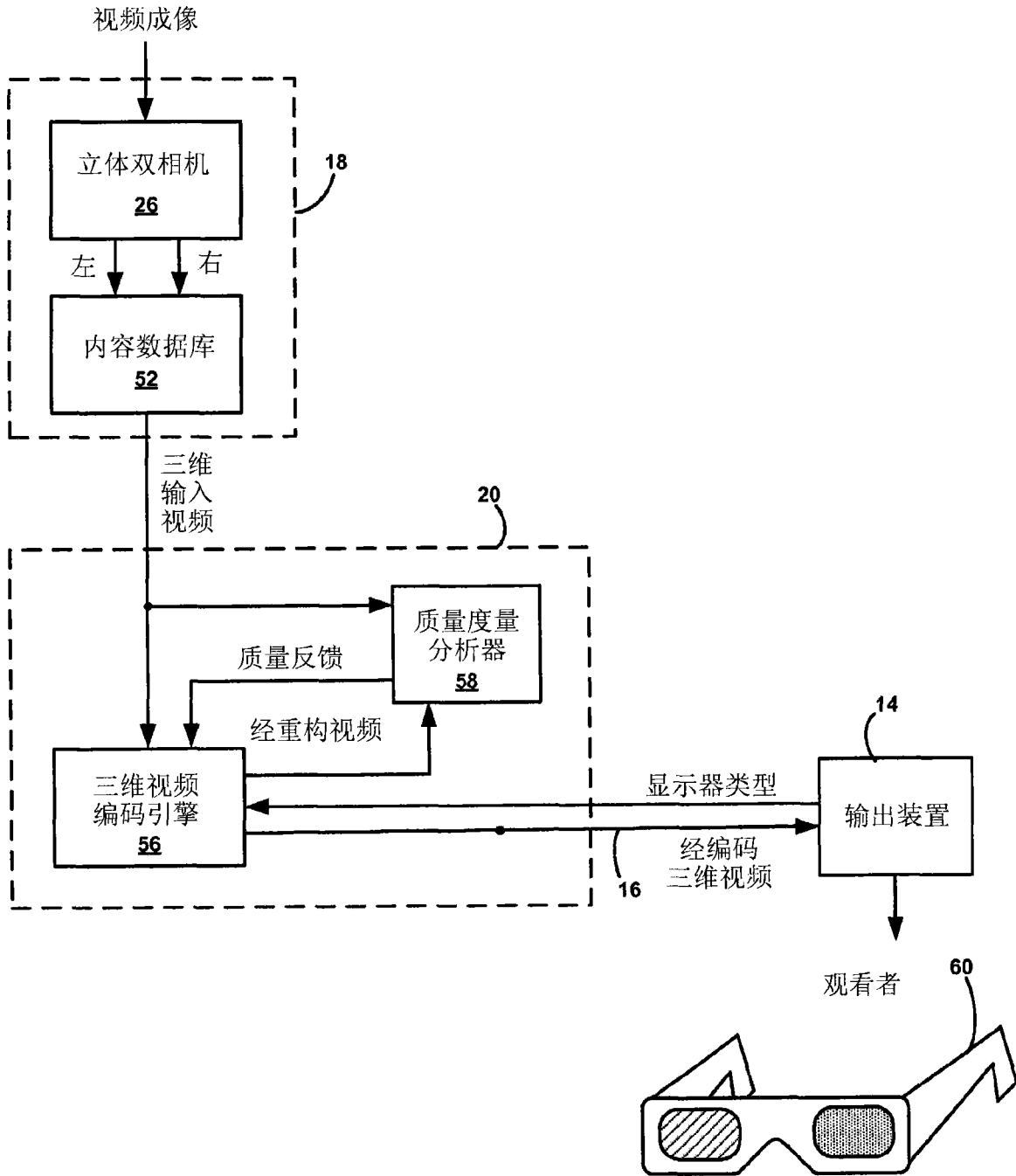


图3A

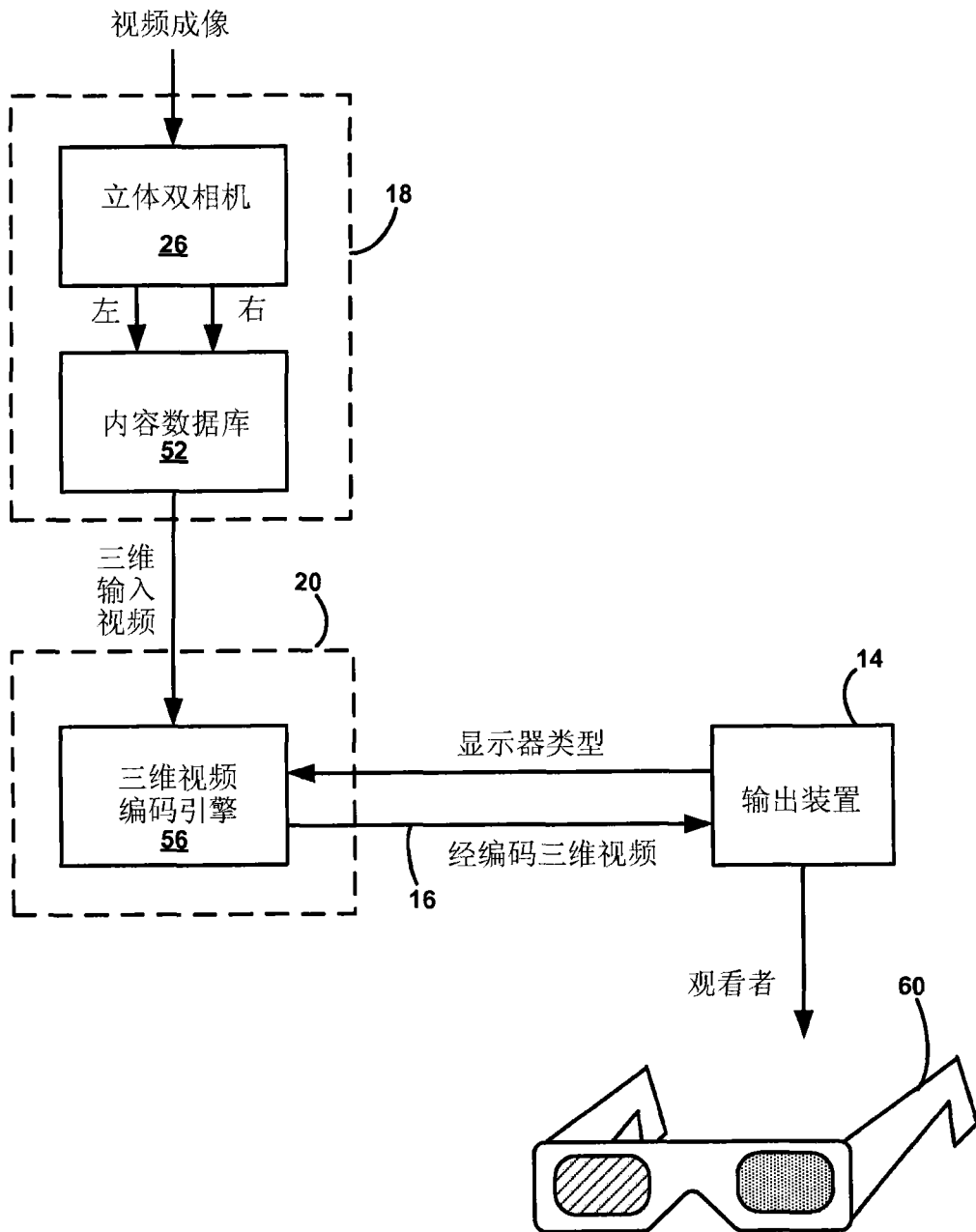


图3B



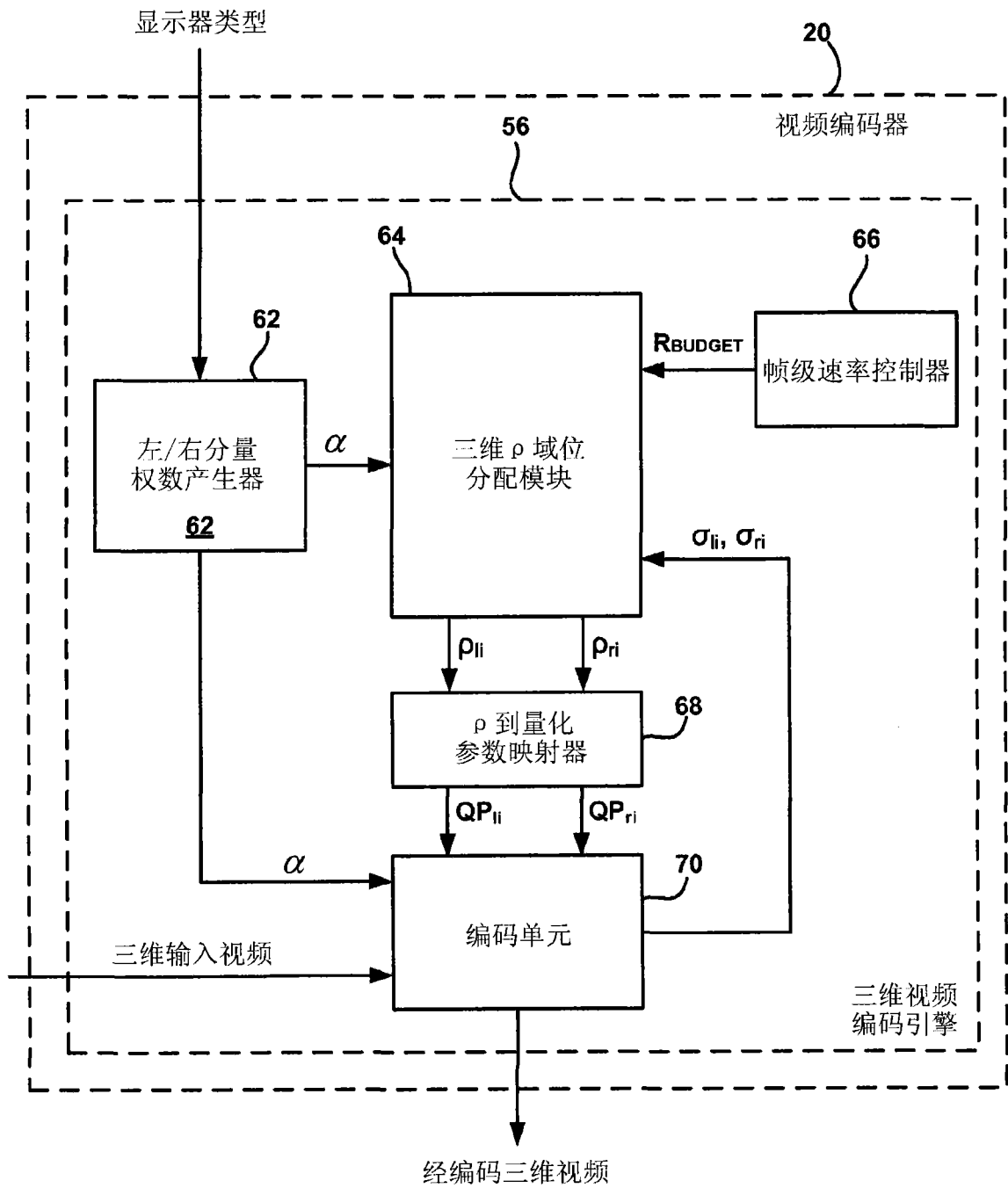


图4

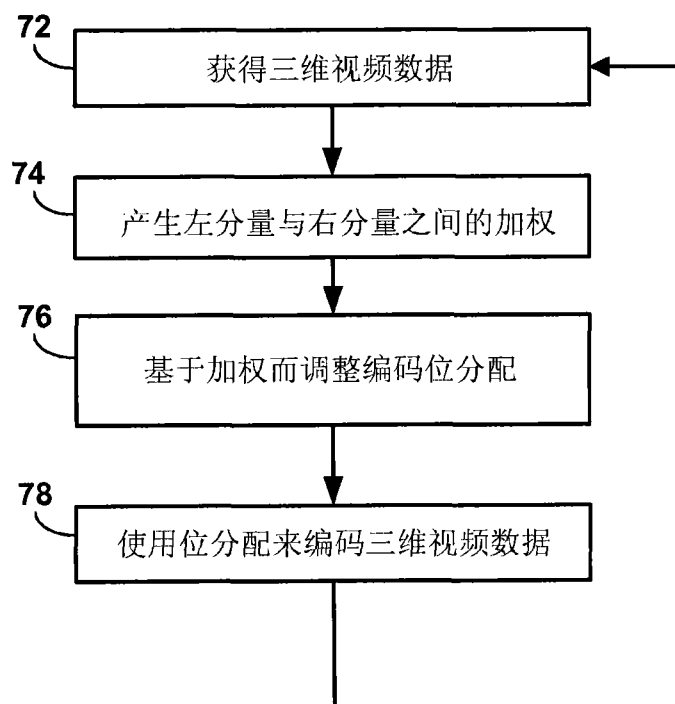


图5

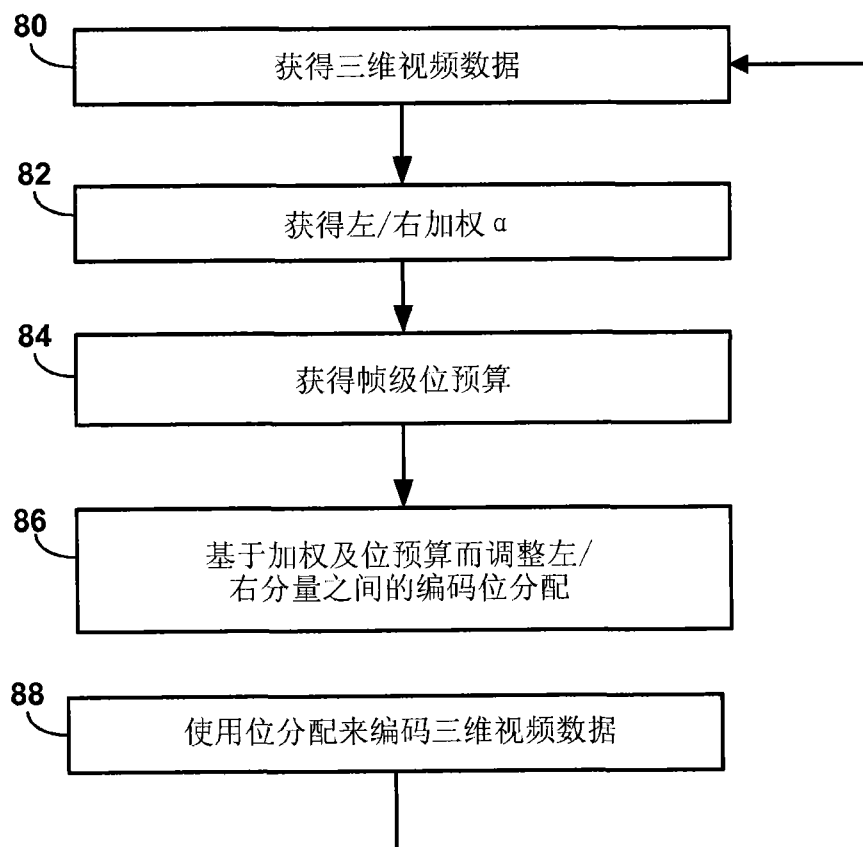


图6

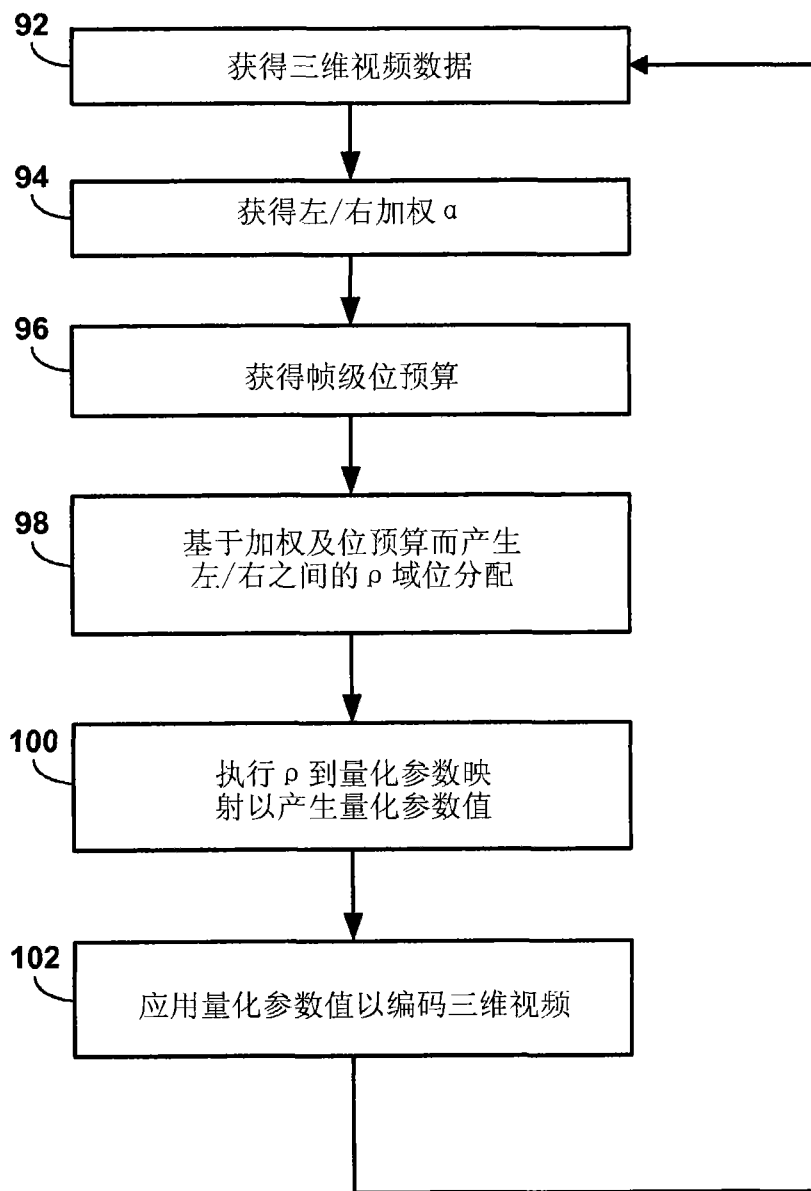


图7