



(10) 授权公告号 CN 108605168 B

(45) 授权公告日 2020.11.03

(21) 申请号 201780010214.9

(22) 申请日 2017.02.15

(65) 同一申请的已公布的文献号  
申请公布号 CN 108605168 A

(43) 申请公布日 2018.09.28

(30) 优先权数据

62/296,528	2016.02.17	US
15/432,660	2017.02.14	US

(85) PCT国际申请进入国家阶段日  
2018.08.07

(86) PCT国际申请的申请数据  
PCT/US2017/017981 2017.02.15

(87) PCT国际申请的公布数据  
W02017/142951 EN 2017.08.24

(73) 专利权人 高通股份有限公司  
地址 美国加利福尼亚州

(72) 发明人 王益魁 伏努·亨利  
马尔塔·卡切维奇

(74) 专利代理机构 北京律盟知识产权代理有限公司 11287

代理人 楊林勳

(51) Int.Cl.

*H04N 21/81* (2011.01)  
*H04N 21/854* (2011.01)  
*H04N 13/189* (2018.01)  
*H04N 13/178* (2018.01)  
*G06F 17/00* (2019.01)

(56) 对比文件

US 2013249900 A1, 2013.09.26

CN 104704830 A.2015.06.10

US 2013278633 A1, 2013.10.24

CN 103327280 A, 2013.09.25

CN 101803394 A, 2010.08.11

Ig-Jae Kim, et al..MART-MAF: Media  
 File Format for AR Tour Guide Service.  
 《2009 11th IEEE International Symposium  
 on Multimedia》.2009.

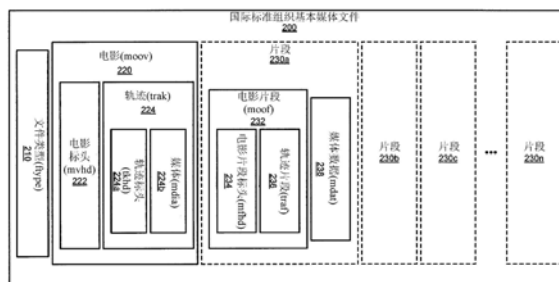
审查员 叶会

权利要求书6页 说明书34页 附图14页

一种处理虚拟现实数据的方法、装置以及非暂时性计算机可读媒体

(57) 摘要

在各种实施方案中,对ISOBMFF的修改和/或添加可指示,已根据ISOBMFF或从ISOBMFF派生的格式而格式化的文件包含虚拟现实内容。所述文件可包含写入到所述文件中的轨迹框中的受限方案信息框。所述受限方案信息框可指示所述轨迹的内容的虚拟现实方案。举例来说,所述受限方案信息框可包含方案类型框,其中所述类型指示虚拟现实方案。作为另一实例,所述受限方案信息框可包含方案信息框,其可提供与所述虚拟现实方案相关的参数。



1. 一种处理虚拟现实数据的方法,其包括:

获得所述虚拟现实数据,其中所述虚拟现实数据表示虚拟环境的360度视图;

将所述虚拟现实数据存储于文件中,其中所述虚拟现实数据根据文件格式存储于所述文件中,其中所述文件格式定义用于将数据存储在该文件中的框,其中框是对象导向式数据结构,其中所述文件格式指定存储所述虚拟现实数据的一或多个框在该文件内的放置且指定存储与所述虚拟现实数据相关联的信息的一或多个框在该文件内的放置,其中与所述虚拟现实数据相关联的所述信息存储于轨迹框内,其中轨迹框存储用于一或多个音频数据或视频数据的数据;及

将样本项存储于所述轨迹框中,其中所述样本项与一或多个样本相关联,其中所述一或多个样本包括所述虚拟现实数据的至少一部分,其中所述样本项包含受限方案信息框,其中所述受限方案信息框将所述一或多个样本的方案定义为虚拟现实方案。

2. 根据权利要求1所述的方法,其进一步包括:

将方案类型框存储于所述样本项中,其中所述方案类型框包含方案类型,且其中所述方案类型定义为所述虚拟现实方案。

3. 根据权利要求1所述的方法,其进一步包括:

将与所述虚拟现实数据相关的参数存储于所述文件内。

4. 根据权利要求3所述的方法,其中所述参数包含以下各者中的一或多个:所述虚拟现实数据包含2维视频数据还是3维视频数据的指示、所述虚拟现实数据为预拼接视频数据还是后拼接视频数据的指示,或针对所述虚拟现实数据的映射类型。

5. 根据权利要求3所述的方法,其进一步包括:

将所述参数存储于方案信息框中。

6. 根据权利要求1所述的方法,其中所述文件格式是基于国际标准组织ISO基本媒体文件格式。

7. 一种用于编码虚拟现实数据的装置,其包括:

经配置以存储所述虚拟现实数据的存储器,其中所述虚拟现实数据表示虚拟环境的360度视图;及

与所述存储器通信的输出引擎,其中所述输出引擎经配置以:

将所述虚拟现实数据存储于文件中,其中所述虚拟现实数据根据文件格式存储到所述文件,其中所述文件格式定义用于将数据存储在该文件中的框,其中框是对象导向式数据结构,其中所述文件格式指定存储所述虚拟现实数据的一或多个框在该文件内的放置且指定存储与所述虚拟现实数据相关联的信息的一或多个框在该文件内的放置,其中与所述虚拟现实数据相关联的所述信息存储于轨迹框内,其中轨迹框存储用于一或多个音频数据或视频数据的数据;及

将样本项存储于所述轨迹框中,其中所述样本项与一或多个样本相关联,其中所述一或多个样本包括所述虚拟现实数据的至少一部分,其中所述样本项包含受限方案信息框,其中所述受限方案信息框定义所述一或多个样本的方案为虚拟现实方案。

8. 根据权利要求7所述的装置,其中所述输出引擎经进一步配置以:

将方案类型框存储于所述样本项中,其中所述方案类型框包含方案类型,且其中所述方案类型定义为所述虚拟现实方案。

9. 根据权利要求7所述的装置,其中所述输出引擎经进一步配置以:

将与所述虚拟现实数据相关的参数存储于所述文件中。

10. 根据权利要求9所述的装置,其中所述参数包含以下各者中的一或多者:所述虚拟现实数据包含2维视频数据还是3维视频数据的指示、所述虚拟现实数据为预拼接视频数据还是后拼接视频数据的指示,或针对所述虚拟现实数据的映射类型。

11. 根据权利要求9所述的装置,其中所述输出引擎经进一步配置以:

将所述参数存储于方案信息框中。

12. 根据权利要求7所述的装置,其中所述文件格式是基于国际标准组织ISO基本媒体文件格式。

13. 一种上面存储有指令的非暂时性计算机可读媒体,所述指令在通过一或多个处理器执行时使得所述一或多个处理器进行以下操作:

获得虚拟现实数据,其中所述虚拟现实数据表示虚拟环境的360度视图;

将所述虚拟现实数据存储于文件中,其中所述虚拟现实数据根据文件格式存储于所述文件中,其中所述文件格式定义用于将数据存储于所述文件中的框,其中框是对象导向式数据结构,其中所述文件格式指定存储所述虚拟现实数据的一或多个框在所述文件内的放置及指定存储与所述虚拟现实数据相关联的信息的一或多个框在所述文件内的放置,其中与所述虚拟现实数据相关联的所述信息存储于轨迹框内,其中轨迹框存储用于一或多个音频数据或视频数据的数据;及

将样本项存储于所述轨迹框中,其中所述样本项与一或多个样本相关联,其中所述一或多个样本包括所述虚拟现实数据的至少一部分,其中所述样本项包含受限方案信息框,其中所述受限方案信息框定义所述一或多个样本的方案为虚拟现实方案。

14. 根据权利要求13所述的非暂时性计算机可读媒体,其进一步包括指令,所述指令在由所述一或多个处理器执行时使得所述一或多个处理器:

将方案类型框存储于所述样本项中,其中所述方案类型框包含方案类型,且其中所述方案类型定义为所述虚拟现实方案。

15. 根据权利要求13所述的非暂时性计算机可读媒体,其进一步包括指令,所述指令在由所述一或多个处理器执行时使得所述一或多个处理器:

将与所述虚拟现实数据相关的参数存储于所述文件内。

16. 根据权利要求15所述的非暂时性计算机可读媒体,其中所述参数包含以下各者中的一或多者:所述虚拟现实数据包含2维视频数据还是3维视频数据的指示、所述虚拟现实数据为预拼接视频数据还是后拼接视频数据的指示,或针对所述虚拟现实数据的映射类型。

17. 根据权利要求15所述的非暂时性计算机可读媒体,其进一步包括指令,所述指令在由所述一或多个处理器执行时使得所述一或多个处理器:

将所述参数存储于方案信息框中。

18. 根据权利要求13所述的非暂时性计算机可读媒体,其中所述文件格式是基于国际标准组织ISO基本媒体文件格式。

19. 一种用于处理虚拟现实数据的设备,其包括:

用于获得所述虚拟现实数据的装置,其中所述虚拟现实数据表示虚拟环境的360度视

图；

用于将所述虚拟现实数据存储于文件中的装置，其中所述虚拟现实数据根据文件格式存储于所述文件中，其中所述文件格式定义用于将数据存储于所述文件中的框，其中框是对象导向式数据结构，其中所述文件格式指定存储所述虚拟现实数据的一或多个框在所述文件内的放置且指定存储与所述虚拟现实数据相关联的信息的一或多个框在所述文件内的放置，其中与所述虚拟现实数据相关联的所述信息存储于轨迹框内，其中轨迹框存储用于一或多个音频数据或视频数据的数据；及

用于将样本项存储于所述轨迹框中的装置，其中所述样本项与一或多个样本相关联，其中所述一或多个样本包括所述虚拟现实数据的至少一部分，其中所述样本项包含受限方案信息框，其中所述受限方案信息框定义所述一或多个样本的方案为虚拟现实方案。

20. 根据权利要求19所述的设备，其进一步包括：

用于将方案类型框存储于所述样本项中的装置，其中所述方案类型框包含方案类型，且其中所述方案类型定义为所述虚拟现实方案。

21. 根据权利要求19所述的设备，其进一步包括：

用于将与所述虚拟现实数据相关的参数存储于所述文件内的装置。

22. 根据权利要求21所述的设备，其中所述参数包含以下各者中的一或多个者：所述虚拟现实数据包含2维视频数据还是3维视频数据的指示、所述虚拟现实数据为预拼接视频数据还是后拼接视频数据的指示，或针对所述虚拟现实数据的映射类型。

23. 根据权利要求21所述的设备，其进一步包括：

用于将所述参数存储于方案信息框中的装置。

24. 根据权利要求19所述的设备，其中所述文件格式是基于国际标准组织ISO基本媒体文件格式。

25. 一种处理虚拟现实数据的方法，其包括：

获得文件，其中所述文件包含所述虚拟现实数据，其中所述虚拟现实数据根据文件格式存储于所述文件中，其中所述文件格式定义用于将数据存储于所述文件中的框，其中框是对象导向式数据结构，其中所述文件格式指定存储所述虚拟现实数据在所述文件内的放置且指定存储与所述虚拟现实数据相关联的信息的一或多个框在所述文件内的放置，其中与所述虚拟现实数据相关联的所述信息存储于轨迹框内，其中轨迹框存储用于一或多个音频数据或视频数据的数据；

处理所述文件内的所述轨迹框中的样本项，其中所述样本项与一或多个样本相关联，其中所述样本项包含受限方案信息框；

基于由所述受限方案信息框提供的信息确定所述一或多个样本的方案为虚拟现实方案，其中所述受限方案信息框将所述一或多个样本的所述方案定义为所述虚拟现实方案，且其中所述一或多个样本包含所述虚拟现实数据的至少一部分；及

根据所述虚拟现实方案处理所述一或多个样本。

26. 根据权利要求25所述的方法，其进一步包括：

处理所述样本项中的方案类型框，其中所述方案类型框包含方案类型，且其中所述方案类型定义为所述虚拟现实方案。

27. 根据权利要求25所述的方法，其进一步包括：

处理与所述虚拟现实数据相关的参数,其中所述参数存储于所述文件中。

28. 根据权利要求27所述的方法,其中所述参数包含以下各者中的一或多者:所述虚拟现实数据包含2维视频数据还是3维视频数据的指示、所述虚拟现实数据为预拼接视频数据还是后拼接视频数据的指示,或针对所述虚拟现实数据的映射类型。

29. 根据权利要求27所述的方法,其中所述参数存储于所述文件中的方案信息框中。

30. 根据权利要求25所述的方法,其中所述文件格式是基于国际标准组织ISO基本媒体文件格式。

31. 一种用于解码虚拟现实数据的装置,其包括:

经配置以存储文件的存储器,其中所述文件包含所述虚拟现实数据,其中所述虚拟现实数据根据文件格式存储于所述文件中,其中所述文件格式定义用于将数据存储于所述文件中的框,其中框是对象导向式数据结构,其中所述文件格式指定存储所述虚拟现实数据的一或多个框在所述文件内的放置且指定存储与所述虚拟现实数据相关联的信息的一或多个框在所述文件内的放置,其中与所述虚拟现实数据相关联的所述信息存储于轨迹框内,其中轨迹框存储用于一或多个音频数据或视频数据的数据;及

与所述存储器通信的输入引擎,其中所述输入引擎经配置以:

处理所述文件内的所述轨迹框中的样本项,其中所述样本项与一或多个样本相关联,其中所述样本项包含受限方案信息框;

基于所述受限方案信息框中的信息确定所述一或多个样本的方案为虚拟现实方案,其中所述受限方案信息框定义所述一或多个样本的所述方案为所述虚拟现实方案,且其中所述一或多个样本包含所述虚拟现实数据的至少一部分;及

根据所述虚拟现实方案处理所述一或多个样本。

32. 根据权利要求31所述的装置,其中所述输入引擎经进一步配置以:

处理所述样本项中的方案类型框,其中所述方案类型框包含方案类型,且其中所述方案类型定义为所述虚拟现实方案。

33. 根据权利要求31所述的装置,其中所述输入引擎经进一步配置以:

处理与所述虚拟现实数据相关的参数,其中所述参数存储于所述文件中。

34. 根据权利要求33所述的装置,其中所述参数包含以下各者中的一或多者:所述虚拟现实数据包含2维视频数据还是3维视频数据的指示、所述虚拟现实数据为预拼接视频数据还是后拼接视频数据的指示,或针对所述虚拟现实数据的映射类型。

35. 根据权利要求33所述的装置,其中所述参数存储于所述文件中的方案信息框中。

36. 根据权利要求31所述的装置,其中所述文件格式是基于国际标准组织ISO基本媒体文件格式。

37. 一种上面存储有指令的非暂时性计算机可读媒体,所述指令在通过一或多个处理器执行时使得所述一或多个处理器进行以下操作:

获得文件,其中所述文件包含虚拟现实数据,其中所述虚拟现实数据根据文件格式存储于所述文件中,其中所述文件格式定义用于将数据存储于所述文件中的框,其中框是对象导向式数据结构,其中所述文件格式指定存储所述虚拟现实数据的一或多个框在所述文件内的放置且指定存储与所述虚拟现实数据相关联的信息的一或多个框在所述文件内的放置,其中与所述虚拟现实数据相关联的所述信息存储于轨迹框内,其中轨迹框存储用于

一或多个音频数据或视频数据的数据；

处理所述文件内的所述轨迹框中的样本项，其中所述样本项与一或多个样本相关联，其中所述样本项包含受限方案信息框

基于由所述受限方案信息框提供的信息确定所述一或多个样本的方案为虚拟现实方案，其中所述受限方案信息框定义所述一或多个样本的所述方案为所述虚拟现实方案，且其中所述一或多个样本包含所述虚拟现实数据的至少一部分；及

根据所述虚拟现实方案处理所述一或多个样本。

38. 根据权利要求37所述的非暂时性计算机可读媒体，其进一步包括指令，所述指令在由所述一或多个处理器执行时使得所述一或多个处理器：

处理所述样本项中的方案类型框，其中所述方案类型框包含方案类型，且其中所述方案类型定义为所述虚拟现实方案。

39. 根据权利要求37所述的非暂时性计算机可读媒体，其进一步包括指令，所述指令在由所述一或多个处理器执行时使得所述一或多个处理器：

处理与所述虚拟现实数据相关的参数，其中所述参数存储于所述文件中。

40. 根据权利要求39所述的非暂时性计算机可读媒体，其中所述参数包含以下各者中的一或多个者：所述虚拟现实数据包含2维视频数据还是3维视频数据的指示、所述虚拟现实数据为预拼接视频数据还是后拼接视频数据的指示，或针对所述虚拟现实数据的映射类型。

41. 根据权利要求39所述的非暂时性计算机可读媒体，其中所述参数存储于所述文件中的方案信息框中。

42. 根据权利要求37所述的非暂时性计算机可读媒体，其中所述文件格式是基于国际标准组织ISO基本媒体文件格式。

43. 一种用于处理虚拟现实数据的设备，其包括：

用于获得文件的装置，其中所述文件包含虚拟现实数据，其中所述虚拟现实数据根据文件格式存储于所述文件中，其中所述文件格式定义用于将数据存储在该文件中的框，其中框是对象导向式数据结构，其中所述文件格式指定存储所述虚拟现实数据的一或多个框在该文件内的放置且指定存储与所述虚拟现实数据相关联的信息的一或多个框在该文件内的放置，其中与所述虚拟现实数据相关联的所述信息存储于轨迹框内，其中轨迹框存储用于一或多个音频数据或视频数据的数据；用于处理所述文件内的所述轨迹框中的样本项的装置，其中所述样本项与一或多个样本相关联，其中所述样本项包含受限方案信息框；

用于基于由所述受限方案信息框提供的信息确定所述一或多个样本的方案为虚拟现实方案的装置，其中所述受限方案信息框将所述一或多个样本的所述方案定义为所述虚拟现实方案，且其中所述一或多个样本包含所述虚拟现实数据的至少一部分；及

用于根据所述虚拟现实方案处理所述一或多个样本的装置。

44. 根据权利要求43所述的设备，其进一步包括：

用于处理所述样本项中的方案类型框的装置，其中所述方案类型框包含方案类型，且其中所述方案类型定义为所述虚拟现实方案。

45. 根据权利要求43所述的设备，其进一步包括：

用于处理与所述虚拟现实数据相关的参数的装置,其中所述参数存储于所述文件中。

46.根据权利要求45所述的设备,其中所述参数包含以下各者中的一或多个:所述虚拟现实数据包含2维视频数据还是3维视频数据的指示、所述虚拟现实数据为预拼接视频数据还是后拼接视频数据的指示,或针对所述虚拟现实数据的映射类型。

47.根据权利要求45所述的设备,其中所述参数存储于所述文件中的方案信息框中。

48.根据权利要求43所述的设备,其中所述文件格式是基于国际标准组织ISO基本媒体文件格式。

## 一种处理虚拟现实数据的方法、装置以及非暂时性计算机可读媒体

### 技术领域

[0001] 本申请涉及虚拟现实 (VR) 视频内容以例如基于 ISO 的媒体文件格式 (ISOBMFF) 和/或从 ISOBMFF 派生的文件格式等一或多个媒体文件格式的存储。举例来说, 本申请涉及用于以向后兼容方式传信 VR 视频内容 (在一些实例中包含一或多个 VR 视频特定细节) 的识别的方法及系统。

### 背景技术

[0002] 视频译码标准包含 ITU-T H.261; ISO/IEC MPEG-1 Visual; ITU-T H.262 或 ISO/IEC MPEG-2 Visual; ITU-T H.263; ISO/IEC MPEG-4 Visual; ITU-T H.264 或 ISO/IEC MPEG-4 AVC, 包含其称为可伸缩视频译码 (SVC) 的可伸缩视频译码扩展及其多视图视频译码扩展, 即多视图视频译码 (MVC) 扩展; 及也称为 ITU-T H.265 及 ISO/23008-2 的高效视频译码 (HEVC), 包含其可伸缩译码扩展 (即, 可伸缩高效视频译码 SHVC) 及多视图扩展 (即, 多视图高效视频译码 MV-HEVC)。

### 发明内容

[0003] 在一些实施例中, 描述数种技术, 所述技术用于在文件中指示文件包含虚拟现实内容, 使得视频播放器装置可恰当地显现和/或忽略虚拟现实内容。

[0004] 根据至少一个实例, 提供一种用于在文件中指示虚拟现实内容的方法, 所述方法包含获得虚拟现实数据。所述虚拟现实数据可表示虚拟环境的 360 度视图。所述方法进一步包含将所述虚拟现实数据存储于文件中。所述虚拟现实数据可根据格式存储于文件中。所述文件格式可指定所述虚拟现实数据在所述文件内的放置并且指定与所述虚拟现实数据相关联的信息在所述文件内的放置。与所述虚拟现实数据相关联的所述信息可存储于轨迹框中。所述方法进一步包含将样本项存储于所述轨迹框中。所述样本项可与一或多个样本相关联。所述样本项可包含受限方案信息框。所述受限方案信息框可将所述一或多个样本的方案定义为虚拟现实方案。所述样本可包含所述虚拟现实数据的至少一部分。

[0005] 在另一实例中, 提供一种设备, 所述设备包含经配置以存储视频数据的存储器, 及处理器。所述处理器经配置以且可获得虚拟现实数据。所述虚拟现实数据可表示虚拟环境的 360 度视图。所述过程经配置以且可将虚拟现实数据存储于文件中。所述虚拟现实数据可根据格式存储于文件中。所述文件格式可指定虚拟现实数据在所述文件内的放置, 并且指定与所述虚拟现实数据相关联的信息的放置。与所述虚拟现实数据相关联的所述信息可存储于轨迹框中。处理器经配置以且可进一步将样本项存储到轨迹框中。所述样本项可与一或多个样本相关联。所述样本项可包含受限方案信息框。所述受限方案信息框可将所述一或多个样本的方案定义为虚拟现实方案。所述样本可包含所述虚拟现实数据的至少一部分。

[0006] 在另一实例中, 提供一种上面存储有指令的计算机可读媒体, 所述指令在由处理



器执行时执行方法,所述方法包含:获得虚拟现实数据。所述虚拟现实数据可表示虚拟环境的360度视图。所述方法进一步包含将所述虚拟现实数据存储于文件中。所述虚拟现实数据可根据格式存储于文件中。所述文件格式可指明虚拟现实数据在所述文件内的放置,并且可指明与所述虚拟现实数据相关联的信息的放置。与所述虚拟现实数据相关联的所述信息可存储于轨迹框中。所述方法进一步包含将样本项写入于所述轨迹框中。所述样本项可与一或多个样本相关联。所述样本项可包含受限方案信息框。所述受限方案信息框可将所述一或多个样本的方案定义为虚拟现实方案。所述样本可包含所述虚拟现实数据的至少一部分。

[0007] 在另一实例中,提供一种设备,所述设备包含用于指示文件中的虚拟现实内容的装置。所述设备进一步包括用于获得虚拟现实数据的装置。所述虚拟现实数据可表示虚拟环境的360度视图。所述设备进一步包含用于将所述虚拟现实数据存储于文件中的装置。所述虚拟现实数据可根据格式存储于文件中。所述文件格式可指定所述虚拟现实数据在所述文件内的放置并且指定与所述虚拟现实数据相关联的信息在所述文件内的放置。与所述虚拟现实数据相关联的所述信息可存储于轨迹框中。所述设备进一步包括用于将样本项存储于所述轨迹框中的装置。所述样本项可与一或多个样本相关联。所述样本项可包含受限方案信息框。所述受限方案信息框可将所述一或多个样本的方案定义为虚拟现实方案。所述样本可包含所述虚拟现实数据的至少一部分。

[0008] 在一些方面中,上文所描述的方法、设备及计算机可读媒体进一步包括将方案类型框存储于所述样本项中。所述方案类型框可包含方案类型。所述方案类型可定义为虚拟现实方案。

[0009] 在一些方面中,上述方法、设备及计算机可读媒体进一步包括将与虚拟现实数据相关的参数存储到文件。在一些方面中,所述参数包含以下各者中的一或多个:所述虚拟现实数据包含2维视频数据还是3维视频数据、所述视频数据为预拼接(pre-stitched)视频数据还是后拼接(post-stitched)视频数据,或针对所述视频数据的映射类型。在一些方面中,存储参数包含将所述参数存储到方案信息框中。

[0010] 在一些方面中,文件格式是基于ISO基本媒体文件格式。

[0011] 根据至少一个实例,提供一种用于处理文件中的现实数据内容的方法,所述方法包含获得所述文件,其中所述文件包含所述虚拟现实数据。所述虚拟现实数据可根据文件格式存储于文件中。所述文件格式可指定所述虚拟现实数据在所述文件内的放置且指定与所述虚拟现实数据相关联的信息在所述文件内的放置。与所述虚拟现实数据相关联的所述信息可存储于轨迹框内。所述方法进一步包含处理所述文件,其中处理所述文件包含处理所述轨迹框中的样本项。所述样本项可与一或多个样本相关联。所述样本项可包含受限方案信息框。所述受限方案信息框可将所述一或多个样本的方案定义为虚拟现实方案。所述一或多个样本可包含虚拟现实数据的至少一部分。所述方法进一步包含根据虚拟现实方案处理一或多个样本。

[0012] 在另一实例中,提供一种设备,所述设备包含经配置以存储文件的存储器,及处理器。所述文件包含虚拟现实数据。所述虚拟现实数据可根据文件格式存储于文件中。所述文件格式可指定所述虚拟现实数据在所述文件内的放置及与所述虚拟现实数据相关联的信息在所述文件内的放置。与所述虚拟现实数据相关联的所述信息可存储于轨迹框中。所述

处理器经配置以且可处理所述文件,其中处理所述文件包含处理所述轨迹框中的样本项。所述样本项可与一或多个样本相关联。所述样本项可包含受限方案信息框。所述受限方案信息框可将所述一或多个样本的方案定义为虚拟现实方案。所述一或多个样本可包含虚拟现实数据的至少一部分。所述处理器经配置以且可根据虚拟现实方案处理一或多个样本。

[0013] 在另一实例中,提供一种上面存储有指令的计算机可读媒体,所述指令在由处理器执行时执行方法,所述方法包含:获得文件,其中所述文件包含虚拟现实数据。所述虚拟现实数据可根据文件格式存储于文件中。所述文件格式可指定所述虚拟现实数据在所述文件内的放置且指定与所述虚拟现实数据相关联的信息在所述文件内的放置。与所述虚拟现实数据相关联的所述信息可存储于轨迹框内。所述方法进一步包含处理所述文件,其中处理所述文件包含处理所述轨迹框中的样本项。所述样本项可与一或多个样本相关联。所述样本项可包含受限方案信息框。所述受限方案信息框可将所述一或多个样本的方案定义为虚拟现实方案。所述一或多个样本可包含虚拟现实数据的至少一部分。所述方法进一步包含根据虚拟现实方案处理一或多个样本。

[0014] 在另一实例中,提供一种设备,所述设备包含用于处理包含虚拟现实内容的文件的装置。所述设备包括用于获得文件的装置,其中所述文件包含虚拟现实数据。所述虚拟现实数据可根据文件格式存储于文件中。所述文件格式可指定所述虚拟现实数据在所述文件内的放置且指定与所述虚拟现实数据相关联的信息在所述文件内的放置。与所述虚拟现实数据相关联的所述信息可存储于轨迹框内。所述设备进一步包含用于处理所述文件的装置,其中处理所述文件包含处理所述轨迹框中的样本项。所述样本项可与一或多个样本相关联。所述样本项可包含受限方案信息框。所述受限方案信息框可将所述一或多个样本的方案定义为虚拟现实方案。所述一或多个样本可包含虚拟现实数据的至少一部分。所述设备进一步包含用于根据虚拟现实方案处理一或多个样本的装置。

[0015] 在一些方面中,上述方法、设备及计算机可读媒体进一步包括处理样本项中的方案类型框。所述方案类型框可包含方案类型。所述方案类型定义为虚拟现实方案。

[0016] 在一些方面中,上述方法、设备及计算机可读媒体进一步包括处理与虚拟现实数据相关的参数。所述参数可存储于文件中。

[0017] 在一些方面中,所述参数包含以下各者中的一或多个:虚拟现实数据包含2维视频数据还是3维视频数据的指示、所述视频数据为预拼接视频数据还是后拼接视频数据的指示,或针对所述视频数据的映射类型。在一些方面中,所述参数存储于文件中的方案信息框中。

[0018] 在一些方面中,文件格式是基于国际标准组织 (ISO) 基本媒体文件格式。

[0019] 根据至少一个实例,提供一种用于在文件中指示虚拟现实内容的方法,所述方法包含获得虚拟现实数据。所述虚拟现实数据可表示虚拟环境的360度视图。所述方法进一步包含将所述虚拟现实数据写入到文件。所述虚拟现实数据可根据格式写入到文件。所述文件格式可指明虚拟现实数据在所述文件内的放置,并且可指明描述虚拟现实数据的信息的放置。描述虚拟现实数据的信息可被写入到轨迹框中。所述方法进一步包含设置所述轨迹框中的值。所述值可指示通过所述轨迹框参考的样本经格式化用于虚拟现实呈现。

[0020] 在另一实例中,提供一种设备,所述设备包含经配置以存储视频数据的存储器,及处理器。所述处理器经配置以且可获得虚拟现实数据。所述虚拟现实数据可表示虚拟环境

的360度视图。所述处理器经配置并且可将虚拟现实数据写入到文件。所述虚拟现实数据可根据格式写入到文件。所述文件格式可指明虚拟现实数据在所述文件内的放置,并且可指明描述虚拟现实数据的信息的放置。描述虚拟现实数据的信息可被写入到轨迹框中。所述处理器经配置并且可设置轨迹框中的值。所述值可指示通过所述轨迹框参考的样本经格式化用于虚拟现实呈现。

[0021] 在另一实例中,提供一种上面存储有指令的计算机可读媒体,所述指令在由处理器执行时执行方法,所述方法包含:获得虚拟现实数据。所述虚拟现实数据可表示虚拟环境的360度视图。所述方法进一步包含将所述虚拟现实数据写入到文件。所述虚拟现实数据可根据格式写入到文件。所述文件格式可指明虚拟现实数据在所述文件内的放置,并且可指明描述虚拟现实数据的信息的放置。描述虚拟现实数据的信息可被写入到轨迹框中。所述方法进一步包含设置所述轨迹框中的值。所述值可指示通过所述轨迹框参考的样本经格式化用于虚拟现实呈现。

[0022] 在另一实例中,提供一种设备,所述设备包含用于获得虚拟现实数据的装置。所述虚拟现实数据可表示虚拟环境的360度视图。所述设备进一步包括用于将所述虚拟现实数据写入到文件的装置。所述虚拟现实数据可根据格式写入到文件。所述文件格式可指明虚拟现实数据在所述文件内的放置,并且可指明描述虚拟现实数据的信息的放置。描述虚拟现实数据的信息可被写入到轨迹框中。所述设备进一步包括用于在所述轨迹框中设置值的装置。所述值可指示通过所述轨迹框参考的样本经格式化用于虚拟现实呈现。

[0023] 在一些方面中,设置所述轨迹框中的所述值包含在媒体处理程序框中设置所述值。所述媒体处理程序框可为视频处理程序框。在一些方面中,媒体处理程序框可为虚拟现实处理程序框。

[0024] 在一些方面中,设置所述轨迹框中的所述值包含在样本项框中设置所述值。

[0025] 在一些方面中,在所述轨迹框中设置所述值包含在轨迹标头框、媒体标头框或视频媒体标头框中设置所述值。

[0026] 在一些方面中,上述方法、设备及计算机可读媒体进一步包括将与虚拟现实数据相关的参数写入到文件。在一些方面中,所述参数包含以下各者中的一或多个:所述虚拟现实数据包含2维视频数据还是3维视频数据、所述视频数据为预拼接视频数据还是后拼接视频数据,或针对所述视频数据的映射类型。在一些方面中,写入参数包含将所述参数写入到方案信息框中。

[0027] 在一些方面中,文件格式是基于ISO基本媒体文件格式。在一些方面中,所述虚拟现实数据包含视频数据及音频数据。

[0028] 此发明内容既不打算识别所要求的主题的关键或基本特征,也不打算孤立使用以确定所要求的主题的范围。应参照此专利的整个说明书的适当部分、任何或所有图式及每一技术方案来理解主题。

[0029] 在参考以下说明书、权利要求书及随附图式之后,前述内容连同其它特征及实施例将变得更显而易见。

## 附图说明

[0030] 下文参考如下图式详细描述本发明的说明性实施例:

- [0031] 图1为说明包含编码装置及解码装置的系统的实例的框图。
- [0032] 图2说明根据ISOBMFF格式化的含有用于视频呈现的数据及元数据的ISO基本媒体文件的实例。
- [0033] 图3A及图3B说明ISO基本媒体文件中的顶层框用以指示文件包含虚拟现实内容的实例。
- [0034] 图4说明电影层级指示用于ISO基本媒体文件400中以指示文件包含虚拟现实内容的实例。
- [0035] 图5说明轨迹层级指示符用于ISO基本媒体文件中以指示文件包含虚拟现实内容的实例。
- [0036] 图6说明处理程序框用以传信轨迹的内容包含虚拟现实视频的ISO基本媒体文件的一个实例。
- [0037] 图7说明新处理程序框已经定义以指示轨迹包含虚拟现实内容的ISO基本媒体文件的实例。
- [0038] 图8说明可包含于ISO基本媒体文件中的媒体框的实例。
- [0039] 图9说明用于产生含有虚拟现实内容的文件的过程的实例。
- [0040] 图10说明用于处理含有虚拟现实内容的文件的过程的实例。
- [0041] 图11说明用于产生含有虚拟现实内容的文件的过程的实例。
- [0042] 图12为说明可实施本发明中所描述的技术中的一或多者的实例编码装置的框图。
- [0043] 图13为说明实例解码装置的框图。

### 具体实施方式

[0044] 下文提供本发明的某些方面及实施例。此些方面及实施例中的一些可独立地应用,且其中的一些可以对所属领域的技术人员将显而易见的组合来应用。在以下描述中,出于解释的目的,阐述众多特定细节以便提供对本发明的实施例的透彻理解。然而,各种实施例可在无此些特定细节的情况下得以实践将是显而易见的。图式及描述不打算为限制性的。

[0045] 随后描述仅提供示范性实施例且并不打算限制本发明的范围、可应用性或配置。确切地说,示范性实施例的随后描述将为所属领域的技术人员提供用于实施示范性实施例的能够实现的描述。应理解,可在不脱离如所附权利要求书中所阐述的本发明的精神及范围的情况下对元件的功能及布置做出各种改变。

[0046] 在以下描述中给定特定细节从而提供实施例的透彻理解。然而,所属领域的普通技术人员应了解,实施例可在无需此些具体细节的情况下实践。举例来说,电路、系统、网络、过程及其它组件可以框图形式中展示为组件以便在不必要细节下不与实施例混淆。在其它情况下,可在无不必要的细节的情况下展示熟知电路、过程、算法、结构及技术以免混淆实施例。

[0047] 而且,应注意,个别实施例可描述为经描绘为流程图、流程框图、数据流程框图、结构图或框图的过程。尽管流程图可能将操作描述为顺序过程,但许多操作可并行地或同时来执行。另外,可再布置操作的次序。过程在其操作完成时终止,但可具有不包含于图中的额外步骤。过程可对应于方法、函数、程序(procedure)、子例程、子程序,等等。当过程对应

于函数时,其终止对应于函数返回到调用函数或主函数。

[0048] 术语“计算机可读媒体”包含但不限于便携式或非便携式存储装置、光学存储装置,及能够存储、含有或携载指令和/或数据的各种其它媒体。计算机可读媒体可包含非暂时性媒体,在所述非暂时性媒体中可存储数据且其不包含载波和/或无线地或经由有线连接传播的暂时电子信号。非暂时性媒体的实例可包含(但不限于)磁盘或磁带;光学存储媒体,例如紧密光盘(CD)或数字多功能盘(DVD);闪存器;存储器或存储器装置。计算机可读媒体可具有存储于其上的代码和/或机器可执行指令,所述代码和/或机器可执行指令可表示程序(procedure)、函数、子程序、程序、例程、子例程、模块、软件套件、类别,或指令、数据结构或程序陈述式的任何组合。可通过传递和/或接收信息、数据、变元、参数或存储器内容将代码段耦合到另一代码段或硬件电路。信息、变元、参数、数据等可经由包含存储器共享、消息传递、令牌传递、网络传输等任何合适的方式传递、转发或传输。

[0049] 此外,实施例可由硬件、软件、固件、中间软件、微码、硬件描述语言或其任何组合实施。当以软件、固件、中间软件或微码实施时,执行必要任务的程序代码或代码段(例如,计算机程序产品)可存储于计算机可读或机器可读媒体中。处理器可执行必要任务。

[0050] 虚拟现实(VR)描述可在看起来真实或物理的方式内交互的计算机产生的三维环境。一般来说,体验虚拟现实环境的用户使用例如头戴显示器(HMD)且任选地也使用衣物(例如,配备有传感器的手套)等电子装备以与虚拟环境交互。随着用户在真实世界中移动,在虚拟环境中显现的图像也改变,从而向用户给予用户正在虚拟环境内移动的感知。在一些状况下,虚拟环境包含与用户的移动相关的声音,从而向用户给予声音来源于特定方向或来源的印象。虚拟现实视频可以极高质量被捕获且显现,从而潜在地提供真实沉浸式虚拟现实体验。虚拟现实应用包含游戏、训练、教育、体育视频及线上购物外加其它。

[0051] 虚拟现实系统通常包含视频捕获装置及视频显示装置,且可能还包含例如服务器、数据存储装置及数据传输装置等其它中间装置。视频捕获装置可包含摄像机集合,即多个摄像机的集合,每一摄像机定向于不同方向且捕获不同视图。少至六个摄像机可用以捕获以摄像机集合的位置为中心的全360度视图。一些视频捕获装置可使用较少摄像机,例如主要捕获侧对侧视图的视频捕获装置。视频大体包含帧,其中帧为场景的电子译码静态图像。摄像机每秒捕获某数目个帧,其通常被称作摄像机的帧速率。

[0052] 为了提供无缝的360度视图,通过摄像机集合中的每一摄像机捕获的视频通常经历图像拼接。360度视频产生的状况下的图像拼接涉及组合或合并来自视频帧重叠或将与其它方式连接的区域中的邻接摄像机的视频帧。结果将为大体球形帧,但类似于墨卡托(Mercator)投影,经合并的数据通常以平面样式表示。举例来说,经合并视频帧中的像素可经映射到立方体形状或某一其它三维平面形状(例如,角锥形、八面体、十面体等)的平面上。视频捕获及视频显示装置通常按光栅原理操作-从而意指视频帧被当作像素栅格-因此,正方形或矩形平面通常用以表示球面环境。

[0053] 映射到平面表示的虚拟现实视频帧可经编码和/或压缩以供存储和/或传输。编码和/或压缩可使用视频编码解码器(例如,H.265/HEVC兼容编码解码器、H.264/AVC兼容编码解码器或其它合适编码解码器)实现,且产生经压缩的视频位流或位流群组。下文进一步详细描述使用视频编码解码器对视频数据进行编码。

[0054] 经编码视频位流可以媒体格式或文件格式经存储和/或封装。所存储位流可(例

如)经由网络传输到接收器装置,所述接收器装置可解码并显现视频以供显示。此接收器装置在本文中可被称作视频显示装置。举例来说,虚拟现实系统可从经编码视频数据产生经囊封文件(例如,使用国际标准组织(ISO)基本媒体文件格式和/或派生的文件格式)。举例来说,视频编码解码器可编码视频数据,且囊封引擎可通过将视频数据囊封于一或多个ISO格式媒体文件中而产生媒体文件。替代地或另外,所存储的位流可直接从存储媒体提供到接收器装置。

[0055] 接收器装置还可实施编码解码器以解码和/或解压缩经编码视频位流。接收器装置可支持用以将视频位流封装到文件中的媒体或文件格式,抽取视频(且可能也抽取音频)数据以产生经编码视频数据。举例来说,接收器装置可用经囊封视频数据解析媒体文件以产生经编码视频数据,且接收器装置中的编码解码器可解码经编码视频数据。

[0056] 接收器装置接着可发送经解码视频信号到显现装置(例如,视频显示装置)。举例来说,显现装置包含头戴显示器、虚拟现实电视机及其它180度或360度显示装置。大体上,头戴显示器能够跟踪穿戴者的头部的移动和/或穿戴者的眼睛的移动。头戴显示器可使用跟踪信息来显现对应于穿戴者正观察的方向的360度视频的部分,使得穿戴者以与穿戴者将体验真实世界相同的方式体验虚拟环境。显现装置可以与捕获视频的帧速率相同的帧速率或以不同帧速率显现视频。

[0057] 文件格式标准可定义用于将视频(及也可能音频)数据封装成一或多个文件及解封装所述一或多个文件的格式。文件格式标准包含国际标准化组织(ISO)基本媒体文件格式(定义于ISO/IEC 14496-12中的ISOBMFF);及从ISOBMFF派生的其它文件格式,包含运动图片专家组(MPEG)MPEG-4文件格式(定义于ISO/IEC 14496-15中)、第三代合作伙伴计划(3GPP)文件格式(定义于3GPP TS 26.244中)及高级视频译码(AVC)文件格式及高效率视频译码(HEVC)文件格式(皆定义于ISO/IEC 14496-15中)。ISO/IEC 14496-12及 14496-15的最新版本的草案本文可分别在[http://phenix.int-evry.fr/mpeg/doc\\_end\\_user/documents/111\\_Geneva/wg11/w15177-v6-w15177.zip](http://phenix.int-evry.fr/mpeg/doc_end_user/documents/111_Geneva/wg11/w15177-v6-w15177.zip)及[http://phenix.int-evry.fr/mpeg/doc\\_end\\_user/documents/112\\_Warsaw/wg11/w15479-v2-w15479.zip](http://phenix.int-evry.fr/mpeg/doc_end_user/documents/112_Warsaw/wg11/w15479-v2-w15479.zip)处得到。

[0058] ISOBMFF用作许多编码解码器囊封格式(例如,AVC文件格式或任何其它合适的编码解码器囊封格式)以及许多多媒体容器格式(例如,MPEG-4文件格式、3GPP文件格式(3GP)、DVB文件格式或任何其它合适多媒体容器格式)的基础。基于ISOBMFF的文件格式可用于连续媒体,其也称作流媒体。

[0059] 除连续媒体(例如,音频及视频)以外,静态媒体(例如,图像)及元数据可存储于遵循ISOBMFF的文件中。根据ISOBMFF结构化的文件可用于许多目的,包含本地媒体文件播放、远程文件的渐进式下载、用作HTTP动态自适应流式传输(DASH)的区段、用作待流式传输的内容的容器(在此状况下容器包含封包化指令)、用于记录接收的实时媒体流,或其它用途。

[0060] ISOBMFF及其派生文件格式(例如,AVC文件格式或其它派生文件格式)在许多多媒体应用中广泛用于媒体内容(例如,包含视频、音频及计时文本)的存储及囊封。然而,ISOBMFF及从ISOBMFF派生的文件格式不包含用于存储虚拟现实(VR)视频的规范。举例来说,如果虚拟现实视频基于ISOBMFF或派生的文件格式存储于文件中,那么播放器装置可将虚拟现实视频当作常规平面视频(例如,可试图处理常规平面视频)(例如,播放器装置可将

虚拟现实信息当作不包含虚拟现实内容)。播放器可因此在显现期间不应用虚拟现实视频的必要投影,从而导致视频在显示时失真并潜在地不可观看。

[0061] 在各种实施方案中,对ISOBMFF的修改和/或添加可指示,已根据所述ISOBMFF 或从所述ISOBMFF派生的格式而格式化的文件包含虚拟现实内容。举例来说,在一些实施方案中,文件可包含文件层级指示,所述指示传信(例如,指示)文件的内容经格式化用于虚拟现实使用状况或实施方案中。作为另一实例,在一些实施方案中,文件可包含电影层级指示,其传信(例如,指示)文件中的电影呈现包含虚拟现实内容。作为另一实例,在一些实施方案中,文件可包含轨迹层级指示,其传信(例如,指示)轨迹包含虚拟现实内容。在各种实施方案中,关于虚拟现实内容的参数也可在文件、电影和/或轨迹层级传信。

[0062] 在此些及其它实施方案中,播放器装置可认识到文件包含虚拟现实内容的时间。在一些状况下,不能够显示虚拟现实内容的播放器装置可忽略和/或跳过虚拟现实媒体。

[0063] 图1为说明包含编码装置104及解码装置112的系统100的实例的框图。编码装置104可为源装置的部分,且解码装置112可为接收装置的部分。源装置和/或接收装置可包含电子装置,例如,移动或静止电话手机(例如,智能电话、蜂窝电话等等)、台式计算机、膝上型计算机或笔记型计算机、平板计算机、机顶盒、电视机、摄像机、显示装置、数字媒体播放器、视频游戏控制台、视频流式传输装置或任何其它合适的电子装置。在一些实例中,源装置及接收装置可包含用于无线通信的一或多个无线收发器。本文中所描述的译码技术适用于各种多媒体应用中的视频译码,包含流式传输视频传输(例如,经由因特网)、电视广播或传输、编码数字视频以供存储于数据存储媒体上、解码存储于数据存储媒体上的数字视频或其它应用。在一些实例中,系统100可支持单向或双向视频传输以支持例如视频会议、视频流式传输、视频播放、视频广播、游戏和/或可视电话的应用。

[0064] 编码装置104(或编码器)可用以使用视频译码标准或协议来编码视频数据(包含虚拟现实视频数据)以产生经编码视频位流。视频译码标准包含ITU-T H.261、ISO/IEC MPEG-1 Visual、ITU-T H.262或ISO/IEC MPEG-2 Visual、ITU-T H.263、ISO/IEC MPEG-4 Visual及ITU-T H.264(也称为ISO/IEC MPEG-4AVC),包含其可伸缩视频译码及多视图视频译码扩展(分别称为SVC及MVC)。已由ITU-T视频译码专家组(VCEG)及ISO/IEC 动画专家组(MPEG)的视频译码联合协作小组(JCT-VC)完成较近期的视频译码标准,即高效率视频译码(HEVC)。对HEVC的各种扩展处理多层视频译码,且也正由JCT-VC开发,包含对HEVC的多视图扩展(称作MV-HEVC)及对HEVC的可伸缩扩展(称作SHVC)或任何其它合适译码协议。

[0065] 本文中所描述的实施方案使用HEVC标准或其扩展来描述实例。然而,本文中所描述的技术及系统也可适用于其它译码标准,例如AVC、MPEG、其扩展或已可用或尚不可用或尚未开发的其它合适译码标准。因此,虽然可参考特定视频译码标准描述本文中所描述的技术及系统,但所属领域的普通技术人员将了解,描述不应解译为仅适用于所述特定标准。

[0066] 视频源102可提供视频数据到编码装置104。视频源102可为源装置的部分,或可为装置的除源装置以外的部分。视频源102可包含视频捕获装置(例如,摄像机、摄像电话、可视电话等等)、含有经存储视频的视频存档、提供视频数据的视频服务器或内容提供者、从视频服务器或内容提供者接收视频的视频馈入接口、用于产生计算机图形视频数据的计算机图形系统、此些源的组合或任何其它合适的视频源。视频源102的一个实例可包含因特网协议摄像机(IP摄像机)。IP摄像机为可用于监测、住家安全或其它合适应用的数字摄像机



的类型。不同于模拟闭路电视 (CCTV) 摄像机, IP摄像机可经由计算机网络及因特网发送及接收数据。

[0067] 来自视频源102的视频数据可包含一或多个输入图片或帧。图片或帧为作为视频的部分的静态图像。编码装置104的编码器引擎106(或编码器) 编码视频数据以产生经编码视频位流。在一些实例中, 经编码视频位流(或“视频位流”或“位流”) 为一系列的一或多个经译码视频序列。经译码视频序列(CVS) 包含一系列接入单元(AU), 其始于具有在基础层中且具有某些性质的随机接入点图片的AU, 直到且不包含具有在基础层中且具有某些性质的随机接入点图片的下一AU。举例来说, 开始CVS的随机接入点图片的某些性质可包含等于1的RASL旗标(例如, NoRaslOutputFlag)。否则, 随机接入点图片(具有等于0的RASL旗标) 并不开始CVS。接入单元(AU) 包含一或多个经译码图片以及对应于共享相同输出时间的经译码图片的控制信息。图片的经译码切片在位流层级囊封到称作网络抽象层(NAL) 单元的数据单元中。举例来说, HEVC视频位流可包含一或多个CVS, 所述一或多个CVS包含NAL单元。两种类别的NAL单元存在于HEVC标准中, 包含视频译码层(VCL) NAL单元及非VCL NAL单元。VCL NAL单元包含经译码图片数据的一个切片或切片区段(如下描述), 且非VCL NAL单元包含关于一或多个经译码图片的控制信息。

[0068] NAL单元可含有形成视频数据(例如, 经编码视频位流、位流的CVS等等) 的经译码表示(例如, 视频中的图片的经译码表示) 的位序列。编码器引擎106通过将每一图片分割成多个切片而产生图片的经译码表示。切片接着分割成明度样本及色度样本的译码树型块(CTB)。明度样本的CTB及色度样本的一或多个CTB连同样本的语法被称为译码树型单元(CTU)。CTU为用于HEVC编码的基本处理单元。CTU可经分裂成具有不同大小的多个译码单元(CU)。CU含有被称为译码块(CB) 的明度及色度样本阵列。

[0069] 明度及色度CB可进一步被分裂成预测块(PB)。PB为使用用于帧间预测的相同运动参数的明度或色度分量的样本的块。所述明度PB及一或多个色度PB连同相关联语法形成预测单元(PU)。在位流中针对每个PU传信运动参数集合, 且所述运动参数集合用于明度PB及一或多个色度PB的帧间预测。CB也可分割成一或多个变换块(TB)。TB表示色彩分量的样本的正方形块, 对所述正方形块应用同一二维变换以用于译码预测残余信号。变换单元(TU) 表示明度及色度样本的TB以及对应语法元素。

[0070] CU的大小对应于译码节点的大小, 且形状可为正方形。举例来说, CU的大小可为  $8 \times 8$  样本、 $16 \times 16$  样本、 $32 \times 32$  样本、 $64 \times 64$  样本或达到相应CTU的大小的任何其它适当大小。短语“ $N \times N$ ”在本文中用于指视频块的就垂直尺寸及水平尺寸而来说像素尺寸(例如, 8像素 $\times$ 8像素)。可按列及行来布置块中的像素。在一些实施例中, 块在水平方向上可不具有与在垂直方向上相同的数目个像素。与CU相关联的语法数据可描述(例如) 将CU分割成一或多个PU。分割模式可在CU经帧内预测模式编码还是经帧间预测模式编码之间有所不同。PU可分割成非正方形形状。与CU相关联的语法数据也可描述(例如) 根据CTU将CU分割成一或多个TU。TU的形状可为正方形或非正方形。

[0071] 根据HEVC标准, 可使用变换单元(TU) 来执行变换。TU可针对不同CU而发生变化。可基于给定CU内的PU的大小而对TU设置大小。TU可与PU大小相同或小于PU。在一些实例中, 可使用被称为残余四分树(RQT) 的四分树结构将对应于CU的残余样本再分成较小单元。RQT的叶节点可对应于TU。可变换与TU相关联的像素差值以产生变换系数。变换系数可接着由编



码器引擎106量化。

[0072] 一旦视频数据的图片被分割成CU,编码器引擎106即使用预测模式来预测每一PU。接着从原始视频数据中减去预测以得到残余数据(下文描述)。对于每一CU,可使用语法数据在位流内部传信预测模式。预测模式可包含帧内预测(或图片内预测)或帧间预测(或图片间预测)。使用帧内预测时,使用(例如)DC预测以发现PU的平均值、使用平坦预测以使平坦表面拟合于PU、使用方向预测以从相邻数据外插或使用任何其它合适类型的预测,从同一图片中的相邻图像数据预测每一PU。使用帧间预测时,使用运动补偿预测从一或多个参考图片(按输出次序在当前图片之前或之后)中的图像数据预测每一PU。可(例如)在CU层级作出使用图片间预测还是图片内预测来译码图片区域的决策。在一些实例中,图片的一或多个切片被指派有切片类型。切片类型包含I切片、P切片及B切片。I切片(帧内,可独立地解码)为仅仅通过帧内预测译码的图片的切片,且因此可独立地解码,这是由于I切片仅仅需要帧内的数据以预测切片的任一块。P切片(单向预测帧)为可通过帧内预测及通过单向帧间预测译码的图片的切片。P切片内的每一块通过帧内预测或帧间预测而译码。当应用帧间预测时,块仅仅通过一个参考图片而预测,且因此参考样本仅仅来自一个帧的一个参考区。B切片(双向预测性帧)为可通过帧内预测及帧间预测译码的图片的切片。B切片的块可从两个参考图片进行双向预测,其中每一图片贡献一个参考区,且两个参考区的样本集合经加权(例如,利用相等权重)以产生双向预测块的预测信号。如上文所解释,一个图片的切片经独立译码。在一些状况下,图片可仅作为一个切片而被译码。

[0073] PU可包含与预测过程有关的数据。举例来说,当使用帧内预测编码PU时,PU可包含描述用于PU的帧内预测模式的数据。作为另一实例,当使用帧间预测编码PU时,PU可包含定义PU的运动向量的数据。定义PU的运动向量的数据可描述(例如)运动向量的水平分量、运动向量的垂直分量、运动向量的分辨率(例如,四分之一像素精度或八分之一像素精度)、运动向量指向的参考图片,和/或运动向量的参考图片列表(例如,列表0、列表1或列表C)。

[0074] 编码装置104可接着执行变换及量化。举例来说,在预测之后,编码器引擎106可计算对应于PU的残余值。残余值可包括像素差值。在预测执行之后可能剩余的任何残余数据使用块变换进行变换,所述块变换可基于离散余弦变换、离散正弦变换、整数变换、小波变换或其它合适的变换函数。在一些状况下,一或多个块变换(例如,大小为 $32 \times 32$ 、 $16 \times 16$ 、 $8 \times 8$ 、 $4 \times 4$ 等等)可应用于每一CU中的残余数据。在一些实施例中,TU可用于通过编码器引擎106实施的变换及量化过程。给定的具有一或多个PU的CU也可包含一或多个TU。如下文进一步详细描述,可使用块变换将残余值变换成变换系数,且接着可使用TU来量化及扫描残余值以产生用于熵译码的串行化变换系数。

[0075] 在一些实施例中,在使用CU的PU进行帧内预测性或帧间预测性译码之后,编码器引擎106可计算CU的TU的残余数据。PU可包括空间域(或像素域)中的像素数据。在应用块变换之后,TU可包括变换域中的系数。如先前所提及,残余数据可对应于未经编码图片的像素与对应于PU的预测值之间的像素差值。编码器引擎106可形成包含CU的残余数据的TU,且接着可变换TU以产生CU的变换系数。

[0076] 编码器引擎106可执行变换系数的量化。量化通过量化变换系数以减少用以表示系数的数据的量而提供进一步压缩。举例来说,量化可减少与系数中的一些或全部相关联的位深度。在一个实例中,具有n位值的系数可在量化期间经降值舍位为m位值,其中n大于

m。

[0077] 一旦执行量化,经译码视频位流包含经量化变换系数、预测信息(例如,预测模式、运动向量等等)、分割信息及任何其它合适的的数据(例如其它语法数据)。接着可通过编码器引擎106对经译码视频位流的不同元素进行熵编码。在一些实例中,编码器引擎106 可利用预定义扫描次序扫描经量化变换系数以产生可经熵编码的串行化向量。在一些实例中,编码器引擎106可执行自适应扫描。在扫描经量化变换系数以形成向量(例如,一维向量)之后,编码器引擎106可熵编码向量。举例来说,编码器引擎106可使用上下文自适应可变长度译码、上下文自适应二进制算术译码、基于语法的上下文自适应二进制算术译码、概率区间分割熵译码或另一合适的熵编码技术。

[0078] 编码装置104的输出110可经由通信链路120将组成经编码视频位流数据的NAL 单元发送到接收装置的解码装置112。解码装置112的输入114可接收NAL单元。通信链路120可包含由无线网络、有线网络或有线网络与无线网络的组合提供的信道。无线网络可包含任何无线接口或数个无线接口的组合,且可包含任何适合的无线网络(例如,因特网或其它广域网、基于包的网路、WiFi™、射频(RF)、UWB、WiFi-Direct、蜂窝、长期演进(LTE)、WiMax™等等)。有线网络可包含任何有线接口(例如,光纤、网络以太网、电力线网络以太网、经由同轴电缆的网络以太网、数字信号线(DSL)等等)。可使用各种装备来实施有线和/或无线网络,所述装备为例如基站、路由器、接入点、桥接器、网关、交换器等等。可根据通信标准(例如,无线通信协议)调制经编码视频位流数据,且将其传输到接收装置。

[0079] 在一些实例中,编码装置104可将经编码视频位流数据存储于存储器108中。输出110可从编码器引擎106或从存储器108检索经编码视频位流数据。存储器108可包含多种分散式或本地接入的数据存储媒体中的任一者。举例来说,存储器108可包含硬盘、存储光盘、闪存器、易失性或非易失性存储器或用于存储经编码视频数据的任何其它合适的数字存储媒体。

[0080] 解码装置112的输入114接收经编码视频位流数据,且可将视频位流数据提供到解码器引擎116,或提供到存储器118以供稍后由解码器引擎116使用。解码器引擎116 可通过熵解码(例如,使用熵解码器)及抽取组成经编码视频数据的一或多个经译码视频序列的元素而解码经编码视频位流数据。解码器引擎116可随后重新按比例调整经编码视频位流数据且对经编码视频位流数据执行反变换。残余数据随后传递到解码器引擎 116的预测级。解码器引擎116接着预测像素的块(例如,PU)。在一些实例中,预测被加到反变换的输出(残余数据)。

[0081] 解码装置112可将经解码视频输出到视频目的地装置122,所述视频目的地装置可包含用于将经解码视频数据显示给内容的消费者的显示器或其它输出装置。在一些方面中,视频目的地装置122可为接收装置的包含解码装置112的部分。在一些方面中,视频目的地装置122可为不同于接收装置的单独装置的部分。

[0082] 补充增强信息(SEI)消息可包含于视频位流中。举例来说,SEI消息可用于携带并非为由解码装置112解码位流所必需的信息(例如,元数据)。此信息用于改进经解码输出的显示或处理(例如,此类信息可由解码器侧实体使用以改进内容的可视性)。

[0083] 在一些实施例中,视频编码装置104和/或视频解码装置112可分别与音频编码装置及音频解码装置集成。视频编码装置104和/或视频解码装置112还可包含实施上文所描

述的译码技术所必要的其它硬件或软件,例如一或多个微处理器、数字信号处理器 (DSP)、专用集成电路 (ASIC)、现场可编程门阵列 (FPGA)、离散逻辑、软件、硬件、固件或其任何组合。视频编码装置104及视频解码装置112可集成为各别装置中的组合式编码器/解码器(编码解码器)的部分。

[0084] 对HEVC标准的扩展包含多视图视频译码扩展(被称作MV-HEVC)及可伸缩视频译码扩展(被称作SHVC)。MV-HEVC及SHVC扩展共享分层译码的概念,其中不同层包含于经编码视频位流中。经译码视频序列中的每一层由唯一层识别符(ID)定址。层ID可存在于NAL单元的标头中以识别NAL单元所相关联的层。在MV-HEVC中,不同层可表示视频位流中的同一场景的不同视图。在SHVC中,提供以不同空间分辨率(或图片分辨率)或不同重建保真度表示视频位流的不同可伸缩层。可伸缩层可包含基础层(其中层ID=0)及一或多个增强层(其中层ID=1,2,...n)。基础层可符合HEVC的第一版本的配置文件,且表示位流中的最低可用层。与基础层相比,增强层具有增加的空间分辨率、时间分辨率或帧速率和/或重建保真度(或质量)。增强层经阶层式组织,且可(或可不)取决于较低层。在一些实例中,可使用单一标准编码解码器来译码不同层(例如,使用 HEVC、SHVC或其它译码标准编码所有层)。在一些实例中,可使用多标准译码解码器来译码不同层。举例来说,可使用AVC来译码基础层,而可使用对HEVC标准的SHVC 和/或MV-HEVC扩展来译码一或多个增强层。一般来说,层包含一组VCL NAL单元及对应的一组非VCL NAL单元。NAL单元被指派有特定层ID值。在层可取决于较低层的意义上,层可为阶层式的。

[0085] 如先前所描述,HEVC位流包含NAL单元的群组,包含VCL NAL单元及非VCL NAL单元。除了其它信息以外,非VCL NAL单元可含有具有与经编码视频位流相关的高层级信息的参数集。举例来说,参数集可包含视频参数集(VPS)、序列参数集(SPS)及图片参数集(PPS)。参数集的目标的实例包含位速率效率、错误复原(error resiliency)及提供系统层接口。每一切片参考单个作用中PPS、SPS及VPS以接入解码装置112可用于解码切片的信息。可针对每一参数集译码识别符(ID),包含VPS ID、SPS ID及PPS ID。SPS包含SPS ID及VPS ID。PPS包含PPS ID及SPS ID。每一切片标头包含PPS ID。使用ID,可识别针对给定切片的作用中参数集。

[0086] VCL NAL单元包含形成经译码视频位流的经译码图片数据。各种类型的VCL NAL单元定义于HEVC标准中,如下表A中所说明。在单层位流中,如第一HEVC标准中所定义,含于AU中的VCL NAL单元具有相同NAL单元类型值,其中NAL单元类型值定义AU的类型及AU内的经译码图片的类型。举例来说,特定AU的VCL NAL单元可包含瞬时解码再新(IDR)NAL单元(值19),从而使得AU为IDR AU且AU的经译码图片为IDR图片。VCL NAL单元的给定类型涉及含于VCL NAL单元(例如,VCL NAL 单元中的图片的切片或切片区段)中的图片或其部分。三类图片经定义于HEVC标准中,包含前置图片、后置图片及帧内随机接入(IRAP)图片(也称作“随机接入图片”)。在多层位流中,AU内的图片的VCL NAL单元具有相同NAL单元类型值及相同类型的经译码图片。举例来说,含有类型IDR的VCL NAL单元的图片称为AU中的IDR图片。在另一实例中,当AU含有为基础层(层ID等于0)处的IRAP图片的图片时,AU为IRAP AU。

[0087] 如上文所论述的经编码视频位流可写入或封装到一或多个文件中以便将位流从编码装置104传送到解码装置112。举例来说,输出110可包含文件写入引擎,其经配置以产生含有位流的一或多个文件。输出110可经由通信链路120传输一或多个文件到解码器装置

112。替代地或另外，一或多个文件可存储于存储媒体（例如，磁带、磁盘或硬盘，或一些其它媒体）上以供稍后传输到解码装置112。

[0088] 解码装置112可包含（例如）在输入114中的文件解析引擎。文件解析引擎可读取经由通信链路120或从存储媒体接收的文件。文件解析引擎可进一步从文件抽取样本，并重建位流以供解码器引擎116解码。在一些状况下，经重建位流可与由编码器引擎 106产生的位流相同。在一些状况下，编码器引擎106可已产生具有用于解码位流的若干可能选项的位流，在所述状况下，经重建位流可包含所有可能选项中的仅一个选项或少于所有可能选项。

[0089] 如上文所论述经编码的视频位流可使用ISOBMFF、从ISOBMFF派生的文件格式、某一其它文件格式和/或包含ISOBMFF的文件格式的组合来写入或封装到一或多个文件中。一或多个文件可使用视频播放器装置播放，可经传输且接着经显示，和/或被存储。

[0090] 图2说明根据ISOBMFF格式化的含有用于视频呈现的数据及元数据的ISO基本媒体文件200的实例。ISOBMFF经设计而以灵活且可扩展格式含有计时媒体信息，所述计时媒体信息促进媒体的互换、管理、编辑及呈现。媒体的呈现对于含有呈现的系统可为“本地的”，或呈现可经由网络或其它流递送机构。

[0091] 如由ISOBMFF规范所定义的“呈现”为常常通过已由视频捕获装置依序捕获相关或由于某一其它原因相关的图片序列。本文中，呈现也可被称作电影或视频呈现。呈现可包含音频。单个呈现可含在一或多个文件中，其中一个文件含有针对整个呈现的元数据。元数据包含信息，例如时序及成帧数据、描述词、指标、参数，及描述呈现的其它信息。元数据自身并不包含视频和/或音频数据。除含有元数据的文件外的文件不需要根据ISOBMFF而格式化，且需要仅经格式化，使得此些文件可由元数据参考。

[0092] ISO基本媒体文件的文件结构为对象导向式结构，且文件中的个别对象的结构可直接从对象的类型来推断。ISO基本媒体文件中的对象由ISOBMFF规范称作“框”。ISO 基本媒体文件构建为可含有其它框的一系列框。框大体包含提供框的大小及类型的标头。大小描述框的整个大小，包含标头、字段及含在所述框内的所有框。具有并非由播放器装置辨识的类型的框通常被忽略并跳过。

[0093] 如通过图2的实例所说明，在文件的顶层，ISO基本媒体文件200可包含文件类型框210、电影框220以及一或多个电影片段框230a、230n。在此层级可被包含但在此实例中并未表示的其它框包含从由空间框、元数据框及媒体数据框外加其它。

[0094] ISO基本媒体文件可包含通过框类型“ftyp”识别的文件类型框210。文件类型框210识别最适合于解析文件的ISOBMFF规范。“最”在此情况下意指，ISO基本媒体文件200可能已根据特定ISOBMFF规范格式化，但很可能与规范的其它迭代兼容。此最合适规范被称作主流品牌。播放器装置可使用主流品牌来确定装置是否能够解码并显示文件的内容。文件类型框210也可包含版本号，所述版本号可用以指示ISOBMFF规范的版本。文件类型框210也可包含兼容品牌的列表，所述列表包含与文件兼容的其它品牌的列表。ISO基本媒体文件可一个以上主流品牌兼容。

[0095] 当ISO基本媒体文件200包含文件类型框210时，存在仅一个文件类型框。ISO基本媒体文件200可省略文件类型框210以便与其它较旧播放器装置兼容。当ISO基本媒体文件200并不包含文件类型框210时，播放器装置可假定默认主流品牌（例如，“mp41”）、次要版本

(例如,“0”)及兼容品牌(例如,“mp41”)。文件类型框210通常尽可能早地放置于ISO基本媒体文件200中。

[0096] ISO基本媒体文件可进一步包含电影框220,所述电影框含有用于呈现的元数据。电影框220通过框类型“moov”识别。ISO/IEC 14496-12规定,呈现可包含仅一个电影框220,不管是含在一个文件还是多个文件中。频繁地,电影框220靠近ISO基本媒体文件的开始。电影框220包含电影标头框222,且可包含一或多个轨迹框224以及其它框。

[0097] 通过框类型“mvhd”识别的电影标头框222可包含为独立于媒体且作为整体与呈现相关的信息。举例来说,电影标头框222可包含信息,例如产生时间、修改时间、时间标度和/或呈现的持续时间外加其它。电影标头框222还可包含识别呈现中的下一轨迹的识别符。举例来说,在所说明的实例中,识别符可指向电影框220所含有的轨迹框224。

[0098] 通过框类型“trak”识别的轨迹框224可含有用于呈现的轨迹的信息。呈现可包含一或多个轨迹,其中每一轨迹独立于呈现中的其它轨迹。每一轨迹可包含对于轨迹中的内容特定的时间及空间信息,且每一轨迹可与媒体框相关联。轨迹中的数据可为媒体数据,在此状况下,轨迹为媒体轨迹;或数据可为用于流式传输协议的封包化信息,在此状况下,轨迹为提示轨迹。举例来说,媒体数据包含视频及音频数据。在所说明的实例中,实例轨迹框224包含轨迹标头框224a及媒体框224b。轨迹框可包含其它框,例如轨迹参考框、轨迹群组框、编辑框、用户数据框、元框及其它框。

[0099] 由框类型“tkhd”识别的轨迹标头框224a可指定含在轨迹框224中的轨迹的特性。举例来说,轨迹标头框224a可包含轨迹的产生时间、修改时间、持续时间、轨迹识别符、层识别符、群组识别符、容量、宽度和/或高度外加其它。对于媒体轨迹,轨迹标头框224a可进一步识别轨迹是否经启用,轨迹是否应作为呈现的部分而播放,或轨迹是否可用以预览呈现,外加其它。轨迹的呈现通常假定为处于呈现的开始处。轨迹框224 可包含此处未说明的编辑列表框,所述编辑列表框可包含显式时刻表图。时刻表图可指定轨迹的偏移时间外加其它,其中偏移指示轨迹的在呈现开始之后的开始时间。

[0100] 在所说明的实例中,轨迹框224还包含由框类型“mdia”识别的媒体框224b。媒体框224b可含有关于轨迹中的媒体数据的对象及信息。举例来说,媒体框224b可含有处理程序参考框,其可识别轨迹的媒体类型以及藉以呈现轨迹中的媒体的过程。作为另一实例,媒体框224b可含有媒体信息框,其可指定轨迹中媒体的特性。媒体信息框可进一步包含样本的表,其中每一样本描述包含(例如)样本的数据的位置的媒体数据(例如,视频或音频数据)区块。样本的数据存储于下文进一步论述的媒体数据框。如同多数其它框一样,媒体框224b也可包含媒体标头框。

[0101] 在所说明的实例中,实例ISO基本媒体文件200还包含呈现的多个片段230a、230b、230c、230n。片段230a、230b、230c、230n并非ISOBMFF框,而是描述电影片段框232 及由电影片段框232参考的媒体数据框238。电影片段框232及媒体数据框238为顶层框,但此处分组在一起以指示电影片段框232与媒体数据框238之间的关系。

[0102] 通过框类型“moof”识别的电影片段框232可通过包含将以其它方式存储于电影框220中的额外信息而扩展呈现。使用电影片段框232,呈现可经递增地建置。电影片段框232可包含电影片段标头框234及轨迹片段框236以及此处未说明的其它框。

[0103] 通过框类型“mfhd”识别的电影片段标头框234可包含序号。播放器装置可使用序

号来验证片段230a包含呈现的下一段数据。在一些状况下,文件的内容或用于呈现的文件可被不按次序地提供到播放器装置。举例来说,网络包可以不同于最初传输包的次序的次序频繁地到达。在此些状况下,序号可在确定片段的正确次序上辅助播放器装置。

[0104] 电影片段框232还可包含通过框类型“traf”识别的一或多个轨迹片段框236。电影片段框232可包含一组轨迹片段(每轨迹零或多个)。轨迹片段可含有零或多个轨迹行程(run),所述行程中的每一者描述所述轨迹的样本的连续行程。除添加样本到轨迹外,轨迹片段可用以将空的时间添加到轨迹。

[0105] 通过框类型“mdat”识别的媒体数据框238含有媒体数据。在视频轨迹中,媒体数据框238将含有视频帧。媒体数据框可替代地或另外地包含音频数据。呈现可包含在一或多个个别文件中的零或多于零个媒体数据框。媒体数据通过元数据描述。在所说明的实例中,媒体数据框238中的媒体数据可通过包含于轨迹片段框236中的元数据来描述。在其它实例中,媒体数据框中的媒体数据可通过电影框220中的元数据来描述。元数据可通过文件200内的绝对偏移参考特定媒体数据,使得媒体数据框238内的媒体数据标头和/或从由空间可被跳过。

[0106] ISO基本媒体文件200中的其它片段230b、230c、230n可含有类似于针对第一片段230a说明的那些的框,和/或可含有其它框。

[0107] 除支持媒体的本地播放外,ISOBMFF包含对经由网络流式传输媒体数据的支持。包含一个电影呈现的一或多个文件可包含称为提示轨迹的额外轨迹,其含有可辅助流式传输服务器形成并传输一或多个文件作为包的指令。举例来说,这些指令可包含供服务器发送的数据(例如标头信息)或对媒体数据的区段的参考。文件可包含用于不同流式传输协议的分离提示轨迹。提示轨迹也可在不需重新格式化文件的情况下添加到文件。

[0108] 用于对媒体数据进行流式传输的一种方法为经由超文本传送协议(HTTP)的动态自适应流式传输,或DASH(定义于ISO/IEC 23009-1:2014)。也被称为MPEG-DASH的 DASH为允许实现使用常规HTTP网页服务器进行媒体内容的高质量流式传输的自适应位速率流式传输技术。DASH通过将媒体内容打断成一系列基于HTTP的小文件区段来操作,其中每一区段含有短时间间隔的内容。使用DASH,服务器可以不同位速率提供媒体内容。正播放媒体的客户端装置可在下载下一区段时在替代性位速率中做出选择,且因此适应于改变的网络条件。DASH使用因特网的HTTP网页服务器基础架构来经由全球信息网递送内容。DASH独立于用以编码并解码媒体内容的编码解码器,且因此可与例如H.264及HEVC外加其它的编码解码器一起操作。

[0109] ISOBMFF规范指定适用于DASH的六种类型流接入点(SAP)。前两种SAP类型(类型1及2)对应于H.264/AVC及HEVC中的瞬时解码再新(IDR)图片。举例来说,IDR图片为图片内(intra-picture, I-picture)图片,其在解码器处完全再新或重新初始化解码过程且开始新的经译码视频序列。在一些实例中,IDR图片及按解码次序在IDR图片之后的任何图片无法依赖于按解码次序在IDR图片之前出现的任何图片。

[0110] 第三SAP类型(类型3)对应于开放GOP(图片群组)随机接入点,因此对应于HEVC 中的断链接入(BLA)或清洁随机接入(CRA)图片。举例来说,CRA图片也为I图片。CRA 图片可能并不再新解码器,且可能不开始新CVS,从而允许CRA图片的前置图片依赖于按解码次序在CRA图片之前出现的图片。随机接入可通过解码以下各者在CRA图片处进行:CRA图片、并不

依赖于按解码次序在CRA图片之前出现的任何图片的与CRA 图片相关联的前置图片,及按解码次序及输出次序两者在CRA之后的所有关联图片。在一些状况下,CRA图片可能并不具有关联前置图片。在一些实施例中,在多层状况下,属于具有大于0的层ID的层的IDR或CRA图片可为P图片或B图片,但此些图片可仅使用来自属于与IDR或CRA图片相同的接入单元的其它图片的层间预测,且具有小于含有IDR或CRA图片的层的层ID。

[0111] 第四SAP类型(类型4)对应于渐次性解码再新(GDR)随机接入点。

[0112] 虽然为灵活且可扩展的且广泛用以存储并传输各种类型的媒体,但ISOBMFF并不包含用于存储虚拟现实视频或识别ISO基本媒体文件的内容为包含虚拟现实内容的机构。播放器装置可因此不能确定文件的内容包含虚拟现实视频。不能显示虚拟现实内容的播放器装置可试图无论如何显示内容,从而导致失真的呈现。

[0113] 在各种实施方案中,ISOBMFF和/或从ISOBMFF派生的文件格式可经修改和/或扩展,使得虚拟现实内容可被识别出。此些实施方案可涉及数个框、品牌值、框中的保留位,和/或可各从独立或组合地识别虚拟现实内容的其它指示符。

[0114] 图3A及图3B说明ISO基本媒体文件300中的顶层框用以指示文件300包含虚拟现实内容的实例。在各种实施方案中,使用顶层框指示文件中300中的全部内容为虚拟现实内容。文件300可包含文件类型框310,其可指定与文件300兼容的ISOBMFF或 ISOBMFF的派生的品牌或特定迭代。文件300还可包含电影框320,其可含有用于呈现的元数据。文件300还可任选地包含一或多个片段330a、330b、330c、330n,如上文所论述。

[0115] 在图3A中的实例中,文件类型框310可用以指示文件300包含虚拟现实内容。举例来说,文件类型框310可用以指定品牌值,所述品牌值指示文件与虚拟现实品牌兼容。在各种实施方案中,文件类型框310中列出的兼容品牌也可用以提供任选品牌指示符,其可用以提供所述虚拟现实相关参数。举例来说,一个兼容品牌值可指示虚拟现实内容为2维(2-D)的,而另一兼容品牌值可指示虚拟现实内容为3维(3-D)的。作为另一实例,兼容品牌值可用以指示映射类型;即,虚拟现实视频的球面表示已被映射到等距矩形(equirectangular)、立方体或角锥体格式,还是某一其它格式从而存储于文件300中。在各种实施方案中,例如视频的维度和/或映射的信息可替代地或另外使用文件类型框310 中的任选字段来指示。

[0116] 在图3B的实例中,已定义了新框类型360。新框类型360为类似于文件类型框310的顶层框。新框类型360在文件中的存在和/或新框类型360中的指示符可用以指示文件300包含虚拟现实内容。举例来说,新框类型360可指定虚拟现实兼容品牌值,和/或包含兼容品牌列表中的与虚拟现实内容兼容的品牌值。新框类型360可进一步包含任选参数,所述参数可指示(例如)虚拟现实内容为2-D还是3-D,和/或存储于文件300中的虚拟现实数据的映射。指定新框类型360可避免对修改文件类型框310的需要,如在图3A 的实例中一样。不可识别新框类型360的播放器装置可忽略所述新框类型。

[0117] 当针对文件的顶层定义的文件类型框310或新框类型360用以指示文件300包含虚拟现实内容,在一些实施方案中,文件300也可能不需要在文件300中的其它框中包含指示符以传信虚拟现实内容的存在。

[0118] 图4说明电影层级指示用于ISO基本媒体文件400中以指示文件400包含虚拟现实内容的实例。文件400可包含文件类型框410,其可指定与文件400兼容的ISOBMFF 或 ISOBMFF的派生的品牌或特定迭代。文件400还可包含电影框420,其可含有用于呈现的元数



据。文件400还可任选地包含一或多个片段430a、430b、430c、430n,如上文所论述。

[0119] 如上文所论述,电影框420可包含电影标头框422及任选地一或多个轨迹框424。在图4的实例中,电影标头框422用以指示由电影框420描述的电影或呈现包含虚拟现实内容。举例来说,电影标头框422中的保留位在设置为一个值时可指示电影内容为虚拟现实视频,且在电影可或可能并非虚拟现实视频时可设置为另一值。在一个说明性实例中,如果保留位中的一者用以传达指示,那么等于1的位指示内容为虚拟现实视频内容,且等于0的位指示内容可能或可能并非虚拟现实视频内容。并未经配置以处理保留位的播放器装置可忽略此些位。

[0120] 电影标头框422中的其它字段和/或保留位可用以提供涉及虚拟现实内容的任选参数。举例来说,电影标头框422可包含指示虚拟现实内容为2-D还是3D的参数。

[0121] 作为另一实例,电影标头框422可包含指示虚拟现实内容经预拼接还是后拼接的参数。“预拼接”意指,针对虚拟现实呈现所捕获的不同视图在存储于文件400中之前经组装成单一表示。“后拼接”意指不同视图个别地存储于文件中400中,且将通过解码器装置组装成单一表示。

[0122] 预拼接虚拟现实视频通常表示为球体形状,且映射到对于存储更方便的另一形状(例如,等距矩形、经映射立方体、经映射角锥体或某一其它形状)。指示所使用的映射类型的参数为可(例如)使用保留位在电影标头框422中传信的参数的另一实例。举例来说,一个保留位可用以传达每一映射类型指示。在各种实施方案中,播放器装置可支持多个映射类型。在此些实施方案中,电影标头框422可包含针对每一个别轨迹和/或针对轨迹群组的映射类型。

[0123] 当电影标头框422用以指示存储于电影框420中的电影呈现包含虚拟现实视频时,在各种实施方案中,电影框420中的其它框也可能不需要传信虚拟现实视频的存在。

[0124] 图5说明轨迹层级指示符用于ISO基本媒体文件500中以指示文件500包含虚拟现实内容的实例。文件500可包含文件类型框510,其可指定与文件500兼容的ISOBMFF 或ISOBMFF的派生的品牌或特定迭代。文件500还可包含电影框520,其可含有用于呈现的元数据。文件500还可任选地包含一或多个片段530a、530b、530c、530n,如上文所论述。

[0125] 电影框520可包含电影标头框522及一或多个轨迹框524,以及此处未说明的其它框。电影标头框522可包含作为整体描述呈现的信息。轨迹框524可包含针对呈现中的轨迹的信息。轨迹框524可包含轨迹标头框524a及零或多于零个媒体数据框524b。

[0126] 在图5的实例中,用于特定轨迹框524的轨迹标头框524a用以指示通过轨迹框524描述的轨迹为虚拟现实轨迹,从而意指通过轨迹参考的样本包含虚拟现实视频数据。举例来说,轨迹中的虚拟现实内容可使用轨迹标头框524a中的保留位来指示。举例来说,当特定保留位设置为一值时,轨迹包含虚拟现实内容,且当位设置为另一值时,轨迹可能或可能不包含虚拟现实内容。在一个说明性实例中,如果保留位中的一者用以传达指示,那么等于1的位指示内容为虚拟现实视频内容,且等于0的位指示内容可能或可能并非虚拟现实视频内容。在一些实施方案中,轨迹标头框524中的虚拟现实内容的传信可取决于哪一内容在电影标头框522中传信。举例来说,当电影标头框622指示电影并不包含虚拟现实内容时,那么轨迹标头框524a中轨迹含有虚拟现实数据的任何指示可被忽略。

[0127] 在各种实施方案中,其它虚拟现实相关参数也可在轨迹标头框524a中传信。举例



来说,保留位或某一其它变量可用以指示轨迹中的虚拟现实视频经预拼接还是后拼接。当轨迹中的视频经预拼接时,额外参数可提供例如摄像机位置(例如,关于视点和/或视角)的信息。当轨迹中的视频经后拼接时,额外参数可提供球体视频表示与用以存储数据于文件500中的表示(例如,等距矩形、立方体映射、角锥体映射或某一其它形状)之间的映射类型。

[0128] 当针对轨迹框524的轨迹标头框524a用以传信轨迹包含虚拟现实内容时,在一些实施方案中,轨迹框524中的其它框可能不需要也传信虚拟现实内容在轨迹中的存在。

[0129] 在各种实施方案中,类似于上文所论述的那些技术的技术可用以指示使用DASH传输的文件中的虚拟现实内容。举例来说,虚拟现实内容可在DASH呈现的媒体呈现层级传信。如通过DASH规范定义的媒体呈现为针对有界或无界媒体呈现的数据的集合(例如,单一动画或连续实况流外加其它实例)。媒体呈现可通过媒体呈现描述来描述,文献含有元数据,所述元数据可由DASH客户端使用以构建适当HTTP统一资源定位符(URL)从而接入媒体呈现的区段。

[0130] 在各种实施方案中,媒体呈现描述可用以指示通过媒体呈现描述来描述的媒体内容包含虚拟现实内容。举例来说,元素可经修改或添加到用于媒体呈现描述的结构描述,其中元素接着传信虚拟现实内容。在各种实施方案中,属性也可经修改或添加到媒体呈现描述以提供关于虚拟现实内容的信息,例如内容为2-D还是3-D,内容经预拼接还是后拼接,和/或内容经后拼接时视频帧的映射。在一些实施方案中,媒体呈现描述中的虚拟现实指示符指示,呈现中的全部内容针对虚拟现实经格式化。

[0131] 在DASH呈现中,用于一个呈现的媒体内容经划分成数个时段。如通过DASH定义的时段为在媒体呈现内的时间间隔。呈现因此由连续时段序列组成。在一时段内,媒体内容通常具有一致编码集合,包含具有一个平均位速率、一种语言、一个说明字幕设置、一个对白字幕设置等。

[0132] 在各种实施方案中,时段的元素和/或属性可用以指示虚拟现实内容。举例来说,元素可经修改或添加到针对时段的结构描述,其中元素接着传信虚拟现实内容。在各种实施方案中,属性也可经修改或添加到时段以提供关于虚拟现实内容的信息,例如内容为2-D还是3-D,内容经预拼接还是后拼接,和/或内容经后拼接时针对视频帧的映射。在一些实施方案中,时段中的虚拟现实指示符指示时段中的内容经格式化用于虚拟现实。

[0133] 在一时段内,内容可经布置成调适集。调适集表示一或多个媒体内容分量的可互换经编码版本的集合。举例来说,时段可包含用于主要视频分量的一个调适集及针对主要音频分量的独立调适集。在此实例中,如果存在其它可用内容,例如说明字幕或音频描述,那么此些描述中的每一者可具有独立调适集。

[0134] 在各种实施方案中,虚拟现实内容可在调适集中传信。举例来说,元素可经修改或添加到调适集的结构描述,其中元素接着传信虚拟现实内容。在各种实施方案中,属性也可经修改或添加到调适集以提供关于虚拟现实内容的信息,例如内容为2-D还是3-D,内容经预拼接还是后拼接,和/或内容经后拼接时视频帧的映射。在一些实施方案中,调适集中的虚拟现实指示符指示调适集中的表示中的每一者包含虚拟现实内容。

[0135] 调适集可含有多个替代表示。表示描述一或若干个媒体内容分量的可递送经编码版本。调适集内的任何单一表示可用以在时段中显现媒体内容分量。经发送的在一个调适中的不同表示可被视为感知上等效的,从而意指客户端装置可在调适集内从一个表示动态

切换到另一表示以便适合于网络条件或其它因素。

[0136] 在各种实施方案中,虚拟现实内容可在表示中传信。举例来说,元素可经修改或添加到针对表示的结构描述,其中元素接着指示虚拟现实内容。在各种实施方案中,属性也可经修改或添加到表示以提供关于虚拟现实内容的信息,例如内容为2-D还是3-D,内容经预拼接还是后拼接,和/或内容经后拼接时视频帧的映射。在一些实施方案中,表示中的虚拟现实指示符指示表示的内容经格式化用于虚拟现实。

[0137] 与媒体内容的流式传输相关的另一格式为描述于RFC 4566中的工作阶段描述协议(SDP)。SDP可用以描述多媒体通信工作阶段。举例来说,这些描述可用于工作阶段发布、工作阶段邀请及参数协商。SDP并非用以递送媒体自身,而是可用于针对媒体类型、格式及关联性质的协商的端点之间。性质及参数集合常常被称作工作阶段配置文件。SDP最初为工作阶段发布协议(SAP)的分量,但发现结合以下各者的其它用途:实时传送协议(RTP)、实时流式传输协议(RTSP)、工作阶段起始协议(SIP),及作为用于描述多播工作阶段的独立格式。

[0138] 在各种实施方案中,虚拟现实内容的指示可包含于工作阶段描述中和/或SDP消息中的媒体描述中。举例来说,可在工作阶段描述和/或媒体描述中添加或修改字段以指示流式传输内容中虚拟现实内容的存在。另外,在一些实施方案中,关于虚拟现实内容的参数也可添加到SDP消息。举例来说,这些参数可包含虚拟现实内容为2-D还是3-D,内容经预拼接还是后拼接,和/或用以存储数据的映射。在此及其它实例中,SDP可用于基于RTP的流式传输、广播和/或遥现(telepresence)或会议应用中以指示媒体内容包含虚拟现实内容。

[0139] 作为另一实例,多媒体广播多播服务(MBMS)可用以指示内容经由3GPP蜂窝网络传输时的虚拟现实内容。MBMS为点对多点接口规范,所述规范可在小区内且又在核心网络内提供广播及多播服务的有效递送。针对MBMS的目标应用包含移动电视、实况视频及音频流线性化(streamlining)、文件递送及紧急警示的递送。

[0140] 在各种实施方案中,虚拟现实内容以及与内容相关的参数的传信可通过添加新特征到MBMS特征要求列表来实现。在各种实施方案中,虚拟现实内容的传信对于其它广播及多播应用可以类似方式实现。

[0141] 在各种实施方案中,当ISO基本媒体文件中的轨迹包含虚拟现实内容时,各种额外或替代性方法可用以将虚拟现实内容传信到播放器装置。图6说明处理程序框624c用以传信轨迹的内容包含虚拟现实视频的ISO基本媒体文件600的一个实例。文件600可包含文件类型框610,其可指定与文件600兼容的ISOBMFF或ISOBMFF的派生的品牌或特定迭代。文件600还可包含电影框620,其可含有用于呈现的元数据。文件600还可任选地包含一或多个片段630a、630b、630c、630n,如上文所论述。

[0142] 电影框620可包含电影标头框622及一或多个轨迹框624,以及此处未说明的其它框。电影标头框622可包含作为整体描述呈现的信息。轨迹框624可包含针对呈现中的轨迹的信息。轨迹框624可包含轨迹标头框624a及零或多于零个媒体数据框624b。

[0143] 媒体数据框624b可包含处理程序框642c外加其它框。也可被称作处理程序参考框的处理程序框642c可指示轨迹的媒体类型。轨迹的媒体类型定义轨迹中的媒体数据经呈现的过程。媒体类型的实例包含视频及音频外加其它。媒体经呈现的方式可包含针对媒体的格式。举例来说,播放器装置用以递送轨迹中的视频数据的格式(例如,纵横比、分辨率、帧

速率等)可存储于视频轨迹中,且通过处理程序框642c的视频处理程序版本来识别。在一些状况下,文件600可包含针对任何类型的元数据流的通用处理程序。在这些状况下,视频内容的特定格式可通过描述内容的样本项来识别。

[0144] 在一些状况下,媒体数据框624b可包含处理程序框642c。处理程序框642c可用以指示由轨迹框624描述的轨迹包含虚拟现实数据。举例来说,当轨迹描述视频数据时,处理程序框642c可具体来说为可通过框类型“vide”识别的视频处理程序框。

[0145] 在各种实施方案中,处理程序框642c可用以指示由媒体数据内容框624b参考的媒体内容包含虚拟现实内容。举例来说,处理程序框642c可包含含在轨迹中的视频内容为虚拟现实视频的任选指示符(例如,在预留位或新变量中)。并未经配置以读取任选指示符的视频播放器可忽略任选指示符。

[0146] 在一些实施方案中,视频处理程序框还可任选地包含描述虚拟现实内容的参数,例如虚拟现实视频为2-D还是3-D,虚拟现实视频经预拼接还是后拼接,和/或针对虚拟现实视频的映射。在各种实施方案中,与虚拟现实内容相关的参数可指示于可在轨迹框524中发现的各种其它框中。举例来说,参数可在轨迹标头框624a中传信。替代地或另外,参数可在媒体标头框(通过框类型“mdhd”识别)中和/或在视频媒体标头框(通过框类型“vmhd”识别)中识别,前述各者在此处并未说明。替代地或另外,参数可指示于样本项中和/或可放置于轨迹框624的顶层处的新定义框中。

[0147] 图7说明新处理程序框724d已经定义以指示轨迹包含虚拟现实内容的ISO基本媒体文件700的实例。文件700可包含文件类型框710,其可指定与文件700兼容的 ISOBMFF或ISOBMFF的派生的品牌或特定迭代。文件700还可包含电影框720,其可含有用于呈现的元数据。文件700还可任选地包含一或多个片段730a、730b、730c、730n,如上文所论述。

[0148] 电影框720可包含电影标头框722及一或多个轨迹框724,以及此处未说明的其它框。电影标头框722可包含作为整体描述呈现的信息。轨迹框724可包含针对呈现中的轨迹的信息。轨迹框724可包含轨迹标头框724a及零或多于零个媒体数据框724b。

[0149] 如上文所论述,在一些状况下,媒体数据框724b可包含处理程序框724d,其可描述用于呈现通过媒体数据框724b描述的媒体内容的格式。在图7的实例中,新处理程序框724d已被定义,其对于虚拟现实视频数据为特定的。举例来说,新处理程序框724d 可通过框类型“vrvd”来识别。在此实例中,并不与虚拟现实内容兼容的视频播放器可能无法识别新处理程序框724d,且因此可忽略新处理程序框724d且跳过通过轨迹框724 参考的任何内容。虚拟现实内容将因此并不通过不经配置以显示虚拟现实视频的播放器显现并显示。

[0150] 在一些实施方案中,新处理程序框还可任选地包含描述虚拟现实内容的参数,例如虚拟现实视频为2-D还是3-D,虚拟现实视频经预拼接还是后拼接,和/或虚拟现实视频的映射。在各种实施方案中,与虚拟现实内容相关的参数可指示于可在轨迹框724中发现的各种其它框中。举例来说,参数可在轨迹标头框724a中传信。替代地或另外,参数可在媒体标头框(通过框类型“mdhd”识别)中和/或在视频媒体标头框(通过框类型“vmhd”识别)中识别,前述各者在此处并未说明。替代地或另外,参数可指示于样本项中和/或可放置于轨迹框724的顶层处的新定义框中。

[0151] 图8说明可包含于ISO基本媒体文件中的媒体框840的实例。如上文所论述,媒体框可包含于轨迹框中,且可含有描述轨迹中的媒体数据的对象及信息。在所说明的实例中,媒

体框840包含媒体信息框842。媒体框840还可包含此处未说明的其它框。

[0152] 媒体信息框842可含有描述关于轨迹中的媒体的特性信息的对象。举例来说,媒体信息框842可包含描述轨迹中的媒体信息的位置的数据信息框。作为另一实例,当轨迹包含视频数据时,媒体信息框842可包含视频媒体标头。视频媒体标头可含有独立于视频媒体的译码的一般呈现信息。当轨迹包含音频数据时,媒体信息框842还可包含声音媒体标头。

[0153] 媒体信息框842还可包含样本表框844,如所说明实例中所提供。通过框类型“stbl”识别的样本表框844可提供媒体样本在轨迹中的位置(例如,文件的位置)以及样本的时间信息。使用通过样本表框844提供的信息,播放器装置可以正确时间次序定位样本,确定样本的类型,和/或确定大小、容器及样本在容器内的偏移外加其它。

[0154] 样本表框844可包含通过框类型“stsd”识别的样本描述框846。样本描述框846可提供关于(例如)用于样本的译码类型的详述信息,及针对所述译码类型需要的任何初始化信息。存储于样本描述框中的信息对于包含样本的轨迹的类型可为特定的。举例来说,在轨迹为视频轨迹时,一个格式可用于样本描述;且在轨迹为提示轨迹时,可使用不同格式。作为另一实例,针对样本描述的格式还可取决于提示轨迹的格式而发生变化。

[0155] 样本描述框846可包含一或多个样本项框848a、848b、848c。样本项类型为摘要类别,且因此通常样本描述框包含特定样本项框,例如视频数据的视觉样本项或音频样本的音频样本项外加其它实例。样本项框可存储针对特定样本的参数。举例来说,对于视频样本来说,样本项框可包含视频样本的宽度、高度、水平分辨率、垂直分辨率、帧计数和/或深度外加其它。作为另一实例,对于音频样本,样本项可包含信道计数、信道布局和/或采样速率外加其它。

[0156] 在所说明的实例中,第一样本项848a包含受限方案信息框860。由框类型“rinf”识别的受限方案信息框可含有理解应用到样本的受限方案及方案的参数两者需要的信息。在一些状况下,文件的作者可要求某些动作形成播放器装置。在此些状况下,文件可包含受限方案信息框,播放器装置可定位并使用所述受限方案信息框来确定用于显现文件的媒体内容的要求。可能无法显现内容的播放器也可使用受限方案信息框来确定其不可显现内容,且因此不应试图处理内容。受限方案信息框通常包含原始样本项类型,即,在通过受限方案信息框描述的任何变换之前的样本项的类型。

[0157] 在各种实施方案中,受限方案可经定义用于虚拟现实内容。在此些实施方案中,受限方案信息框860可被添加到包含虚拟现实数据的样本项848a。受限方案的类型可指定于通过框类型“schm”识别的方案类型框862中。举例来说,对应于“vrvd”的编码可用以识别针对虚拟现实内容的受限方案。

[0158] 受限方案信息框860在所说明的实例中包含通过框类型“schi”识别的方案信息框864。方案信息框864可存储用于特定方案的信息。举例来说,当受限方案用于虚拟现实内容时,方案信息框864可包含用于虚拟现实内容的参数。举例来说,这些参数可包含虚拟现实视频为2-D还是3-D,虚拟现实视频经预拼接还是后拼接,和/或针对虚拟现实视频的映射。在各种实施方案中,方案信息框可经定义用于虚拟现实内容,具体来说用于含有针对虚拟现实内容的参数。

[0159] 使用说明于图8中的技术,无新框需要添加到ISOBMFF规范,所述新框可能不被旧版播放器装置理解。甚至在新框情况下,旧版播放器装置可试图播放装置不可识别的内容,

且当此内容为虚拟现实媒体时,结果可为失真的呈现。在避免添加新框时,可产生用于虚拟现实内容的文件,其中文件很可能仅包含旧版播放器装置可识别的框。旧版播放器装置可进一步确定,装置不能实施通过受限方案信息框864描述的受限方案,且因此不试图显示虚拟现实内容。

[0160] 技术另外为旧版播放器及播放器两者提供能够显现虚拟现实内容的灵活性。举例来说,旧版播放器可确定播放器是否理解通过受限方案信息框识别的虚拟现实方案。当播放器装置未能符合受限方案时,播放器装置可选择根本不显现轨迹中的内容,或可能能够改为处理原始的未经变换样本。受限方案机构因此可使得播放器装置能够检查文件以确定对显现位流的需求,且可阻止旧版播放器装置解码并显现装置可能无法处理的文件。

[0161] 在各种实施方案中,虚拟现实内容可替代地或另外地包含于视频位流中的补充增强信息(SEI)消息中。SEI消息可因此指示位流包含虚拟现实内容。在各种实施方案中,SEI消息可指示文件层级、电影层级和/或轨迹层级的虚拟现实内容。在各种实施方案中,SEI消息还可包含参数,所述参数描述虚拟现实视频的性质(例如,视频为2-D还是3-D,经预拼接还是后拼接等)。

[0162] 在各种实施方案中,对ISOBMFF规范的扩展包含轨迹层级指示,以指示轨迹包含虚拟现实视频。在各种实施方案中,后解码器要求机构可用于此轨迹层级指示,(例如)使用受限方案信息(rinf)框。在各种实施方案中,扩展可包含对ISOBMFF规范的章节8.15的修改。以下段落提供ISO/IEC 14496-12的章节8.15的本文,包含新章节及子章节 8.15.5(“用于虚拟现实视频的方案”)、8.15.5.1(“通用”)、8.15.5.2(“VR视频框”)、8.15.5.2.1(“定义”)、8.15.5.2.2(“语法”)及8.15.5.2.3(“语义”)。使用带下划线文本来指示对章节8.15的本文的改变(对本文的改变的实例)。

[0163] 8.15对媒体的后解码器要求

[0164] 8.15.1通用

[0165] 为了处置文件作者需要对播放器或呈现器进行某些动作的情形,此子条款指定使得播放器能够简单地检查文件以找出此些要求从而显现位流且阻止旧版播放器解码并显现需要进一步处理的文件的机构。机构适用于任何类型的视频编解码器。确切地说,所述机构适用于AVC且对于此状况来说,特定传信定义于AVC文件格式(ISO/IEC 14496-15)中,其允许文件作者列出出现的SEI消息ID且区分显现过程的所要求动作及非要求动作。

[0166] 所述机构类似于内容保护变换,其中样本项隐藏于指示经加密或囊封的媒体的通用样本项“encv”、“enca”等后方。用于受限视频的类似机构使用通过通用样本项“resv”的变换。当内容应仅由正确地呈现内容的播放器解码时,可应用所述方法。

[0167] 8.15.2变换

[0168] 方法应用如下:

[0169] 1) 样本项的四字符码通过意指受限视频的新样本项码“resv”来替换。

[0170] 2) 受限方案信息框经添加到样本描述,所有其它框未经修改。

[0171] 3) 原始样本项类型存储于含在受限方案信息框中的原始格式框内。

[0172] RestrictedSchemeInfoBox与ProtectionSchemeInfoBox完全相同地格式化,除了使用识别符“rinf”而非“sinf”外(参见下文)。

[0173] 原始样本项类型含在位于受限方案信息框中的原始格式框中(以相同于用于经加

密媒体的保护方案信息框的方式)。

[0174] 限制的准确本质定义于SchemeTypeBox中,且所述方案需要的数据再次类似于保护信息而存储于SchemeInformationBox中。

[0175] 请注意,可同时应用限制及保护。变换的次序遵循样本项的四字符码。举例来说,如果样本项类型为“resv”,那么撤销以上变换可引起样本项类型“encv”,从而指示媒体受到保护。

[0176] 请注意,如果文件作者仅想要提供咨询信息而不阻止旧版播放器播放文件,那么受限方案信息框可放置于样本项内部而不应用四字符码变换。在此状况下,不必要包含原始格式框。

[0177] 8.15.3受限方案信息框

[0178] 8.15.3.1定义

[0179] 框类型:“rinf”

[0180] 容器:受限样本项或样本项

[0181] 必选:是

[0182] 数量:仅一个

[0183] 受限方案信息框含有理解所应用的限制方案及其参数两者需要的所有信息。其还记录媒体的原始(未经变换)样本项类型。受限方案信息框为容器框。其在使用指示受限流的码的样本项(即,“resv”)中为必选的。

[0184] 当用于受限样本项中时,此框必须含有原始格式框以记录原始样本项类型及方案类型框。方案信息框取决于限制方案可被需要。

[0185] 8.15.3.2语法

```
aligned(8) class RestrictedSchemeInfoBox(fmt) extends Box('rinf') {  
    OriginalFormatBox(fmt)    original_format;  
[0186]    SchemeTypeBox            scheme_type_box;  
    SchemeInformationBox      info;                //任选  
}
```

[0187] 8.15.4针对立体视频布置的方案

[0188] 8.15.4.1通用

[0189] 当经立体译码的视频帧经解码时,经解码帧含有形成立体像对(帧封装)的两个空间封装构成帧的表示,或立体像对的仅一个视图(不同轨迹中的左视图及右视图)。归因于经立体译码视频的限制含在立体视频框中。

[0190] 使用方案类型“stvi”(立体视频)。

[0191] 8.15.4.2立体视频框

[0192] 8.15.4.2.1定义

[0193] 框类型:“stvi”

[0194] 容器:方案信息框(“schi”)

[0195] 必选:是(当方案类型为“stvi”时)

[0196] 数量:一个

[0197] 立体视频框用以指示经解码帧含有形成立体像对的两个空间封装构成帧的表示或含有立体像对的两个视图中的一者。立体视频框在方案类型为“stvi”时应存在。

[0198] 8.15.4.2.2语法

```
aligned(8) class StereoVideoBox extends extends FullBox('stvi', version = 0, 0) {  
    template unsigned int(30) reserved = 0;  
    unsigned int(2)      single_view_allowed;  
    unsigned int(32)      stereo_scheme;  
[0199]    unsigned int(32)      length;  
    unsigned int(8)[length] stereo_indication_type;  
    Box[] any_box; //任选  
}
```

[0200] 8.15.4.2.3语义

[0201] single\_view\_allowed为整数。零值指示内容仅可显示于立体显示器上。当 (single\_view\_allowed&1) 等于1, 允许在单像单一视图显示器上显示右视图。当 (single\_view\_allowed&2) 等于2时, 允许在单像单一视图显示器上显示左视图。

[0202] stereo\_scheme为整数, 其指示所使用的立体布置方案及根据所使用的方案的立体指示类型。stereo\_scheme的以下值经指定如下:

[0203] 1: 如由ISO/IEC 14496-10[ISO/IEC 14496-10]的帧封装布置补充增强信息消息指定的帧封装方案

[0204] 2: 如ISO/IEC 13818-2[ISO/IEC 13818-2:2000/修订4]的附录L中所指定的布置类型方案

[0205] 3: 如针对帧/服务兼容服务及2D/3D混合服务两者在ISO/IEC 23000-11中所指定的立体方案。

[0206] 保留stereo\_scheme的其它值。

[0207] 长度指示stereo\_indication\_type字段的字节的数目。

[0208] stereo\_indication\_type指示根据所使用的立体指示方案的立体布置类型。stereo\_indication\_type的语法及语义取决于stereo\_scheme的值。针对stereo\_scheme的以下值的stereo\_indication\_type的语法及语义指定如下:

[0209] stereo\_scheme等于1: 长度的值应为4, 且stereo\_indication\_type应为无正负号的int(32), 其含有来自ISO/IEC 14496-10[ISO/IEC 14496-10] (“frame\_packing\_arrangement\_type的定义”)的表D-8的frame\_packing\_arrangement\_type值。

[0210] stereo\_scheme等于2: 长度的值应为4, 且stereo\_indication\_type应为无正负号的int(32), 其含有ISO/IEC 13818-2[ISO/IEC 13818-2:2000/修正4] (“arrangement\_type的定义”)的表L-1的类型值。

[0211] stereo\_scheme等于3: 长度的值应为2, 且stereo\_indication\_type应含有无正负号Int(8)的两个语法元素。第一语法元素应含有来自ISO/IEC 23000-11:2009的表4的立体组成类型。第二语法元素的最低有效位应含有如ISO/IEC 23000-11:2009的8.4.3中指定

的is\_left\_first的值,而其它位经保留且应设置为0。

[0212] 当使用立体视频框时,以下情形适用:

[0213] 在轨迹标头框中

[0214] 宽度及高度指定单一视图在解封装之后的视觉呈现大小。

[0215] 在样本描述框中

[0216] frame\_count应为1,这是因为解码器物理上输出单一帧。换言之,包含于帧封装图片内的构成帧不由frame\_count记录。

[0217] 宽度及高度记录帧封装图片的像素计数(且不记录帧封装图片内单一视图的像素计数)。

[0218] 当视图显示于单像单一视图显示器上时,像素纵横比框记录每一视图的像素纵横比。举例来说,在许多空间帧封装布置中,像素纵横比框因此指示2:1或1:2的像素纵横比,这是由于帧封装视频的一个视图的空间分辨率相较于同一格式的单一视图视频的所述空间分辨率通常沿着一个坐标轴线被减半。

[0219] 8.15.5针对虚拟现实视频的方案

[0220] 8.15.5.1通用

[0221] 虚拟现实(VR)视频的经解码图片需要专用显现过程。关于VR视频的信息含在VR视频框中。

[0222] 使用方案类型“vrvd”(VR视频)。

[0223] 8.15.5.2 VR视频框

[0224] 8.15.5.2.1定义

[0225] 框类型:“vrvd”

[0226] 容器:方案信息框(“schi”)

[0227] 必选:是(当方案类型为“vrvd”时)

[0228] 数量:一个

[0229] VR视频框用以指示含在轨迹中的视频为VR视频。VR视频框在方案类型为“vrvd”时应存在。

[0230] 8.15.5.2.2语法

[0231] aligned(8) class VrVideoBox extends extends FullBox('vrvd', version = 0, 0) {  
template unsigned int(28) reserved = 0;

unsigned int(4) vr\_mapping\_type;

[0232] Box[] any\_box; //任选

}

[0233] 8.15.5.2.3语义

[0234] vr\_mapping\_type为整数,所述整数指示从球体视频到矩形格式的映射类型。零值指示等距矩形映射。值1指示立方体映射。值2指示角锥体映射。将其它值保留。

[0235] 图9说明如本文中所描述的用于产生含有虚拟现实内容的文件的过程900的实例。在902处,过程900包含获得虚拟现实数据,其中所述虚拟现实数据表示虚拟环境的360度视图。在一些实施方案中,虚拟现实数据包含虚拟现实视频。在一些实施方案中,虚拟现实



视频可经预拼接。在一些实施方案中,虚拟现实视频中的帧可能未经组装,且可能需要后拼接。

[0236] 在904处,过程900包含将虚拟现实数据存储在文件,其中所述虚拟现实数据根据文件格式存储,其中文件格式指定虚拟现实数据在文件内的放置且指定与虚拟现实数据相关联的信息在文件内的放置,其中与虚拟现实数据相关联的信息存储于轨迹框中。在各种实施方案中,文件格式为ISOBMFF或从ISOBMFF派生的文件格式。在一些实施方案中,与虚拟现实数据相关联的信息可包含(例如)帧速率、分辨率、文件内或视频和/或音频样本的其它文件内的位置,和/或其它信息。

[0237] 在906处,过程900包含将样本项存储于轨迹框中,其中样本项与一或多个样本相关联,其中样本项包含受限方案信息框,其中受限方案信息框将一或多个样本的方案定义为虚拟现实方案,且其中一或多个样本包含虚拟现实数据的至少一部分。在一些实施方案中,受限方案信息框可包含方案类型框。在此些实施方案中,虚拟现实方案可使用方案类型框来定义。在一些实施方案中,受限方案信息框可包含方案信息框。在此些实施方案中,与虚拟现实数据中的视频相关的任选参数可使用方案信息框写入到文件中。举例来说,任选参数包含(例如)虚拟现实数据为2-D视频或3-D视频,视频经预拼接或需要后拼接,和/或针对视频的映射。

[0238] 在一些实施方案中,过程900还可包含将与虚拟现实视频相关的参数(例如,上述任选参数)存储在文件。在一些实施方案中,参数存储于方案信息框中。

[0239] 图10说明如本文中所描述的用于处理含有虚拟现实内容的文件的过程1000的实例。在1002处,过程包含获得文件,其中文件包含虚拟现实数据,其中虚拟现实数据根据文件格式存储于文件中,其中文件格式指定虚拟现实内容在文件内的放置及与虚拟现实数据相关联的信息在文件内的放置,其中与虚拟现实数据相关联的信息存储于轨迹框内。在各种实施方案中,文件格式是基于ISOBMFF格式。

[0240] 在1004处,过程包含处理文件,其中处理文件包含处理轨迹框中的样本项,其中样本项与一或多个样本相关联,其中样本项包含受限方案信息框,其中受限方案信息框将一或多个样本的方案定义为虚拟现实方案,且其中一或多个样本包含虚拟现实数据的至少一部分。在一些实施方案中,处理文件进一步包含处理样本项中的方案类型框。在此些实施方案中,方案类型框包含定义为虚拟现实方案的方案类型。

[0241] 在一些实施方案中,处理文件包含处理与虚拟现实数据相关的参数。参数也存储于文件中。在各种实施方案中,所述参数可包含以下各者中的一或多个:虚拟现实数据包含2维视频数据还是3维视频数据的指示、所述视频数据为预拼接视频数据还是后拼接视频数据的指示,或针对所述视频数据的映射类型。

[0242] 在1006处,过程包含根据虚拟现实方案处理一或多个样本。在各种实施方案中,处理样本可包含产生视频帧以供能够显示虚拟现实内容的装置显示。

[0243] 图11说明如本文中所描述的用于产生含有虚拟现实内容的文件的过程1100的实例。在1102处,过程1100包含获得虚拟现实数据,其中所述虚拟现实数据表示虚拟环境的360度视图。虚拟现实数据可包含视频数据及音频数据。

[0244] 在1104处,过程1100包含将虚拟现实数据写入到文件,其中根据文件格式写入虚拟现实数据,其中文件格式指明虚拟现实数据及描述虚拟现实数据的信息在文件内的放

置,其中描述虚拟现实数据的信息被写入到轨迹框中。

[0245] 在1106处,过程1100包含设置轨迹框中的值,其中所述值指示通过轨迹框参考的样本针对虚拟现实呈现格式化。虚拟现实呈现可包含将存储于文件中的视频数据映射到球形360度格式。视频数据接着可使用经配置用于检视360度视频的装置来显示。

[0246] 在一些实施方案中,指示样本经格式化用于虚拟现实呈现的值可在媒体处理程序框中设置。在这些实施方案中,媒体处理程序框可为视频处理程序框。在一些实施方案中,媒体处理程序框更具体来说为虚拟现实视频处理程序框。

[0247] 在一些实施方案中,值可被写入到样本项框中。在这些实施方案中,值可被写入到轨迹标头框、媒体标头框或视频媒体标头框中外加其它实例。

[0248] 在各种实施方案中,与虚拟现实数据相关的参数也可写入到文件。举例来说,此些参数可包含虚拟现实数据包含2维视频数据还是3维视频数据,视频数据经预拼接还是后拼接,或针对视频数据的映射类型外加其它实例。

[0249] 在各种实施方案中,文件格式可基于ISO基本媒体文件格式。

[0250] 在一些实例中,过程900、1000及1100可由例如系统100等计算装置或设备来执行。举例来说,过程900、1000和/或1100可由展示于图1中的系统100和/或存储器108 或输出110来执行。在一些状况下,计算装置或设备可包含处理器、微处理器、微计算机,或经配置以进行过程900、1000或1100的步骤的装置的其它组件。在一些实例中,计算装置或设备可包含经配置以捕获包含视频帧的视频数据(例如,视频序列)的摄像机。举例来说,计算装置可包含摄像机装置(例如,IP摄像机或其它类型的摄像机装置),所述摄像机装置可包含视频编码解码器。在一些实例中,捕获视频数据的摄像机或其它捕获装置与计算装置分离,在此状况下,计算装置接收所捕获视频数据。计算装置可进一步包含经配置以传达视频数据的网络接口。网络接口可经配置以传达基于因特网协议(IP)的数据。

[0251] 过程900、1000及1100被说明为逻辑流程图,所述过程的操作表示可以硬件、计算机指令或其组合实施的一系列操作。在计算机指令的上下文中,操作表示存储于一或多个计算机可读存储媒体上的计算机可执行指令,所述指令当通过一或多个处理器执行时执行所叙述的操作。通常,计算机可执行指令包含执行特定功能或实施特定数据类型的例程、程序、对象、组件、数据结构等等。描述操作的次序并不打算被解译为限制,且任何数目个经描述操作可按任何次序和/或并行地组合以实施过程。

[0252] 另外,过程900、1000及1100可在经配置有可执行指令的一或多个计算机系统的控制下执行,且可被实施为共同于一或多个处理器上执行的代码(例如,可执行指令、一或多个计算机程序或一或多个应用程序)、通过硬件实施,或其组合。如上文所提到,代码,例如以包括可由一或多个处理器执行的多个指令的计算机程序的形式,可存储于计算机可读或机器可读存储媒体上。计算机可读或机器可读存储媒体可为非暂时性的。

[0253] 编码装置1204及解码装置1312的特定细节分别展示于图12及图13中。图12为说明可实施本发明中所描述的技术中的一或多者的实例编码装置1204的框图。编码装置1204可(例如)产生本文中所描述的语法结构(例如,VPS、SPS、PPS或其它语法元素的语法结构)。编码装置1204可执行视频切片内的视频块的帧内预测及帧间预测译码。如先前所描述,帧内译码至少部分地依赖于空间预测以减少或去除给定视频帧或图片内的空间冗余。帧间译码至少部分地依赖于时间预测以减少或去除视频序列的邻近或周围帧内的时间冗余。帧内模

式(I模式)可指若干基于空间的压缩模式中的任一者。帧间模式(例如,单向预测(P模式)或双向预测(B模式))可指若干基于时间的压缩模式中的任一者。

[0254] 编码装置1204包含分割单元35、预测处理单元41、滤波器单元63、图片存储器 64、求和器50、变换处理单元52、量化单元54及熵编码单元56。预测处理单元41包含运动估计单元42、运动补偿单元44及帧内预测处理单元46。对于视频块重构建,编码装置1204还包含反量化单元58、反变换处理单元60及求和器62。滤波器单元63打算表示一或多个环路滤波器,例如解块滤波器、自适应环路滤波器(ALF)及样本自适应偏移(SAO)滤波器。虽然滤波器单元63在图12中展示为环路滤波器,但在其它配置中,滤波器单元63可实施为后环路滤波器。后处理装置57可对由编码装置1204产生的经编码视频数据执行额外处理。本发明的技术可在一些情况下由编码装置1204实施。然而,在其它情况下,本发明的技术中的一或多个者可由后处理装置57实施。

[0255] 如图12中所展示,编码装置1204接收视频数据,且分割单元35将数据分割成视频块。分割还可包含分割成切片、切片区段、小块或其它较大单元以及例如根据LCU及CU的四分树结构的视频块分割。编码装置1204通常说明将视频块编码于待编码的视频切片内的组件。切片可划分成多个视频块(且可能划分成被称作小块的视频块集合)。预测处理单元41可基于错误结果(例如,译码速率及失真等级,等等)选择多个可能译码模式中的一者(例如,多个帧内预测译码模式中的一者或多个帧间预测译码模式中的一者)以用于当前视频块。预测处理单元41可将所得经帧内或帧间译码块提供到求和器50以产生残余块数据并提供到求和器62以重构建经编码块以用作参考图片。

[0256] 预测处理单元41内的帧内预测处理单元46可执行当前视频块相对于在与待译码的当前块相同的帧或切片中的一或多个相邻块的帧内预测译码,以提供空间压缩。预测处理单元41内的运动估计单元42及运动补偿单元44执行当前视频块相对于一或多个参考图片中的一或多个预测性块的帧间预测性译码,以提供时间压缩。

[0257] 运动估计单元42可经配置以根据视频序列的预定模式来确定用于视频切片的帧间预测模式。预定模式可将序列中的视频切片指定为P切片、B切片或GPB切片。运动估计单元42及运动补偿单元44可高度集成,但出于概念目的而单独说明。由运动估计单元42执行的运动估计为产生运动向量的过程,所述运动向量估计视频块的运动。运动向量(例如)可指示当前视频帧或图片内的视频块的预测单元(PU)相对于参考图片内的预测性块的位移。

[0258] 预测性块为就像素差而言被发现紧密地匹配待译码的视频块的PU的块,所述像素差可由绝对差和(SAD)、平方差和(SSD)或其它差度量确定。在一些实例中,编码装置1204可计算存储于图片存储器64中的参考图片的子整数像素位置的值。举例来说,编码装置1204可内插所述参考图片的四分之一像素位置、八分之一像素位置或其它分率像素位置的值。因此,运动估计单元42可执行相对于全像素位置及分率像素位置的运动搜索,且输出具有分率像素精确度的运动向量。

[0259] 运动估计单元42通过将PU的位置与参考图片的预测性块的位置比较而计算经帧间译码切片中的视频块的PU的运动向量。参考图片可从第一参考图片列表(列表0)或第二参考图片列表(列表1)选择,所述列表中的每一者识别存储于图片存储器64中的一或多个参考图片。运动估计单元42将经计算运动向量发送到熵编码单元56及运动补偿单元44。

[0260] 由运动补偿单元44执行的运动补偿可涉及基于通过运动估计确定的运动向量提

取或产生预测性块,可能执行内插达子像素精确度。在接收到当前视频块的PU的运动向量之后,运动补偿单元44可在参考图片列表中定位运动向量指向的预测性块。编码装置1204通过从正经译码的当前视频块的像素值减去预测性块的像素值从而形成像素差值来形成残余视频块。像素差值形成用于块的残余数据,且可包含明度差分量及色度差分量两者。求和器50表示执行此减法运算的一或多个组件。运动补偿单元44还可产生与视频块及视频切片相关联的语法元素,以供解码装置1312用于解码视频切片的视频块。

[0261] 作为如上文所描述由运动估计单元42及运动补偿单元44执行的帧间预测的替代,帧内预测处理单元46可对当前块进行帧内预测。特定来说,帧内预测处理单元46可确定帧内预测模式以用以编码当前块。在一些实例中,帧内预测处理单元46可(例如)在单独的编码遍次期间使用各种帧内预测模式来编码当前块,且帧内预测单元处理46(或在一些实例中,模式选择单元40)可从所测试的模式选择待使用的适当帧内预测模式。举例来说,帧内预测处理单元46可使用对各种所测试帧内预测模式的速率-失真分析来计算速率-失真值,且可在所测试模式间选择具有最优选速率-失真特性的帧内预测模式。速率-失真分析大体上确定经编码块与原始未经编码块(其经编码以产生经编码块)之间的失真(或误差)量,以及用以产生经编码块的位速率(即,位的数目)。帧内预测处理单元46可根据各种经编码块的失真及速率计算比率以确定哪一帧内预测模式展现所述块的最优选速率-失真值。

[0262] 在任何状况下,在选择用于块的帧内预测模式后,帧内预测处理单元46可将指示用于块的选定帧内预测模式的信息提供到熵编码单元56。熵编码单元56可编码指示选定帧内预测模式的信息。编码装置1204可将各种块的编码上下文的定义以及待用于上下文中的每一者的最可能的帧内预测模式、帧内预测模式索引表及经修改帧内预测模式索引表的指示包含于经传输位流配置数据中。位流配置数据可包含多个帧内预测模式索引表及多个经修改帧内预测模式索引表(也被称作码字映射表)。

[0263] 在预测处理单元41经由帧间预测或帧内预测产生当前视频块的预测性块之后,编码装置1204通过从当前视频块减去预测性块而形成残余视频块。残余块中的残余视频数据可包含于一或多个TU中且被应用于变换处理单元52。变换处理单元52使用例如离散余弦变换(DCT)或概念上类似变换的变换将残余视频数据变换成残余变换系数。变换处理单元52可将残余视频数据从像素域转换到变换域(例如,频域)。

[0264] 变换处理单元52可将所得变换系数发送到量化单元54。量化单元54量化变换系数以进一步减少位速率。所述量化过程可减小与所述系数中的一些或所有相关联的位深度。可通过调整量化参数来修改量化程度。在一些实例中,量化单元54可接着执行对包含经量化变换系数的矩阵的扫描。替代性地,熵编码单元56可执行扫描。

[0265] 在量化之后,熵编码单元56对经量化变换系数进行熵编码。举例来说,熵编码单元56可执行上下文自适应可变长度译码(CAVLC)、上下文自适应二进制算术译码(CABAC)、基于语法的上下文自适应二进制算术译码(SBAC)、概率区间分割熵(PIPE)译码或另一熵编码技术。在由熵编码单元56进行熵编码之后,可将经编码位流传输到解码装置1312,或加以存档以供稍后由解码装置1312传输或检索。熵编码单元56还可对正经译码的当前视频切片的运动向量及其它语法元素进行熵编码。

[0266] 反量化单元58及反变换处理单元60分别应用反量化及反变换以重构建像素域中的残余块以供稍后用作参考图片的参考块。运动补偿单元44可通过将残余块添加到参考图

片列表内的参考图片中的一者的预测性块来计算参考块。运动补偿单元44还可将一或多个内插滤波器应用到经重建残余块,以计算子整数像素值以用于运动估计中。求和器62将经重建残余块添加到由运动补偿单元44所产生的运动补偿预测块以产生用于存储于图片存储器64中的参考块。参考块可由运动估计单元42及运动补偿单元44 用作参考块以帧间预测后续视频帧或图片中的块。

[0267] 以此方式,图12的编码装置1204表示经配置以产生经编码视频位流的语法的视频编码器的实例。编码装置1204可(例如)产生如上文所描述的VPS、SPS及PPS参数集。编码装置1204可执行本文中所描述的技术中的任一者,包含上文关于图12及图13描述的过程。大体上已关于编码装置1204描述了本发明的技术,但如上文所提及,本发明的技术中的一些也可由后处理装置57实施。

[0268] 图13为说明实例解码装置1312的框图。解码装置1312包含熵解码单元80、预测处理单元81、反量化单元86、反变换处理单元88、求和器90、滤波器单元91及图片存储器92。预测处理单元81包含运动补偿单元82及帧内预测处理单元84。在一些实例中,解码装置1312可执行大体上可逆于关于来自图12的编码装置1204所描述的编码遍次的解码遍次。

[0269] 在解码过程期间,解码装置1312接收表示由编码装置1204发送的经编码视频切片的视频块及相关联语法元素的经编码视频位流。在一些实施例中,解码装置1312可从编码装置1204接收经编码视频位流。在一些实施例中,解码装置1312可从网络实体79(例如,服务器、媒体感知网络元件(MANE)、视频编辑器/编接器(splicer)或经配置以实施上文所描述的技术中的一或多者的其它此类装置)接收经编码视频位流。网络实体79 可或可不包含编码装置1204。在网络实体79将经编码视频位流传输到解码装置1312之前,本发明中所描述的技术中的一些可由网络实体79实施。在一些视频解码系统中,网络实体79及解码装置1312可为独立的装置的部分,而在其它情况下,关于网络实体 79描述的功能性可由包括解码装置1312的同一装置执行。

[0270] 解码装置1312的熵解码单元80对位流进行熵解码以产生经量化系数、运动向量及其它语法元素。熵解码单元80将运动向量及其它语法元素转递到预测处理单元81。解码装置1312可接收视频切片层级和/或视频块层级的语法元素。熵解码单元80可处理并解析例如VPS、SPS及PPS的一或多个参数集中的固定长度语法元素及可变长度语法元素两者。

[0271] 当将视频切片译码为经帧内译码(I)切片时,预测处理单元81的帧内预测处理单元 84可基于经传信帧内预测模式及来自当前帧或图片的先前经解码块的数据而产生当前视频切片的视频块的预测数据。当视频帧被译码为经帧间译码(即,B、P或GPB)切片时,预测处理单元81的运动补偿单元82基于从熵解码单元80接收的运动向量及其它语法元素而产生当前视频切片的视频块的预测性块。所述预测性块可从参考图片列表内的参考图片中的一者产生。解码装置1312可基于存储于图片存储器92中的参考图片使用默认构建技术来构建参考帧列表(列表0及列表1)。

[0272] 运动补偿单元82通过解析运动向量及其它语法元素来确定当前视频切片的视频块的预测信息,且使用所述预测信息产生正经解码的当前视频块的预测性块。举例来说,运动补偿单元82可使用参数集中的一或多个语法元素来确定用于译码视频切片的视频块的预测模式(例如,帧内或帧间预测)、帧间预测切片类型(例如,B切片、P切片或GPB 切片)、切片的一或多个参考图片列表的构建信息、切片的每一经帧间编码视频块的运动向量、切片

的每一帧间译码视频块的帧间预测状态及用以解码当前视频切片中的视频块的其它信息。

[0273] 运动补偿单元82还可执行基于内插滤波器的内插。运动补偿单元82可使用如在编码视频块期间由编码装置1204使用的内插滤波器来计算参考块的子整数像素的内插值。在此状况下,运动补偿单元82可从所接收语法元素确定由编码装置1204使用的内插滤波器,且可使用内插滤波器来产生预测性块。

[0274] 反量化单元86反量化或解量化位流中所提供且由熵解码单元80解码的经量化变换系数。反量化过程可包含使用由编码装置1204针对视频切片中的每一视频块计算的量化参数来确定量化程度及同样应应用的反量化程度。反变换处理单元88将反变换(例如,反DCT或其它合适的反变换)、反整数变换或概念上类似的反变换过程应用于变换系数以便在像素域中产生残余块。

[0275] 在运动补偿单元82基于运动向量及其它语法元素而产生当前视频块的预测性块之后,解码装置1312通过将来自反变换处理单元88的残余块与由运动补偿单元82产生的对应预测性块求和而形成经解码视频块。求和器90表示执行此求和运算的一或多个组件。如果需要,还可使用环路滤波器(在译码环路中或在译码环路后)以使像素转变平滑,或以其它方式改进视频质量。滤波器单元91打算表示一或多个环路滤波器,例如,解块滤波器、自适应环路滤波器(ALF)及样本自适应偏移(SAO)滤波器。虽然滤波器单元91在图13中展示为环路滤波器,但在其它配置中,滤波器单元91可实施为后环路滤波器。给定帧或图片中的经解码视频块随后存储于图片存储器92中,图片存储器92存储用于后续运动补偿的参考图片。图片存储器92还存储经解码视频以供稍后呈现于显示装置(例如,图1中所展示的视频目的地装置122)上。

[0276] 在前述描述中,参考其具体实施例描述申请的方面,但所属领域的技术人员将认识到本发明不限于此。因此,尽管已在本文中详细地描述申请的说明性实施例,但应理解,本发明概念可以其它方式不同地具体化并使用,且打算解译所附权利要求书以包含除现有技术所限制外的此些变化。上文所描述的发明的各种特征及方面可单独或联合地使用。另外,实施例可用于超出本文所描述的那些环境及应用的任何数目个环境及应用中,而不脱离本说明书的更广精神及范围。因此,本说明书及图式被视为说明性而非限制性的。出于说明的目的,以特定次序描述方法。应了解,在替代实施例中,可以与所描述的次序不同的次序来执行所述方法。

[0277] 在组件被描述为“经配置以”执行某些操作的情况下,此配置可(例如)通过设计电子电路或其它硬件以执行操作、通过编程可编程电子电路(例如,微处理器或其它适合的电子电路)以执行操作或其任何组合来实现。

[0278] 结合本文中所公开的实施例而描述的各种说明性逻辑块、模块、电路及算法步骤可实施为电子硬件、计算机软件或两者的组合。为了清楚地说明硬件与软件的此可互换性,各种说明性组件、块、模块、电路及步骤已在上文大体按其功能性加以描述。将此类功能性实施为硬件还是软件取决于特定应用及强加于整个系统的设计约束。所属领域的技术人员可针对每一特定应用以变化方式实施所描述功能性,但此些实施决策不应被解译为导致脱离本发明的范围。

[0279] 本文中描述的技术可以硬件、软件、固件或其任何组合来实施。此些技术可实施于

多种装置中的任一者中,例如,实施于通用计算机、无线通信装置手机或具有多种用途(包含在无线通信装置手机及其它装置中的应用)的集成电路装置中。可将描述为模块或组件的任何特征一起实施于集成式逻辑装置中或分离地实施为离散但可互操作的逻辑装置。如果以软件实施,那么所述技术可至少部分由包括包含指令的程序代码的计算机可读数据存储媒体实现,所述指令在经执行时执行上文所描述的方法中的一或多者。计算机可读数据存储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包装材料。计算机可读媒体可包括存储器或数据存储媒体,例如,随机接入存储器(RAM),例如,同步动态随机接入存储器(SDRAM);只读存储器(ROM);非易失性随机接入存储器(NVRAM);电可擦除可编程只读存储器(EEPROM);闪存器;磁性或光学数据存储媒体;等等。所述技术另外或替代地可至少部分由计算机可读通信媒体实现,所述计算机可读通信媒体携带或传达呈指令或数据结构的形式且可由计算机接入、读取和/或执行的程序代码,例如,传播的信号或波。

[0280] 程序代码可由可包含例如以下各者的一或多个处理器的处理器执行:一或多个数字信号处理器(DSP)、通用微处理器、集成电路专用集成电路(ASIC)、现场可编程逻辑阵列(FPGA)或其它等效积体或离散逻辑电路。此处理器可经配置以执行本发明中所描述的技术中的任何者。通用处理器可为微处理器;但在替代例中,处理器可为任何常规处理器、控制器、微控制器或状态机。处理器还可实施为计算装置的组合,例如DSP与微处理器的组合、多个微处理器、结合DSP核心的一或多个微处理器,或任何其它此配置。因此,如本文中所使用的术语“处理器”可指前述结构中的任一者、前述结构的任何组合或适合于实施本文中描述的技术的任何其它结构或设备。此外,在一些方面中,本文中描述的功能性可提供于经配置用于编码及解码的专用软件模块或硬件模块内,或并入于组合的视频编码器-解码器(编码解码器)中。

[0281] 本文中所论述的译码技术可体现于实例视频编码及解码系统中。系统包含通过目的地装置提供待稍后经解码的经编码视频数据的源装置。确切地说,源装置经由计算机可读媒体将视频数据提供到目的地装置。源装置及目的地装置可包括广泛范围的装置中的任一者,包含台式计算机、笔记型(即,膝上型)计算机、平板计算机、机顶盒、电话手持机(例如所谓的“智能”电话)、所谓的“智能”平板计算机、电视机、摄像机、显示装置、数字媒体播放器、视频游戏控制台、视频流式传输装置等等。在一些状况下,源装置及目的地装置可经装备以用于无线通信。

[0282] 目的地装置可经由计算机可读媒体接收待解码的经编码视频数据。计算机可读媒体可包括能够将经编码视频数据从源装置移动到目的地装置的任何类型的媒体或装置。在一个实例中,计算机可读媒体可包括通信媒体以使得源装置能够实时地将经编码视频数据直接传输到目的地装置。可根据通信标准(例如,无线通信协议)调制经编码视频数据,且将经编码视频数据传输到目的地装置。通信媒体可包括任何无线或有线通信媒体,例如,射频(RF)频谱或一或多个物理传输线。通信媒体可形成基于包的网路(例如局域网、广域网或全域网路例如因特网)的一部分。通信媒体可包含路由器、交换器、基站或可用于促进从源装置到目的地装置的通信的任何其它装备。

[0283] 在一些实例中,经编码数据可从输出接口输出到存储装置。类似地,可由输入接口从存储装置接入经编码数据。存储装置可包含多种分布式或本地接入的数据存储媒体中的任一者,例如,硬盘、蓝光光盘、DVD、CD-ROM、闪存器、易失性或非易失性存储器或用于存储



经编码视频数据的任何其它合适的数字存储媒体。在另一实例中,存储装置可对应于文件服务器或可存储由源装置产生的经编码视频的另一中间存储装置。目的地装置可经由流式传输或下载从存储装置接入所存储的视频数据。文件服务器可为能够存储经编码视频数据且将所述经编码视频数据传输到目的地装置的任何类型的服务器。实例文件服务器包含网页服务器(例如,用于网站)、FTP服务器、网络连接存储(NAS)装置或本地磁盘驱动器。目的地装置可经由任何标准数据连接(包含因特网连接)而接入经编码的视频数据。此数据连接可包含适于接入存储于文件服务器上的经编码视频数据的无线信道(例如,Wi-Fi连接)、有线连接(例如,DSL、缆线调制解调器等),或两者的结合。来自存储装置的经编码视频数据的传输可为流式传输、下载传输,或其组合。

[0284] 本发明的技术不必限于无线应用或设置。所述技术可应用于视频译码以支持多种多媒体应用中的任一者,例如,空中电视广播、有线电视传输、卫星电视传输、因特网流式传输视频传输(例如,经由HTTP的动态自适应流式传输(DASH))、经编码到数据存储媒体上的数字视频、存储于数据存储媒体上的数字视频的解码或其它应用。在一些实例中,系统可经配置以支持单向或双向视频传输以支持例如视频流式传输、视频播放、视频广播和/或可视电话的应用。

[0285] 在一个实例中,源装置包括视频源、视频编码器及输出接口。目的地装置可包含输入接口、视频解码器及显示装置。源装置的视频编码器可经配置以应用本文中所公开的技术。在其它实例中,源装置及目的地装置可包含其它组件或布置。举例来说,源装置可从外部视频源(例如,外部摄像机)接收视频数据。同样,目的地装置可与外部显示装置介接,而非包含集成式显示装置。

[0286] 以上实例系统仅为一个实例。用于并行地处理视频数据的技术可由任何数字视频编码和/或解码装置来执行。尽管本发明的技术通常由视频编码装置执行,但所述技术还可由视频编码器/解码器(通常被称为“编码解码器”)执行。此外,本发明的技术还可由视频预处理器执行。源装置及目的地装置仅为源装置产生经译码视频数据以供传输到目的地装置的此类译码装置的实例。在一些实例中,源装置及目的地装置可以大体上对称的方式操作,使得所述装置中的每一者包含视频编码及解码组件。因此,实例系统可支持视频装置之间的单向或双向视频传输,例如用于视频流式传输、视频播放、视频广播或可视电话。

[0287] 视频源可包含视频捕获装置,例如摄像机、含有先前捕获的视频的视频存档和/或用以从视频内容提供者接收视频的视频馈入接口。作为另一替代,视频源可产生基于计算机图形的数据作为源视频,或实况视频、存档视频及计算机产生的视频的组合。在一些状况下,如果视频源为摄像机,那么源装置及目的地装置可形成所谓的摄像机电话或可视电话。然而,如上文所提及,本发明中所描述的技术一般可适用于视频译码,且可适用于无线和/或有线应用。在每一状况下,可由视频编码器编码所捕获、经预捕获或计算机产生的视频。经编码视频信息可随后由输出接口输出到计算机可读媒体上。

[0288] 如所提及,计算机可读媒体可包含瞬时媒体,例如无线广播或有线网络传输;或存储媒体(即,非暂时性存储媒体),例如硬盘、闪存驱动器、紧密光盘、数字视频光盘、蓝光光盘或其它计算机可读媒体。在一些实例中,网络服务器(未图示)可从源装置接收经编码视频数据,且(例如)经由网络传输将经编码视频数据提供到目的地装置。类似地,媒体产生设施(例如光盘冲压设施)的计算装置可从源装置接收经编码视频数据且生产含有经编码视



频数据的光盘。因此,在各种实例中,计算机可读媒体可理解为包含各种形式的一或多个计算机可读媒体。

[0289] 目的地装置的输入接口从计算机可读媒体接收信息。计算机可读媒体的信息可包含由视频编码器定义的语法信息(其也由视频解码器使用),所述语法信息包含描述块及其它经译码单元(例如,图片群组(GOP))的特性和/或处理的语法元素。显示装置将经解码视频数据显示给用户,且可包括多种显示装置中的任一者,例如阴极射线管(CRT)、液晶显示器(LCD)、等离子显示器、有机发光二极管(OLED)显示器或另一类型的显示装置。已描述了本发明的各种实施例。

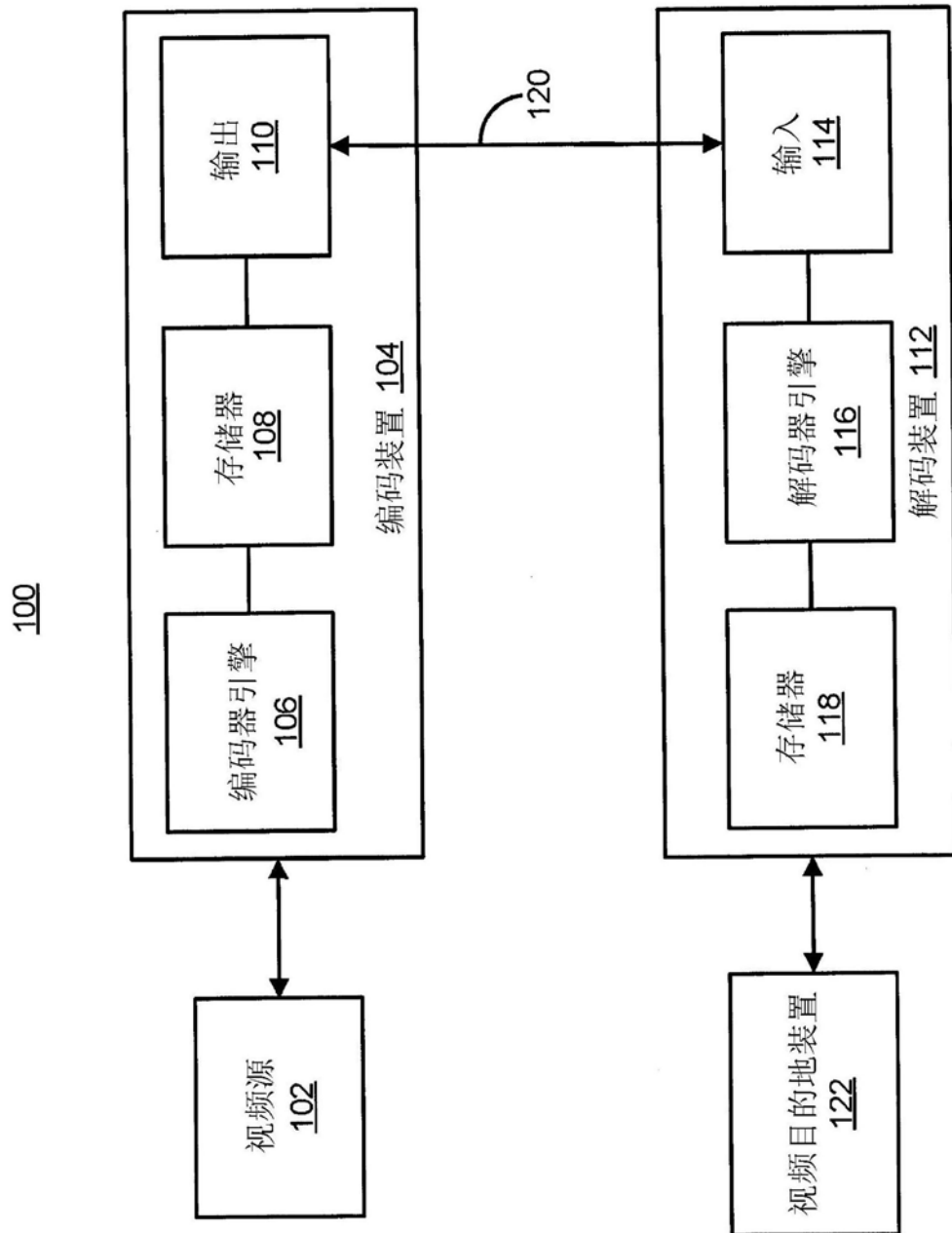


图1

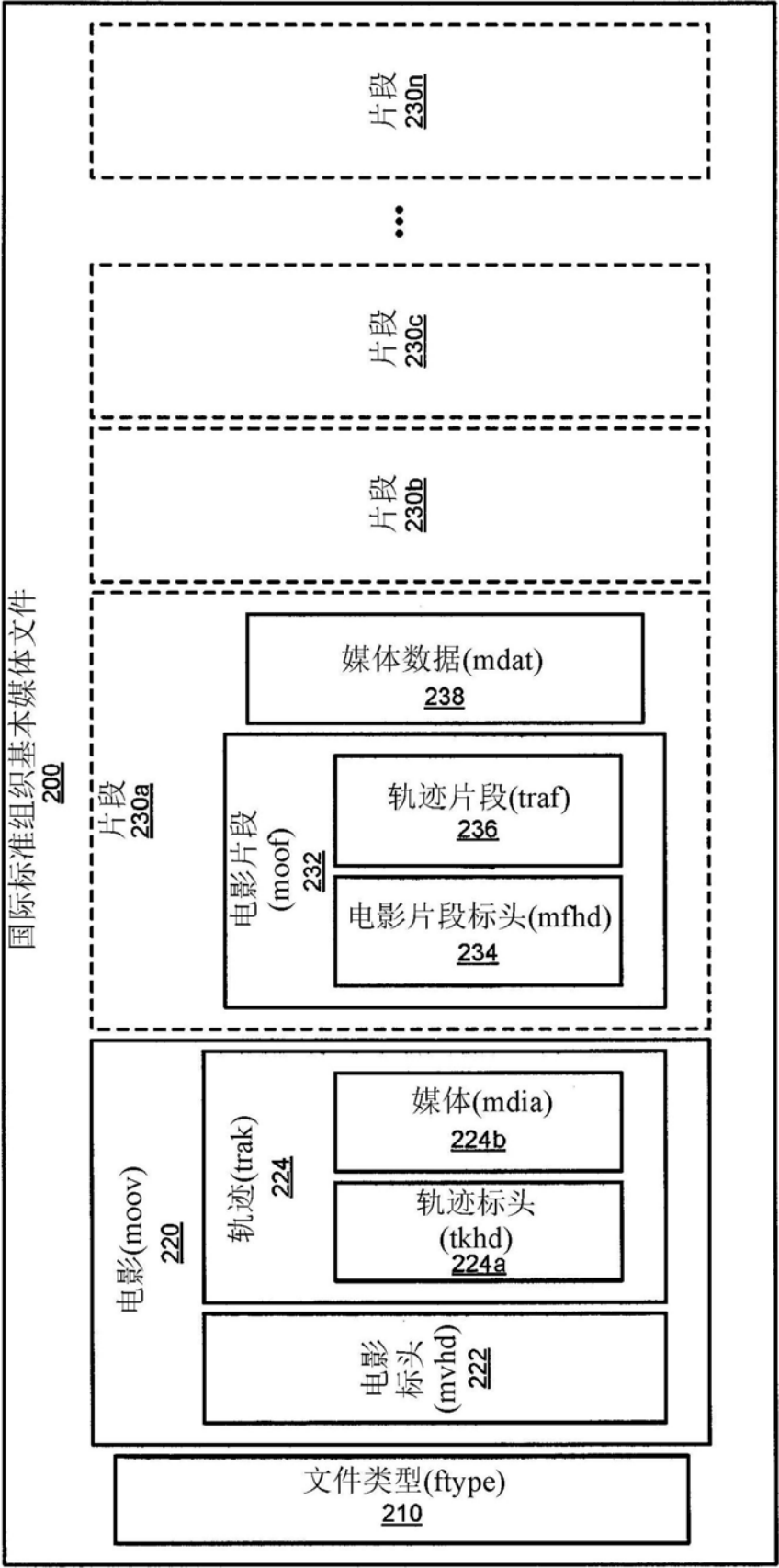


图2

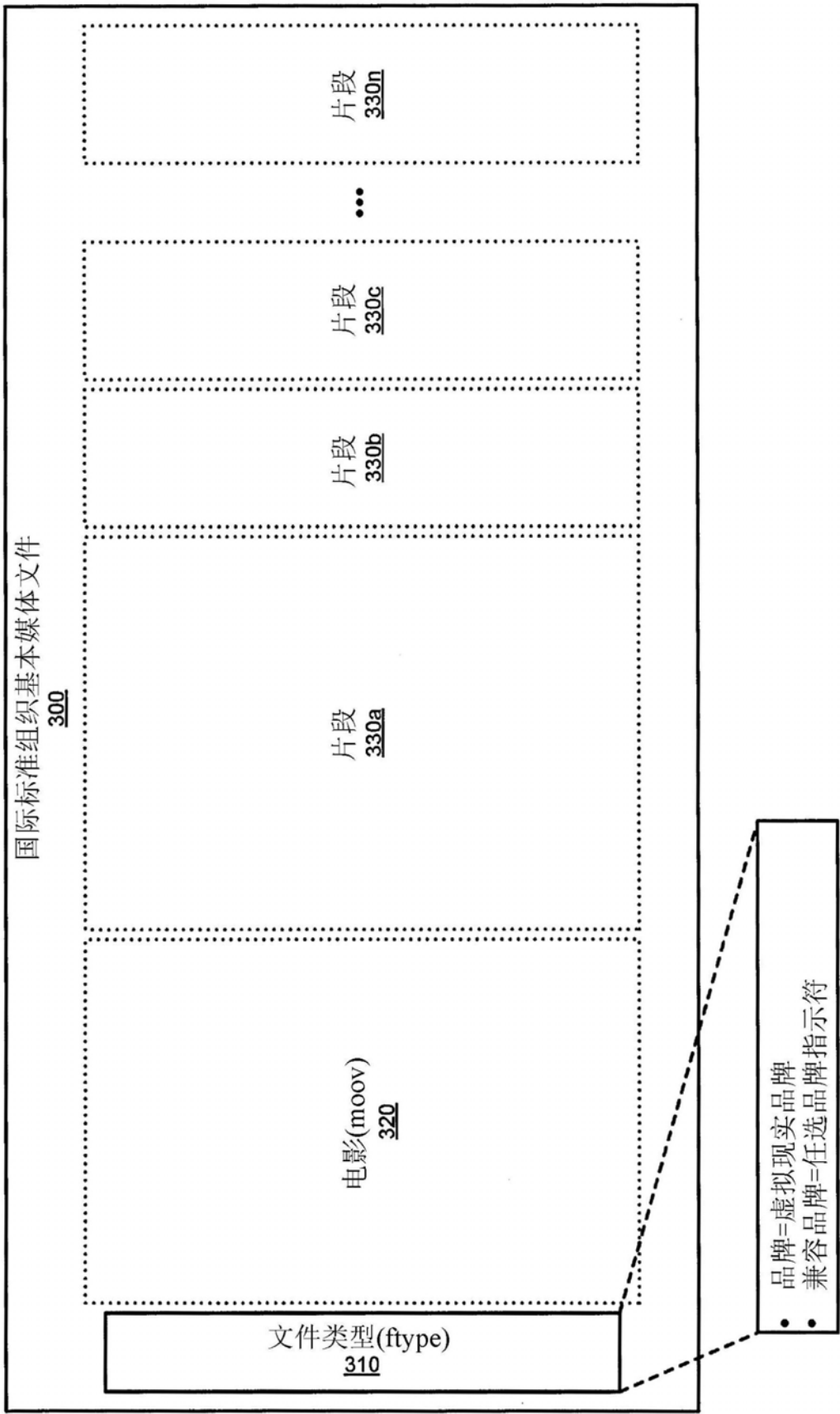


图3A

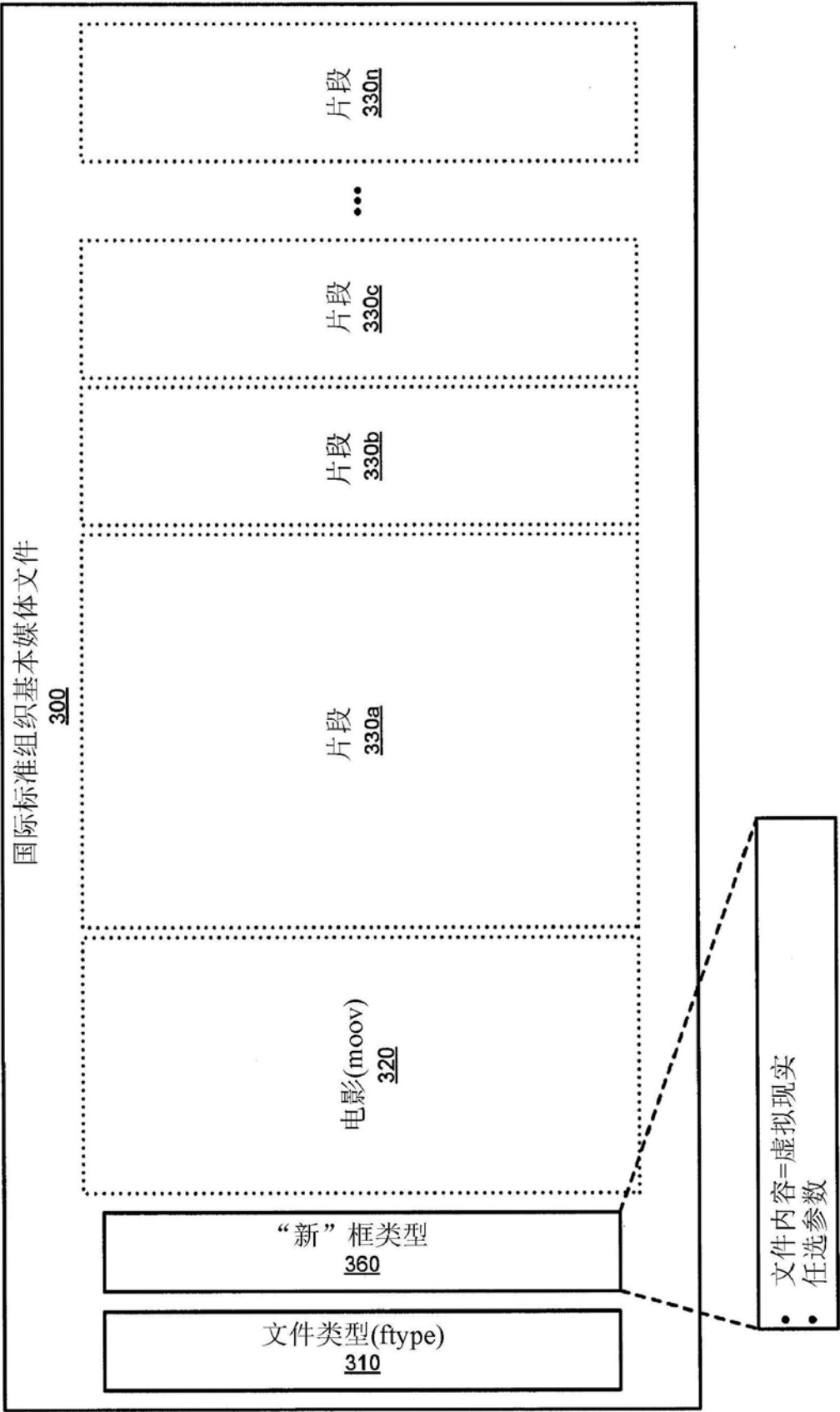


图3B

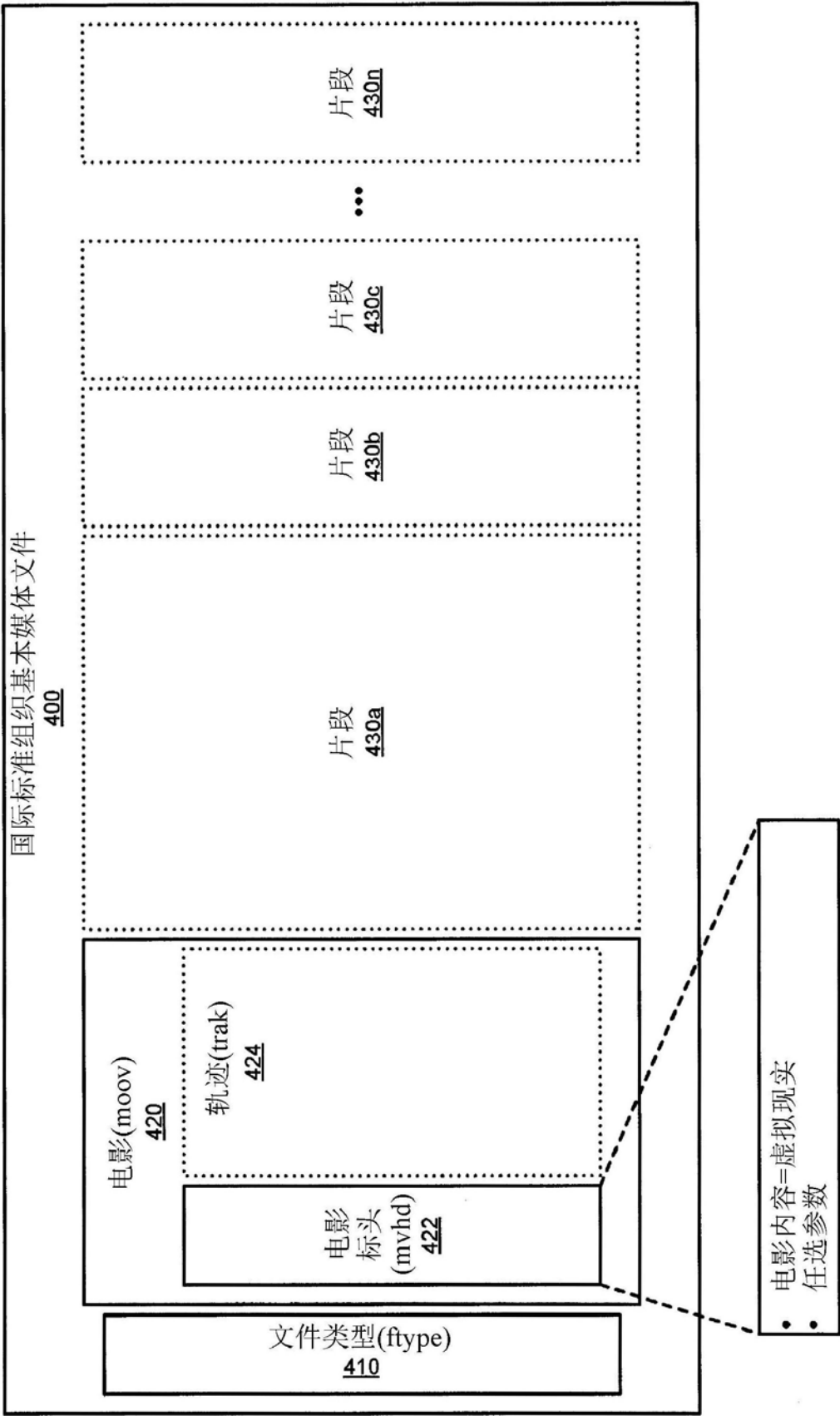


图4

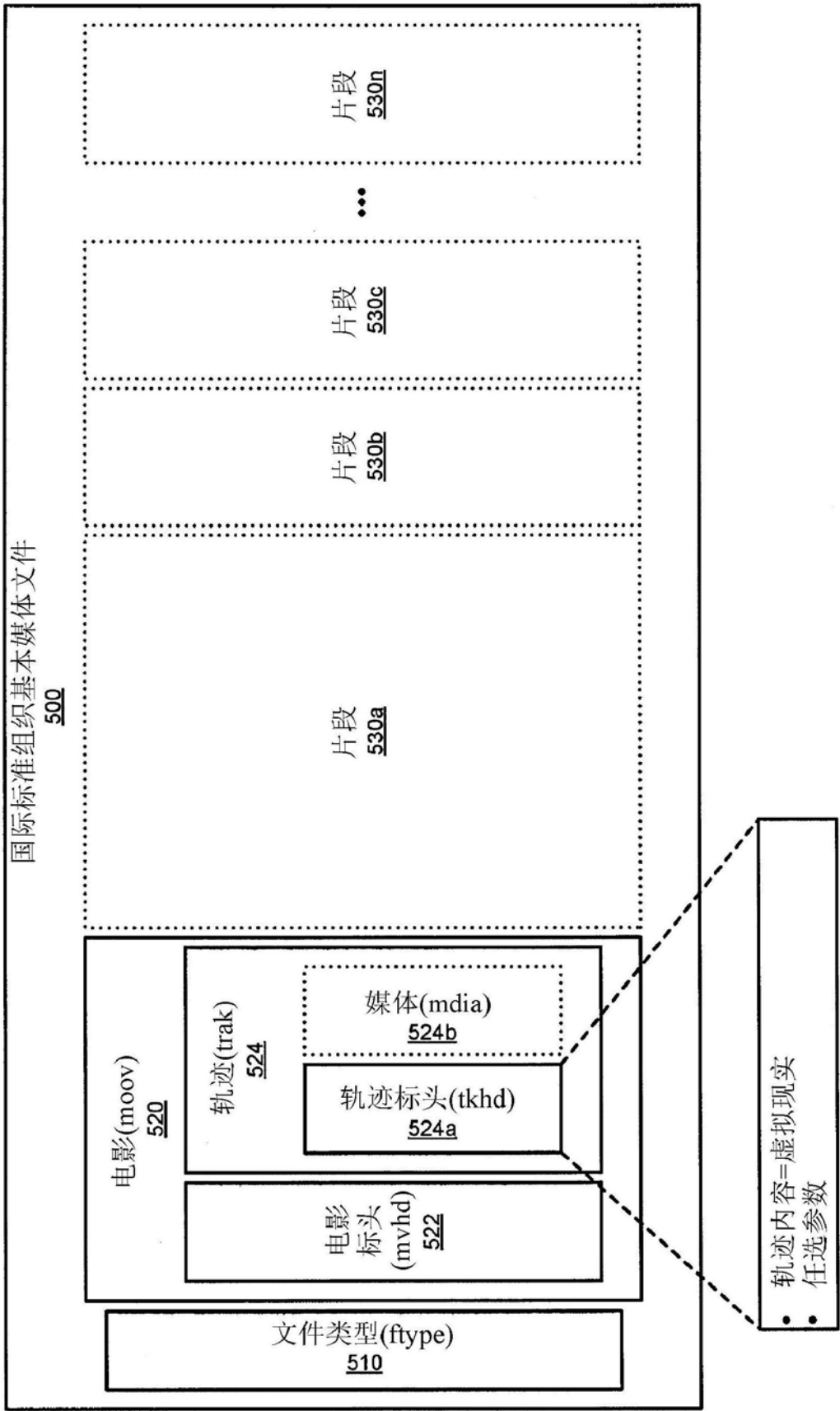


图5

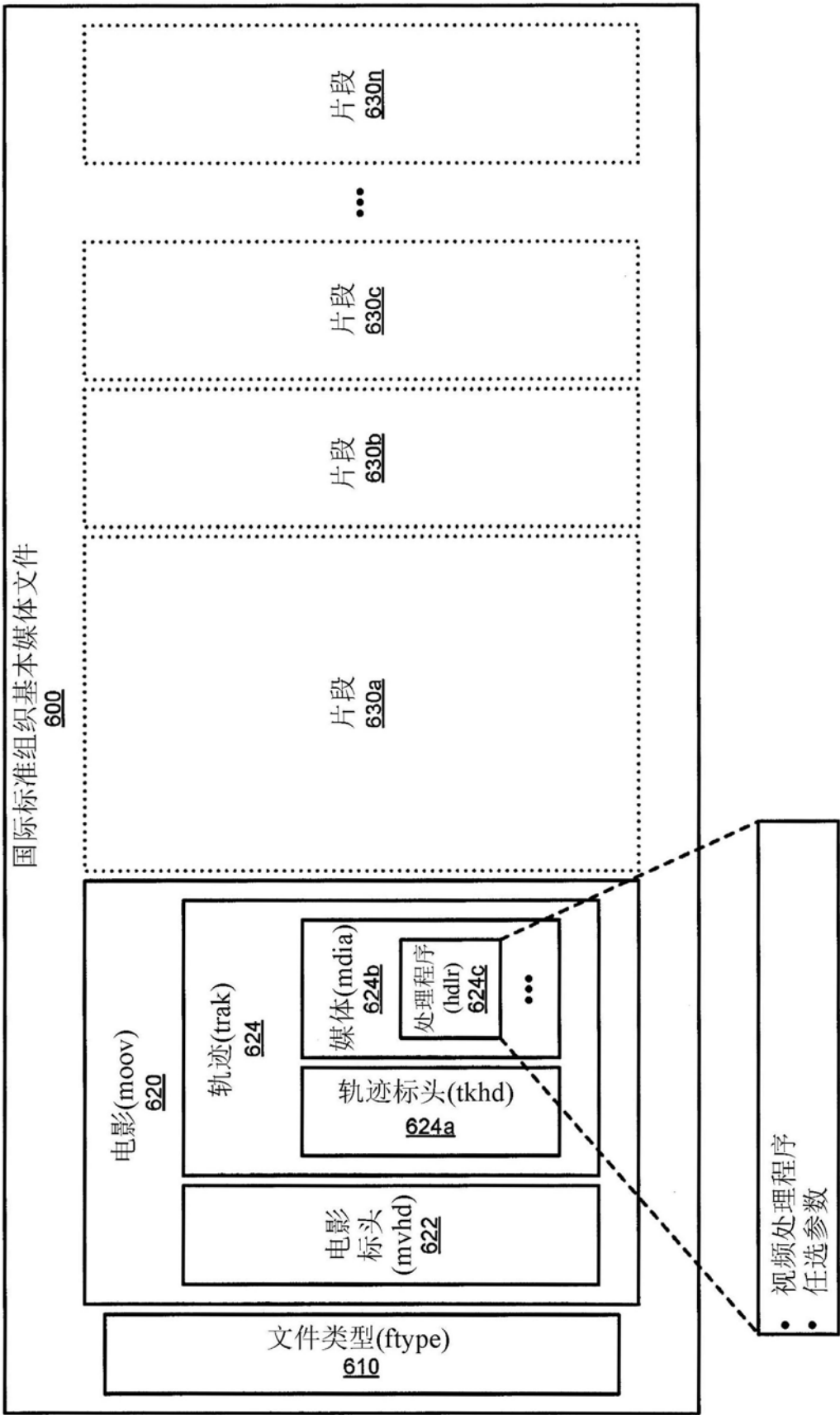


图6



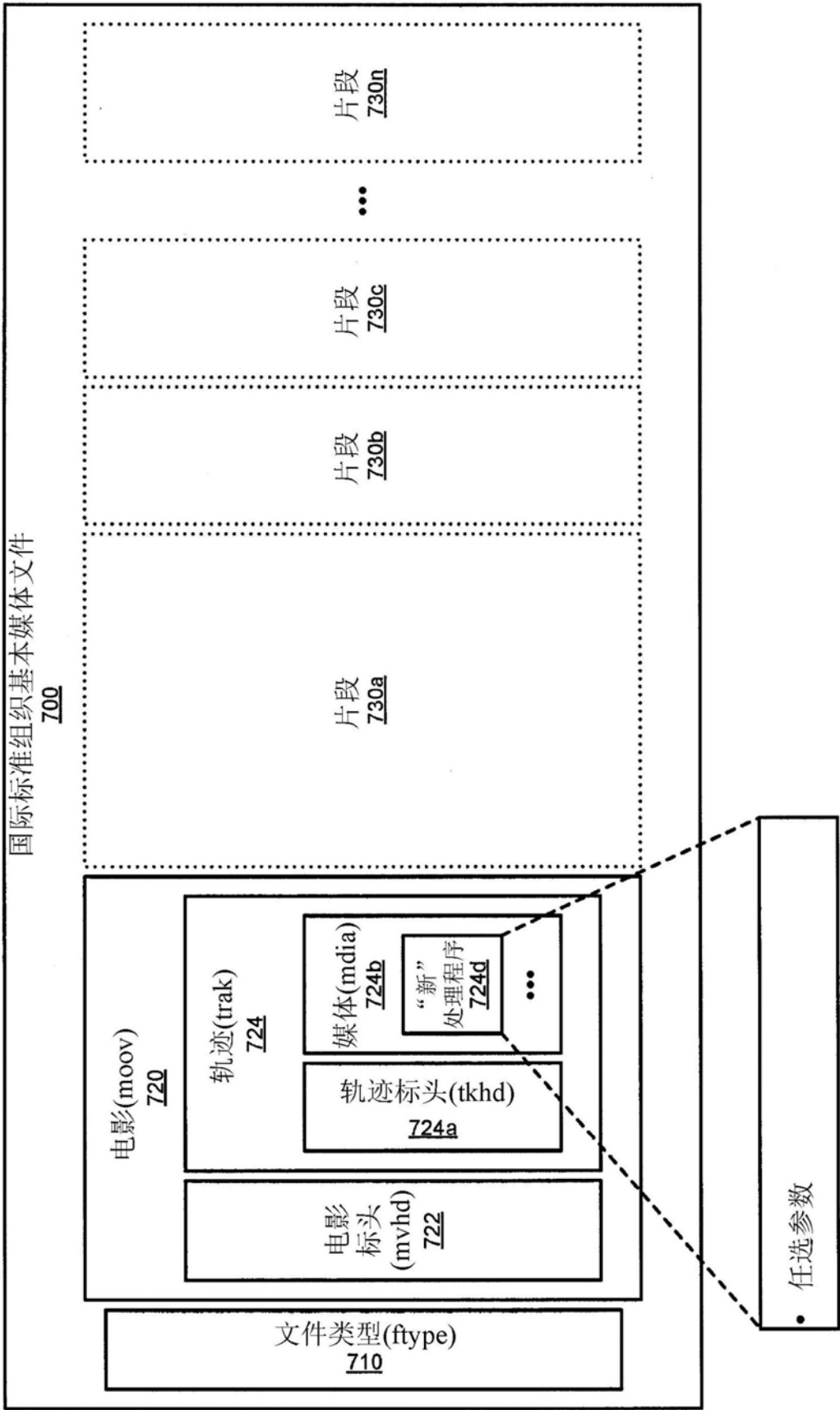


图7

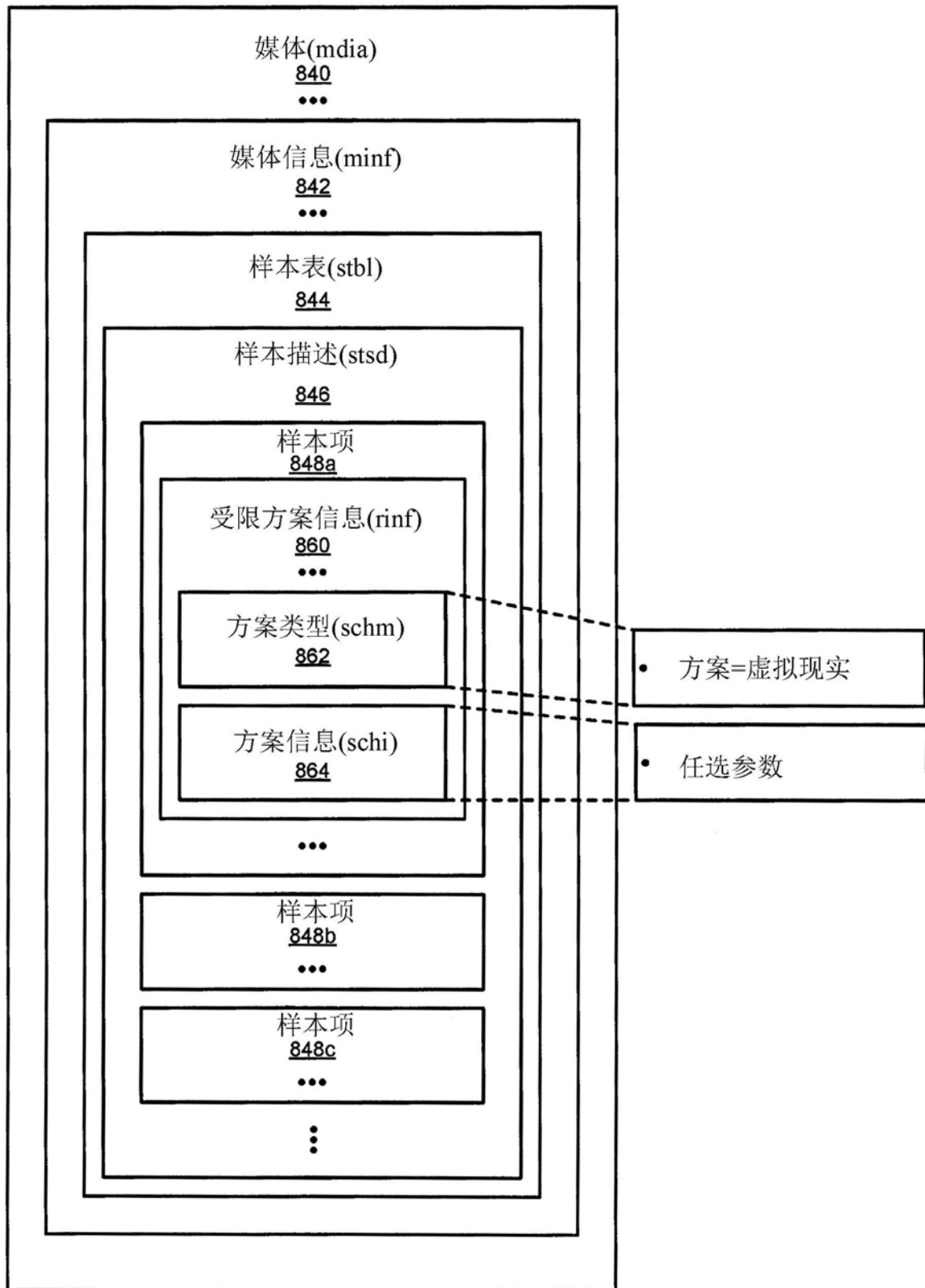


图8

900

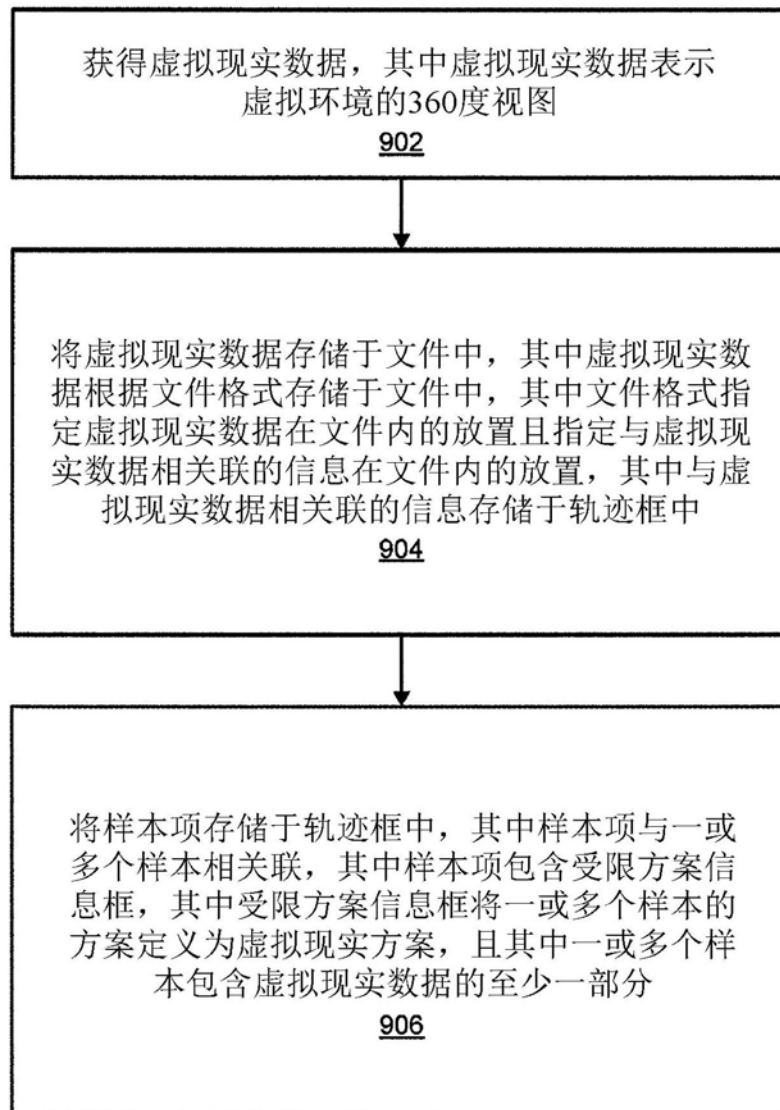


图9

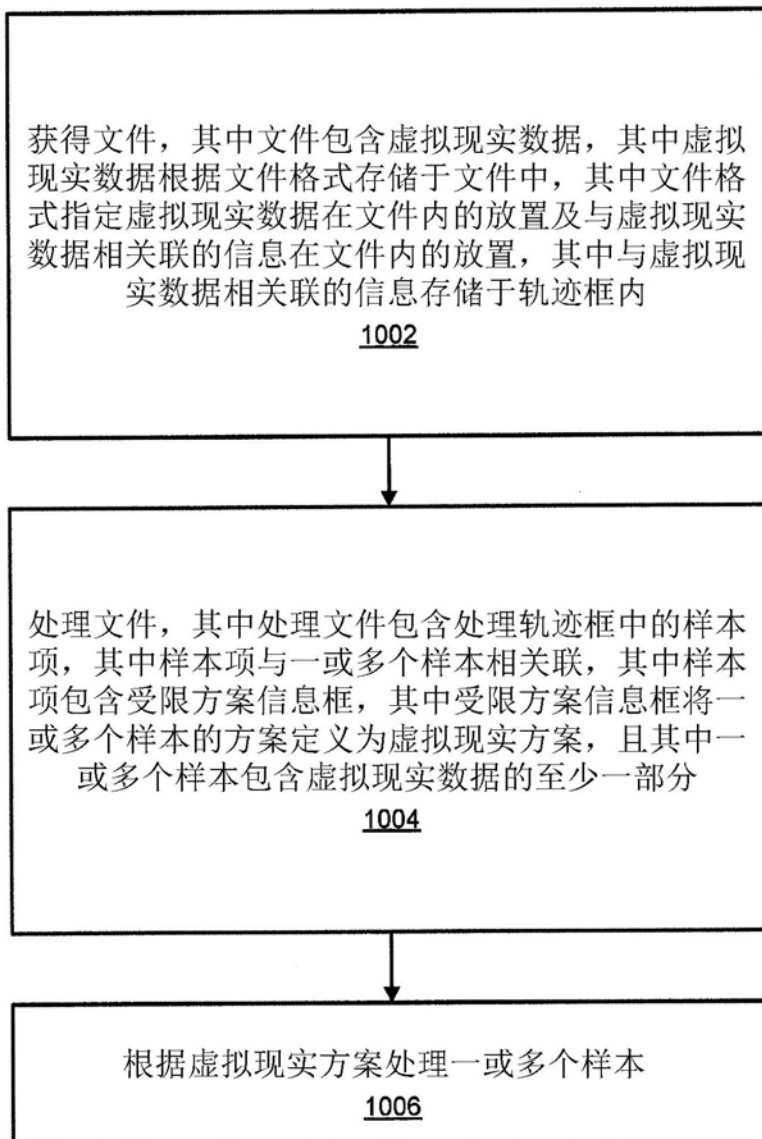

1000 

图10

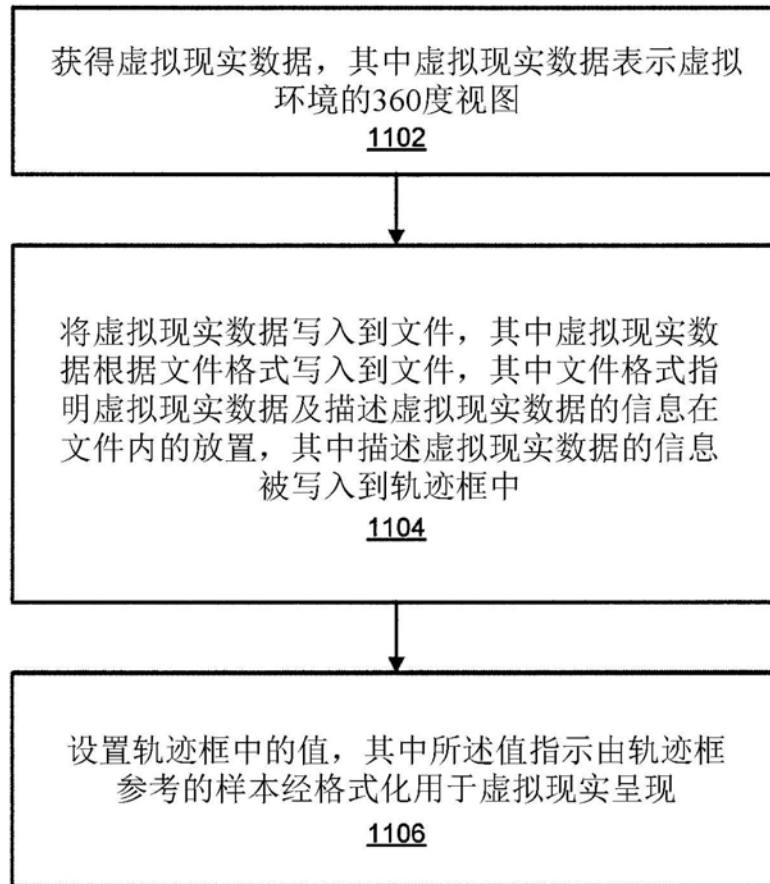

1100 

图11

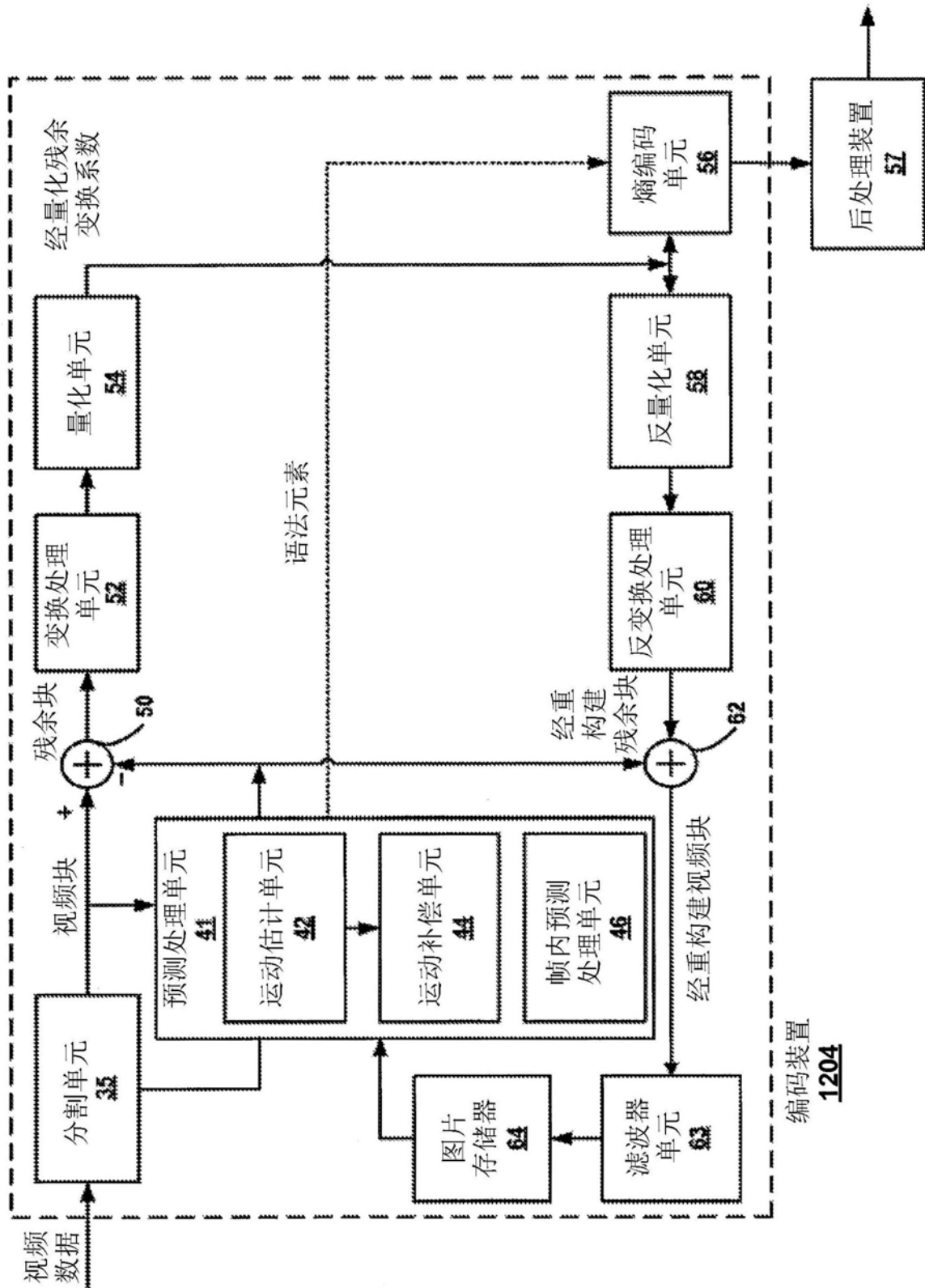


图12

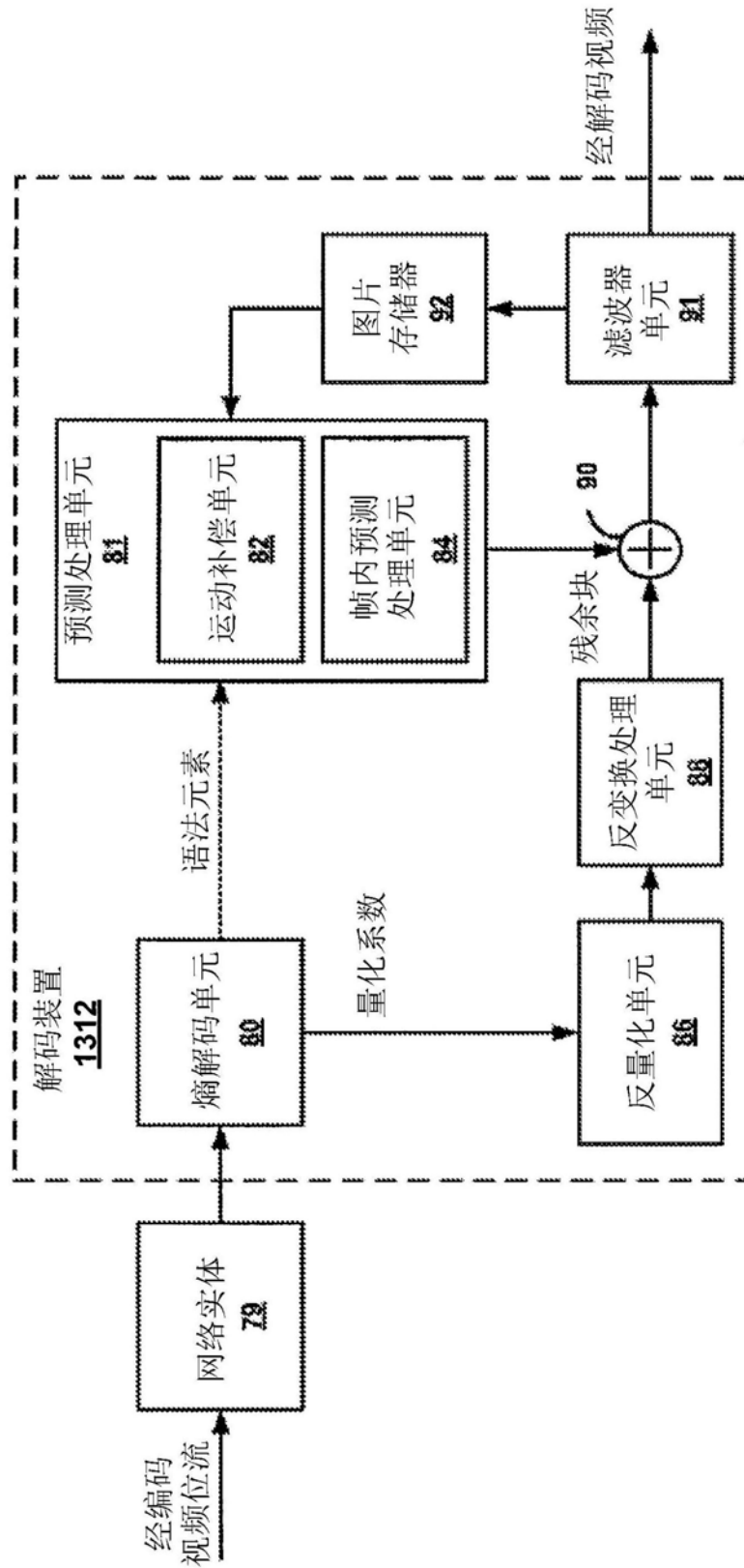


图13