

(12) 发明专利

(10) 授权公告号 CN 101711483 B
(45) 授权公告日 2012. 05. 30

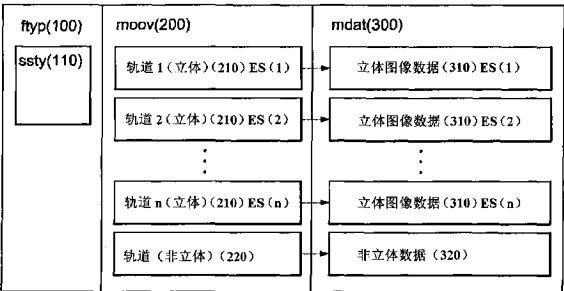
(21) 申请号 200880019052. 6
(22) 申请日 2008. 06. 05
(30) 优先权数据
10-2007-0055620 2007. 06. 07 KR
10-2007-0075212 2007. 07. 26 KR
(85) PCT申请进入国家阶段日
2009. 12. 07
(86) PCT申请的申请数据
PCT/KR2008/003145 2008. 06. 05
(87) PCT申请的公布数据
W02008/150111 EN 2008. 12. 11
(73) 专利权人 升级芯片技术公司
地址 韩国首尔
专利权人 庆熙大学校算学协力团
(72) 发明人 金奎宪 李润辰 朴光勋 徐德荣
千胜文 吴镕协 丁泰燮 边大燮
(74) 专利代理机构 北京润平知识产权代理有限公司 11283
代理人 周建秋 王凤桐

(51) Int. Cl.
H04N 13/00 (2006. 01)
(56) 对比文件
CN 1613263 A, 2005. 05. 04,
JP 2005-26800 A, 2005. 01. 27,
JP 2006-13759 A, 2006. 01. 12,
KR 100716142 B1, 2007. 05. 02,
CN 1947431 A, 2007. 04. 11,
审查员 胡雅琴

权利要求书 2 页 说明书 14 页 附图 6 页

(54) 发明名称
用于立体图像编码数据文件的格式

(57) 摘要
提供了一种用于构建立体图像编码数据文件的方法。所述立体图像编码数据文件包括文件类型声明单元, 该单元用于指示所述文件是否是立体图像; 元数据单元, 包括一个或多个轨道容器, 该轨道容器用于容纳立体图像编码数据的元数据; 和图像数据单元, 包括一个或多个立体图像数据容器, 该立体图像数据容器用于容纳立体图像编码数据的图像信息。



CN 101711483 B

1. 一种用于再现立体图像编码数据的方法,其中所述立体图像编码数据包括文件类型声明单元、元数据单元和图像数据单元,所述元数据单元包括一个或多个立体轨道容器,所述立体轨道容器包括媒体容器,所述媒体容器包括处理器参考框和媒体信息容器,所述媒体信息容器包括立体报头框和采样表框,所述图像数据单元包括一个或多个立体图像数据容器,所述方法包括:

从所述文件类型声明单元中获取关于所述立体图像编码数据的类型以及构成该立体图像编码数据的元素流的数量的信息;

从所述处理器参考框中获取关于所述元素流的类型定义的信息;

从所述立体报头框中获取关于由所述元素流所表示的立体图像的尺寸的信息;

从所述立体报头框中获取关于获取所述立体图像的左右照相机的信息以及获取关于显示所述立体图像的显示设备的信息;

从所述采样表框中获取关于构建编码帧的方法的信息;

从所述立体图像数据容器中获取关于所述立体图像编码数据的图像信息;以及
利用上述信息来再现所述立体图像。

2. 根据权利要求1所述的方法,其中所述立体轨道容器的数量和所述立体图像数据容器的数量与所述元素流的数量相同。

3. 根据权利要求1所述的方法,其中,所述获取所述立体图像的左右照相机的信息包括:获取关于获取所述立体图像的左照相机和右照相机之间的距离的信息。

4. 根据权利要求1所述的方法,其中,所述获取关于显示所述立体图像的显示设备的信息包括:获取关于显示所述立体图像的栅栏式显示设备的栅栏式样的距离的信息,和/或指示所述栅栏式样的间隔的信息。

5. 根据权利要求1所述的方法,其中,该方法还包括从所述采样表框中获取关于对相应媒体流的描述进行定义的信息。

6. 根据权利要求5所述的方法,其中,该方法还包括从所述采样表框中获取元素流类型信息,该元素流类型信息用于指示构建编码帧的方法。

7. 根据权利要求6所述的方法,

其中所述元素流的数量为1,

其中所述元素流类型信息所指示的编码帧为第一类型至第五类型中的一种类型,

其中在第一类型中,左图像和右图像在时间轴方向以帧为单元交替排列,

其中在第二类型中,左图像和右图像并排排列,

其中在第三类型中,左图像和右图像按从上到下的方式排列,

其中在第四类型中,左图像和右图像的垂直像素线交替排列,和

其中在第五类型中,左图像和右图像的水平像素线交替排列。

8. 根据权利要求7所述的方法,

其中所述元素流类型信息指示所述第二类型至第五类型中的一种类型,且

其中该方法还包括从所述采样表框中获取关于构成编码帧的左图像和右图像的帧率的信息,和/或视差信息。

9. 根据权利要求8所述的方法,其中关于所述帧率的信息包括关于左图像的帧率是否与右图像的帧率相同的信息,以及在左图像的帧率和右图像的帧率相互不同的情况下,在

显示立体图像时,用于将左图像的帧率和右图像的帧率相互匹配的信息。

10. 根据权利要求 8 所述的方法,其中所述视差信息包括关于左图像和右图像之间是否存在视差的信息,以及当左图像和右图像之间存在视差时,用于修正该视差的信息。

11. 根据权利要求 6 所述的方法,

其中所述元素流的数量为 2,且

其中由所述元素流类型信息所指示的编码帧为左图像、右图像、参考图像和差分图像之一。

用于立体图像编码数据文件的格式

技术领域

[0001] 本发明涉及数据文件格式,特别是,涉及用于存储或发送立体图像编码数据(即,被编码的立体图像数据)的文件格式,或为存储或发送立体图像编码数据而构造文件的方法。

背景技术

[0002] 双目立体图像(在下文中称作“立体图像”)表示通过使用独立的左右照相机对对象拍照而获得的左右图像对。虽然该左右图像是通过拍摄同一个对象而获得的,但是其视点不同。因此,由于对象的表面特性、光源位置等,图像信息也会不同。对象的左右图像之间的图像信息差称为视差。

[0003] 立体图像通常表示使用左右照相机所获得的图像。从广义上讲,立体图像还包括通过向平面图像使用预定转换算法而产生的三维图像。该立体图像通常可用于向所显示的对象添加三维效果。

[0004] 有多种方法可通过使用立体图像来向平板显示设备所再现的图像添加三维效果,该平面显示设备例如是液晶显示屏(LCD)和等离子显示屏(PDP)。其中一种方法使用了栅栏式(barrier type)显示设备。由于栅栏式显示设备可以同时显示平面图像和立体图像,因此,栅栏式显示设备被追捧为下一代显示设备之一。

[0005] 在栅栏式显示设备中,在平面显示设备的前表面中附有或包含栅栏极化板。该栅栏极化板包括线形的栅栏式样(pattern)。通过该栅栏式样,左眼只能观看到所显示图像的左部分。通过该栅栏式样,右眼只能观看到所显示图像的右部分。存在多种类型的栅栏式样。通常,包括垂直和水平线型。因此,可将栅栏式样分为条型、锯齿型和斜线型。这些类型的栅栏式样能对所显示的图像的三维效果造成差异。

[0006] 另一方面,可根据现有编码标准进行编码的静止图片或运动图片(在整个说明书中,图像都同时包括静止图像和运动图像)的平面图像数据大多可分为两类并存储。一类是直接和图像的像素值有关的图像信息。另一类是元数据,是用于对图像信息进行解码和显示所需的附加信息。虽然图像信息根据对图像进行编码所用的国际标准的类型而有不同,但是图像信息可以包括纹理信息和运动信息,该纹理信息例如是亮度和色度。另外,图像信息还可以进一步包括背景和物体的形状信息。元数据包括除图像信息外,用于再现和显示图像信息所需的附加数据。

[0007] 可以任意地将图像信息与元数据区分开。区分方法可以根据国际标准或数据分类标准的内容。在本说明书中,“图像数据”通常同时表示图像信息和元数据。在一些情况下,图像数据可能仅指图像数据。本说明书各部分中所包含的图像数据的意思需要分别根据内容来进行分析。例如,图1的图像数据单元中的“图像数据”仅指图像信息。但是,本发明的发明名称中的图像数据同时表示图像数据和元数据。

[0008] 图1是结构图,用于表示现有的用于存储平面图像编码数据的文件格式。参考图1,现有的文件格式10包括基本报头单元12和图像数据单元14。图像数据单元14包括图

像编码数据的图像信息,例如纹理信息、形状信息和 / 或运动信息。基本报头单元 12 包括除图像数据单元 14 中的图像信息以外的附加数据。但是,现有的图像数据的文件格式 10 适合存储和 / 或发送平面图像编码数据,该现有的文件格式 10 不适合存储和 / 或发送立体图像编码数据。与平面图像不同,立体图像通过使用左右照相机来获得左右图像对,并通过使用多种方式将所获得的左右图像对进行结合来对立体图像进行编码。另外,使用特定的显示设备,例如栅栏式显示屏,来再现该立体图像。

发明内容

[0009] 由于与现有的平面图像不同,立体图像包含左右图像对,因此可以以多种方式来构建所要编码的帧(编码帧)。例如,可以通过将左右图像对进行结合来构建编码帧。有多种方式来将左右图像进行结合。有多种方法可通过左右图像对来设置两个或多个编码帧。由于存在多种方法可通过使用左右图像对来构建编码帧,因此对图像进行编码会产生图像数据和元数据的各种值、类型和特征。但是,前述的文件格式不适于系统地构建和存储各种类型的信息和衍生数据。

[0010] 因此,本发明提供一种方法,用于构建文件格式或文件,该文件格式或文件可有效并系统地存储立体图像编码数据。

[0011] 立体图像编码数据是通过由单独的左右照相机对所获得的图像进行编码而获得的。左右照相机的特征,例如左右照相机的距离和帧率的差异,会对所再现的三维图像的图像质量或三维效果有所影响。另外,该立体图像编码数据可通过使用专门设计的显示设备来再现或以多种方式显示。显示设备或显示方法的特征都对三维图像的图像质量或三维效果有所影响。因此,为了再现最适于显示设备的三维图像,在立体图像编码数据的图像数据中必须包含有关拍照的照相机和 / 或显示设备的信息以及有关显示方法的信息。使用现有的文件格式很难满足这一需求。

[0012] 因此,本发明还提供一种方法,用于构建立体图像编码数据的文件格式或文件,该方法能够通过反映拍照的照相机和 / 或显示设备或显示方法的特征来显示逼真的三维图像。

[0013] 另一方面,在建立了国际多媒体标准的运动图像专家组(MPGE)中,定义了国际标准组织(ISO)基本媒体文件格式。在联合图像专家组(JPEG)2000 的第 12 部分和 ISO/IEC 15444-12 中所公开的 ISO 基本媒体文件格式提供了用于以后应用的基本文件格式。另外,在 MPEG 中,定义了适用于相应应用目的的多媒体应用文件格式(MAF)。在 MAF 与 ISO 基本媒体文件格式相兼容的情况下,可以获得各种使用了立体图像的业务。

[0014] 因此,本发明还提供一种方法,用于构建与 ISO 基本媒体文件格式相兼容的立体图像编码数据文件及其格式。

[0015] 根据本发明的一方面,提供一种立体图像编码数据文件的格式,该格式包括:文件类型声明单元,用于指示所述文件是否是立体图像;元数据单元,包括一个或多个轨道(track)容器,该轨道容器用于容纳立体图像编码数据的元数据;和图像数据单元,包括一个或多个立体图像数据容器,该立体图像数据容器用于容纳立体图像编码数据的图像信息。

[0016] 在本发明的上述方面中,所述文件类型声明单元可以包括第一信息和第二信息,

该第一信息用于指示该文件是否与立体图像有关,该第二信息用于指示构成该文件的元素流 (ES) 的数量。在这种情况下,轨道容器的数量和立体图像数据容器的数量可以与该第二信息相同。

[0017] 此外,所述轨道容器可以包括处理器 (handler) 参考容器和媒体信息容器,该处理器参考容器用于指示相应 ES 的类型,该媒体信息容器用于容纳相应 ES 的元数据。

[0018] 在这种情况下,所述媒体信息容器可以包括立体报头容器,该用于容纳可指示编码帧尺寸的信息。此外,所述立体报头容器可以包括容器,该容器用于容纳可指示用于获取立体图像的左右照相机之间距离的信息,和 / 或容器,该容器用于容纳可指示用于显示立体图像的栅栏式显示设备的栅栏式样的信息和 / 或可指示栅栏式样的间隔的信息。

[0019] 此外,媒体信息容器可包括采样描述容器,该容器用于对相应 ES 的描述进行定义。在这种情况下,该采样描述容器可以包括 ES 类型信息,该 ES 类型信息用于指示构建编码帧的方法。

[0020] 例如,当所述文件类型声明单元的第二信息表明 ES 的数量为 1 时,由 ES 类型信息所指示的编码帧可以为第一类型到第五类型中的一种类型。在第一种类型中,左图像和右图像在时间轴的方向上以帧为单元交替排列。在第二种类型中,左图像和右图像并排排列。在第三种类型中,左图像和右图像按从上到下的方式排列。在第四种类型中,左图像和右图像的垂直像素线交替排列。在第五种类型中,左图像和右图像的水平像素线交替排列。在这种情况下,ES 类型信息可以指示第二类型到第五类型中的一种类型,且采样描述容器可以进一步包括有关构成编码帧的左图像和右图像的帧率的信息和 / 或视差信息。

[0021] 在此,有关帧率的信息可以包括有关左图像的帧率是否与右图像的帧率相同的信息以及在左图像和右图像的帧率不同的情况下,用于当显示立体图像时,将左图像和右图像的帧率相互匹配的信息。视差信息可以包括有关左图像和右图像间是否存在视差的信息以及在左图像和右图像间存在视差的情况下,用于修正该视差的信息。

[0022] 此外,当文件类型声明单元的第二信息表明 ES 的数量为 2 时,由 ES 类型信息所指示的编码帧可以是左图像、右图像、参考图像和差分图像中的一个。

[0023] 如下文所述,由于根据本发明实施例的文件格式具有分层结构以及能够系统地存储立体图像特殊元数据的结构,因此能够有效地构建和存储立体图像编码数据。此外,由于根据本发明实施例的文件格式具有能够包括关于拍摄照相机和 / 或用于获得立体图像的显示设备的特征的信息的结构,因此能够通过使用所存储和编码的立体图像数据来显示逼真的三维图像。此外,根据本发明实施例的用于存储立体图像编码数据的文件格式能够与作为国际标准的 ISO 基本媒体文件格式相兼容。

附图说明

[0024] 图 1 是结构图,表示现有用于存储平面图像编码数据的文件格式。

[0025] 图 2 表示整个合成图像的结构,其中左右图像并排排列组成编码帧。

[0026] 图 3 表示整个合成图像的结构,其中左右图像的像素线交替排列组成编码帧。

[0027] 图 4 表示整个合成图像的结构,其中左右图像以帧为单元顺序排列组成编码帧。

[0028] 图 5 表示包括左右图像的编码帧的结构。

[0029] 图 6 表示包括参考图像和差分图像的编码帧的结构。

- [0030] 图 7 表示包括参考图像和多个差分图像的编码帧的结构。
- [0031] 图 8 是结构图,表示根据本发明实施例用于存储立体图像编码数据的文件格式。
- [0032] 图 9 是结构图,表示图 8 的立体轨道容器的结构。
- [0033] 图 10 表示图 8 和图 9 中所示的文件格式的分层结构。
- [0034] 图 11 表示图 8 的 ssty 框的语法示例。
- [0035] 图 12 表示图 9 的 hdlr 框的语法示例。
- [0036] 图 13 表示图 9 的立体报头框的语法示例。
- [0037] 图 14 表示图 9 的立体照相机信息框的语法示例。
- [0038] 图 15 表示图 9 的立体显示信息框的语法示例。
- [0039] 图 16 至图 19 表示 mpss 框的语法示例。

具体实施方式

[0040] 在下文中,将结合附图详细地描述本发明的示例实施例。以下实施例应理解为仅用于描述,而不是用于进行限制。虽然通过特定术语对本发明实施例进行描述,但是这些描述仅是为了进行表述,应当理解,可以进行改变和变化而不偏离本发明的精神。类似地,虽然本发明是参考所附附图进行特定表示和描述的,但是本领域技术人员应当理解,可以对其进行形式和细节上的各种改变,而不偏离本发明的精神和范围。

[0041] 在对本发明实施例进行描述之前,将描述对根据本发明实施例的立体图像编码数据文件格式进行定义所考虑的因素。这些考虑因素基于立体图像不同于平面图像的特点。

[0042] 第一个考虑因素涉及使用左右图像构建编码帧的方法。构建编码帧的方法对立体图像编码数据的结构具有直接影响。例如,用于构建图像编码数据的元素流 (ES) 的数量取决于构建编码帧的方法。即使是在 ES 数量相同的情况下,也有多种构建编码帧的方法。

[0043] 首先,可使用左右图像来生成编码帧。在下文中,通过使用左右图像而生成的帧称作“整体合成图像”或“合成图像”。通过对整体合成图像进行编码而生成的立体图像数据是使用 ES 进行构建的。有多种方法可使用左右图像对来构建整体合成图像。图 2 至图 4 示出了构建整体合成图像的示例方法。

[0044] 在一种构建整体合成图像的方法中,左右图像并排排列。图 2 表示了这种方法。参考图 2,在编码帧,例如整体合成图像 22 中,左右图像并排排列。可替换地,在编码帧,例如整体合成图像 24 中,左右图像按从上到下的方式排列。在这种情况下,组成整体合成图像 22 或 24 的左右图像的位置可以互换。

[0045] 在另一种构建整体合成图像的方法中,将左右图像以区为单元插入。图 3 表示了这种排列方式。参考图 3,整体合成图像 32 可以是左图像的垂直像素线与右图像的垂直像素线相互交替排列的帧,或是左图像的水平像素线与右图像的水平像素线相互交替排列的帧。构成整体合成图像 32 或 34 的左右图像的像素线的位置可以互换。

[0046] 而在另一种构建整体合成图像的方法中,左右图像以帧为单元顺序排列。图 4 表示了这种排列方式。参考图 4,整体合成图像 40 是通过在时间轴的方向,以帧为单元交替排列左右图像来构建的。在该整体合成图像 40 的情况中,左图像的像素和右图像的像素不在一个编码帧中同时存在。

[0047] 下面参考图 5 和 6,将描述通过使用左右图像对来生成两个编码帧的情况。在两个

编码帧的情况中,使用两个 ES 来构建通过对两个帧进行编码所产生的图像数据。

[0048] 参考图 5,如图所示,左右图像 52a 和 52b 是编码帧。当对帧 52a 和 52b 进行编码时,使用两个元素流 ES1 和 ES2 来构建图像编码数据,这两个元素流表示各自的图像。另一方面,参考图 6,可以使用参考图像 54a 和差分图像 54b 来构建编码帧。在这种情况下,左右图像中的一个是要作为参考图像 54a 的编码帧。利用与参考图像的差分(差异)构建的差分图像 54b 是另一个编码帧。

[0049] 图 7 表示了有三个编码帧的情况。参考图 7,连续的 $(n+1)/2$ 个帧中的左右图像中的一个作为参考图像 62 的编码帧。除该参考图像以外的其他图像是作为差分图像 62a 至 62n 的编码帧。当对编码帧进行编码时,使用 $(n+1)$ 个元素流 ES1 至 ES $(n+1)$ 来构建图像编码数据。

[0050] 可以使用现有的图像编码方法来对前述一个或多个编码帧或编码帧序列进行编码。现有的图像编码方法包括静态图像(例如, JPEG)编码方法,或对运动图片(例如, MPEG-1、MPEG-2、MPEG-4、H. 264/AVC、VC-1 等)进行编码的方法。之后,通过使用现有图像编码方法所编码的图像数据可直接发送至支持该编码方法的显示设备进行再现。可替换地,可在存储介质中存储该图像数据,并由显示设备再现。

[0051] 如上所述,在立体图像的情况中,有多种方法来构建编码帧。可以使用两个或多个 ES 来构建立体图像编码数据。即使是在有相同数量 ES 的情况中,也有多种方法来构建编码帧。因此,衍生数据或再现所述图像数据所需的数据可以是变化的。用于存储该立体图像编码数据的文件格式应当适于存储构建编码帧的方法和该方法的衍生数据。

[0052] 对存储立体图像编码数据文件格式进行定义的第二个考虑因素是使用以预定间隔相互分开的左右照相机,以获得立体图像。这是由于必须向显示设备提供左右照相机的信息,以有效地再现和/或提高所再现的三维图像的图像质量或三维效果。因此,立体图像编码数据可以额外地包括左右照相机的信息。考虑到该额外包括的有关左右照相机的信息,因此必须定义用于存储立体图像编码数据的文件格式。

[0053] 有关左右照相机的信息具有多种类型。例如,该多种类型的信息包括有关左右照相机之间的距离的信息;使用左右照相机所采集的左右图像每秒的帧数(帧/秒, fps),即帧率;有关左右图像同步的信息和/或有关左右照相机类型的信息。此外,在一些情况下,该多种类型的信息还可以包括左右图像间的视差信息。

[0054] 对存储立体图像编码数据的文件格式进行定义的第三个考虑因素是使用与现有显示设备不同的专门显示设备来再现立体图像(例如,栅栏式显示设备)。这是由于所再现的图像数据必须适于显示设备,以通过使用专门的显示设备来再现三维图像。此外,由于有关显示设备特征的信息会对三维图像的图像质量或三维效果产生影响,因此必须考虑该信息或额外所需的信息,以对立体图像编码数据文件的格式进行定义。

[0055] 有关显示设备的信息具有多种类型。例如,在再现设备是栅栏式显示设备的情况下,该各种类型的信息包括有关最适于再现立体图像编码数据的栅栏式样的信息。如上所述,该栅栏式样是以垂直或水平线的方式位于栅栏极化板上的。微小的(minute)线形会对三维图像的图像质量产生影响。此外,有关栅栏式样的间隔(该间隔基于栅栏式样在显示设备上的位置)的信息(有关该间隔是否不论其位置都是常量、或者该间隔取决于其位置的信息)也会对三维图像的图像质量产生影响。

[0056] 图 8 和图 9 是结构图,表示根据本发明实施例的用于存储立体图像编码数据的文件格式。图 9 是表示图 8 的立体轨道容器 210 的结构的结构图。另外,图 10 表示图 8 和图 9 所示的文件格式的分层结构。参考图 8 至图 10 可知,根据本发明实施例的文件格式是基于 ISO 基本媒体文件格式的。

[0057] 首先参考图 8 和图 10,根据本发明实施例的文件格式主要包括文件类型声明单元 (ftyp) 100、元数据单元 (moov) 200 和图像数据单元 (mdat) 300。

[0058] 文件类型声明单元 100 用于表示为立体图像使用了相应的文件。在为立体图像使用了文件的情况下,文件类型声明单元 100 可以包括有关组成该立体图像的 ES 的数量的信息。如图 8 和图 10 所示,文件类型声明单元 100 是 ftyp 容器的子分类,其包括一框,该框用于包括指示该文件是否是立体类型的信息和 / 或有关组成该立体图像的 ES 的数量的信息。该框可以是立体类型框 (ssty) 110,如图 8 和图 10 所示。这样,立体图像解码器就能够识别出该文件是否与立体图像有关和 / 或识别出组成该立体图像的 ES 的数量。这在下面进行简述。

[0059] ssty(Stereoscopic Type)

[0060] -Box Type : 'ssty'

[0061] -Container : File Type Box ('ftyp')

[0062] -Mandatory : Yes

[0063] -Quantity : Exactly one

[0064] 通过前面的叙述可知,在立体图像编码数据的情况下, ssty 框 110 是必不可少的部分。在 ftyp 容器中仅有一个 ssty 框。图 11 表示了 ssty 框 110 的语法示例。在图 11 中,元素 'Stereoscopic_Type' 指示文件是否是立体文件。例如,该元素的值可按表 1 进行分配。另外,元素 'StereoScopic_ES_Count' 指示组成该立体文件的 ES 的数量。

[0065] 表 1

[0066] [表 1]

[0067]

值	内容
0	文件不是立体数据文件
1	文件是立体数据文件

[0068] 参考图 8 和图 10,作为元数据单元 200 的 moov 容器包括一个或多个轨道容器 210 或 220,该轨道容器 210 或 220 用于存储文件的元数据。在文件是立体图像文件的情况下,该 moov 容器包括与组成该文件的 ES 的数量相对应的立体轨道容器 210,例如,用于元素流 ES1 的立体轨道容器轨道 1 (立体),用于元素流 ES2 的立体轨道容器轨道 2 (立体), 和立体轨道容器轨道 (n) (立体) (此处, n 为大于或等于一的整数)。另一方面,在文件不是立体图像文件的情况下, moov 容器包括非立体的轨道容器 220,例如,用于平面图像和音频或文本文件的元数据的轨道容器轨道 (非立体)。由于本发明涉及立体图像,因此在下文中,将参考图 9 和图 10 对立体轨道容器 210 的结构进行描述。

[0069] 立体轨道容器 210 包括媒体容器 (媒体) 211。将该媒体容器 211 定义为其包括有

关存储在容器（该容器被称之为轨道）的媒体流的信息。该媒体容器 211 包括处理器参考框（hdlr）212 和媒体信息容器（minf）（未示出）。该媒体信息容器（minf）可以是用于包括有关 ES 所表示的图像的尺寸的信息的框（该框可以是立体报头框（sshd）213，其名称可变）和采样表框（stbl）216。

[0070] 所述处理器参考框 212 包括有关 ES 流类型定义的信息。在 ES 是通过对立体图像进行编码所得数据的情况下，可以例如将处理器参考框 212 中所包含的信息值表示为“ssvi”。该处理器参考框 212 被表示如下。

[0071] hdlr(Handler Reference)

[0072] -Box Type: ‘hdlr’

[0073] -Container:Media Box(‘media’)

[0074] -Mandatory:Yes

[0075] -Quantity:Exactly one

[0076] 通过前面的叙述可知，hdlr 框 212 是不可缺少的部分。在媒体容器 211 中仅有一个处理器参考框 212。图 12 表示 hdlr 框 211 的语法示例。在图 12 中，使用元素 ‘handler_type’ 来定义媒体数据的流类型。表 2 表示流类型的示例，在该示例中，对现有流的定义包括对本发明的立体图像流的定义。

[0077] 表 2

[0078] [表 2]

[0079]

值	内容
ssvi	立体可视数据
soun	音频数据
vide	可视数据
text	文本数据
hint	提示数据

[0080] 立体报头框 213 包括有关由 ES 所表示的图像的尺寸的信息。例如，立体报头框 213 可以包括有关 ES 所表示的立体合成图像的宽度和 / 或高度的信息。图 13 表示立体报头框 213 的语法示例。在图 13 中，元素 ‘StereoScopic_CompoundImageWidth’ 表示立体合成图像的宽度，而元素 ‘StereoScopic_CompoundImageHeight’ 表示立体合成图像的高度。该立体报头框 213 被表示如下。

[0081] sshd(stereoscopic header)

[0082] -Box Type: ‘sshd’, ‘vmhd’, ‘smhd’, ‘hmhd’

[0083] -Container:MediaInformation Box(‘minf’)

[0084] -Mandatory:Yes(must be present)

[0085] -Quantity:Exactly One

[0086] 通过前面的叙述可知, sshd 框 213 是不可缺少的部分。在 minf 容器(未示出)中仅有一个立体报头框 213。该 minf 容器可以除 sshd 框 213 以外,进一步包括报头框,以用于另一种类型的媒体。表 3 表示了包括在 minf 容器中的报头框的值的示例。

[0087] 表 3

[0088] [表 3]

[0089]

值	内容
sshd	立体可视媒体报头
smhd	音频媒体报头
vmhd	可视媒体报头
hmhd	提示媒体报头
nmhd	空白媒体报头

[0090] 参考图 9 和图 10, 立体报头框 213 进一步包括框, 该框用于包括有关获取立体图像的左右照相机的信息; 以及框, 该框用于包括有关显示立体图像的显示设备的信息。所述框可以是立体照相机信息框(ssci)214 和立体显示信息框(ssdi)215。所述框的名称可变。

[0091] 所述立体照相机信息框(ssci)214 可以包括有关左右照相机的信息, 例如, 有关左右照相机之间的距离的信息。所述立体照相机信息框 214 被概述如下。

[0092] ssci(stereoscopic Camera Information)

[0093] -Box Type: 'ssci'

[0094] -Container: Stereoscopic Header Box('sshd')

[0095] -Mandatory: No

[0096] -Quantity: Zero or One

[0097] 通过上述概述可知, ssci 框 214 是可选的部分。在立体报头框 213 中包含 ssci 框 214 的情况下, 在作为容器的 sshd 框 213 中仅存在一个 sshd 框 214。图 12 表示 ssci 框 214 的语法示例。在图 14 中, 元素 'StereoScopicCamera_Left_Right-Distance' 表示左右照相机之间的距离。

[0098] 所述立体显示信息框 215 可以包括有关显示设备的信息, 例如, 有关栅栏式样类型的信息和 / 或有关栅栏式样的间隔的信息。立体显示信息框 215 被概述如下。

[0099] ssdi(stereoscopic Display Information)

[0100] -Box Type: 'ssdi'

[0101] -Container: Stereoscopic Header Box('sshd')

[0102] -Mandatory: No

[0103] -Quantity: Zero or One

[0104] 通过上述概述可知, ssdi 框 215 是可选的部分。在 sshd 框 213 中包含 ssdi 框 215 的情况下, 作为容器的 sshd 框 213 中仅存在一个 ssdi 框 215。图 15 表示 ssdi 框 215 的语

法示例。在图 15 中,元素 ‘StereoScopic_Barrier_Pattern’ 表示栅栏式样的类型。例如,可按表 4 来分配类型值。此外,元素 ‘StereoScopic_Barrier_Distance’ 表示栅栏式样的间隔。当该间隔值为 0 时,表示非固定比率。当该间隔值为 1 时,表示固定比率。此处,固定比率表示不论栅栏式样在显示设备上的位置如何,栅栏式样的间隔都是常量。而非固定比率表示栅栏式样的间隔取决于其在显示设备上的位置(例如,在中央或边缘部分)。

[0105] 表 4

[0106] [表 4]

[0107]

值	内容
00	条型
01	锯齿型
10	斜线型

[0108] 参考图 9 和图 10,采样表框 216 用作时间 / 空间映射的容器,该框包括采样描述框 (stsd) 217。该采样描述框 217 用于对轨道容器 210 中所定义的媒体流 (ES) 的描述进行定义,该框 217 包括用于指示立体可视采样项的框。该框可以称作 mpss 框 218。该框并不限于此。所述采样描述框 217 除 mpss 框 218 以外,还可进一步包括 mp4v 框,用于指示可视采样项;以及 mp4a 框,用于指示音频采样项;等等。

[0109] 所述 mpss 框 218 是用于公开有关组成立体图像编码数据的 ES 的详细信息的框容器。该 mpss 框 218 被概述如下。

[0110] mpss(stereoscopic Visual Sample Entry)

[0111] -Box Type : ‘mpss’, ‘mp4v’, ‘mp4a’

[0112] -Container :Stereoscopic Table Box(‘stbl’)

[0113] -Mandatory :Yes

[0114] -Quantity :Exactly One

[0115] 通过上述概述可知, mpss 框 218 是不可缺少的部分。在 stbl 容器 217 中仅有一个 mpss 框 218。所述 stbl 容器 217 除 mpss 框 218 以外,还可进一步包括另一种类型的媒体的采样项。表 5 示出了包含在 stbl 容器 217 中的采样项的示例。

[0116] 表 5

[0117] [表 5]

[0118]

值	内容
mpss	立体可视采样项
mp4v	可视采样项
mp4a	音频采样项

[0119] 所述 mpss 框 218 包括有关构建编码帧的方法的信息、各种类型的衍生信息等。该 mpss 框 218 中所包括的信息可以根据构建立体图像编码数据的 ES 的数量和 / 或与 ES 相对应的编码帧的类型而变化。特别是, mpss 框 218 可以包括有关编码帧类型的信息 (构建方法)、有关左右图像帧率、构建编码帧的图像的尺寸、构建编码帧的区线数量的信息和 / 或构建编码帧的左右图像的视差信息。在下文中, 将基于立体图像编码数据的 ES 的数量来详细描述 mpss 框 218 中所包括的信息内容。

[0120] 首先, 描述存在 ES 的情况。在一个 ES 的情况下, 可采用图 2 至图 4 的方法之一用作构建编码帧的方法。在图 2 至图 4 中示出了五种构建编码帧的方法。mpss 框 218 中所包括的信息必须能够支持上述五种类型。因此, mpss 框 218 包括用于指示构成 ES 的编码帧的类型的信息。将帧的类型表示为 ‘StereoScopic_CompositionType’。可使用例如表 6 中的三个比特来分配类型值。表 6 示出了一个示例。

[0121] 表 6

[0122] [表 6]

[0123]

值	内容
000	左右图像在时间轴方向以帧为单元交替排列 (参考图 4)
001	左右图像并排排列 (参考图 2 左图)
010	左右图像按从上到下的方式排列 (参考图 2 右图)
011	左右图像的垂直像素线交替排列 (参考图 3 左图)
100	左右图像的水平像素线交替排列 (参考图 3 右图)

[0124] 在编码帧是图 2 和图 3 中所示的帧 22、24、32 或 34 的情况下, mpss 框 218 可进一步包括有关编码帧尺寸的信息。例如, 在编码帧是图 2 左图所示的帧的情况下, mpss 框 218 可以包括有关图像宽度的信息。在编码帧是图 2 右图所示的帧的情况下, mpss 框 218 可以包括有关图像高度的信息。在编码帧是图 3 左图所示的帧的情况下, mpss 框 218 可以包括有关以区为单元插入的垂直线的宽度的信息。在编码帧是图 3 右图所示的帧的情况下, mpss 框 218 可以包括有关以区为单元插入的水平线的宽度的信息。

[0125] 可将有关编码帧的信息表示为 ‘width_or_height’。例如, 如果表 6 中所公开的 StereoScopic_CompositionType 的值为 0b001, 则 ‘width_or_height’ 的值可以表示图像宽度。如果 StereoScopic_CompositionType 的值为 0b010, 则 ‘width_or_height’ 的值可以表示以区为单元插入的垂直线的高度。如果 StereoScopic_CompositionType 的值为 0b100, 则 ‘width_or_height’ 的值可以表示以区为单元插入的水平线的高度。

[0126] 此外, 在编码帧是图 2 和图 3 所示的帧 22、24、32 或 34 的情况下, mpss 框 218 可以包括有关线的数量的信息, 该线构成组成编码帧的合成图像的奇偶线区。例如, 在帧是图 2 中所示的帧 22 或 24 的情况下, mpss 框 218 可以包括有关构成奇数线区的线的数量的信息, 和 / 或有关构成偶数线区的线的数量的信息。

[0127] 可将有关构成奇数线区的线的数量的信息表示为 ‘odd_field_count’。可将有关构成偶数线区的线的数量的信息表示为 ‘even_field_count’。例如,如果表 6 中所公开的 StereoScopic_CompositionType 的值为 0b001 或 0b010,则 ‘odd_field_count’ 和 ‘even_field_count’ 的值都为 0。如果 StereoScopic_CompositionType 的值为 0b011 或 0b100 时, ‘odd_field_count’ 和 ‘even_field_count’ 的值则可以分别表示奇数线的数量和偶数线的数量。

[0128] 所述 mpss 框 218 可以进一步包括有关奇数线区的帧率与偶数线区的帧率是否相同的信息,和有关在奇数和偶数线区的帧率不同的情况下,所使用的同步方法的信息。此处,在两个图像的帧率相互不同的情况下,有关同步方法的信息可以是有关参考图像的信息,该参考图像在显示立体图像时,将帧率相互匹配。也就是说,有关同步方法的信息可以是有关参考图像的信息。可将有关帧率和 / 或同步方法的信息表示为 ‘StereoScopic_ES_FrameSync’,并按照表 7 所示使用两个比特来进行分配。表 7 表示有一个 ES 的示例情况。

[0129] 表 7

[0130] [表 7]

[0131]

值	内容
00	左图像 (奇数线区) 的帧率与右图像 (偶数线区) 的帧率相同
01	左图像的帧率与右图像的不同, 且左图像 (或奇数线区) 为参考图像
10	左图像的帧率与右图像的不同, 且右图像 (或偶数线区) 为参考图像

[0132] 所述 mpss 框 218 可以进一步包括有关视差存在的信息,也就是,奇数线和偶数线区之间的图像信息的差异 (例如,Y/Cb/Cr 值或 R/G/B 值) 和在存在视差的情况下的视差值 (有关视差的信息)。在此,视差值表示有关一个图像 (或区) 相对于另一个图像 (或区) 的差异值的信息。使用该视差信息来修正所显示的立体图像的三维效果。

[0133] 可将包含在视差信息中的有关视差存在的信息表示为 ‘StereoScopic_ImageInformationDifference’,并按表 8 所示使用两个比特来进行分配。表 8 表示了有一个 ES 的示例情况。

[0134] 表 8

[0135] [表 8]

[0136]

值	内容
00	左右图像 (奇数线区和偶数线区) 间的视差为零
01	视差不为零,且左图像 (或奇数线区) 为参考图像

10	视差不为零,且右图像（或偶数线区）为参考图像
----	------------------------

[0137] 可将包含在视差信息中的视差值表示为图像信息的差异。有多种方法来表示图像信息。通常的方法是 Y/Cb/Cr 或 R/G/B 方法。因此,可以使用以下方法来表示视差值。

[0138] Y_or_R_difference :Y 或 R 值的图像信息的差异

[0139] Cb_or_G_difference :Cb 或 G 值的图像信息的差异

[0140] Cr_or_B_difference :Cr 值或 B 值的图像信息的差异

[0141] 下面将描述具有两个 ES 的情况。在两个 ES 的情况中,可例如按照图 5 或 6 所描述的方法之一来构造编码帧。在两个 ES 的情况中,moov 容器 200 包括两个轨道容器,为轨道 1 和轨道 2 容器。每个轨道容器都可以包括相应 ES 的元数据信息。在下文中,将描述一个 ES 的情况与两个 ES 的情况之间的不同。

[0142] 在有两个立体图像编码数据的 ES 的情况下, mpss 框 218 包括有关构成相应 ES 的编码帧类型的信息。参考图 5 和图 6,由于编码帧的类型可以包括左图像、右图像、参考图像和差分图像,因此, mpss 框 218 包含有关帧类型的信息。将编码帧的类型表示为 ‘StereoScopic_ES_Type’。可按表 9 使用两个比特来分配类型值。表 9 示出了一个示例。

[0143] 表 9

[0144] [表 9]

[0145]

值	内容
00	左图像
01	右图像
10	参考图像
11	差分图像

[0146] 所述 mpss 框 218 可以进一步包括有关左图像的帧率是否与右图像的帧率相同的信息和有关在左右图像的帧率相互不同的情况下的同步方法的信息。只有在编码帧是图 5 所示的帧的情况下（由左右图像所构建的帧）, mpss 框 218 才包括有关帧率的信息。在编码帧是图 6 所示的帧的情况下, mpss 框 218 并不包括有关帧率的信息。可将有关帧率和 / 或同步方法的信息表示为 ‘StereoScopic_ES_FrameSync’, 并按照表 10 所示使用两个比特来进行分配。在此,表 10 表示存在两个 ES 的示例情况。

[0147] 表 10

[0148] [表 10]

[0149]

值	内容
00	左图像的帧率与右图像的相同,或不需要有关帧率的信息

01	左图像的帧率与右图像的不同,且相应 ES 的帧为参考图像
10	左图像的帧率与右图像的不同,且相应 ES 的相对帧为参考图像

[0150] 所述 mpss 框 218 可以进一步包括有关视差存在的信息,也就是,左右图像的图像信息之间的差异(例如,Y/Cb/Cr 值或 R/G/B 值)和在存在视差的情况下的视差值(有关视差的信息)。只有在编码帧是图 5 所示的帧的情况下(由左右图像所构建的帧),mpss 框 218 才包括视差信息。在编码帧是图 6 所示的帧的情况下,mpss 框 218 不包括视差信息。可将该视差信息表示为‘StereoScopic_ImageInformationDifference’,并按表 11 使用两个比特来进行分配。在此,表 10 表示存在两个 ES 的示例情况。

[0151] 表 11

[0152] [表 11]

[0153]

值	内容
00	左右图像间的视差为零,或不考虑
01	视差不为零,且相应 ES 的帧为参考图像
10	视差不为零,且相应 ES 的相对帧为参考图像

[0154] 作为图像信息差异的视差值可以不包含在相应 ES 的 mpss 框 218 中,而是包含在与相应 ES 相对的另一个 ES 的 mpss 框中。在这种情况下,有关视差存在的信息和有关视差值的信息可以分布于两个 ES 中。

[0155] 在用于表示编码帧类型的立体 ES 类型对应于图 6 所示的图像的情况下,将该编码帧划分为参考图像和差分图像。因此,在‘StereoScopic_ES_Type’指示参考图像或差分图像的情况下,帧率信息和视差信息对于 ES 而言并非是必须的。因此,在有两个 ES 情况下,如果编码帧是图 6 所示的图像,则 mpss 框 218 不包括所述信息。

[0156] 下面,将描述有三个或更多 ES 的情况。在有三个或更多 ES 的情况中,编码帧如图 7 所示。图 7 的帧与图 6 的帧相同,其帧都是由参考图像和差分图像所构建的。因此,在三个或更多 ES 的情况下,mpss 框 218 中所包括的信息与两个 ES 且编码帧的类型是图 6 所示的图像的情况相同。因此省略对该信息的描述。

[0157] 在图 16 至图 19 中示出了包括前述信息的 mpss 框 218 的语法示例。虽然图 16 至图 19 中所示的语法应当以一个语法进行完全表示,但是由于本说明书空间的限制而将其分开表示。因此,图 16 中所示的语法与图 17 中所示的语法顺序相连。随后,图 18 和图 19 的语法在图 17 的语法之后。由于已经详细描述了这些语法,因此将省略对这些语法的描述。

[0158] 接下来,参考图 8,作为图像数据单元的 mdat 容器 (mdat) 300 包括编码帧的图像编码信息。该 mdat 容器包括一个或多个立体图像数据容器 (立体图像数据) 310。每个立体图像数据容器 310 都对应于元数据单元 200 中所包括的每个轨道容器 (轨道) 210。因此,图像数据单元 300 包括与 ES 数量相对应的多个立体图像数据容器 310。由于每个立体图像数据容器 310 中所包含的图像数据类型都与现有的图像数据相类似,因此,在下文中将省

略对这些图像数据类型的描述。

[0159] 虽然以特定方式表示了本发明,并结合其典型实施例进行了描述,但是本领域技术人员应当理解,可以对其形式和细节进行各种变化,而不偏离本发明如所附权利要求所定义的精神和范围。

[0160] 工业实用性

[0161] 本发明涉及立体图像编解码。

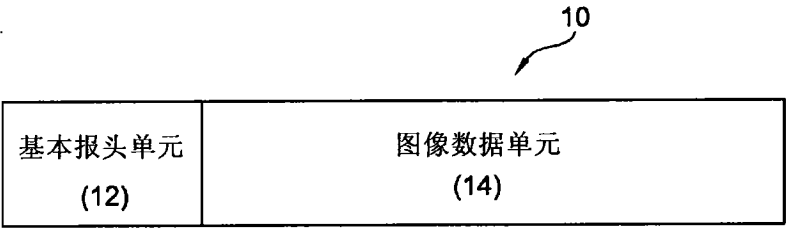


图 1

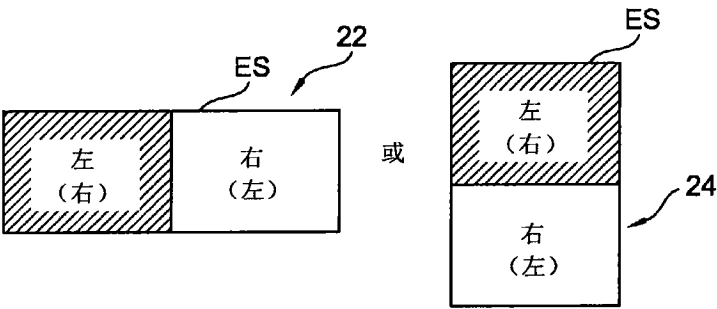


图 2

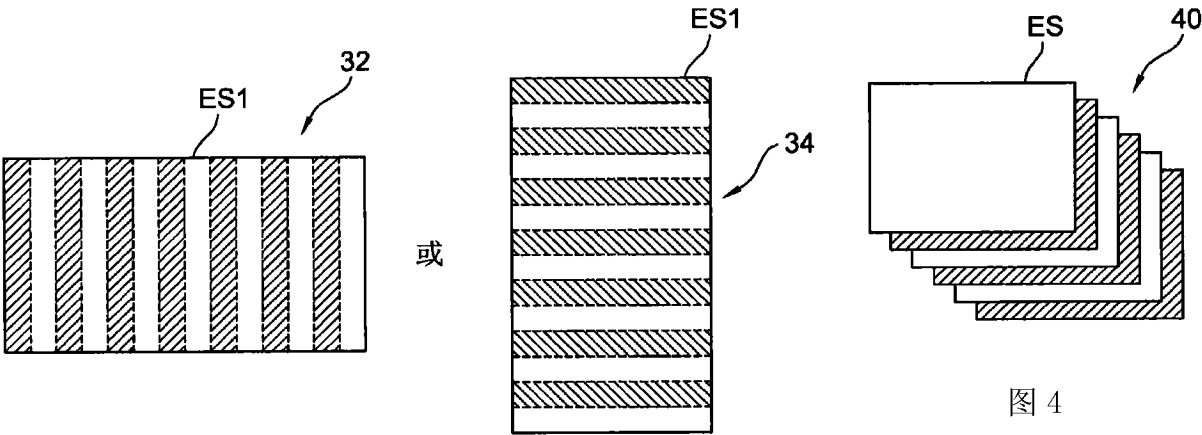


图 4

图 3

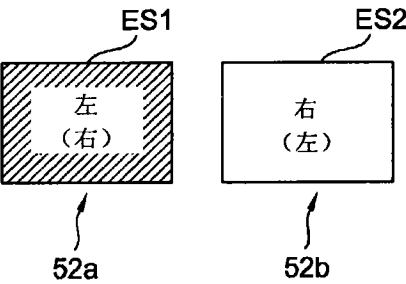


图 5

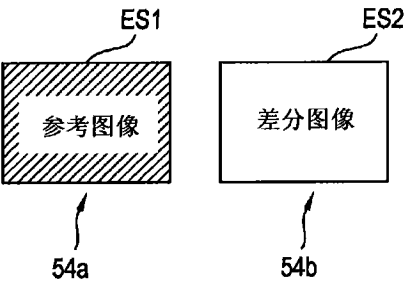


图 6

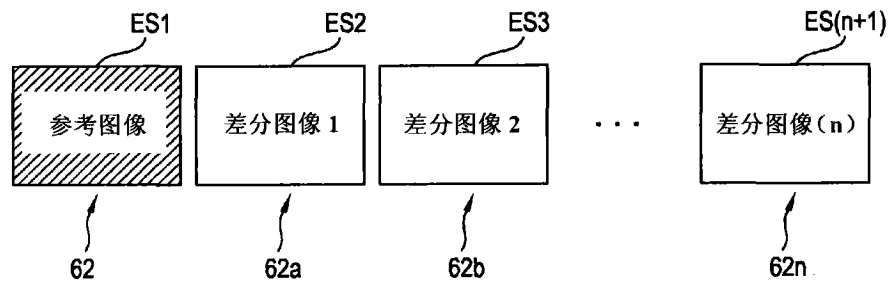


图 7

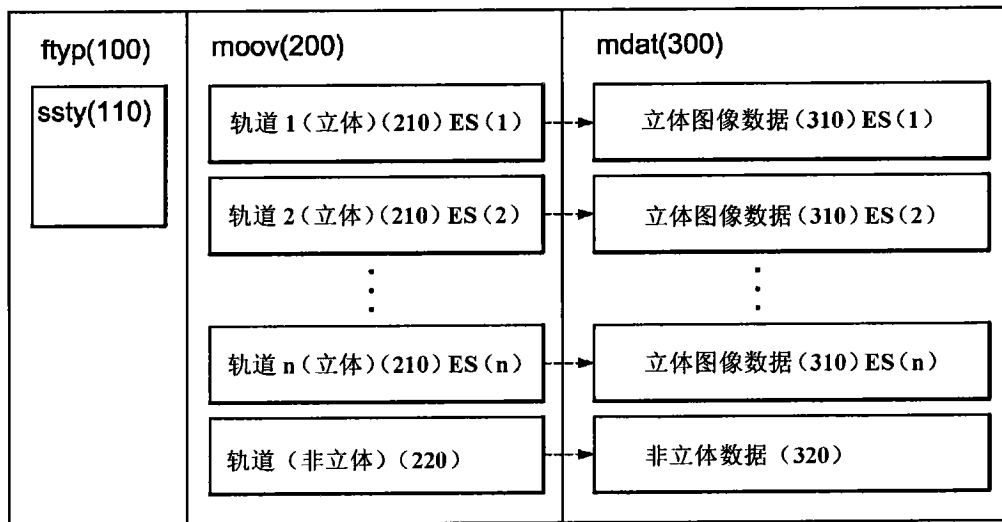


图 8

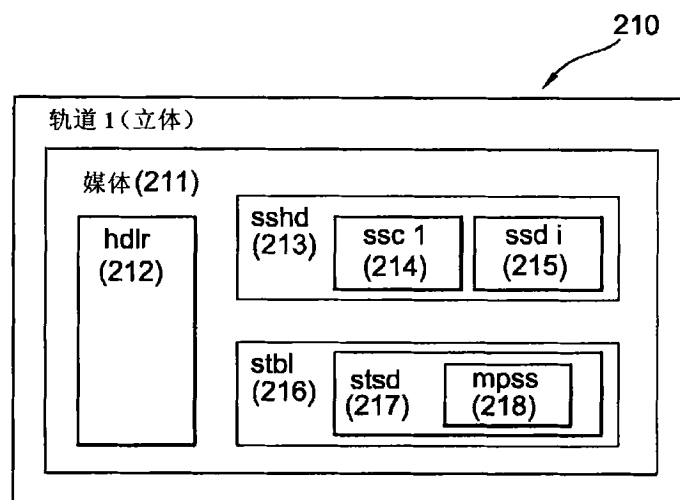


图 9

ftyp					文件类型和兼容性 major_brand= "isom" or "mp??"
***	ssty				声明立体类型
moov					容器，用于时间资源的所有元数据
	mvhd				影片报头，全部声明 包括 edts 或 stbl 框的时间标尺值
	trak				容器，用于各个轨道或流 指示主要的 AV 数据或辅助数据
		tkhd			轨道报头，有关轨道的全部信息 包含数据的图像尺寸或数据的大小
		edts			用于数据的编辑列表容器
			elst		编辑列表 包含对于一个时间数据的 sync 信息（起始时间&持续时间）
		mdia			容器，用于轨道中的媒体信息
			mdhd		媒体报头，有关媒体的全部信息
			hdlr		处理器，声明媒体（处理器）类型 "soun" 用于音频数据 "vide" 用于可视数据 "text" 用于时间文本数据 "ssvi" 用于立体数据 （可选的-已经在 linf 框中声明）
			minf		媒体信息容器
***				sshd	立体媒体报头，全部信息（可选的）
***				ssci	立体照相机信息
***				sedi	立体显示信息
				vmhd	视频媒体报头，全部信息（可选的）
				smhd	声音媒体报头，全部信息（可选的）
				hmhd	提示媒体报头，全部信息（仅提示轨道）
				nmhd	空闲媒体报头，全部信息（可选的）
				stbl	采样表框，用于时间/空间映射的容器 包含用于时间媒体采样的时间/空间映射（仅时间文本数据需要）
				stsd	采样描述（编解码类型、初始化等） "mp4v" 用于可视数据 "mp4a" 用于音频数据 "mpss" 用于立体数据
				stts	（解码）时间到采样 包含每个采样的（文本行）解码时间
				stsc	采样到块，部分的数据偏移信息 （可选的-通常一个块仅包含一个采样）
				stsz	采样尺寸
				stz2	压缩的采样尺寸
				stco	块偏移，部分的数据偏移信息 指示每个块或采样的实际文件偏移
				co64	64 位的块偏移（如果文件尺寸过大）
mdat					媒体数据容器 主要的 AV 和辅助数据

图 10

Syntax

```
aligned (8) class StereoScopicTypeBox extend Box ('asty', version = 0, 0) {  
    unsigned int (1)  StereoScopic_Type;  
  
    unsigned int (4)  StereoScopic_ES_Count;  
    unsigned int (3)  reserved;  
}
```

图 11

Syntax

```
aligned (8) class HandlerBox extend Box ('hdlr', version = 0, 0) {  
    unsigned int (32) handler-type;  
}
```

图 12

Syntax

```
aligned (8) class StereoScopicHeaderBox extend Box ('sshd', version = 0, 0) {  
    unsigned int (16)  StereoScopic_CompoundImageWidth;  
    unsigned int (16)  StereoScopic_CompoundImageHeight;  
}
```

图 13

Syntax

```
aligned (8) class StereoScopicCameraInformationBox extend Box ('ssci', version = 0, 0) {  
    unsigned int (32) StereoScopic_Camera_Left_Right_Distance;  
}
```

图 14

Syntax

```
aligned (8) class StereoScopicDisplayInformationBox extend Box ('ssdi', version = 0, 0) {  
    unsigned int (2) StereoScopic_Barrier_Pattern;  
    unsigned int (1) StereoScopic_Barrier_Distance;  
    unsigned int (5) reserved;  
}
```

图 15

Syntax

```
aligned (8) class ESDBox extends Box ('esds', version = 0, 0) {  
    ES_Descriptor ES;}  
aligned (8) abstract class SampleEntry (unsigned int (32) format) extends Box (format) {  
    unsigned int(16)      data-reference-index; }  
class VisualSampleEntry() extends SampleEntry ('mp4v') {  
    // 14496-1 참조  
}
```

图 16

```
class StereoScopicVisualSampleEntry(unsigned int(4) StereoScopic_ES_count) extends  
VisualSampleEntry ('mpss') {  
    if ( StereoScopic_ES_count == 1 ) {  
        unsigned int (3)   StereoScopic_CompositionType;  
        unsigned int (1)   reserved;  
        unsigned int (2)   StereoScopic_ES_FrameSync;  
        unsigned int (2)   StereoScopic_ImageInformationDifference;  
  
        if ( StereoScopic_CompositionType != 0 ) {  
            int (16)   width_or_height;  
            int (16)   odd_field_count;  
            int (16)   even_field_count;  
        }  
  
        if ( StereoScopic_ImageInformationDifference != 0 ) {  
            int (16)   Y_or_R_difference;  
            int (16)   Cb_or_G_difference;  
            int (16)   Cr_or_B_difference;  
        }  
    }  
}
```

图 17

```

} else if ( StereoScopic_ES_count == 2 ) {
    unsigned int (2)    StereoScopic_ES_Type;
    unsigned int (2)    reserved;
    unsigned int (2)    StereoScopic_ES_FrameSync;
    unsigned int (2)    StereoScopic_ImageInformationDifference;

    if ( StereoScopic_ImageInformationDifference == '0b10' ) {
        int (16)    Y_or_R_difference;
        int (16)    Cb_or_G_difference;
        int (16)    Cr_or_B_difference;
    }
} else {
    unsigned int (2)    StereoScopic_ES_Type;
    unsigned int (6)    reserved;
}
}

```

图 18

```

class SampleDescriptionBox (unsigned int (32) handle-type) extend Box ('std', 0, 0) {
    int i;
    unsigned int (32)    entry-count;
    for ( i=0; i<entry-count; i++) {
        switch (handler-type) {
            case 'soun' : // Audio Stream
                AudioSampleEntry();
                break;
            case 'vide' : // Visual Stream
                VisualSampleEntry();
                break;
            case 'ssvi' : // StereoScopic Visual Stream
                StereroScopicVisualSampleEntry();
                break;
            case 'hint' : // Hint track
                HintSampleEntry();
                break;
            default :
                MpegSampleEntry();
                break;
        }
    }
}

```

图 19