

(12) 发明专利

(10) 授权公告号 CN 101595475 B

(45) 授权公告日 2012. 12. 12

(21) 申请号 200680033481. X

(51) Int. Cl.

(22) 申请日 2006. 07. 14

G06F 17/30(2006. 01)

G06K 9/60(2006. 01)

(30) 优先权数据

60/699, 535 2005. 07. 15 US

11/398, 195 2006. 04. 04 US

(56) 对比文件

US 2004/0006575 A1, 2004. 01. 08, 全文 .

WO 2004/036916 A1, 2004. 04. 29, 参见说明书第 5-8 页, 图 2-4.

(85) PCT 申请进入国家阶段日

2008. 03. 12

CN 1334677 A, 2002. 02. 06, 全文 .

(86) PCT 申请的申请数据

PCT/US2006/027602 2006. 07. 14

审查员 吴鑫

(87) PCT 申请的公布数据

W02007/011836 EN 2007. 01. 25

(73) 专利权人 索尼株式会社

地址 日本东京都

专利权人 索尼电子有限公司

(72) 发明人 M·Z·维沙拉姆 A·塔巴塔拜

(74) 专利代理机构 北京东方亿思知识产权代理

有限责任公司 11258

代理人 宋鹤

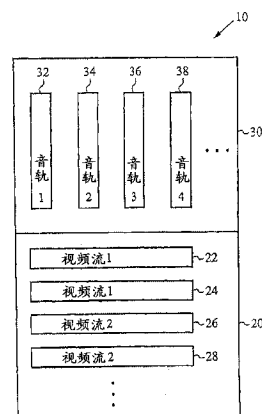
权利要求书 4 页 说明书 15 页 附图 5 页

(54) 发明名称

可伸缩视频编码(SVC)文件格式

(57) 摘要

通过提供用于存储和访问目前 SVC 标准定义的视频内容的扩充来修改目前现有的 ISO/AVC 文件格式。确切地说, 进行对 AVC 文件格式的扩充提供能够存储和访问可伸缩视频数据的新 SVC 文件格式。将可伸缩视频数据作为 SVC 文件格式的媒体数据部分内的单个音轨来存储。对 SVC 文件格式的元数据部分内的描述项和框定义了新扩充。这些扩充提供从媒体数据部分中存储的可伸缩视频数据的单个音轨提取子流或层的方式。



1. 一种从装置上实现的修改的文件格式中提取数据流的系统,所述系统包括:
用于接收对特定数据流的请求的装置,其中所述请求包括一个或多个装置要求;
用于将所述请求与所述修改的文件格式的媒体数据部分中存储的特定可伸缩数据流关联的装置;
用于确定与所述特定可伸缩数据流对应的一个或多个音轨的装置,其中所述一个或多个音轨存储在所述修改的文件格式的元数据部分中,而且其中每个音轨包括一个或多个元数据框;
用于根据所述一个或多个装置要求来确定所述特定可伸缩数据流的子层数据流的装置,其中使用所述一个或多个元数据框来确定所述子层数据流;以及
用于从所存储的可伸缩数据流中提取所述子层数据流的装置。
2. 如权利要求 1 所述的系统,其特征在于,还包括用于发送所提取的子层数据流的装置。
3. 如权利要求 1 所述的系统,其特征在于,还包括用于在确定所述子层数据流之前将所确定的一个或多个音轨解码的装置。
4. 如权利要求 1 所述的系统,其特征在于,所述一个或多个元数据框包括对伸缩视频编码 (SVC) 文件格式标准的扩充。
5. 如权利要求 1 所述的系统,其特征在于,所述可伸缩数据流包括可伸缩视频流。
6. 如权利要求 5 所述的系统,其特征在于,所述可伸缩视频流包括可伸缩视频编码 (SVC) 基本流。
7. 如权利要求 1 所述的系统,其特征在于,所述修改的文件格式包括修改的可伸缩视频编码 (SVC) 文件格式。
8. 如权利要求 1 所述的系统,其特征在于,所述可伸缩数据流包括单个编码的音轨。
9. 如权利要求 1 所述的系统,其特征在于,所述可伸缩数据流包括一系列访问单元。
10. 如权利要求 9 所述的系统,其特征在于,还包括用于将所述一个或多个元数据框配置成根据所述一个或多个装置要求来定义多个描述项的其中一个描述项的装置。
11. 如权利要求 9 所述的系统,其特征在于,还包括用于将所述一个或多个元数据框配置成根据所选的描述项来定义访问单元的子集的装置,其中所定义的访问单元的子集包括所述子层数据流。
12. 如权利要求 11 所述的系统,其特征在于,所述一个或多个元数据框包括 SVC 样本组描述框,所述 SVC 样本组描述框配置成定义所述一个描述项。
13. 如权利要求 12 所述的系统,其特征在于,所述一个或多个元数据框包括 SVC 样本至分组框,所述 SVC 样本至分组框用于定义访问单元的子集并将其分组。
14. 如权利要求 1 所述的系统,其特征在于,所述一个或多个元数据框包括对伸缩视频编码 (SVC) 标准的扩充。
15. 如权利要求 1 所述的系统,其特征在于,所述可伸缩数据流包括多个子层数据流。
16. 如权利要求 1 所述的系统,其特征在于,还包括用于将所述可伸缩数据流的已编码版本存储在所述媒体数据部分中的装置。
17. 如权利要求 16 所述的系统,其特征在于,根据伸缩视频编码 (SVC) 标准将所述可伸缩数据流编码。

18. 如权利要求 1 所述的系统,其特征在于,还包括用于将所述一个或多个元数据框配置成定义与所述子层数据流关联的提示音轨的装置。

19. 如权利要求 18 所述的系统,其特征在于,所述子层数据流包括高级视频编码 (AVC) 兼容的基层流。

20. 一种从装置上实现的修改的文件格式中提取数据流的方法,所述方法包括:

a. 接收对特定数据流的请求,其中所述请求包括一个或多个装置要求;

b. 将所述请求与所述修改的文件格式的媒体数据部分中存储的特定可伸缩数据流关联;

c. 确定与所述特定可伸缩数据流对应的一个或多个音轨,其中所述一个或多个音轨存储在所述修改的文件格式的元数据部分中,而且其中每个音轨包括一个或多个元数据框;

d. 根据所述一个或多个装置要求来确定所述特定可伸缩数据流的子层数据流,其中使用所述一个或多个元数据框来确定所述子层数据流;以及

e. 从所存储的可伸缩数据流中提取所述子层数据流。

21. 如权利要求 20 所述的方法,其特征在于,还包括发送所提取的子层数据流。

22. 如权利要求 20 所述的方法,其特征在于,还包括在确定所述子层数据流之前将所确定的一个或多个音轨解码。

23. 如权利要求 20 所述的方法,其特征在于,所述一个或多个元数据框包括对伸缩视频编码 (SVC) 文件格式标准的扩充。

24. 如权利要求 20 所述的方法,其特征在于,所述可伸缩数据流包括可伸缩视频流。

25. 如权利要求 24 所述的方法,其特征在于,所述可伸缩视频流包括可伸缩视频编码 (SVC) 基本流。

26. 如权利要求 20 所述的方法,其特征在于,所述修改的文件格式包括修改的可伸缩视频编码 (SVC) 文件格式。

27. 如权利要求 20 所述的方法,其特征在于,所述可伸缩数据流包括单个编码的音轨。

28. 如权利要求 20 所述的方法,其特征在于,所述可伸缩数据流包括一系列访问单元。

29. 如权利要求 28 所述的方法,其特征在于,还包括将所述一个或多个元数据框配置成根据所述一个或多个装置要求来定义多个描述项的其中一个描述项。

30. 如权利要求 28 所述的方法,其特征在于,还包括将所述一个或多个元数据框配置成根据所选的描述项来定义访问单元的子集,其中所定义的访问单元的子集包括所述子层数据流。

31. 如权利要求 30 所述的方法,其特征在于,所述一个或多个元数据框包括 SVC 样本组描述框,所述 SVC 样本组描述框配置成定义所述一个描述项。

32. 如权利要求 31 所述的方法,其特征在于,所述一个或多个元数据框包括 SVC 样本至分组框,所述 SVC 样本至分组框用于定义访问单元的子集并将其分组。

33. 如权利要求 20 所述的方法,其特征在于,所述一个或多个元数据框包括对伸缩视频编码 (SVC) 标准的扩充。

34. 如权利要求 20 所述的方法,其特征在于,所述可伸缩数据流包括多个子层数据流。

35. 如权利要求 20 所述的方法,其特征在于,还包括将所述可伸缩数据流的已编码版本存储在所述媒体数据部分中。

36. 如权利要求 35 所述的方法,其特征在于,根据伸缩视频编码 (SVC) 标准将所述可伸缩数据流编码。

37. 如权利要求 20 所述的方法,其特征在于,还包括将所述一个或多个元数据框配置成定义与所述子层数据流关联的提示音轨。

38. 如权利要求 37 所述的方法,其特征在于,所述子层数据流包括高级视频编码 (AVC) 兼容的基层流。

39. 一种从装置上实现的修改的文件格式中提取数据流的方法,所述方法包括:

- a. 接收对特定数据流的请求,其中所述请求包括一个或多个装置要求;
- b. 将所述请求与所述修改的文件格式的媒体数据部分中存储的特定可伸缩数据流关联,所述可伸缩数据流包括一系列访问单元和多个子层数据流;
- c. 确定与所述特定可伸缩数据流对应的一个或多个音轨,其中所述一个或多个音轨存储在所述修改的文件格式的元数据部分中,而且其中每个音轨包括一个或多个元数据框,其中所述一个或多个元数据框包括对伸缩视频编码 (SVC) 文件格式标准的扩充;
- d. 根据所述一个或多个装置要求来确定所述特定可伸缩数据流的子层数据流,其中使用所述一个或多个元数据框来确定所述子层数据流;
- e. 从所存储的可伸缩数据流中提取所述子层数据流;
- f. 发送所提取的子层数据流;
- g. 在确定所述子层数据流之前将所确定的一个或多个音轨解码;
- h. 将所述一个或多个元数据框配置成根据所述一个或多个装置要求来定义多个描述项的其中一个描述项;
- i. 将所述一个或多个元数据框配置成根据所选的描述项来定义访问单元的子集,其中所定义的访问单元的子集包括所述子层数据流;
- j. 将所述可伸缩数据流的已编码版本存储在所述媒体数据部分中;
- k. 将所述一个或多个元数据框配置成定义与所述子层数据流关联的提示音轨,其中所述子层数据流包括高级视频编码 (AVC) 兼容的基层流。

40. 一种从装置上实现的修改的文件格式中提取数据流的系统,所述系统包括:

- 用于接收对特定数据流的请求的装置,其中所述请求包括一个或多个装置要求;
- 用于将所述请求与所述修改的文件格式的媒体数据部分中存储的特定可伸缩数据流关联的装置,所述可伸缩数据流包括一系列访问单元和多个子层数据流;
- 用于确定与所述特定可伸缩数据流对应的一个或多个音轨的装置,其中所述一个或多个音轨存储在所述修改的文件格式的元数据部分中,而且其中每个音轨包括一个或多个元数据框,其中所述一个或多个元数据框包括对伸缩视频编码 (SVC) 文件格式标准的扩充;
- 用于根据所述一个或多个装置要求来确定所述特定可伸缩数据流的子层数据流的装置,其中使用所述一个或多个元数据框来确定所述子层数据流;
- 用于从所存储的可伸缩数据流中提取所述子层数据流的装置;
- 用于发送所提取的子层数据流的装置;
- 用于在确定所述子层数据流之前将所确定的一个或多个音轨解码的装置;
- 用于将所述一个或多个元数据框配置成根据所述一个或多个装置要求来定义多个描述项的其中一个描述项的装置;

用于将所述一个或多个元数据框配置成根据所选的描述项来定义访问单元的子集的装置,其中所定义的访问单元的子集包括所述子层数据流;

用于将所述可伸缩数据流的已编码版本存储在所述媒体数据部分中的装置;

用于将所述一个或多个元数据框配置成定义与所述子层数据流关联的提示音轨的装置,其中所述子层数据流包括高级视频编码(AVC)兼容的基层流。

可伸缩视频编码 (SVC) 文件格式

[0001] 相关申请

[0002] 本申请要求相同发明人于 2005 年 7 月 15 日提交的美国临时申请序列号 60/699, 535、标题为“可伸缩视频编码 (Scalable Video CodingSVC) 文件格式”的优先权。本申请通过引用与 2005 年 7 月 15 日提交的美国临时申请序列号 60/699, 535、标题为“可伸缩视频编码 (SVC) 文件格式”结合。

技术领域

[0003] 本发明涉及视频编码的领域。更具体地说,本发明涉及 SVC 编码并扩充目前的 AVC 文件格式以便使用可伸缩视频编码来支持视频编码的数据的存储的领域。

背景技术

[0004] 文件格式是用于将信息编码以用于在计算机文件中存储的特定方式。存储文件的格式的一种便利方式是在文件系统中明确地存储有关格式的信息。此方法将元数据与主数据和文件名都分开。

[0005] ISO 基本媒体文件格式设计为包含定时媒体信息或媒体数据流 (例如电影)。存储的媒体信息可以在本地传输或经由网络或其他流送机构来传输。这些文件具有逻辑结构、时间结构和物理结构。文件的逻辑结构包括一组时间并行的音轨 (track)。文件的时间结构为这些音轨提供时间上的数据样本序列,并通过可选的编辑列表将那些序列映射到整个媒体数据流的时间线。文件的物理结构将逻辑、时间和结构分解所需的数据与媒体数据样本本身分开。此结构信息被汇聚在元数据框中,并且可能地通过元数据分段框在时间上延伸。该元数据框将数据样本的逻辑和定时关系归档,并还包括指向存储数据样本所在位置的指针。

[0006] 每个媒体数据流包含在专用于该媒体类型 (音频、视频等) 的音轨中,并进一步地由样本项来设定参数。样本项包括确切媒体类型的“名称”,例如将媒体数据流解码所需的解码器的类型和解码所需的其他参数。对于多种媒体类型,有定义的样本项格式。

[0007] 对于元数据的支持采用两种形式。首先,将定时的元数据存储于适合的音轨中,并将其与它所描述的媒体数据同步。其次,存在对附着于媒体数据流或个别音轨的非定时元数据的普遍支持。也以元数据框的形式在文件级使用这些规范化的元数据结构。在此情况中,元数据框是对存储的媒体数据流的主访问部件。

[0008] 在一些情况中,音轨内的数据样本具有不同特征或需要专门标识音轨内的数据样本。一种此类特征是同步点,通常称为视频 I-帧。这些点由每个音轨中的特殊表来标识。更一般性地,以此方式来归档音轨样本之间的相关特性。还存在样本组的概念。样本组能够将一些数据样本所共有的任意特征归档在音轨中。在高级视频编码 (AVC) 文件格式中,使用样本组来支持分层和子序列的概念。

[0009] AVC 文件格式定义用于根据 AVC 标准编码的视频流的存储格式。AVC 文件格式扩展 ISO 基本媒体文件格式。AVC 文件格式使 AVC 视频流能够结合其他媒体流 (例如音频)

一起使用,使用提示音轨进行格式化以通过流传输服务器来送交,以及继承 ISO 基本媒体文件格式的所有使用实例和特征。

[0010] 图 1 图示 AVC 文件格式 10 的示范配置,AVC 文件格式 10 包括媒体数据部分 20 和元数据部分 30。每个数据流被存储在媒体数据部分 20 中。可以将多个数据流以一种文件格式来存储。如图 1 所示,四个数据流 22、24、26 和 28 存储在媒体数据部分 20 中。对于 AVC 文件格式的媒体数据部分中存储的每个数据流,有元数据部分中存储的一个对应音轨。在图 1 中,音轨 32 对应于数据流 22,音轨 33 对应于数据流 24,音轨 36 对应于数据流 26,以及音轨 38 对应于数据流 28。一般来说,对应于数据部分中存储的 N 个数据流,有元数据部分中存储的 N 个音轨。

[0011] H. 264 或 MPEG-4 部分 10 规范是采取常常公知为联合视频组 (JVT) 的协同工作关系的 ITU-T 视频编码专家组 (VCEG) 与 ISO/TEC 运动图片专家组 (MPEG) 一起编制的一种高压压缩数字视频编解码器标准。ITU-T H. 264 标准和 ISO/IEC MPEG-4 部分 10 标准 (规范为 ISO/IEC 14496-10) 在技术上是完全相同的,该技术也公知为 AVC (高级视频编码的简称)。应该注意 H. 264 是与 H. 26x 视频标准的 ITU-T 线相关的名称,而在 ITU-T 中作为称为 H. 26L 的项目完成早期开发之后,AVC 与完成该标准工作的伙伴项目的 ISO/IEC MPEG 端相关。通常将该标准称为 H. 264/AVC (或 AVC/H. 264 或 H. 264/MPEG-4 AVC 或 MPEG-4/H. 264 AVC) 来强调其共有的继承性。有时候,根据开发它的 JVT 组织,也曾称为“JVT 编解码器”。

[0012] 目前 JVT 正在致力于开发一种公知为可伸缩视频编解码器 (SVC) 的新编解码器,它是对现有 AVC 编解码器的扩充。SVC 的开发工作最初在 MPEG 领域中作为 MPEG-21 标准的一部分于 2003 年独立启动。但是,在 2004 年的开发期间,它与 JVT 团体的活动合并以着重于开发与现有 AVC 编解码器后向兼容的编码技术。因此,它目前在 MPEG 和 ITU-T 中由 JVT 团体联合开发。可伸缩视频编解码器 (SVC) 活动的目标是解决对空间、时间和指令 (SNR) 级别中的可伸缩性的需求和提供空间、时间和指令 (SNR) 级别中的可伸缩性。

[0013] 现有文件格式 (ISO/MP4 和 AVC) 未提供容易且清楚的机制以从以该文件格式存储的媒体数据中提取空间、时间和 SNR (质量) 层的不同变化。因此,必须通过分析编码的媒体流来提取此信息,这是非常低效且缓慢的。因此,有需要增强并定义新扩充来支持新出现的视频编码标准 (SVC) 的存储以及解决目前文件格式存储方法的现有局限性。这些新扩充定义用于一组图片和每个样本内存在的相关性的结构化和分组机制,以便获取提供空间、时间和质量方面灵活性的灵活的流结构。SVC 标准提出将整个可伸缩媒体数据作为一个可伸缩比特流来编码,从其中可以提取时间、空间和质量层的变化。

[0014] 在 AVC 标准中,将每个视频流作为符合特定帧速率、分辨率和质量的独立流来编码,此后将其作为符合特定帧速率、分辨率和质量的独立流来解码。根据 SVC 标准,从单个编码的视频流 (称为 SVC 基本流) 中,可以提取多种不同类型的视频,例如低分辨率的视频流、标准分辨率的视频流或高分辨率的视频流。要支持文件格式中的此类可伸缩视频流的存储和提取,需要修改文件格式。

[0015] SVC 标准目前正在开发中,并且因为 SVC 标准定义视频编解码器的新设计,所以还需要适当定义的新文件标准以便能够存储和提取新的 SVC 视频流。为了支持新的 SVC 视频流,正在开发的 SVC 文件格式将 AVC 文件格式扩充成支持 SVC 视频流的存储。但是,仍需要开发定义对存储的可伸缩视频的访问的特定扩充。

发明内容

[0016] 描述一种系统和方法,用于将目前的 ISO/MP4/AVC 文件格式扩充成存储例如使用 MPEG-4 :部分 10/Amd-1 可伸缩视频编解码器 (SVC) 标准编码的视频内容,在 MPEG/ITU-T 中,MPEG-4 :部分 10/Amd-1 可伸缩视频编解码器 (SVC) 标准的开发目前正在进行中。确切地说,进行对 AVC 文件格式的扩充以便提供能够存储和访问可伸缩视频数据的新 SVC 文件格式。可以将可伸缩视频数据作为 SVC 文件格式的媒体数据部分内的单个音轨来存储。对 SVC 文件格式的元数据部分内的描述项和框定义了新扩充。这些扩充提供从媒体数据部分中存储的可伸缩视频数据的单个音轨提取子流或层的方式。

[0017] 在一方面中,修改的文件格式包括用于存储可伸缩数据流的媒体数据部分,以及包括与媒体数据部分中存储的可伸缩数据流关联的至少一个音轨的元数据部分,其中每个音轨包括一个或多个元数据框,该一个或多个元数据框用于定义可伸缩数据流的子层数据流并将其分组。该可伸缩数据流可以包括可伸缩视频流。该可伸缩视频流可以包括可伸缩视频编码 (SVC) 基本流。该修改的文件格式可以包括修改的可伸缩视频编码 (SVC) 文件格式。该可伸缩数据流可以包括单个编码的音轨。该可伸缩数据流可以包括一系列访问单元。该一个或多个元数据框可以配置成根据从能够处理子层数据流的终端用户装置接收的一个或多个装置要求来定义该子层数据流。该一个或多个元数据框还可以配置成根据一个或多个装置要求来定义多个描述项的其中一个描述项。该一个或多个元数据框还可以配置成根据所选的描述项来定义访问单元的子集,其中所定义的访问单元的子集包括子层数据流。该一个或多个元数据框可以包括 SVC 样本组描述框,该 SVC 样本组描述框配置成定义一个描述项。该一个或多个元数据框可以包括 SVC 样本至分组框 (SVC Sample To Group Box),该 SVC 样本至分组框用于定义访问单元的子集并将其分组。该一个或多个元数据框可以包括对伸缩视频编码 (SVC) 标准的扩充。该可伸缩数据流可以包括多个子层数据流。该一个或多个元数据框可以配置成定义与子层数据流关联的提示音轨。该子层数据流可以包括高级视频编码 (AVC) 兼容的基层流。

[0018] 在另一方面中,描述了配置成利用修改的文件格式的文件服务器。该文件服务器包括用于根据修改的文件格式来存储和提取数据的存储器,其中该修改的文件格式包括用于存储可伸缩数据流的媒体数据部分,以及包括与元数据部分中存储的可伸缩数据流关联的至少一个音轨的元数据部分,其中每个音轨包括一个或多个元数据框,该一个或多个元数据框用于定义可伸缩数据流的子层数据流并将其分组。该文件服务器还包括处理模块,该处理模块配置成向存储器提供控制指令并从可伸缩数据流中提取子层数据流。该可伸缩数据流可以包括一系列访问单元。该文件服务器还可以包括网络接口模块,该网络接口模块配置成从终端用户装置接收一个或多个装置要求并传送所定义的子层数据流。该一个或多个元数据框还可以配置成根据一个或多个装置要求来定义多个描述项的其中一个描述项。该一个或多个元数据框还可以配置成根据所选的描述项来定义访问单元的子集,其中所定义的访问单元的子集包括子层数据流。该一个或多个元数据框可以包括 SVC 样本组描述框,该 SVC 样本组描述框配置成定义一个描述项。该一个或多个元数据框可以包括 SVC 样本至分组框,该 SVC 样本至分组框用于定义访问单元的子集并将其分组。该可伸缩数据流可以包括可伸缩视频流。该可伸缩数据流可以包括可伸缩视频编码 (SVC) 基本流。

该修改的文件格式可以包括修改的可伸缩视频编码 (SVC) 文件格式。该可伸缩数据流可以包括单个编码的音轨。该可伸缩数据流可以包括多个子层数据流。

[0019] 在另一方面中,描述了配置成利用修改的文件格式的系统。该系统包括用于传送一个或多个装置要求的终端用户装置和配置成接收一个或多个装置要求并利用修改的文件系统的文件服务器。该文件服务器包括存储器和处理模块。该存储器配置成根据修改的文件格式来存储和提取数据,其中该修改的文件格式包括用于存储可伸缩数据流的媒体数据部分,以及包括与元数据部分中存储的可伸缩数据流关联的至少一个音轨的元数据部分,其中每个音轨包括一个或多个元数据框,该一个或多个元数据框用于根据一个或多个装置要求来定义可伸缩数据流的子层数据流并将其分组。该处理模块配置成向存储器提供控制指令并从媒体数据部分中提取子层数据流。该文件服务器还可以包括网络接口模块,该网络接口模块配置成从终端用户装置接收一个或多个装置要求并传送所定义的子层数据流。该可伸缩数据流可以包括一系列访问单元。该一个或多个元数据框还可以配置成根据一个或多个装置要求来定义多个描述项的其中一个描述项。该一个或多个元数据框还可以配置成根据所选的描述项来定义访问单元的子集,其中所定义的访问单元的子集包括子层数据流。该一个或多个元数据框可以包括 SVC 样本组描述框,该 SVC 样本组描述框配置成定义一个描述项。该一个或多个元数据框可以包括 SVC 样本至分组描述框,该 SVC 样本至分组描述框用于定义访问单元的子集并将其分组。该可伸缩数据流可以包括可伸缩视频流。该可伸缩视频流可以包括可伸缩视频编码 (SVC) 基本流。该修改的文件格式可以包括修改的可伸缩视频编码 (SVC) 文件格式。该一个或多个元数据框可以包括对伸缩视频编码 (SVC) 标准的扩充。该一个或多个元数据框可以包括对伸缩视频编码 (SVC) 标准的扩充。该可伸缩数据流可以包括单个编码的音轨。该可伸缩数据流可以包括多个子层数据流。该一个或多个元数据框可以配置成定义与子层数据流关联的提示音轨。该子层数据流可以包括高级视频编码 (AVC) 兼容的基层流。

[0020] 在另一方面中,描述了从修改的文件格式中提取数据流的方法。该方法包括接收对特定数据流的请求,其中该请求包括一个或多个装置要求,将该请求与修改的文件格式的媒体数据部分中存储的特定可伸缩数据流关联,确定与该特定可伸缩数据流对应的一个或多个音轨,其中该一个或多个音轨存储在修改的文件格式的元数据部分中,每个音轨包括一个或多个元数据框,该一个或多个元数据框根据一个或多个装置要求来确定特定可伸缩数据流的子层数据流,其中使用该一个或多个元数据框来确定子层数据流,并从存储的可伸缩数据流中提取子层数据流。该方法还可以包括传送所提取的子层数据流。该方法还可以包括在确定子层数据流之前将所确定的一个或多个音轨解码。该一个或多个元数据框可以包括对伸缩视频编码 (SVC) 文件格式标准的扩充。该可伸缩数据流可以包括可伸缩视频流。该可伸缩视频流可以包括可伸缩视频编码 (SVC) 基本流。该修改的文件格式可以包括修改的可伸缩视频编码 (SVC) 文件格式。该可伸缩数据流可以包括单个编码的音轨。该可伸缩数据流可以包括一系列访问单元。该方法还可以包括将一个或多个元数据框配置成根据一个或多个装置要求来定义多个描述项的其中一个描述项。该方法还可以包括将一个或多个元数据框配置成根据所选的描述项来定义访问单元的子集,其中所定义的访问单元的子集包括子层数据流。该一个或多个元数据框可以包括 SVC 样本组描述框,该 SVC 样本组描述框配置成定义一个描述项。该一个或多个元数据框可以包括 SVC 样本至分组描述框,

该 SvC 样本至分组描述框用于定义访问单元的子集并将其分组。该一个或多个元数据框可以包括对伸缩视频编码 (SVC) 标准的扩充。该可伸缩数据流可以包括多个子层数据流。该方法还可以包括将可伸缩数据流的已编码版本存储在媒体数据部分中。该可伸缩数据流可以根据伸缩视频编码 (SVC) 标准来进行编码。该方法还可以包括将一个或多个元数据框配置成定义与子层数据流关联的提示音轨。该子层数据流可以包括高级视频编码 (AVC) 兼容的基层流。

附图说明

- [0021] 图 1 图示 AVC 文件格式的示范配置。
- [0022] 图 2 图示包括配置成实现修改的文件格式的文件服务器的示范网络的框图。
- [0023] 图 3 图示图 2 的文件服务器的内部组件的示范框图。
- [0024] 图 4 图示 SVC 基本流的示范配置。
- [0025] 图 5 图示修改的 SVC 文件格式的示范配置。
- [0026] 图 6 示出根据 SVC 标准的 SVC 访问单元的示范结构。
- [0027] 图 7 图示实现修改的 SVC 文件格式的示范方法。
- [0028] 修改的文件格式的实施例是参考多个附图来描述的。在一个或多个附图中公开并示出适合且仅完全相同的部件的情况下,将使用相同的引用号来表示此类完全相同的部件。

具体实施方式

[0029] 图 2 图示包括配置成实现修改的文件格式的文件服务器的示范网络的框图。文件服务器 50 经由网络 70 耦合到回放装置 60。网络是能够传送数据的任何便利的网络,包括有线或无线网络。回放装置 60 是能够接收和处理所传送的数据的任何便利的装置。

[0030] 图 3 图示图 2 的文件服务器 50 的内部组件的示范框图。该文件服务器 50 是可配置成实现修改的文件格式的任何便利的计算装置。该文件服务器 50 包括处理模块 82、主机存储器 84、视频存储器 86、海量存储装置 88 以及接口电路 90,所有这些部件由常规双向系统总线 92 耦合在一起。该接口电路 90 包括用于通过网络 70 发送和接收通信的物理接口电路 (图 3)。该接口电路 90 在文件服务器 50 内的网络接口卡上实现。但是,对于本领域技术人员来说,应该显见到接口电路 90 可以在文件服务器 50 内以任何其他适合的方式来实现,包括将该接口电路内置于主板本身上。海量存储装置 88 可以包括采用磁、光或磁光存储技术中的任何一种或多种技术或任何其他可用海量存储技术的固定和可移动的媒体。系统总线 92 使得能够访问存储器 84 和 88 的任何部分以及能够实现 CPU 82、主机存储器 84、视频存储器 86 和海量存储装置 88 之间及其中间的数据传输。主机存储器 84 用作处理模块 82 使用的系统主存储器。

[0031] 文件服务器 50 还耦合到多个外围输入和输出装置,包括键盘 94、鼠标 96 和关联的显示器 98。键盘 94 耦合到 CPU 82 以使用户能够将数据和控制命令输入到文件服务器 50 中。鼠标 96 耦合到键盘 94 或耦合到 CPU 82,以便作为光标控制装置来操纵显示器 98 上的图形图像。文件服务器 50 包括用于将数据转换成适于显示的信号的图形电路 100。要理解的是图 3 所示的文件服务器 50 的配置仅是出于举例的目的,并且可以采用任何其他便利的

方式来配置文件服务器 50。

[0032] 因为 SVC 文件格式是 AVC 文件格式的扩充,所以 SVC 文件格式也包括元数据部分和媒体数据部分。媒体数据部分存储未改变的编码的 SVC 基本流,例如视频数据流,其中 SVC 基本流由一系列访问单元组成。图 4 图示 SVC 基本流 40 的示范配置。SVC 基本流 40 包括一系列称为访问单元 (AU) 的连续单元。接收 SVC 基本流的解码器将每个访问单元解码成图片,由此产生视频序列。

[0033] 元数据部分存储有关媒体数据部分中存储的每个 SVC 基本流的信息。此类信息包括但不限于,视频流的类型、视频流的分辨率、视频流的帧速率、视频流中每个访问单元的存储地址、视频流内的随机访问点以及有关何时要将每个访问单元解码的定时。

[0034] 可伸缩视频流使得能够与多种网络相适应。例如,如果以每秒 60 帧的速率将可伸缩视频流编码,但是回放装置仅支持每秒 30 帧,则仅将可伸缩视频流的一部分传送到该回放装置。又如,如果编码的可伸缩视频流的质量非常高,如 10Mbps,但是传送该可伸缩视频流所要通过的网络仅支持 1Mbps,则同样仅传送该可伸缩视频流的一部分以与该网络所支持的传输速度相匹配。以此方式,基于网络或回放装置的要求以从文件格式中提取编码的可伸缩视频流的全部或仅部分。

[0035] 基于如果将整个可伸缩视频流解码,则结果是完整分辨率、完整帧速率和高质量,所以可伸缩视频流是完全可伸缩的。但是,如果网络或回放装置不支持整个可伸缩视频流的完整分辨率、完整帧速率或高质量,则仅从文件格式的数据部分提取该可伸缩视频流的若干部分,并通过网络传送这些部分,由此节省了带宽。

[0036] 为了以 SVC 文件格式来进行存储,根据 SVC 标准定义的格式将可伸缩视频流编码。将编码的可伸缩视频流存储在 SVC 文件格式的数据部分中。将音轨存储在 SVC 文件格式的元数据部分中,其中这些音轨包含与媒体数据部分中存储的可伸缩视频流对应的元数据信息。这些音轨包含用于提取 SVC 文件格式的数据部分中存储的全部或部分编码的可伸缩视频流的信息。在每个音轨内,定义了新的元数据框。这些新元数据框定义用于从可伸缩视频流提取多种类型的视频流的参数。用于定义提取的视频流的参数包括但不限于分辨率、帧速率和质量。例如,一个提取的视频流对应于低分辨率要求,第二个提取的视频流对应于标准分辨率要求,以及第三个提取的视频流对应于高分辨率要求。

[0037] 图 5 图示修改的 SVC 文件格式 110 的示范配置。修改的 SVC 文件格式 110 包括媒体数据部分 120 和元数据部分 130。媒体数据部分 120 包括一个或多个可伸缩数据流、即 SVC 基本流。元数据部分 130 包括一个或多个音轨。在一个实施例中,对应于媒体数据部分 120 中存储的每个 SVC 基本流,元数据部分 130 中有一个音轨。如图 5 所示,媒体数据部分 120 包括两个 SVC 基本流 122 和 124。元数据部分 130 中的音轨 132 对应于媒体数据部分 120 中存储的 SVC 基本流 122。元数据部分 130 中的音轨 134 对应于媒体数据部分 120 中存储的 SVC 基本流 124。每个 SVC 基本流包括一系列访问单元。如图 5 所示,SVC 基本流 122 扩展为显示它的访问单元的一部分。

[0038] 在修改的 SVC 文件格式的元数据部分内提供新的扩充,以便定义并提取存储的可伸缩视频流内的子层,其中按指定的装置要求(例如分辨率、帧速率和/或质量)来确定特定的子层数据流。定义一个或多个元数据框来标识和提取与每个子层数据流对应的特定访问单元。每个音轨配置成包括这些一个或多个元数据框。图 5 示出一个实施例,其中定

义了元数据框 140 和元数据框 142 以便定义并提取构成该特定子层数据流的特定访问单元。在此实施例中,元数据框 140 称为 SampleGroupDescription 框,以及元数据框 142 称为 SampleToGroup 框。

[0039] 参考图 3 和 5,回放装置 60 向文件服务器 50 发送对特定视频流的请求,在本例中是对文件格式 110 的媒体数据部分 120 中存储的 SVC 基本流 122 的请求。该请求包含回放装置 60 的必要规范,例如所支持的特定分辨率和帧速率。当接收到此请求时,文件服务器 50 将所请求的视频流 SVC 基本流 122 与元数据部分 130 中其对应的音轨 132 匹配。使用文件格式解码器,仅将匹配音轨 132 解码。文件格式解码器并不将元数据部分 130 中存储的其他音轨解码,该文件格式解码器也不将媒体数据部分 120 中存储的包含 SVC 基本流 122 的任何编码的可伸缩视频流解码。此外,对媒体数据部分 120 中存储的媒体数据编码不同于元数据部分 130 中存储的元数据,因此需要不同的解码器。

[0040] 一旦确定音轨 132 并将其解码,则访问元数据框 140。元数据框 140 利用一个或多个装置要求来确定匹配描述项。匹配描述项定义与这些装置要求对应的参数值和 group_description_index 值。元数据框 142 使用 group_description_index 的值来标识并提取对应 SVC 基本流 122 中的特定访问单元。图 5 图示元数据框 142 确定并提取访问单元 1、4、7 等的示范例子。所提取的访问单元保持未改变和编码状态。一旦提取,则将访问单元作为满足最初提供的装置要求的子层数据流传送到终端用户装置。

[0041] 在上面结合图 5 描述的实施例中,将每个可伸缩视频流作为单个视频流存储在文件格式的数据部分,并将与该单个视频流关联的对应一个音轨存储在文件格式的元数据部分中。定义新的扩充,以提供从单个视频流中提取感兴趣的子流或层的方式。或者,可以将多于一个音轨与单个视频流关联。

[0042] 作为此备选实施例的实例,假定为不同的终端用户装置创建了单独的音轨。这些装置的每个装置仍可能具有它们可以支持的内部帧速率、空间分辨率、质量等的各自内部子集。一种可伸缩视频流支持多种装置要求,例如空间分辨率 (QCIF、CIF、SD 和 HD)、帧速率 (7.5、15、30、60fps) 和对应于上面这些的多种质量。仍将流存储在修改的文件格式的媒体数据部分中。在元数据部分中,可以有例如三个音轨。这三个音轨的每个音轨指代相同的可伸缩视频流。但是,每个音轨对整个可伸缩流的子集起作用。例如,将音轨 1 定制为用于便携式播放器 (小尺寸屏幕、低帧速率等),将音轨 2 定制为用于标准 TV、计算机等 (SD 尺寸屏幕、中等帧速率等) 以及将音轨 3 定制为用于 HD 播放器 (大尺寸屏幕、高帧速率等)。每个音轨仍包括定义描述项的元数据框,因为每个音轨仍支持一定数量的变化 (可伸缩性)。在此实例中,音轨 1 支持 QCIF 和 CIF、7.5 和 15fps。音轨 2 支持全部音轨 1 以及 SD 和 30fps。音轨 3 支持全部音轨 2 以及 HD 和 60fps。

[0043] 在此备选实施例中,终端用户装置所请求的特定视频流不对应于单个音轨。设想不同的方法来在此情况中确定适合的音轨。首先,文件服务器将音轨的一部分解码以确定它定义的装置要求。其次,文件服务器在高级别上分析这些音轨来标识要使用哪个音轨。第三,音轨具有提供有关它们描述的内容的更多信息的音轨头。在此特定情况中,文件服务器将一些高级别信息从每个音轨中解码来决定要使用哪个音轨。一般来说,可以使用将终端用户装置请求中接收到的装置要求与适合的音轨匹配的任何方法。

[0044] 在一个实施例中,通过定义两个新框 SampleGroupDescription 框和

SVCSampleToGroup 框来支持提取 SVC 基本流的若干部分。这两个新框定义用于标识要提取可伸缩视频流内的哪些访问单元的特征和字段。在定义唯一的描述项的 SampleGroupDescription 框内定义每个新描述项,其中每个描述项指代可以从 SVC 基本流中提取的不同类型的视频流。例如,如果将 SVC 基本流编码成支持三种不同的分辨率:低、标准和高,则有三个描述项,每个支持的分辨率对应于一个描述项。每个描述项定义如下所述对应于特定分辨率(低、标准或高)的描述项字段。如果将 SVC 基本流编码成支持三种不同的分辨率:低、标准和高,并且还支持两种不同的帧速率:高和低,则有六个不同的描述项,每个分辨率和帧速率的组合对应于一个描述项。

[0045] SVC 相关性描述项或简称为“描述项”将 SVC 基本流内存在的多种可能的空间-时间组合归档并描述它们。使用分组类型“svcd”来定义每个描述项。依据这些描述项的相关性来将其排序。第一个描述项归档基层描述,而后续描述项归档增强层描述。使用 group_description_index 来对排序的描述项建立索引。还使用 group_description_index 来将多种 SVC NAL 单元分组。每个 SVC NAL 单元指代描述项的其中之一。

[0046] 使用索引 group_description_index 将 SVC 基本流中的每个 SVCNAL 单元分组成排序的描述项列表。指代特定索引的 NAL 单元可能需要所有或一些指代所有较低索引的 NAL 单元来实现适合的解码操作,但是不需要指代较高索引值的任何 NAL 单元。换言之,仅在较低层方向上存在相关性。

[0047] 在一个实施例中,文件服务器基于描述项内存在的描述字段的值(例如分辨率和或时间速率)来确定适合的解码操作所需的索引的子集。在另一个实施例中,终端用户装置确定索引的子集。

[0048] 用于指代排序的描述项列表的分组类型是“svcd”。分组类型用于将此列表中存在项链接到 SVCSampleToGroup 框,SVCSampleToGroup 框包括所有 SVC 样本的分组,下文将对此进行更详细的描述。

[0049] 下列的是示范 SVC 相关性描述项语法:

```
[0050] class SVCDependencyDescriptionEntry() extends
VisualSampleGroupEntry('svcd')
```

```
[0051] {
```

```
[0052]     unsigned int(8)ProfileIndication;
```

```
[0053]     unsigned int(8)profile_compatibility;
```

```
[0054]     unsigned int(8)LevelIndication;
```

```
[0055]     unsigned int(8)temporal_level;
```

```
[0056]     unsigned int(8)dependency_id;
```

```
[0057]     unsigned int(8)temporalFrameRate;
```

```
[0058]     unsigned int(16)visualWidth;
```

```
[0059]     unsigned int(16)visualHeight;
```

```
[0060]     unsigned int(16)baseBitRate;
```

```
[0061]     unsigned int(16)maxBitRate;
```

```
[0062]     unsigned int(16)avgBitRate;
```

```
[0063]     unsigned int(8)progressiveRefinementLayerFlag;
```

```

[0064]         unsigned int(32)reserved = 0 ;
[0065] // 当稍后定义 ScalabilityInformationSEIBox() 时可以跟随可选的框或字段 ;
[0066] // 可选
[0067] }
[0068] class ScalabilityInformationSEIBox extends Box(' seib' )
[0069] {
[0070] // 包含 JSVM 2.0 文献编码描述 (N7084) 中的第 4.1 节可伸缩信息 SEI 消息语法
中定义的所有字段
[0071] }

```

[0072] 变量 profileIndication 包括 AVC/SVC 视频规范中定义的分布 (profile) 代码。变量 profile_compatibility 是定义为刚好与 AVC/SVC 视频规范中定义的序列参数集中的 profile_DC 与 level_DC 之间出现的字节相同的字节。变量 levelIndication 包括 AVC/SVC 视频规范中定义的级别代码。每个项中存在的 profileIndication 和 levelIndication 字段指示字段提供特定层所兼容的分布和级别值。变量 temporal_level 取 AVC/SVC 视频规范中定义的可伸缩扩充 NAL 单元中存在的 temporal_level 语法元素的值。其非负整数指示样本沿时间提供的时间级别。将最低时间级别编号为零,将时间方向上的增强层编号为 1 或更高数值。在 AVC NAL 单元的情况中,temporal_level 字段取零的缺省值。在 SVC 可伸缩扩充 NAL 单元中,如果 extension_flag 等于 0,则 SPS 中存在指定将 simple_priority_id 到 temporal_level 的映射的参数,并相应地对其进行映射。

[0073] 变量 dependency_id 取 AVC/SVC 视频规范中定义的可伸缩扩充 NAL 单元中存在的 dependency-id 语法元素的值。Dependency_id 是非负整数,其零值表明 NAL 单元对应于最低空间分辨率,而所有较高值表明增强层提供空间分辨率和 / 或质量 (例如粗略颗粒可伸缩性 (coarse grain scalability)) 中的提高。dependency_id 也控制空间可伸缩性。在 AVC NAL 单元的情况中,dependency_id 字段取零的缺省值。在 SVC 可伸缩扩充 NAL 单元中,如果 extension_flag 等于 0,则 SPS 中存在指定 simple_priority_id 到 dependency_id 的映射的参数,并相应地对其进行映射。变量 temporalFrameRate 指示与项中的时间级别字段关联的时间帧速率。变量 visualWidth 对 SVC 流的此层中的像素中的编码的图片的宽度赋值。变量 visualHeight 对 SVC 流的此层中的像素中的编码的图片的高度赋值。变量 baseBitRate 赋予在无任何渐进精细化 (progressive refinement) 的情况下由此层提供的最低质量的比特率 (以比特数 / 秒为单位)。在计算中将落在相关性层次结构内的此级别和较低级别中的 NAL 单元纳入考虑。变量 maxBitRate 赋予任何一秒的窗口上由此层提供的最高速率 (以比特数 / 秒为单位)。在计算中将落在相关性分层结构内的此级别和较低级别中的 NAL 单元纳入考虑。变量 avgBitRate 赋予平均比特率 (以比特数 / 秒为单位)。在计算中将此级别和较低级别中的 NAL 单元纳入考虑。当变量 progressiveRefinementLayerFlag 为真时,它指示此层包含渐进精细化单元,并且是 FGS 可伸缩的。

[0074] SVCSampleToGroup 框用于根据时间、空间和质量要求方面施加的约束,从 SVC 文件格式中存储的 SVC 基本流提取 SVC 可伸缩子流。SVCSampleToGroup 框提供用于 SVC 样本的每个 NAL 单元的分组信息。分组信息通过 group_description_index 来提供,group_description_index 将每个 NAL 单元与 SVCDependencyDescriptionEntry 中存在的其描述

信息关联。group_description_index 指代特定的描述项。group_description_index 的范围是 1 至 SampleGroupDescription 框中的样本组项的数量。对这些访问单元内的每个 NAL 单元赋予特定的 group_description_index 值。终端用户装置指定的要求最终确定特定的 group_description_index 值。从文件格式的媒体数据部分中存储的可伸缩视频流中提取这些访问单元内被赋予所确定的 group_description_index 值的所有 NAL 单元,或这些访问单元内其 group_description_Index 值低于所确定的 group_description_index 值的那些 NAL 单元。将那些提取的访问单元传送到终端用户装置。以此方式,可以根据终端用户装置指定的要求来提取单个可伸缩视频流的若干部分,并将其传送到终端用户装置。

[0075] 下列的是示范 SVCSampleToGroup 框语法:

```
[0076] aligned(8)class SVCSampleToGroupBox extends FullBox(' svcg' , version
= 0,0)
[0077] {
[0078]     unsigned int(32)grouping_type ;// 分组类型“svcd”
[0079]     unsigned int sample_count ;// 由样本至块框计算
[0080]     for(int i = 1 ;i <= sample_count ;i++)
[0081]     {
[0082]         unsigned int(8)numberOfNalUnits ;
[0083]         for(int j = 1 ;j <= numberOfNalUnits ;j++)
[0084]         {
[0085]             unsigned int(8)         group_description_index ;
[0086]             unsigned int(1)         isDiscardableNalUnitFlag ;
[0087]             unsigned int(1)         isPRNalUnitFlag ;
[0088]             unsigned int(2)         quality_level ;
[0089]             unsigned int(4)         reserved = 0 ;
[0090]         }
[0091]     }
[0092] }
```

[0093] 变量 grouping_type 是标识所采用的样本分组的类型的整数。它将 SVCSampleToGroup 框链接到它关联的具有分组类型值“svcd”的样本组描述表。变量 sample_count 表示 SVC 基本流的媒体音轨中存在的样本的数量,它是使用样本至块 (Sample to Chunk) 框计算的推导值。变量 numberOfNalUnits 是赋予 SVC 样本中存在的 NAL 的数量的整数。变量 group_description_index 是赋予描述此组中的 NAL 单元的样本组项的索引的整数。该索引的范围为 1 至特定 SampleGroupDescription 框中的样本组项的数量,或取值 0 以指示该特定的 SVC 样本是此类型的无组的成员。变量 isDiscardableNalUnitFlag 是一个标记,在 AVC 视频规范的 NAL 单元语义中指定了其语义。变量 isPRNalUnitFlag 是一个标记,如果等于 1,则该标记指示 NAL 单元是渐进精细化 NAL 单元。变量 quality_level 指定 NAL 单元头或 quality_level_list[] 中指定的当前 NAL 单元的质量级别。如果不存在,则将其值推导为零。

[0094] ISO 基本语法中定义的子样本信息框无法容易地用于定义用来提取 SVC 基本流的

选定部分的 SampleToGroup 框。首先,子样本信息框没有映射到相关描述项的“grouping_type”指示符。其次,子样本信息框对于达到此目的是不够的,因为子样本信息框需要称为“sample_count”的(32)位特项。这起初是用来指示一组样本(a runof samples),但是在 SVC 文件格式中,每个 SVC 样本可以具有可变数量的 subsample_count,从而利用了对每个 SVC 样本强制的此 sample_count。第三,新的文件格式无需告知每个子样本大小,因为使用 NAL 单元的计数,并且在它们之前已经指定每个 NAL 单元的长度。

[0095] 图 6 图示根据 SVC 标准的 SVC 样本或 SVC 访问单元的示范结构。SVC 样本在外部成帧,并且具有该外部成帧提供的大小。SVC 访问单元由一组 NAL 单元组成。每个 NAL 单元是附有长度来表示的,该长度指示后续 NAL 单元的以字节为单位的长度。该长度字段配置为 1、2 或 4 个字节的。该配置大小是在解码器配置记录中指定的。每个 NAL 单元包含 ISO/TEC AVC/SVC 视频规范中指定的 NAL 单元数据。

[0096] SVC 解码器配置记录包含每个样本中使用的长度字段的大小,以便指示它包含的 NAL 单元的长度以及初始参数集。SVC 解码器配置记录是外部成帧的,这意味着它的大小是由包含它的结构提供的。SVC 解码器配置记录还包括版本字段。版本号的改变指示对 SVC 解码器配置记录的不兼容更改。

[0097] 当用来提供参数集基本流的配置或结合参数集基本流使用的视频基本流的配置时,该配置记录未包含任何序列参数集或图片参数集,例如变量 numOfSequenceParameterSets 和 numOfPictureParameterSets 具有值。

[0098] SVCProfileIndication、SVCLevelIndication 和指示分布兼容性的标记的值对于可伸缩视频流的所有参数集有效。级别指示指示等于或大于所含的参数集中指示的最高级别的能力的级别。如果将所有所含的参数集设置了该标记,则每个分布兼容性标记被设置。分布指示指示整个流所遵循的分布。在 SVCDependencyDescriptionEntry 框中归档每个层的个别分布和级别。

[0099] 下列的是示范 SVC 解码器配置记录语法:

```
[0100] aligned(8)class SVCDecoderConfigurationRecord{
[0101]     unsigned int(8)configurationVersion = 1;
[0102]     unsigned int(8)SVCProfileIndication;
[0103]     unsigned int(8)profile_compatibility;
[0104]     unsigned int(8)SVCLevelIndication;
[0105]     bit(6)reserved = '111111' b;
[0106]     unsigned int(2)lengthSizeMinusOne;
[0107]     bit(3)reserved = '111' b;
[0108]     unsigned int(5)numOfSequenceParameterSets;
[0109]     for(i = 0;i < numOfSequenceParameterSets;i++){
[0110]         unsigned int(16)sequencePararneterSetLength;
[0111]         bit(8*sequenceParameterSetLength)
sequenceParameterSetNALUnit;
[0112]     }
[0113]     unsigned int(8)numOfPictureParameterSets;
```



```

[0114]         for(i = 0, i < numOfPictureParameterSets ;i++) {
[0115]             unsigned int(16)pictureParameterSetLength ;
[0116]                 bit(8*pictureParameterSetLength)
pictureParameterSetNALUnit ;
[0117]         }
[0118]     }

```

[0119] 变量 SVCProfileIndication 包含 SVC 规范中的定义的分部码。变量 profile_compatibility 是定义为与 AVC 视频规范中定义的序列参数集中的 profile_DC 与 level_IDC 之间出现的字节相同的字节。变量 SVCLevelIndication 包含 AVC 规范中的定义的级别码。变量 lengthSizeMinusOne 指示关联的流的 SVC 视频样本或 SVC 参数集样本中的 NALUnitLength 字段的长度（以字节为单元）减一。例如，以 0 的值指示一个字节的长度。此字段的值是分别对应于以 1、2 或 4 个字节编码的长度的 0、1 或 3 的其中之一。变量 numOfSequenceParameterSets 指示用作用于将 SVC 基本流解码的序列参数集的初始设置的序列参数集的数量。变量 sequenceParameterSetLength 指示 AVC 规范中定义的序列参数集 NAL 单元的长度（以字节为单位）。变量 sequenceParameterSetNALUnit 包含 AVC 规范中指定的序列参数集 NAL 单元。序列参数集按参数集标识符的升序出现，且允许空缺。变量 numOfPictureParameterSets 指示用作用于将 SVC 基本流解码的图片参数集的初始设置的图片参数集的数量。变量 pictureParameterSetLength 指示 AVC 规范中定义的图片参数集 NAL 单元的长度（以字节为单位）。变量 pictureParameterSetNALUnit 包含 AVC 规范中指定的图片参数集 NAL 单元。图片参数集按参数集标识符的升序出现，且允许空缺。

[0120] 如本文描述的，可伸缩 SVC 视频流作为单个音轨存储。如果可伸缩 SVC 视频流具有属于 AVC 兼容的基层，则使用新扩充将每个 SVC 样本中存在的那些 AVC 兼容 NAL 单元分组在一起，如前文描述的。为了找出属于 AVC 兼容的项，使用 SVCDependencyDescriptionEntries 中存在的分布和级别指示符来分析 SVCSampleToGroup 框，并从属于 AVC 兼容的每个 SVC 样本中仅提取那些 NAL 单元。

[0121] 修改的 SVC 文件格式从 ISO 基本媒体文件格式导出。因此，存在修改的 SVC 文件格式与 ISO 基本媒体文件格式中的术语的对应。例如，修改的 SVC 文件格式中使用的术语流和访问单元分别对应于 ISO 基本媒体文件格式中的术语音轨和样本。

[0122] 在 ISO 基本媒体文件格式规范的术语中，SVC 音轨（视频和参数集音轨）是视频或可视音轨。因此，它们使用 HandlerBox 中的“vide”的 handler_type、视频媒体头“vmhd”以及如下文定义的 VisualSampleEntry 的派生。

[0123] 如下定义了 SVC 视频基本流的样本项和样本格式。定义包括：

[0124] 框类型：“avc1”、“avcC”、“svc1”、“svcC”

[0125] 容器：样本表框（“stb1”）

[0126] 强制性：Avc1（如果基层是 AVC）或 svc1 框是强制的。

[0127] 数量：可能存在一个或多个样本项

[0128] 为了保持与 AVC 的后向兼容性，定义了两种类型的可视样本项。第一，如果 SVC 基本流包含 AVC 兼容的基层，则使用 AVC 可视样本项（“avc1”）。这里，该项最初包含 AVC 配置框，后跟 SVC 配置框，如下文定义的。在 AVC 文件格式规范中，AVC 配置框

归档 AVCDerConfigurationRecord 定义的有 AVC 兼容的基层的分布、级别和参数集信息。SVC 配置框归档下文定义的 SVCDecoderConfigurationRecord 所定义的有 AVC 兼容增强层的分布、级别和参数集信息。如果 SVC 基本流未包含 AVC 基层,则使用 SVC 可视样本项 (“svc1”)。SVC 可视样本项包含 SVC 配置框,如下文定义的。这包括 SVCDecoderConfigurationRecord,也如下文定义的。按 ISO 基本媒体文件格式规范所允许的,使用多个样本描述来指示使用不同配置或参数集的视频部分。

[0129] 下列的是示范 AVC 配置框和 SVC 配置框语法:

```
[0130] // 可视序列
[0131] class AVConfigurationBox extends Box('avcC') {
[0132]     AVDecoderConfigurationRecord() AVConfig;
[0133] }
[0134] class SVCConfigurationBox extends Box('svcC') {
[0135]     SVCDecoderConfigurationRecord() SVCConfig;
[0136] }
[0137] // 如果基层属于 AVC 兼容的,则使用它
[0138] class AVCSampleEntry() extends VisualSampleEntry('avc1') {
[0139]     AVConfigurationBox avcconfig;
[0140]     SVCConfigurationBox svcconfig;
[0141]     MPEG4BitRateBox(); // 可选
[0142]     MPEG4ExtensionDescriptorsBox(); // 可选
[0143] }
[0144] // 如果基层不属于 AVC 兼容的,则使用它
[0145] class SVCSampleEntry() extends VisualSampleEntry('svc1') {
[0146]     SVCConfigurationBox svcconfig;
[0147]     MPEG4BitRateBox(); // 可选
[0148]     MPEG4ExtensionDescriptorsBox(); // 可选
[0149] }
```

[0150] 基本类 VisualSampleEntry 中的变量 Compressorname 指示所使用的压缩器的名称,并且建议值“\012AVC 编码”或“\012SVC 编码”(012 是 10,作为字节的字符串的长度)。如果使用单独的参数集流,则变量 numOfSequenceParameterSets 和 numOfPictureParameterSets 必须均为零。

[0151] SVC 视频基本流中的样本的格式通过 SVC 基本流的解码器专用配置来进行配置。SVC 样本包含有关对应于图 7 所示的主要编码的图片存在的所有可伸缩级别的所有 NAL 单元。

[0152] 下列的是示范 SVC 样本语法:

```
[0153] aligned(8) class SVCSample
[0154] {
[0155]     unsigned int PictureLength = sample_size; // 来自 SampleSizeBox 的
SVCSample 太小
```

```

[0156]   for(i = 0 ;i < PictureLength ;)// 直到图片结束为止
[0157]           unsigned int((SVCDecoderConfigurationRecord.
LengthSizeMinusOne+1)*8)
[0158]           NALUnitLength ;
[0159]           bit(NALUnitLength*8)NALUnit ;
[0160]           i+= (SVCDecoderConfigurationRecord. LengthSizeMinusOne+1)+NALUnit
Length ;
[0161]           }
[0162]   }

```

[0163] 变量 NALUnitLength 指示以字节为单位测量的 NAL 单元的大小。该长度字段包括一个字节的 NAL 头和 EBSP 有效负载的大小,但是不包括该长度字段本身。变量 NALUnit 包含单个 NAL 单元。NAL 单元的语法按 ISO/TEC AVC/SVC 视频规范中定义的,并且包括一个字节的 NAL 头和可变长度的封装字节流有效负载。

[0164] AVC 文件格式中指定的 AVC 参数集基本流的定义也适用于此情况中用于将 SVC 参数集作为单独的基本流存储。

[0165] VisualSampleEntry 中的宽度和高度归档根据通过将整个可伸缩比特流解码获得的像素的正确剪切的正确最大空间分辨率。为了获得每个层的个别宽度和高度,估算组描述项。除非这里另行指定,否则 AVC 文件格式规范中指定的所有其他定义均适用。

[0166] 一般来说,新的 SVC 文件格式扩充支持完全可伸缩且分层的可伸缩流的访问和调整。该分组方法使得能够创建多个组描述项,其中每个描述项描述完全可伸缩流情况中的可伸缩层的特性及其可能的精细化。每个描述项归档时间帧速率(时间可伸缩性)、空间尺寸(空间可伸缩性)、可从此层获取的比特率的范围、指示此层是否是精细颗粒可伸缩的、分布和级别指示、以及相关性的信息。相关性分层结构容易地由组描述项的索引来维护,其中每个较高的索引指示它与它以下的项描述的所有或一些较低层相关。SampleToGroup 框将 SVC 样本的每个 NAL 单元映射到它的 group_description_index。这样能够实现读取、分析和跳过任何不需要的数据的高效方法。如果期望整个可伸缩样本,则读取完整的 SVC 样本。如果仅期望特定的可伸缩层,则在分析时跳过未映射到期望的层的那些 NAL 单元(VCL 或其他)。

[0167] 该修改的文件格式定义一种访问和提取文件格式中存储的整个可伸缩层或其一部分的机制。为了通过网络来传输以及通过网络可能地调整编码的数据,使用 SEI 消息中的可伸缩性信息或使用通过网络丢弃非期望的 NAL 单元的备选规则。一种可能性是作为 RTP 映射进程的一部分来定义规则以启用此类备选功能。

[0168] 修改的文件格式与现有 AVC 文件格式规范达到最充分程度地后向兼容。当存储 AVC 兼容的流时 DecoderConfiguration、样本语法结构和基本流结构中没有改变。文件类型指示使用其 compatible_brands 列表中的牌子“avc1”来告知 AVC 兼容的基层流的存在。通过读取每个 group_description_entry 中存在的分布/级别指示符来检测 AVC 兼容的流的存在。或者,也为 AVC 兼容的基层流创建单独的“提示”音轨。

[0169] 在一个实施例中,使用实时传输协议(RTP)通过网络传送所提取的访问单元。RTP 具有添加到有效负载(在此情况中为提取的访问单元)中的它自己的头。提示音轨包括预

先生成的 RTP 头和指向可伸缩数据的指针。当文件服务器传送提取的访问单元时,访问适合的提示音轨以便检索预先生成的 RTP 头和指针,从而消除生成 RTP 头的附加开销。

[0170] AVC 提示音轨中的每个样本存储指代可伸缩视频流中的 AVC 基层兼容的 NAL 单元的信息。此样本内的所有 NAL 单元具有相同的时间戳。

[0171] 下列的是 AVC 提示音轨的样本语法:

```
[0172] aligned(8)class AVCHintsample{
[0173]     unsigned int(32)    sample_index;
[0174]     unsigned int(8)     nalunitcount;
[0175]     unsigned int(8)     reserved;
[0176] }
```

[0177] 变量 `sample_index` 指示包含 AVC 基层 NAL 单元的 SVC 样本的样本数量。变量 `nalunitcount` 指示自属于 AVC 兼容的 SVC 样本的起始处起的连续 NAL 单元的数量。

[0178] 图 7 图示实现修改的 SVC 文件格式的示范方法。在步骤 200,配置成实现修改的 SVC 文件格式的文件服务器从终端用户装置接收请求。该请求标识要传送的特定数据流的名称。该特定的数据流对应于修改的文件格式中存储的特定可伸缩数据流。该请求还包含终端用户装置的装置要求,例如支持的分辨率和帧速率。在步骤 210,文件服务器确定与指定的数据关联的音轨。在步骤 220,文件服务器将步骤 210 中确定的音轨解码。在步骤 230,使用解码的音轨内的一个或多个元数据框来确定与装置要求关联的描述项。在一个实施例中,解码的音轨包括用于确定关联的描述项的 `SampleGroupDescription` 框。在此实施例中,该描述项定义与装置要求对应的参数值。在步骤 240,使用一个或多个元数据框确定访问单元;以及特定可伸缩数据流内的每个访问单元内的特定 NAL 单元。根据步骤 230 中确定的描述项来确定访问单元内的特定 NAL 单元。在一个实施例中,解码的音轨包括用于确定特定访问单元的 `SampleToGroup` 框。在步骤 250,从特定可伸缩数据流提取步骤 240 中确定的特定访问单元。所提取的访问单元是特定可伸缩数据流的子层数据流。该子层数据流与步骤 200 中接收的装置要求匹配,并因此被终端用户装置所支持。在步骤 260,将子层数据流传送到终端用户装置。

[0179] 已结合细节就特定实施例描述了本发明,以帮助理解本发明的结构和操作的原理。本文中对特定实施例及其细节的此类引用无意将所附权利要求的范围局限于此。对于本领域人员来说,显然在不背离本发明的精神和范围的前提下可以在用于说明而选择的这些实施例中进行修改。

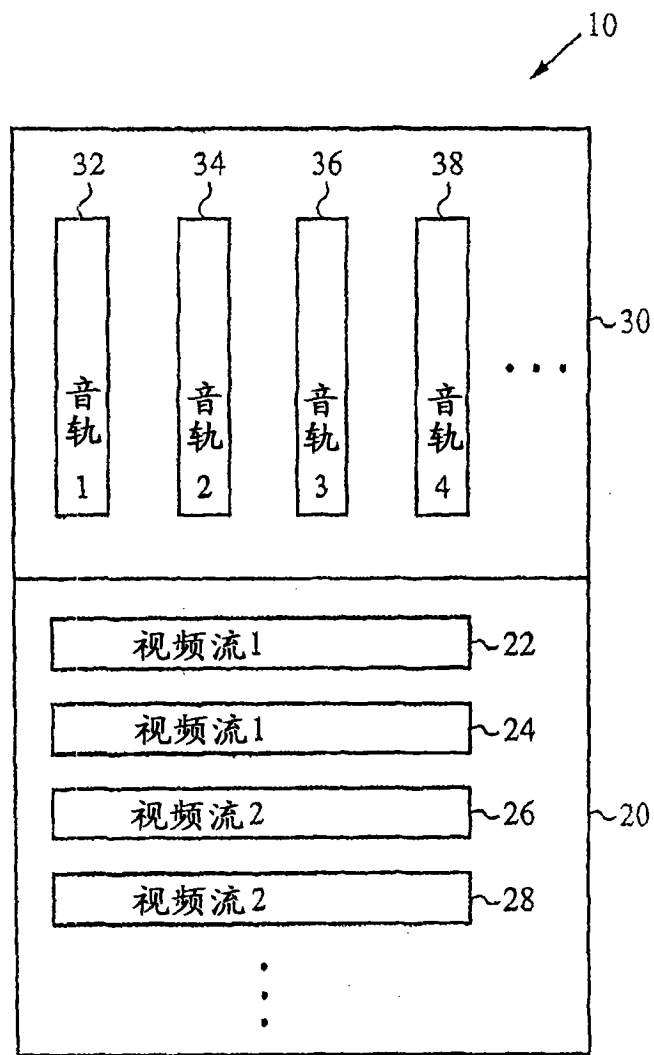


图 1

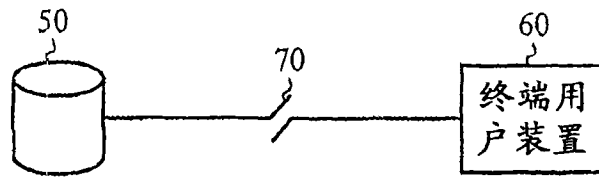


图 2



图 4

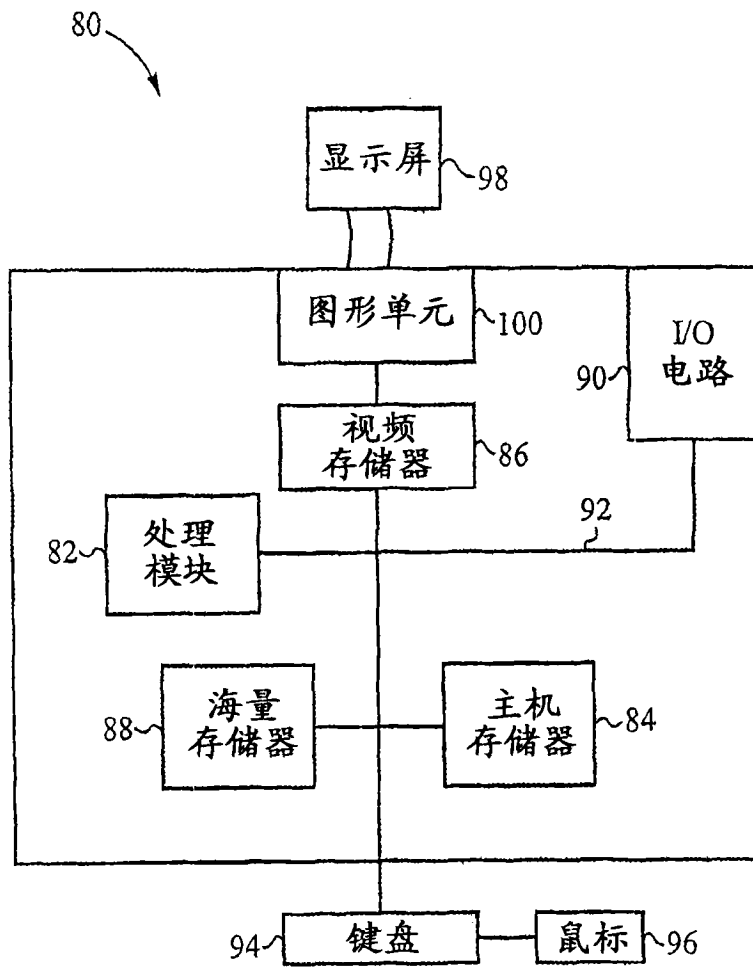


图 3

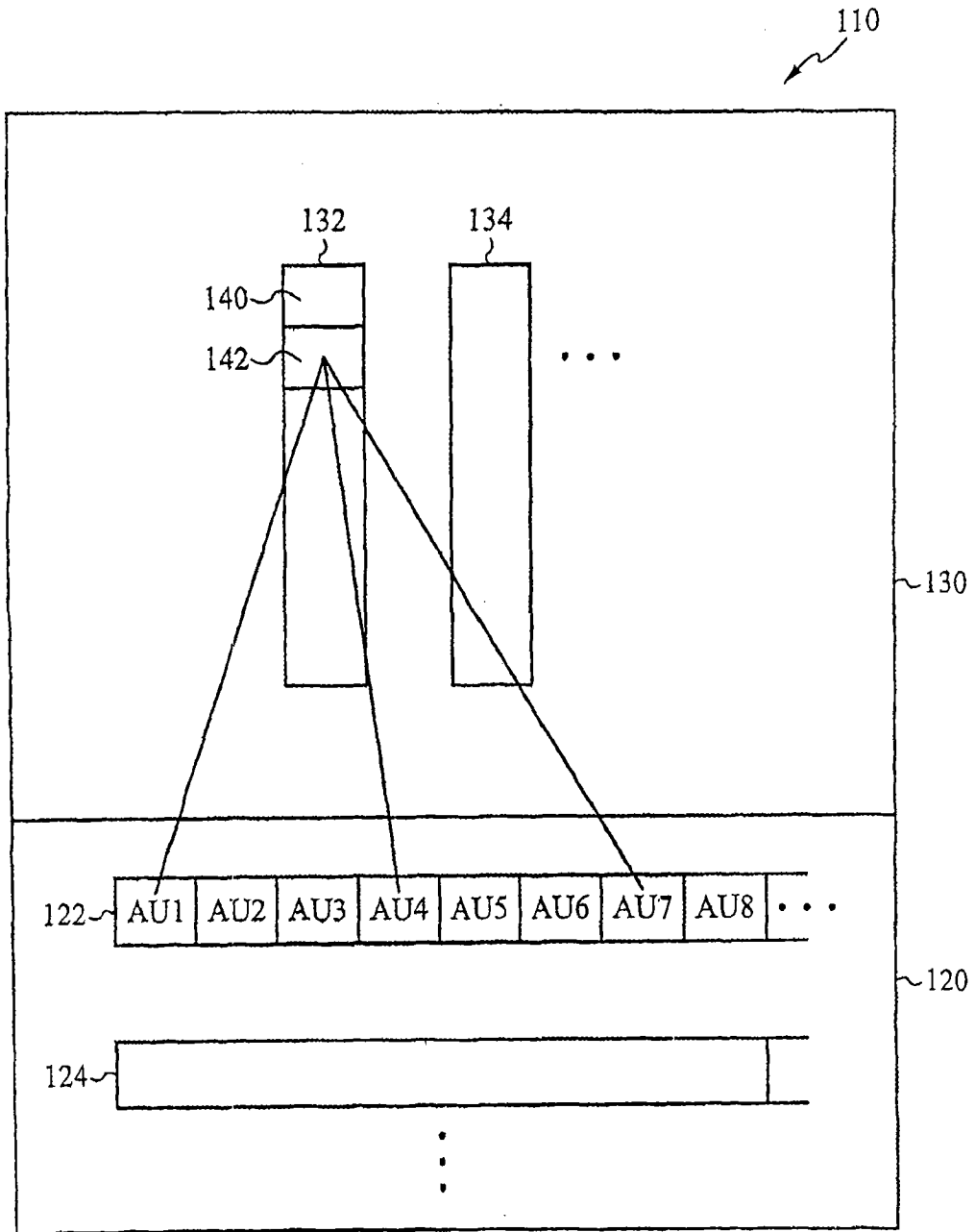


图 5



图 6

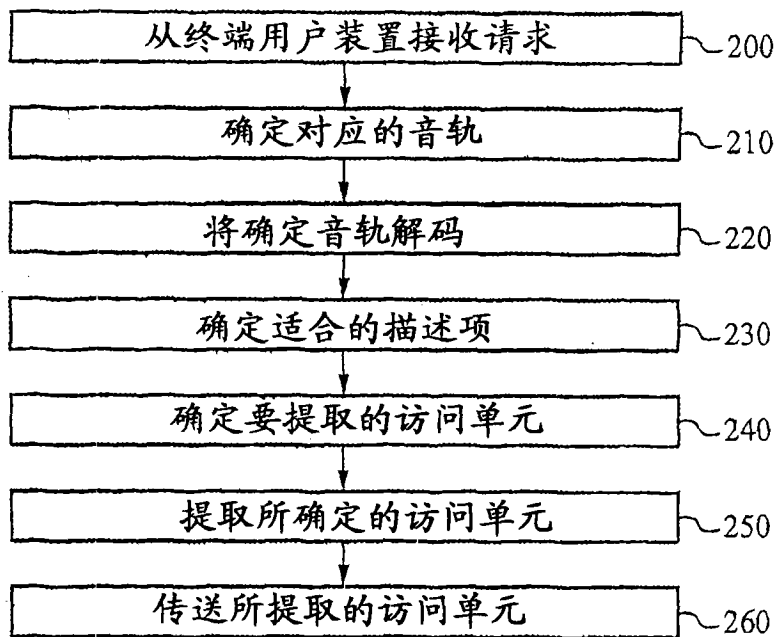


图 7