



(12) 发明专利申请

(10) 申请公布号 CN 102222077 A

(43) 申请公布日 2011. 10. 19

(21) 申请号 201010165733. 2

(22) 申请日 2010. 04. 14

(71) 申请人 中兴通讯股份有限公司
地址 518057 广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦法务部

(72) 发明人 孙健 李加周 阮亚平 林亚

(74) 专利代理机构 北京安信方达知识产权代理有限公司 11262

代理人 王艺 龙洪

(51) Int. Cl.
G06F 17/30(2006. 01)

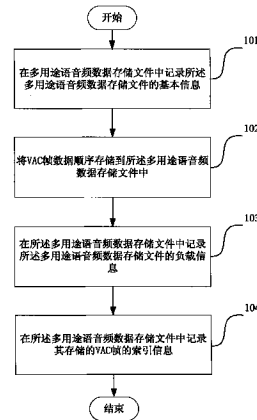
权利要求书 2 页 说明书 6 页 附图 3 页

(54) 发明名称

一种语音频数据的存储方法和装置

(57) 摘要

本发明公开一种语音频数据的存储方法和装置,所述方法包括:在多用途语音频数据存储文件中记录所述多用途语音频数据存储文件的基本信息;将多用途语音频编解码器(VAC)帧数据顺序存储到所述多用途语音频数据存储文件中;在所述多用途语音频数据存储文件中记录所述多用途语音频数据存储文件的负载信息;在所述多用途语音频数据存储文件中记录其存储的VAC帧的索引信息。本发明能够实现多用途语音频数据方便的回放功能,并且可以高效获取文件信息,实现定位播放和分级播放。



1. 一种语音频数据的存储方法,包括:
在多功能语音频数据存储文件中记录所述多功能语音频数据存储文件的基本信息;
将多功能语音频编解码器(VAC)帧数据顺序存储到所述多功能语音频数据存储文件中;
在所述多功能语音频数据存储文件中记录所述多功能语音频数据存储文件的负载信息;
在所述多功能语音频数据存储文件中记录其存储的VAC帧的索引信息。
2. 如权利要求1所述的方法,其特征在于,
所述基本信息包括以下一种或多种信息:文件类型、类型版本、特征信息、厂家标识、创建时间;
所述特征信息包括以下信息中的一种或多种:
 - (1) 配置信息;
 - (2) 分级信息;
 - (3) 选项信息;
 - (4) 属性信息。
3. 如权利要求1或2所述的方法,其特征在于,
所述负载信息包括以下一种或多种信息:文件总时长、存储的VAC总帧数、编码模块标识、VAC编码码流的平均比特率、存储的VAC帧中出现过的特征信息。
4. 如权利要求3所述的方法,其特征在于,
在将VAC帧数据顺序存储的步骤中,记录存储的总VAC帧数、编码模块标识、VAC帧数据的总大小和出现过的特征信息;
在记录负载信息的步骤中,计算出文件总时长和VAC编码码流的平均比特率,将所述文件总时长、存储的VAC总帧数、编码模块标识、VAC编码码流的平均比特率和存储的VAC帧中出现过的特征信息记录到所述多功能语音频数据存储文件中。
5. 如权利要求3所述的方法,其特征在于,
所述方法还包括:将所述多功能语音频数据存储文件中存储的所有VAC帧在逻辑上划分为一个或多个块;所述块由固定个数的VAC帧组成;
所述负载信息还包括:一个块中包含的VAC帧的个数;
所述索引信息包括:所述多功能语音频数据存储文件中所有块的块索引。
6. 如权利要求5所述的方法,其特征在于,所述块索引包括:当前块的块偏移信息。
7. 如权利要求6所述的方法,其特征在于,
所述块索引还包括:当前块中所有VAC帧的特征信息。
8. 如权利要求7所述的方法,其特征在于,
在将VAC帧数据顺序存储的步骤中,记录每个VAC帧的特征信息和每个块的偏移信息;
在记录负载信息的步骤中,将一个块中包含的VAC帧的个数记录到所述多功能语音频数据存储文件中;
在记录索引信息的步骤中,在每个块的块索引中顺序记录当前块的块偏移信息和当前块包含的所有的VAC帧的特征信息,并将存储的所有块的块索引记录到所述多功能语音频

数据存储文件中。

9. 一种语音频数据的存储装置,其特征在于,包括:基本信息记录模块、VAC 帧数据存储模块、负载信息记录模块和索引信息记录模块;

所述基本信息记录模块用于在多功能语音频数据存储文件中记录所述多功能语音频数据存储文件的基本信息;

所述 VAC 帧数据存储模块用于将 VAC 帧数据顺序存储到所述多功能语音频数据存储文件中;

所述负载信息记录模块用于在所述多功能语音频数据存储文件中记录所述多功能语音频数据存储文件的负载信息;

所述索引信息记录模块用于在所述多功能语音频数据存储文件中记录其存储的 VAC 帧的索引信息。

10. 如权利要求 9 所述的装置,其特征在于,还包括块划分模块,

所述块划分模块用于将所述多功能语音频数据存储文件中存储的所有 VAC 帧在逻辑上划分为一个或多个块;所述块由固定个数的 VAC 帧组成;

所述索引信息记录模块还用于在每个块的块索引中顺序记录当前块的块偏移信息和当前块包含的所有 VAC 帧的特征信息,并将存储的所有块的块索引作为索引信息,记录到所述多功能语音频数据存储文件中。

一种语音频数据的存储方法和装置

技术领域

[0001] 本发明涉及多媒体通信和传输领域,尤其涉及一种语音频数据的存储方法和装置。

背景技术

[0002] 在多媒体通信领域,语音频通信一直是居于主导地位的通信方式。由于通信双方终端设备的差距和通信链路实际状况的千差万别,已有的语音频编解码方式无法为所有的用户提供优异的语音频服务。

[0003] 目前提出一种多用途语音频编解码器 (Versatile Audio Codec,简称 VAC) 的方案。该方案应用控制方法和信号分类方法相结合的方式,可以实现根据用户设置、应用场景、客户端反馈以及音频数据等信息,灵活地按不同技术方案对语音频帧编码、传输、解码。另外,还可以包含多个编解码器,可以支持从窄带到全频带、可分级的音频编解码功能。

[0004] 然而,在诸如通话录制,音视频录像等应用中,需要将语音频数据存储在物理介质上,以便随后的访问和查看。而且,为了使用户可以方便的获取文件信息,以及有效、快速的进行各种访问操作。

[0005] 因此,迫切需要一种存储方法来合理地组织多用途语音频数据的存储结构和信息。

发明内容

[0006] 本发明要解决的技术问题就是提出一种语音频数据的存储方法和装置,能够合理地组织多用途语音频数据的存储结构和信息。

[0007] 为了解决上述技术问题,本发明提供一种语音频数据的存储方法,包括:

[0008] 在多用途语音频数据存储文件中记录所述多用途语音频数据存储文件的基本信息;

[0009] 将多用途语音频编解码器 (VAC) 帧数据顺序存储到所述多用途语音频数据存储文件中;

[0010] 在所述多用途语音频数据存储文件中记录所述多用途语音频数据存储文件的负载信息;

[0011] 在所述多用途语音频数据存储文件中记录其存储的 VAC 帧的索引信息。

[0012] 进一步地,上述方法还可具有以下特点:

[0013] 所述基本信息包括以下一种或多种信息:文件类型、类型版本、特征信息、厂家标识、创建时间;

[0014] 所述特征信息包括以下信息中的一种或多种:

[0015] (1) 配置信息;

[0016] (2) 分级信息;

[0017] (3) 选项信息;

[0018] (4) 属性信息。

[0019] 进一步地,上述方法还可具有以下特点:

[0020] 所述负载信息包括以下一种或多种信息:文件总时长、存储的 VAC 总帧数、编码模块标识、VAC 编码码流的平均比特率、存储的 VAC 帧中出现过的特征信息。

[0021] 进一步地,上述方法还可具有以下特点:

[0022] 在将 VAC 帧数据顺序存储的步骤中,记录存储的总 VAC 帧数、编码模块标识、VAC 帧数据的总大小和出现过的特征信息;

[0023] 在记录负载信息的步骤中,计算出文件总时长和 VAC 编码码流的平均比特率,将所述文件总时长、存储的 VAC 总帧数、编码模块标识、VAC 编码码流的平均比特率和存储的 VAC 帧中出现过的特征信息记录到所述多用途语音频数据存储文件中。

[0024] 进一步地,上述方法还可具有以下特点:

[0025] 所述方法还包括:将所述多用途语音频数据存储文件中存储的所有 VAC 帧在逻辑上划分为一个或多个块;所述块由固定个数的 VAC 帧组成;

[0026] 所述负载信息还包括:一个块中包含的 VAC 帧的个数;

[0027] 所述索引信息包括:所述多用途语音频数据存储文件中所有块的块索引。

[0028] 进一步地,上述方法还可具有以下特点:

[0029] 所述块索引包括:当前块的块偏移信息。

[0030] 进一步地,上述方法还可具有以下特点:

[0031] 所述块索引还包括:当前块中所有 VAC 帧的特征信息。

[0032] 进一步地,上述方法还可具有以下特点:

[0033] 在将 VAC 帧数据顺序存储的步骤中,记录每个 VAC 帧的特征信息和每个块的偏移信息;

[0034] 在记录负载信息的步骤中,将一个块中包含的 VAC 帧的个数记录到所述多用途语音频数据存储文件中;

[0035] 在记录索引信息的步骤中,在每个块的块索引中顺序记录当前块的块偏移信息和当前块包含的所有的 VAC 帧的特征信息,并将存储的所有块的块索引记录到所述多用途语音频数据存储文件中。

[0036] 为了解决上述技术问题,本发明提供一种语音频数据的存储装置,包括:基本信息记录模块、VAC 帧数据存储模块、负载信息记录模块和索引信息记录模块;

[0037] 所述基本信息记录模块用于在多用途语音频数据存储文件中记录所述多用途语音频数据存储文件的基本信息;

[0038] 所述 VAC 帧数据存储模块用于将 VAC 帧数据顺序存储到所述多用途语音频数据存储文件中;

[0039] 所述负载信息记录模块用于在所述多用途语音频数据存储文件中记录所述多用途语音频数据存储文件的负载信息;

[0040] 所述索引信息记录模块用于在所述多用途语音频数据存储文件中记录其存储的 VAC 帧的索引信息。

[0041] 进一步地,上述装置还可包括块划分模块,

[0042] 所述块划分模块用于将所述多用途语音频数据存储文件中存储的所有 VAC 帧在

逻辑上划分为一个或多个块；所述块由固定个数的 VAC 帧组成；

[0043] 所述索引信息记录模块还用于在每个块的块索引中顺序记录当前块的块偏移信息和当前块包含的所有的 VAC 帧的特征信息，并将存储的所有块的块索引作为索引信息，记录到所述多用途语音频数据存储文件中。

[0044] 本发明通过记录多用途语音频数据存储文件的基本信息，VAC 帧数据，文件的负载信息，VAC 帧的索引信息，实现多用途语音频数据方便的回放功能，并且可以高效获取文件信息，实现定位播放和分级播放。

附图说明

[0045] 图 1 是本发明实施例的存储方法流程图；

[0046] 图 2 是本发明存储文件的基本信息示意图；

[0047] 图 3 是本发明存储文件的 VAC 帧数据示意图；

[0048] 图 4 是本发明存储文件的负载信息示意图；

[0049] 图 5 是本发明存储文件的 VAC 帧索引信息示意图。

具体实施方式

[0050] 目前已有的 AMR(Adaptive Multi-Rate, 自适应多速率), AAC(Advanced Audio Coding, 先进语音频编解码), MP3(MPEG Audio Layer3, MPEG 语音频层 3) 等音频编码数据的存储方法要么没有索引, 无法实现定位回放; 要么索引设计不太合理, 无法对帧数据的特性很好的描述。特别是对于 VAC 这种分层多码率的码流, 现有的存储方法必须遍历读取所有数据, 经解析后才能选择目标数据, 无法实现高效的码流选择和分级播放。

[0051] 本发明可解决现有技术中存储多用途语音帧时无法实现分级播放的问题。

[0052] 下面结合附图及具体实施例对本发明进行详细说明。

[0053] 如图 1 所示, 本发明实施例的语音频数据的存储方法, 包括:

[0054] 步骤 101, 在多用途语音频数据存储文件中记录所述多用途语音频数据存储文件的基本信息。

[0055] 该基本信息提供了当前文件的总体描述信息, 如文件类型, 类型版本, 特征信息(profile), 厂家标识, 文件创建时间等。

[0056] 步骤 102, 将 VAC 帧数据顺序存储到所述多用途语音频数据存储文件中。

[0057] 具体地, 将获取到的 VAC 帧逐字节顺序存储到文件中。

[0058] 其中, 多用途语音频数据即指 VAC 帧数据。

[0059] 步骤 103, 待所有的 VAC 帧数据存储结束后, 在所述多用途语音频数据存储文件中记录所述多用途语音频数据存储文件的负载信息。

[0060] 该负载信息提供了当前文件的补充描述, 可包括如下信息中的一种或多种: 文件总时长, 文件中存储的 VAC 总帧数, VAC 编码码流的平均比特率等统计信息。

[0061] 此外, 该负载信息还可包括: 编码模块标识, 文件中存储的 VAC 编码帧中出现过的特征信息(即 profile 列表)。

[0062] 步骤 104, 在所述多用途语音频数据存储文件中记录其存储的 VAC 帧的索引信息。

[0063] 为了在回放媒体文件时支持定位拖动, 即支持方便快捷的随机访问文件指定时间

点的 VAC 帧的功能,需要记录 VAC 帧的索引。

[0064] 进一步地,为了提高索引效率并减少索引大小,可以将文件中顺序存储的所有 VAC 帧逻辑上划分为一个或多个块 (block),并对这些块进行索引,即每个帧条目录录某个块的位置偏移信息。

[0065] 此外,为了对分级多码率数据进行选择性读取和分级播放,可以在索引信息中记录各 VAC 帧的特征信息,这样在回放时,不必逐个访问 VAC 帧数据,而只需根据索引中的指引即可得到想要的目标数据。

[0066] 进一步地,在索引内部记录各个块实际包含的 VAC 帧的特征信息 (profile),以优化块内部 VAC 帧操作的效率,实现便捷的分级回放。

[0067] 其中,块由固定个数的 VAC 帧组成,其数值在文件负载信息中定义,即在步骤 103 中,将一个块中包含的 VAC 帧的个数记录到所述多用途语音频数据存储文件中。该参数可在文件生成时自行配置。

[0068] 在一个块的索引内部顺序记录该块包含的所有的 VAC 帧的特征信息。当用户欲选择某一特定属性的码流时,在索引处可以方便查找,实现分级回放功能。

[0069] 参见图 2 ~ 5,为一个多用途语音频数据存储文件中存储内容的实施例。

[0070] 如图 2 所示,一个多用途语音频数据存储文件可以包括:文件的基本信息,VAC 帧数据,文件的负载信息,VAC 帧的索引信息。其中,基本信息可包括文件类型字段,类型版本字段,特征信息 (profile) 字段,厂家标识字段,创建时间字段和其他一些字段。其中,文件类型字段标识当前文件由多用途语音频数据存储而成。类型版本字段标识了当前文件类型的版本。特征信息字段标识了本文件中的所有多用途语音频数据的特征信息,可包括以下信息中的一种或多种:配置信息、分级信息、选项信息、属性信息。厂家标识字段用于标识当前文件的生成者。创建时间字段标识了当前文件的创建时间。

[0071] 如图 3 所示,VAC 帧数据部分由一个个 VAC 帧顺序排列而成。

[0072] 如图 4 所示,负载信息可包括文件总时长字段,文件总帧数,一个块中包含的 VAC 帧的个数 (即每块帧个数 M),VAC 编码码流的平均比特率,编码模块标识字段,文件中存储的 VAC 编码帧中出现过的特征信息 (即特征列表)。文件总时长字段标识了整个文件回放时总的持续时间。文件总帧数标识了文件中存储的 VAC 帧总数。一个块中包含的 VAC 帧的个数字段标识了索引中的一个条目指向的块中 VAC 帧的个数。VAC 编码码流的平均比特率标识了文件的平均比特率。编码模块标识字段标识了文件中存储的 VAC 编码帧中出现过的编码模块标识。文件中存储的 VAC 编码帧中出现过的特征列表字段标识了文件中所有 VAC 帧的特征信息,利用该信息,一个访问者可以根据自身的终端特性选择是否全部解析或部分解析文件。

[0073] 如图 5 所示,文件索引信息包含一个个顺序排列的块索引信息。一个块索引信息包括块偏移量字段和该块内部所有 VAC 帧的特征信息字段。块索引字段用于在文件中按时间点定位和随机访问。块偏移量字段标识了定位到的块在文件中的位置偏移。一个块内部 VAC 帧的特征信息字段用于在一个块内部选择特定类型的帧数据。

[0074] 综上,参考图 1 ~ 图 5,存储多用途语音频数据时,首先写入文件的基本信息,接着存储一个个的 VAC 帧数据。在此过程中,记录存储的总 VAC 帧数,记录存储 VAC 帧数据的总大小,记录出现过的特征信息和编码模块标识,记录每个 VAC 帧的特征信息和每个块的偏

移信息。帧数据存储结束后,计算出文件总时长和 VAC 编码码流的平均比特率,记录在负载信息中,并写入文件。最后根据每 VAC 帧的特征信息和块的偏移信息生成索引信息,并写入文件。

[0075] 其中,根据每 VAC 帧的特征信息和块的偏移信息生成索引信息,写入文件具体可以是:在每个块的块索引中顺序记录当前块的块偏移信息和当前块包含的所有的 VAC 帧的特征信息,并将存储的所有块的块索引记录到该文件中。

[0076] 本发明实施例的语音频数据的存储装置,包括:基本信息记录模块、VAC 帧数据存储模块、负载信息记录模块和索引信息记录模块;

[0077] 所述基本信息记录模块用于在多用途语音频数据存储文件中记录所述多用途语音频数据存储文件的基本信息;

[0078] 所述 VAC 帧数据存储模块用于将 VAC 帧数据顺序存储到所述多用途语音频数据存储文件中;

[0079] 所述负载信息记录模块用于在所述多用途语音频数据存储文件中记录所述多用途语音频数据存储文件的负载信息;

[0080] 所述索引信息记录模块用于在所述多用途语音频数据存储文件中记录其存储的 VAC 帧的索引信息。

[0081] 进一步地,该装置还包括块划分模块,用于将所述多用途语音频数据存储文件中存储的所有 VAC 帧在逻辑上划分为一个或多个块;所述块由固定个数的 VAC 帧组成;

[0082] 所述索引信息记录模块还用于在每个块的块索引中顺序记录当前块的块偏移信息和当前块包含的所有的 VAC 帧的特征信息,并将存储的所有块的块索引作为索引信息,记录到所述多用途语音频数据存储文件中。

[0083] 所述负载信息记录模块还用于记录存储的总 VAC 帧数、VAC 帧数据的总大小、出现过的特征信息,出现过的编码模块标识,计算出文件总时长和 VAC 编码码流的平均比特率,将所述文件总时长、存储的 VAC 总帧数、一个块中包含的 VAC 帧的个数、VAC 编码码流的平均比特率、存储的 VAC 帧中出现过的特征信息记录到所述多用途语音频数据存储文件中。

[0084] 下面通过一个应用示例来进一步说明本发明:

[0085] 假定有一组多用途语音频码流,其 VAC 帧每 10 帧的特征信息(比如,该特征信息为分级信息)顺序增加 1,即 1-10 帧的特征信息值为 1,11-20 帧的特征信息值为 2,....90-100 帧的特征信息值为 10。将其存储为文件时,首先写入文件基本信息,文件类型即为“VACFF”,文件类型版本为 1.0,特征信息为 0xFF(即所存储的 VAC 帧特征信息值不完全相同),厂家标识为“ZTEMM”,创建时间为“20100322”。

[0086] 其中,特征信息字段可用于指示本文件中所有多用途语音频数据的特征信息是否相同,比如,如上所述的当所存储的 VAC 帧特征信息不完全相同,可写入 0xFF(或其它标识特征信息不同的值),如果所存储的 VAC 帧特征信息全部相同,则可将该相同的特征信息的值写入该字段。另外,当写文件基本信息时无法获知所存储的 VAC 帧特征信息是否相同时,也可以直接写入某个指定值,比如 0xFF(或其它指定值),此时,0xFF 即表示无法确定所存储的 VAC 帧特征信息是否完全相同。

[0087] 接着顺序将 VAC 帧存入文件中,并记录各种信息。

[0088] VAC 帧数据存储结束后,接着写入负载信息,总时长为 20ms(每个 VAC 帧时长为

20ms)*100 = 2000ms,总帧数为 100 帧,每块帧个数为 50(1s 数据为一个块),平均比特率为

[0089] $\text{Sum}(\text{sizeof}(\text{VAC1})+\text{sizeof}(\text{VAC2})+\dots+\text{sizeof}(\text{VAC100}))/2\text{bps}$

[0090] sizeof(VACN) 表示第 N 个 VAC 帧的 bit 数,Sum() 表示将括号内的数相加。

[0091] 接着写入编码模块标识,如果所有的帧都是同一个编码模块生成的,该字段记为 1。

[0092] 接着写入这组码流中出现过的特征信息值,1 2 3 4 5 6 7 8 9 10。

[0093] 最后写入索引信息,如第一个块在文件中的偏移量为 30,该块索引记录数据为 30,1,1,1,1,1,1,1,1,1,1(第一个 10 帧的特征信息值),2,2,2,2,2,2,2,2,2,2(第二个 10 帧特征信息值),...5,5,5,5,5,5,5,5,5,5(第五个 10 帧的特征信息值)。

[0094] 利用该块索引信息可以方便的实现文件中时间点的定位:如需定位到第 0.5s 的数据,因为第一个块的起始时间为 0s,第二个块的起始时间为 1s,则首先定位欲读取的数据位于第一个块,然后定位到第 25 个 VAC 帧。通过当前块的偏移位置和该块前 24 个 VAC 帧的长度之和,计算出第 25 个 VAC 帧的偏移位置。

[0095] 利用该块索引信息还可以方便的实现数据的分级回放:如某个客户因某种原因只能播放分级为 2 的数据,则通过本索引信息,发现只有第一个块的 10 帧信息满足要求,则只需根据索引读取相应的数据即可。此时,不再需要在文件中遍历查找帧数据,大大提供了读取效率。

[0096] 综上所述,本发明提供一种多用途语音频数据的存储方法和装置。按照本发明可以生成文件的基本信息,VAC 帧数据,文件的负载信息和 VAC 帧的索引信息。使用该方法生成的文件结构简单,方便读取和访问,可以满足多用途语音频的多种应用。

[0097] 当然,本发明还可有其它多种实施例,在不背离本发明精神及其实质的情况下,熟悉本领域的技术人员当可根据本发明作出各种相应的改变和变形,但这些相应的改变和变形都应属于本发明所附的权利要求的保护范围。

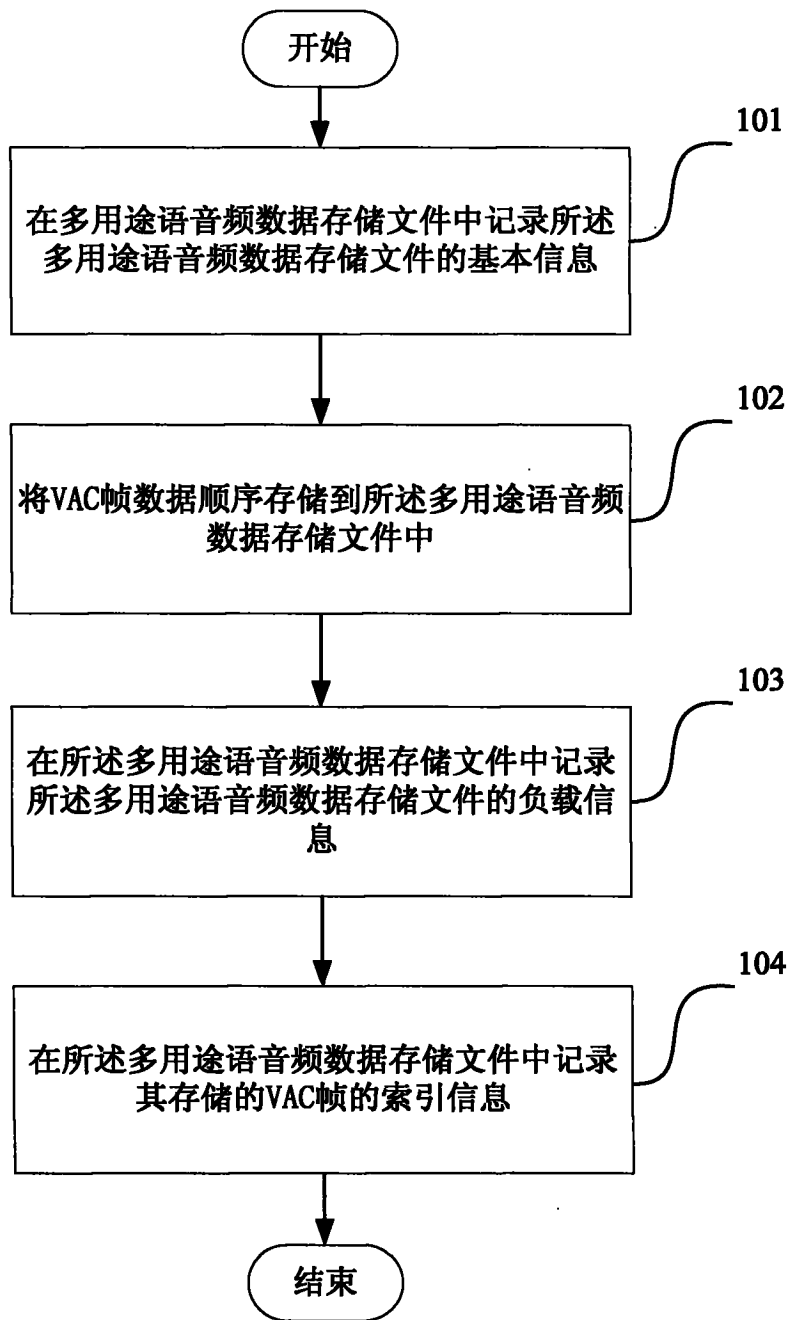


图 1

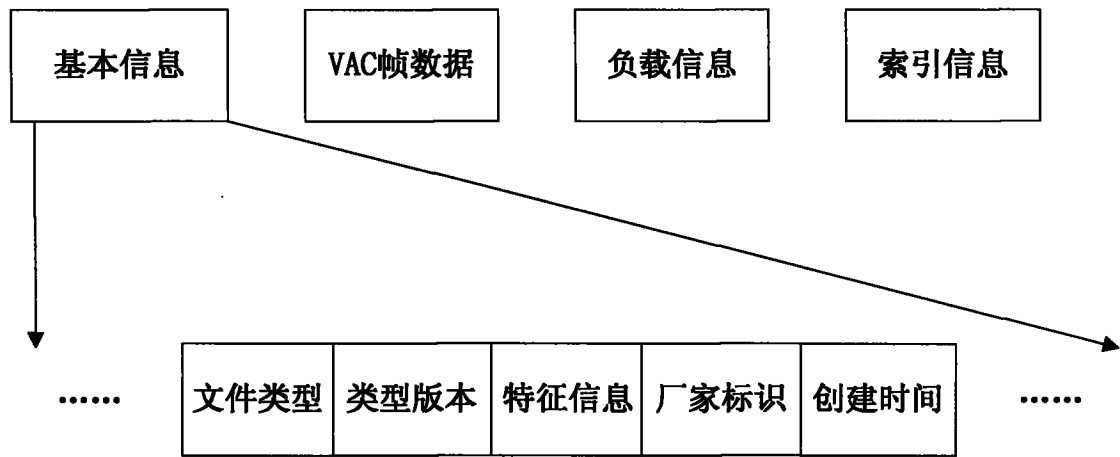


图 2

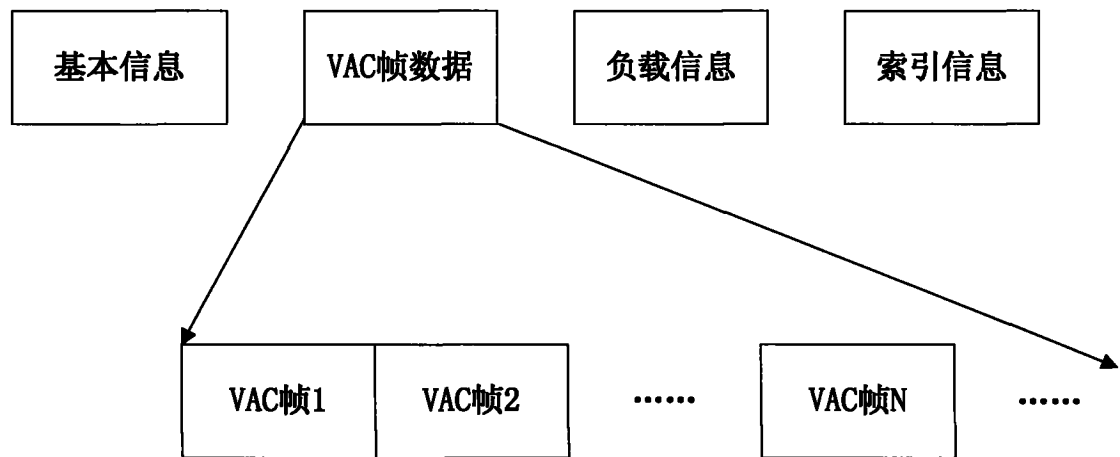


图 3

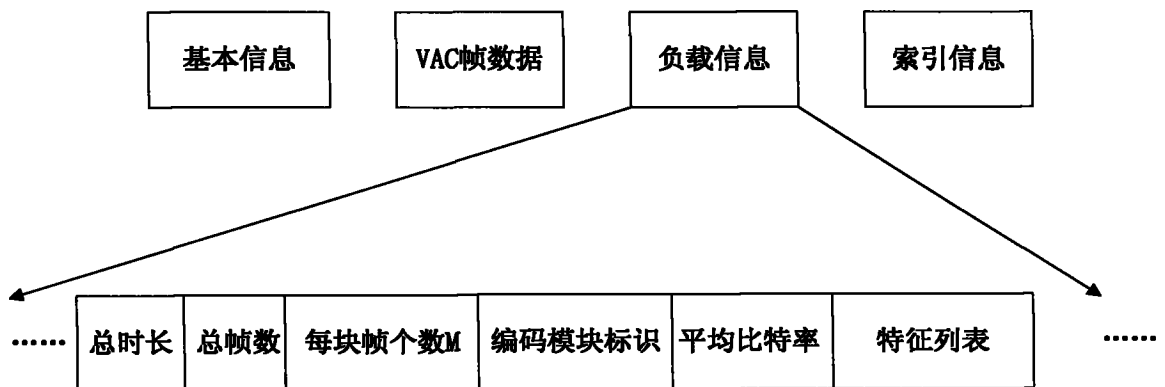


图 4

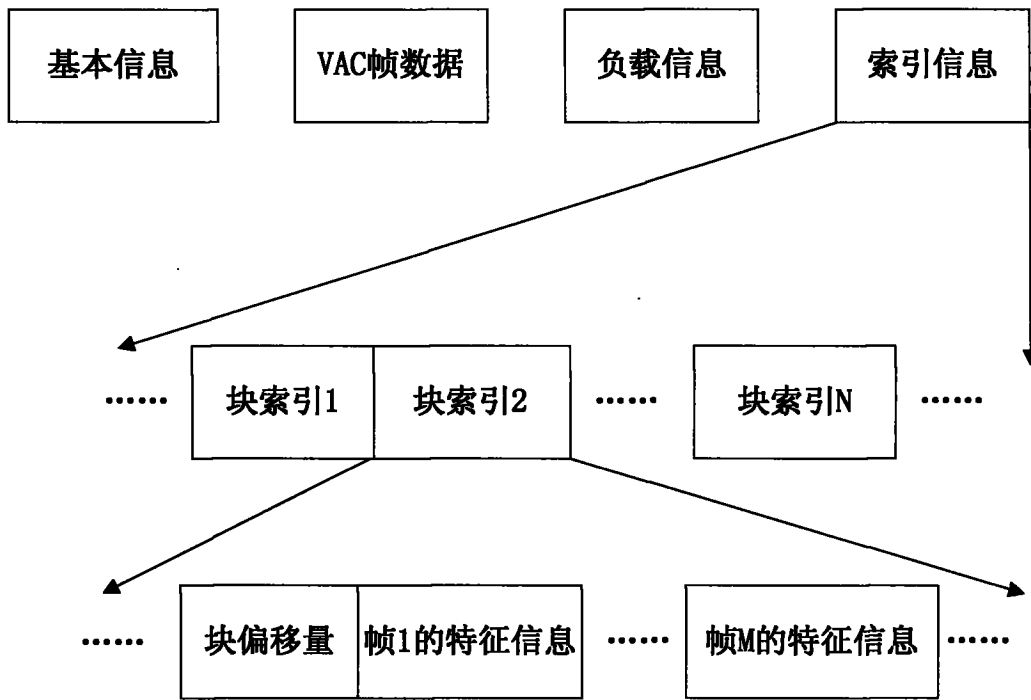


图 5