

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
G06F 17/30 (2006.01)



[12] 发明专利说明书

专利号 ZL 200480012356.1

[45] 授权公告日 2009年7月29日

[11] 授权公告号 CN 100520773C

[22] 申请日 2004.7.30

[21] 申请号 200480012356.1

[30] 优先权

[32] 2004.5.3 [33] US [31] 10/836,261

[86] 国际申请 PCT/US2004/024963 2004.7.30

[87] 国际公布 WO2005/111869 英 2005.11.24

[85] 进入国家阶段日期 2005.11.7

[73] 专利权人 微软公司

地址 美国华盛顿州

[72] 发明人 B·D·曼德斯 O·罗森布鲁姆

V·萨多弗斯基

[56] 参考文献

US 6133908A 2000.10.17

US 2001/0014891A1 2001.8.16

CN 1470984A 2004.1.28

CN 1372669A 2002.10.2

审查员 沈乐平

[74] 专利代理机构 上海专利商标事务所有限公司

代理人 张政权

权利要求书 2 页 说明书 8 页 附图 5 页

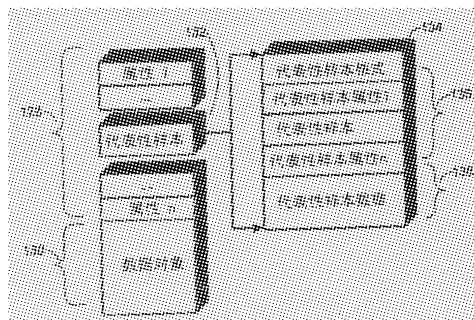
[54] 发明名称

用于封装媒体对象的代表性样本的系统和方法

[57] 摘要

一种系统和相关技术涉及从诸如数字视频便携式摄像机、音频播放器或其它来源或设备等媒体设备中检索具有相关联特性的媒体对象列表。依照各实施例，可从源设备中检索所检索的列表或媒体对象组及其属性或特性，例如歌曲列表以及歌曲长度、编解码器、艺术家和其它信息，以及每一对象的媒体或内容的代表性样本。因此，标识歌曲的标题可以伴随着该歌曲的代表性样本的短声音剪辑，或者标识数字视频影片的标题可以伴随着该影片的介绍部分的短视频剪辑。依照本发明的一方面，该代表性样本可在源设备中生成，并被封装到连同对象标识符一起传递到诸如访问视频摄像机来下载的用户计算机等启动器设备的一组特性中。以单块二进制格式储存而没有预定义内部样本的媒体对象却

可以被访问和采样，以允许用户通过经由底层媒体传输协议传递的特性中的样本的封装来观看或收听代表性内容。



1. 一种用于生成媒体对象的代表性样本的系统，包括：
访问接口，用于访问一组媒体对象，所述一组媒体对象的每一个包括媒体对象数据和一组媒体对象特性；以及
样本引擎，所述样本引擎与所述访问接口通信，以访问所述一组媒体对象并为所述媒体对象的至少一个生成代表性样本，所述代表性样本被封装在所述至少一个对象的一组媒体对象特性中。
2. 如权利要求 1 所述的系统，其特征在于，所述一组媒体对象包括视频媒体对象、音频媒体对象和图像媒体对象的至少一个。
3. 如权利要求 1 所述的系统，其特征在于，所述访问接口包括在应答器设备上的储存所述一组媒体对象的媒体目录。
4. 如权利要求 1 所述的系统，其特征在于，所述样本引擎通过选择性地检测所述媒体对象数据的样本部分来生成所述代表性样本。
5. 如权利要求 4 所述的系统，其特征在于，所述选择性检测包括检测峰值音频振幅和检测视频运动中的改变的至少其中一项。
6. 如权利要求 1 所述的系统，其特征在于，所述代表性样本至少包括代表性样本数据和一组代表性样本特性。
7. 如权利要求 6 所述的系统，其特征在于，所述一组代表性样本特性是由一组代表性样本特性参数来约束的。
8. 如权利要求 1 所述的系统，其特征在于，所述一组媒体对象特性被发送到启动器设备。
9. 如权利要求 8 所述的系统，其特征在于，所述启动器设备被配置成显示或可听见地输出所述代表性样本，由此以指示所述媒体对象的内容。
10. 如权利要求 1 所述的系统，其特征在于，所述一组媒体对象包括二进制对象。
11. 一种用于生成媒体对象的代表性样本的方法，包括：访问一组媒体对象，所述一组媒体对象的每一个包括媒体对象数据和一组媒体对象特性；以及为所述媒体对象的至少一个生成代表性样本，所述代表性样本被封装在该媒体对象的所述一组媒体对象特性中。

12. 如权利要求 11 所述的方法, 其特征在于, 所述一组媒体对象包括视频媒体对象、音频媒体对象和图像媒体对象的至少一个。

13. 如权利要求 11 所述的方法, 其特征在于, 所述访问包括访问应答器设备上储存所述一组媒体对象的媒体目录。

14. 如权利要求 11 所述的方法, 其特征在于, 所述生成包括选择性地检测所述媒体对象数据的样本部分。

15. 如权利要求 14 所述的方法, 其特征在于, 所述选择性地检测包括检测峰值音频振幅和检测视频运动中的改变的至少一项。

16. 如权利要求 11 所述的方法, 其特征在于, 所述代表性样本至少包括代表性样本数据和一组代表性样本特性。

17. 如权利要求 16 所述的方法, 其特征在于, 所述一组代表性样本特性是由一组代表性样本特性参数来约束的。

18. 如权利要求 11 所述的方法, 其特征在于, 还包括将所述一组媒体对象特性发送到启动器设备。

19. 如权利要求 18 所述的方法, 其特征在于, 还包括显示或可听见地输出所述代表性样本, 由此以指示所述媒体对象的内容。

20. 如权利要求 11 所述的方法, 其特征在于, 所述一组媒体对象包括二进制对象。

用于封装媒体对象的代表性样本的系统和方法

发明领域

本发明涉及计算领域，尤其涉及一种平台，它从诸如音频播放器或数码便携式摄像机等源设备中检索媒体对象列表以及该媒体对象内容的代表性样本，以允许用户即使在原始内容不包含任何预定义内部样本时也能够观看或访问歌曲、视频或其它样本。

发明背景

数字媒体设备的容量和类型的稳定增长驱动了对更复杂且更灵活的数字媒体互换协议和服务的需求。例如，用户通常希望共享或交换诸如数字照片、数字视频和其它媒体或内容等数字内容，以及操纵和管理它们所拥有的内容。在典型的情形中，用户可能希望观看其便携式媒体或其它设备上的视频、照片、歌曲或其它媒体的集合。

如图 1 所示，例如，当用户希望检索或上传储存在其数码便携式摄像机上的一组数字视频时，他们可访问该设备上的目录以标识所储存并可访问的视频。该数字视频内容一般是以基于对象的格式来储存的，其中每一完整的视频摄制包含在媒体对象中，该媒体对象具有一组相关联的属性或特性，诸如视频的长度、该视频中使用的编解码器、日期和时间信息、作者或所有者信息、或其它数据或特性。视频帧本身可以用编码像素值和其它信息的二进制格式来储存。一般而言，二进制媒体对象作为单个、单块数据结构来储存，而没有构成部分。相反，对象的特性可标识帧长度、帧大小、色彩深度和读取或消费该内容所需的其它特性。一般而言，基于对象的存储可使得媒体对象的检索或两个或多个媒体设备之间的数据同步变得更有效率。

当用户希望例如在连接的个人计算机或其它客户机上观看储存在其数码便携式摄像机、摄像机或其它设备上的媒体对象列表时，可向他们呈现视频对象的列表以及日期和其它特性。然而，按照名称或其它字段排列的内容的表格或清单可能不用每一视频或其它对象包含哪一内容的概念来提示用户。例如，当数字设备包含大

量的媒体对象时，标识内容可能会变得甚至更困难。然而，用户经常希望诸如在打印、储存或操纵该内容之前被给予对给定对象的内容的指示。想到那些目标中的某一些之后，某些文件或媒体格式准许例如生成数字照片的缩略图表示。

然而，音频、视频或其它媒体内容通常不能容易地用样本形式来表示，并且该内容通常没有为样本选择的目的而结构化。一个问题是代表性样本必须在它被访问之前按照类型来标识。必须对样本给予媒体类型，以准许例如在通过用户界面向用户显示或播放该样本对象之前应用适当的解码器。如果没有对二进制样本的类型的指示，为样本的目的而对该内容进行译码是不可能的。

此外，即使有了标识的类型，例如指示视频样本，对该样本准许的特性范围同样必须被定义为准许回放。例如，视频帧的像素大小或帧回放速率可能需要被约束在便携式设备或其它回放机器的回放能力之内。例如，即使内容的前五秒例如被选为默认样本，该样本的格式、大小或其它特性也可能与便携式设备或其它回放客户机不兼容，使得可靠的回放可能不被保证或是不可能的。

在许多情况下，视频剪辑的前几帧或数字音频的前几秒也可能不是非常能够作为整体来代表媒体内容的，因为视频和音频可能实际上在内容的长度上变化。存在数字媒体存储和管理中的其它问题。

发明概述

克服本领域中的这些和其它问题的本发明的一方面涉及一种用于封装媒体对象的代表性样本的系统和方法，其中可访问源视频、音频、图像或其它内容来生成一组描述该对象的特性。在各实施例中，构成该特性集合的该组元数据可被扩展以包括媒体对象的代表性样本，代表性对象可储存在该媒体对象的该组特性中，但是也具有其自己的另一组特性。该代表性样本特性可例如定义样本对象的类型以及诸如帧宽度、帧长度、歌曲样本点或准许生成可对每一媒体对象灵活地标识的样本表示的其它参数等特性。样本然后可例如在该对象之前发送，以查看媒体列表或其它视图中的内容。由于代表性样本可从底层的对象中提取并被集成到该对象的特性组中，因此任何媒体对象可具有为其生成的相关联样本，即使该对象的二进制内容不包含或不能储存这一表示。在各实施例中，可使用各种算法来确定合适的样本选择。

附图简述

图1示出了依照已知技术的媒体对象内容图。

图 2 依照本发明的实施例示出了其中可操作用于封装媒体对象的代表性样本的系统的系统的环境。

图 3 依照本发明的实施例示出了媒体对象表示的关联。

图 4 依照本发明的实施例示出了将媒体对象的代表性样本封装在与该对象相关联的特性组中。

图 5 依照本发明的实施例示出了样本特性值标识符的数据结构。

图 6 依照本发明的实施例示出了总的代表性样本处理的流程图。

实施例的详细描述

图 2 依照本发明的实施例示出了其中可操作用于对象特性的优化检索的系统和方法的体系结构。如图所示，用户可例如通过用户界面 102 操作客户机计算机或其它启动器 104 来启动媒体对象检索活动，包括媒体对象特性的流传送或累积检索。依照本发明的一方面，检索操作可以在储存在媒体播放器或其它应答器 118 上的媒体对象和相关联的数据上执行。

应答器 118 可以是或包括，例如，诸如 MP3 (MPEG-1、音频层 3)、Windows™ 媒体音频播放器、高级音频编码 (AAC) 或其它便携式或台式音频设备等便携式音频播放器或回放设备；诸如储存 JPG (联合图像专家组) 或其它文件或对象的数码相机等数码相机；诸如数字视频 (DV)、miniDV、数字多功能盘 (DVD) 或其它视频捕捉或回放设备等数码便携式摄像机；诸如配备了数码相机的蜂窝设备等启用媒体的蜂窝电话；个人数字助理或其它平台、硬件或设备。

应答器 118 可以主存或包含一组媒体对象 126，它编码其上所储存的音频、视频、图像或其它媒体或内容。该组媒体对象 126 可包括，例如，诸如音乐、语音录音或其它音频内容等一组音频对象；诸如从数码相机、照片共享或其它因特网站点或其它来源下载的数字照片；诸如 DV 兼容视频剪辑等一组视频对象；或其它对象、媒体或内容。应答器 118 可以在媒体目录 124 中储存或用媒体目录 124 来组织该组媒体对象 126，媒体目录 124 可以是或包括文件分层结构或索引格式，用于将该组媒体对象 126 储存在硬盘、闪存、诸如 CD-ROM 或 CD-ROM R/W、DVD 等光驱、或其它磁、光、电或其它介质存储，或从这些存储中访问该组媒体对象 126。应答器 118 可通过诸如通用串行总线 (USB) 2.0 连接、FireWire™ (火线) 连接或其它有线或无线端口、接口或连接等连接器 112 与启动器通信。应答器 118 可以在嵌入式设备逻辑 122 的控制下操作，嵌入式设备逻辑 122 可以是、包括或接口到微处理

器、微控制器、数字信号处理器或其它控制电路、以及例如固件或其它编程。应答器可包括诸如 LCD（液晶显示）屏等接口 120，以显示媒体对象数据和其它信息，诸如歌曲列表、照片幻灯片演示或其它数据或显示。

在依照所示实施例的本发明的操作中，用户可例如操纵用户界面 102，例如通过媒体引擎 106 来选择媒体对象检索功能，媒体引擎 106 可以与诸如 Windows™ 家族操作系统或其它等操作系统 108 通信或集成。当期望对象枚举时，媒体引擎 106 可生成媒体枚举请求 114，该请求可通过操作系统 108，并例如通过媒体传输协议 110 被传递到应答器 118。在各实施例中，媒体传输协议 110 可以是或包括，例如上述美国专利申请第 10/429,116 号中描述的协议或其它协议或标准。

在一方面，如此生成的媒体枚举请求 114 可以不限于检索机制，通过该机制，只能全局地检索一媒体对象标识符列表，从而需要进一步的请求和额外开销来查看耦合到那些对象的特性或属性。相反，依照本发明的实施例，媒体枚举请求 114 可以被配置并被启用以用一种命令结构来请求应答器 118 中的某些或全部对象集合的对象特性。

如图 3 所示，在一方面，作为媒体枚举请求 114 或其它访问活动的对象的该组媒体对象 126 可由媒体对象数据 130 以及相关的一组媒体对象特性 128 组成。在一方面，该组媒体对象特性 128 可以描述例如用于指定图像、视频或音频内容的它们所相关的媒体对象数据 130 的类型、格式、内容和其它属性或特征，以及诸如帧宽度、帧高度、视频样本长度、艺术家数据、音轨号、歌曲标题、像素色深度或媒体对象数据 130 中编码的媒体内容的其它逻辑、物理或其它特性等特性。储存在该组媒体对象特性 128 中的特性的数目可以例如依照当前由那些特性特征化的媒体或内容的类型而变化。

图 4 依照本发明的实施例示出了将媒体对象数据 130 的代表性样本 132 封装在该对象的特性内。在所示的实施例中，该组媒体对象特性 128 可被扩展，以包括代表性样本 132，作为媒体对象数据 130 的另一属性或特性。代表性样本 132 本身可具有与其相关联的另一组特性和资源，包括以嵌套的方式将代表性样本 132 特征化为特性的一组代表性样本特性 136。

代表性样本 132 可表示媒体对象数据 130 的一部分，它总结、反映、特征化或相关到该内容。当被完全提取时，代表性样本 132 可例如包含较长的数字视频影片或录像的相对较短的视频剪辑或尾片，诸如 5 秒或另一长度。代表性样本 132 同样可以是或包括较长的歌曲或其它音频录音的相对较短的样本，例如 5 秒或另一

长度，或可以是或包括一组图像样本，诸如数字照片的四分之一部分，或其它媒体或内容的样本、节或部分。为代表性样本 132 选择的媒体对象数据 130 的该部分可以用各种不同的方法来隔离。例如，在各实施例中，设备逻辑 122 或其它样本引擎或其它控制智能可处理媒体对象数据以检测合适的样本部分或选择的标识特征。

例如，在音频对象的情况下，设备逻辑 122 或其它逻辑可检查歌曲或音频剪辑的长度以检测具有最大峰值振幅的段。那些节可对应于例如演奏大量乐器、演唱副歌或存在更活跃或更有特征的声音质量的歌曲的一部分。在视频对象的情况下，设备逻辑 122 或其它控制器可例如使用帧增量技术检查视频帧数据，以确定动态对比度或外观运动最大的一组连续帧。在图像对象的情况下，可提取图像的一部分，诸如中心或左象限或右象限，或具有高亮度值的部分。其它选择规则、试探法和其它算法是可能的。

依照本发明另一方面的实施例，代表性样本数据 138 可以是或包括与底层媒体对象数据 130 不同类型的数据。例如个人的数字图像可与联系人或电子邮件对象中该个人的联系人信息相关联，或者涉及录音艺术家或乐队的标志或象征可作为歌曲或其它媒体的代表性样本数据嵌入。依照其它实施例，代表性样本数据 138 可以是或包括一种类型的代表性样本数据的多个片断，或者一种以上数据或媒体类型的组合。

当内容被选为代表性样本数据 138 时，该数据可与指定该样本数据的类型、属性、特性或其它元数据的一组代表性样本特性 136 相关联。代表性样本特性 136 可以例如指示该样本包含音频或视频内容，以及视频或音频样本的长度，或图像样本的像素色深度或音频样本。代表性样本数据 138 也可具有与其相关联的代表性样本格式 134，例如用于指示将以 16 或 24 位格式表示色彩值，音频数据将以 Windows™ Audio 或其它格式或方案来编码。

在如图 5 所示的本发明的实施例中，该组代表性样本特性 136 和特征化该代表性样本数据 138 的对应的代表性样本特性值 140（例如，5 秒的样本长度）可经受检查，以确保观察到诸如对像素宽度、样本长度和其它参数的限制等约束。这些约束例如对于确保在给定便携式媒体播放器、台式计算机或其它回放或演示设备上的可靠回放可能是必需的。例如，数码视频摄像机可准许仅某一帧或像素大小的样本，以准许在整体 LCD 或其它观看屏幕上回放。依照所示出的实施例，代表性样本特性值 140 可以与查找表或其它资源相关联，可对照该查找表或资源来检查值的有效性。

查找表或其它资源在另一方面可包含一组代表性样本特性参数 142, 它们定义了代表性样本特性值 140 有效的值或范围。例如, 该组代表性样本特性参数 142 可包括最小或最大允许值, 以及例如在各实施例中准许的中间步长值。该组代表性样本特性参数 142 同样可包含准许值的枚举列表, 诸如用于像素平铺块大小的 120×120 、 160×160 或 200×200 。其它参数和约束是可能的。依照本发明的各实施例, 一个或多个代表性样本特性值 140 可以基于该组代表性样本特性参数 142 来预填充。在各实施例中, 该组代表性样本特性参数 142 可以是自包含的, 或连同每一代表性样本特性一起储存并与其相关联, 或可以是全局可访问的以供在应答器 118 中的查找表或其它资源中查找。其它配置是可能的。

由于依照本发明的实施例生成的代表性样本 132 在格式上是不固定的, 而是相反可被修改成适合不同的媒体或主机设备, 因此检查该组代表性样本特性 136 与设备或其它限制的符合性的能力可有助于确保可检索、访问和观看或可听见地输出代表性样本 132 而没有无法预料的错误。

因此, 依照本发明的一般方面的各实施例, 代表性样本数据 138 的数据组、代表性样本特性组 136 和代表性样本格式 134 可以与封装代表性内容的代表性样本 132 相关联, 作为较大媒体对象数据 130 的特性组 128 中的特性。由于代表性样本 132 被编码为特性对象, 在各实施例中特性对象可在媒体传输协议 110 中理解并集成到媒体传输协议 110 中, 因此不需要应答器 118 或其它资源的任何修改或特殊编程来生成代表性样本 134。即使媒体对象数据 130 本身既不是明确包含也不虑及内部样本, 情况也是如此。

同样, 由于代表性样本 132 储存在单独的对象中, 因此代表性样本 132 可以在底层媒体内容之前被发送到启动器 104 或其它请求设备以供独立的观看或可听见的或其它输出。用户因此可以获取媒体对象组 126 内一个或多个对象内的媒体内容的性质或类型的指示, 而无须检索或打开对象本身。

图 6 依照本发明的各实施例示出了总体代表性样本和相关的处理。在步骤 602, 处理开始。在步骤 604, 启动器 104 可启动对媒体对象列表 116 的媒体枚举请求, 或可启动其它访问活动, 诸如自己生成或更新媒体对象的应答器 118。在步骤 606, 可枚举涉及储存在应答器 118 中的一个或多个媒体对象数据 130 的一组媒体对象特性 128。在步骤 608, 可启动代表性样本 132 的构建, 例如作为形成涉及媒体对象数据 130 的媒体对象特性组 128 的一部分的元数据、属性或特性。在步骤 610, 可向处理的媒体对象数据 130 应用样本选择算法, 诸如检测视频运动速率或

音频峰值振幅的算法。

在步骤 612, 可例如通过提取视频对象的第二个完整的分钟或通过对围绕音频或音乐对象的峰值音量或振幅的 20 秒音频进行采样, 来捕捉或生成代表性样本数据 138。在步骤 614, 一组代表性样本特性 136, 诸如样本长度、编解码器、像素色深度或其它参数或特征可与代表性样本数据 138 相关联。在步骤 616, 可执行一组代表性样本特性参数 142 的查找或其它检索或关联, 以检测该组代表性样本特性 136 的约束或标准, 例如最大帧宽度或最大样本持续时间。在步骤 618, 可用所处理的媒体对象的样本的值来填充该组代表性样本特性 136, 例如用于指示由该组代表性样本特性参数 142 或其它约束的样本长度、创建日期或时间、样本对象字节大小、像素色深度或其它属性或特性。不符合代表性样本特性参数 142 中编码的约束的特性可以例如被减小到最大允许值、设为默认值或用其它方式来处理。

在步骤 620, 可将包含所生成的该组代表性样本特性 136 的完成的代表性样本 132 插入到与如此生成其样本的媒体对象数据 130 相关联的该组媒体对象特性 128 中, 代表性样本本身作为特性或元数据。在步骤 622, 包含媒体对象组 126 中的一个或多个媒体对象, 包括具有如此生成的代表性样本的那些媒体对象的媒体对象特性组 128 的媒体对象列表 116 可被发送到启动器 104,

在步骤 624, 可在启动器 104 的界面上查看媒体对象列表 116, 包括包含这类样本的每一媒体对象的代表性样本 132 的观看、音频回放或其它演示。在步骤 626, 可在适当时编辑或操纵媒体对象列表 116, 例如, 用于对媒体对象列表 116 中的对象标题或其它特性进行排序, 或检索、删除或访问或处理应答器 118 上的媒体对象组 126 中的底层媒体对象。在步骤 628, 处理可返回到前一处理点、跳至后一处理点、重复或结束。

本发明的以上描述是说明性的, 本领域的技术人员可以明白配置和实现上的修改。例如, 尽管本发明一般是按照其中为媒体对象生成一个代表性样本 132 的实施例来描述的, 但是在各实施例中, 可以生成一个以上代表性样本 132, 例如来自一组数字照片的多张照片。又如, 尽管本发明一般是按照从对一组相关联的媒体对象排序的便携式媒体设备中检索的代表性样本 132 的检索来描述的, 但是在各实施例中, 可询问非便携式设备, 诸如其它台式客户机或其它机器来检索样本。

同样, 尽管代表性样本对象一般被说明性地描述为或是包括音频、视频或图形内容, 然而在各实施例中, 媒体对象组同样可包含文本、字符、数字或其它信息, 例如个人数字助理中的联系人列表, 或其它数据或媒体。此外, 尽管在各实施例中,

检索代表性样本 132 的启动器设备一般被描述为台式客户机或工作站,然而在各实施例中,启动器可以是或包括诸如启用媒体的蜂窝电话、数码相机、音频播放器或其它设备或硬件等便携式设备。便携式设备此外还在需要时可彼此询问以检索代表性样本。

出于同样的原因,尽管本发明在各实施例中一般被描述为涉及诸如数字视频便携式摄像机或其它来源或主机设备等本身可生成代表性样本 132 的应答器 118,然而在各实施例中,启动器 114 或诸如远程服务器等其它来源可从检索自源设备的媒体对象中生成样本构造。被描述为单个的其它硬件、软件或其它资源在各实施例中可以是分布式的,并且类似地,在各实施例中,被描述为分布式的资源可以被组合。因此,本发明的范围仅由所附权利要求书来限定。

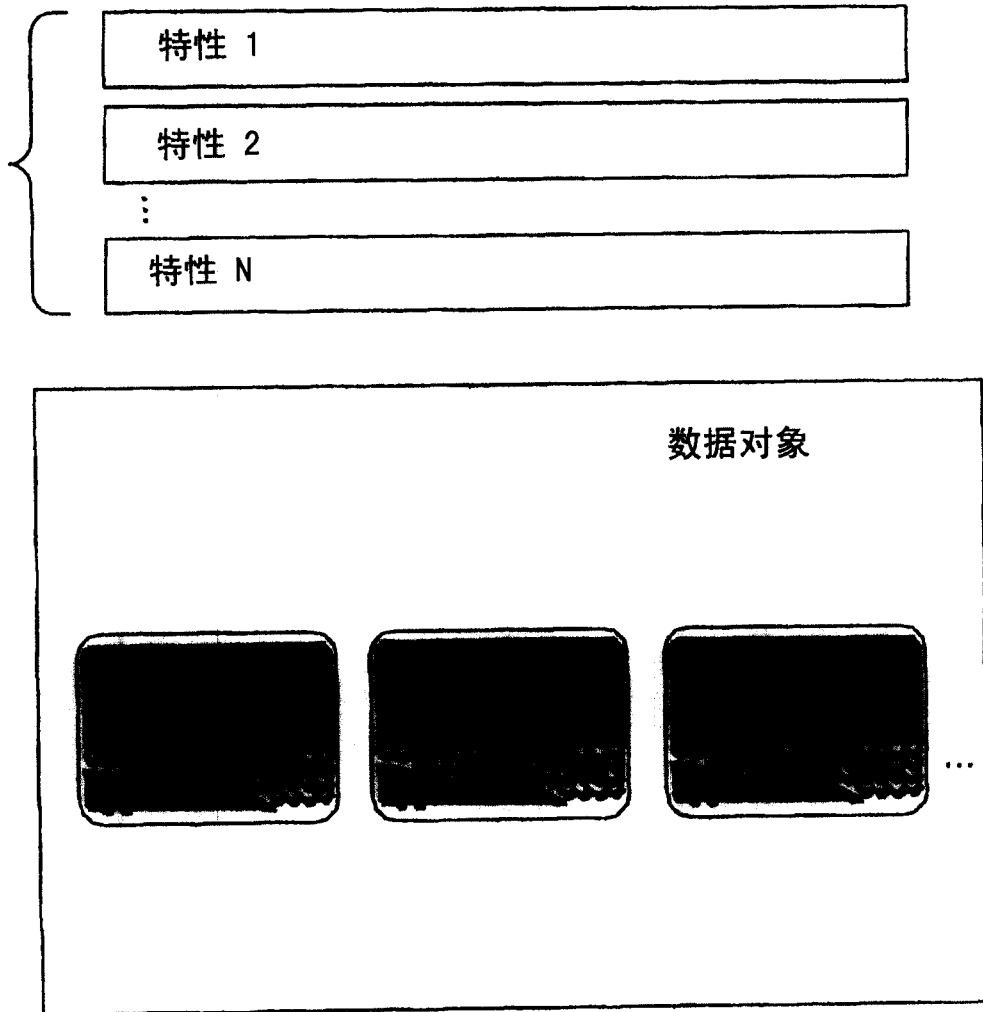


图 1

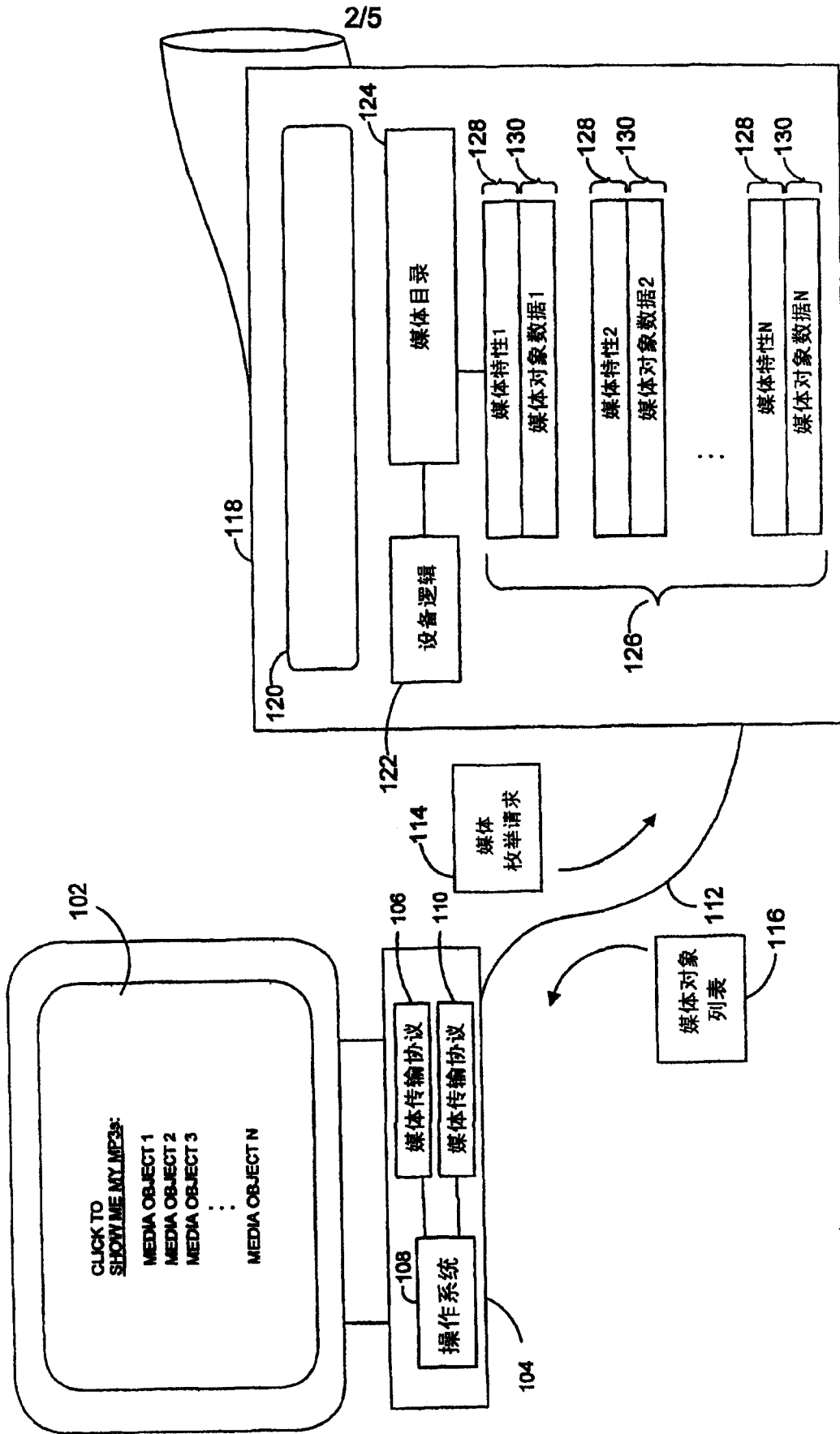


图 2

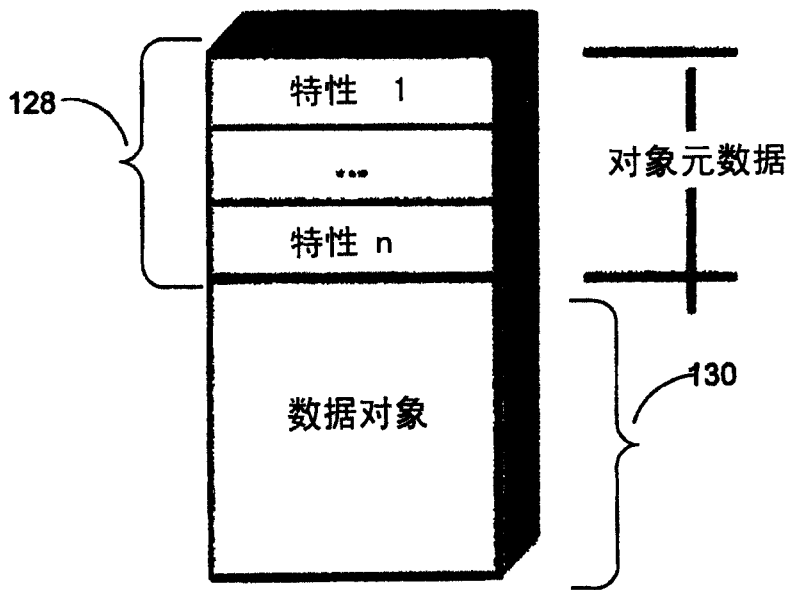


图 3

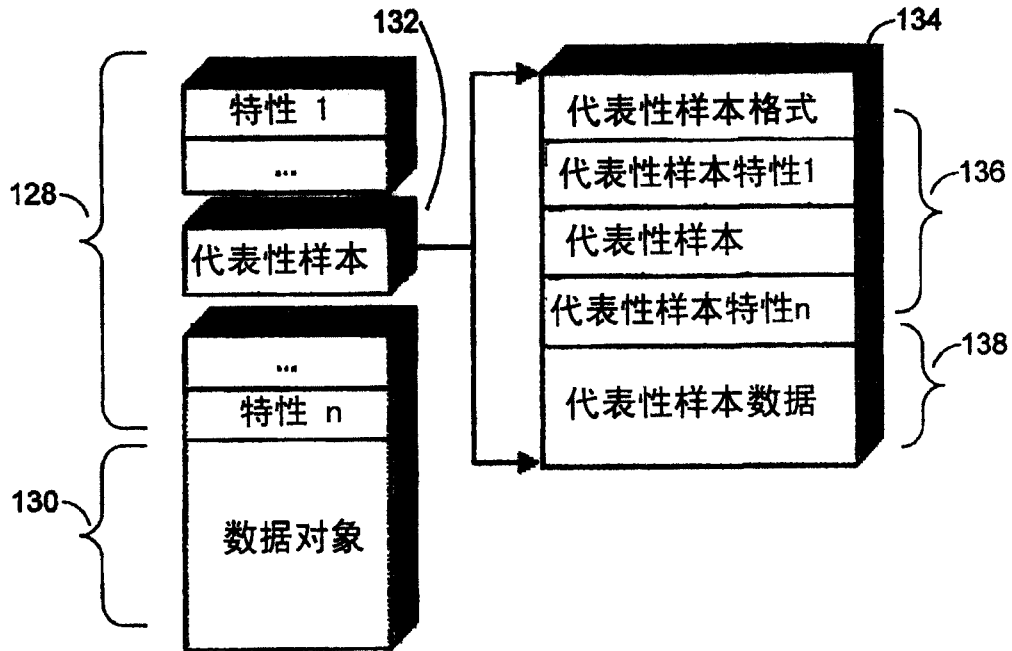


图 4

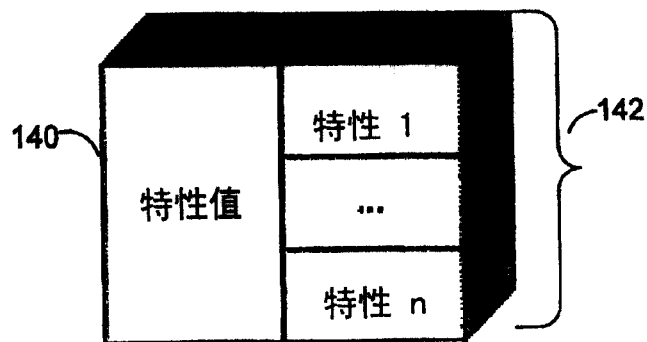


图 5

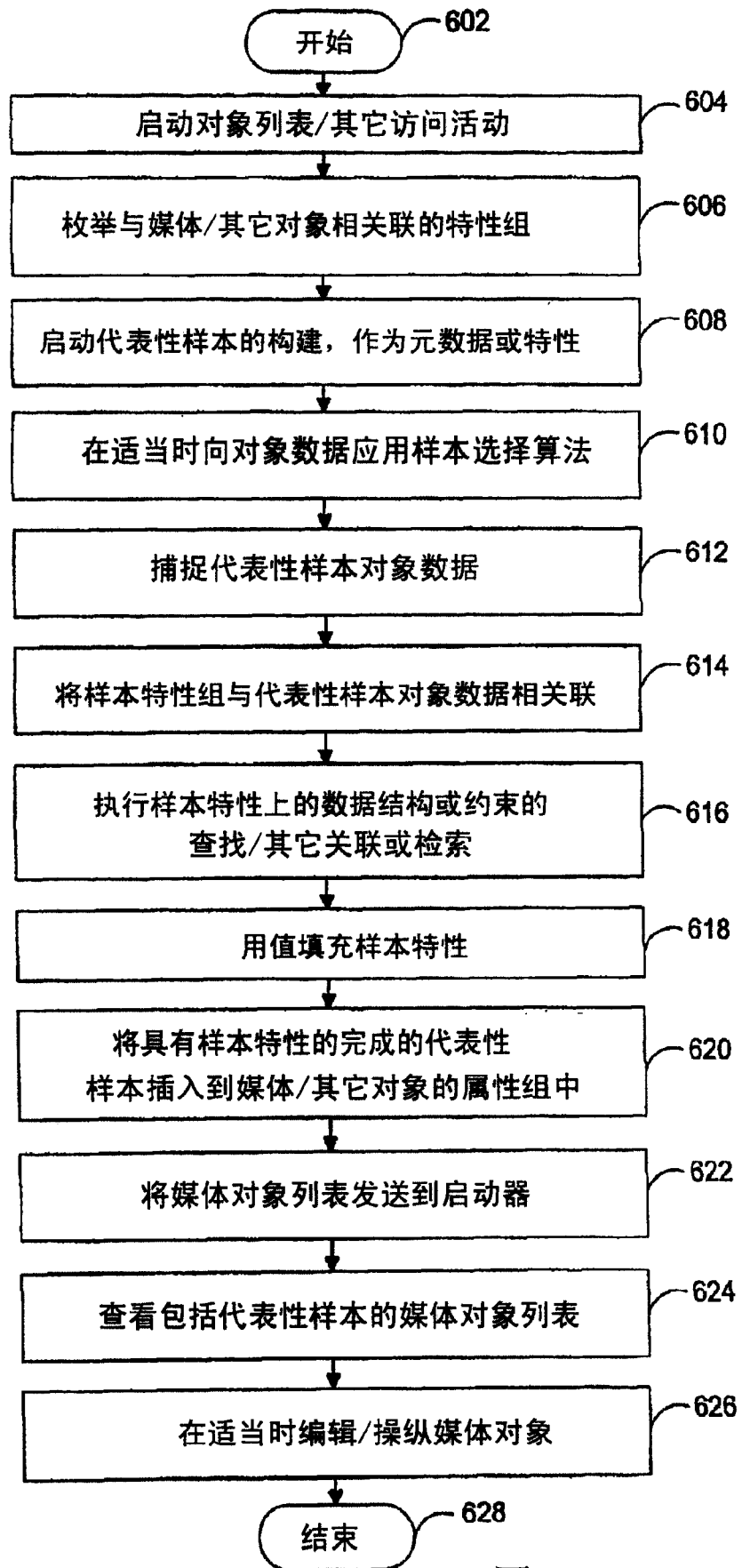


图 6