



(12) 发明专利

(10) 授权公告号 CN 102906816 B

(45) 授权公告日 2015. 09. 09

(21) 申请号 201180025388. 5

(22) 申请日 2011. 05. 23

(30) 优先权数据

12/786, 483 2010. 05. 25 US

(85) PCT国际申请进入国家阶段日

2012. 11. 22

(86) PCT国际申请的申请数据

PCT/US2011/037530 2011. 05. 23

(87) PCT国际申请的公布数据

W02011/149825 EN 2011. 12. 01

(73) 专利权人 高智 83 基金会有限责任公司

地址 美国内华达州

(72) 发明人 A·迪弗

(74) 专利代理机构 中科专利商标代理有限责任

公司 11021

代理人 王波波

(51) Int. Cl.

G11B 27/034(2006. 01)

G11B 27/28(2006. 01)

G11B 27/32(2006. 01)

G06F 17/30(2006. 01)

H04N 5/76(2006. 01)

(56) 对比文件

CN 101427250 A, 2009. 05. 06, 说明书第 3

页.

W0 2007073347 A1, 2007. 06. 28, 全文.

审查员 郭晓冰

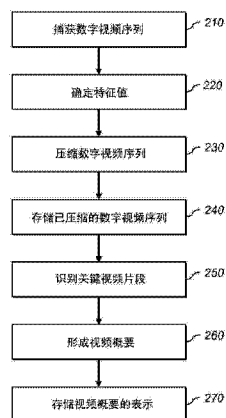
权利要求书2页 说明书18页 附图4页

(54) 发明名称

视频概要方法

(57) 摘要

一种用于形成数字视频序列的视频概要的方法,包括:使用数字视频设备捕获数字视频序列,其中,所述数字视频序列包括多个视频帧;使用处理器在所述数字视频序列被捕获时自动地分析所述数字视频序列中的至少一个视频帧以确定一个或多个特征值;压缩所述数字视频序列;将已压缩数字视频序列存储在处理器可访问的存储器中;使用处理器自动地分析所述特征值,以识别一个或多个关键视频片段,而不对所存储的已压缩数字视频序列进行解压缩,每个关键视频片段与来自所述数字视频序列的一组视频帧相对应;对所述关键视频片段进行组装以形成视频概要;以及将所述视频概要的表示存储在处理器可访问的存储器中。



1. 一种用于形成数字视频序列的视频概要的方法,所述数字视频序列是使用数字视频捕获设备捕获的,所述方法包括:

使用所述数字视频设备捕获数字视频序列,其中,所述数字视频序列包括多个视频帧和一个或多个音频采样;

使用处理器在所述数字视频序列被捕获时自动地分析所述数字视频序列中的至少一个所述视频帧以确定一个或多个特征值;

压缩所述数字视频序列,其中,仅在已经确定了相应特征值之后压缩所分析的视频帧;

将已压缩数字视频序列存储在处理器可访问的存储器中;以及

在存储已压缩数字视频序列之后,使用处理器;

从所存储的已压缩数字视频序列中提取音频采样;

自动地分析所述特征值和所提取的音频采样,以识别一个或多个关键视频片段,而不对所存储的已压缩数字视频序列的视频帧进行解压缩,每个关键视频片段与来自所述数字视频序列的一组视频帧相对应;

对所述关键视频片段进行组装以形成视频概要;以及

将所述视频概要的表示存储在处理器可访问的存储器中。

2. 如权利要求1所述的方法,还包括:使用处理器在所述数字视频被捕获时自动地分析一个或多个数字视频捕获设备设置,以确定一个或多个特征值。

3. 如权利要求1所述的方法,还包括:使用处理器在所述数字视频被捕获时自动地分析加速计所生成的数据,以确定一个或多个特征值。

4. 如权利要求1所述的方法,其中,所确定的特征值被存储在所述处理器可访问的存储器中,作为与所存储的已压缩数字视频序列相关联的元数据。

5. 如权利要求1所述的方法,其中,所确定的特征值被存储在所述处理器可访问的存储器中,作为与所存储的已压缩数字视频序列相关联的单独文件。

6. 如权利要求1所述的方法,其中,自动地分析所述特征值的步骤还包括:根据用户输入识别一个或多个关键视频片段。

7. 如权利要求1所述的方法,其中,将所述视频概要的表示存储在处理器可访问的存储器中还包括:对所存储的已压缩数字视频序列的至少一部分进行解压缩,以提取与所述视频概要相对应的视频帧,并对所提取的视频帧进行压缩,以形成已压缩视频概要;以及将已压缩视频概要存储在处理器可访问的存储器中。

8. 如权利要求7所述的方法,还包括:提取与所述视频概要相对应的音频采样;以及将所提取的音频采样包括在已压缩视频概要中。

9. 如权利要求7所述的方法,其中,所提取的视频帧被再采样,以产生不同空间分辨率的所提取的视频帧。

10. 如权利要求1所述的方法,其中,所述视频概要的所存储的表示是以适合于使用视频播放器进行播放的格式进行存储的数字视频文件。

11. 如权利要求1所述的方法,其中,将所述视频概要的表示存储在处理器可访问的存储器中还包括:生成元数据,所述元数据提供关于所述数字视频序列中的与所述视频概要相对应的视频帧的指示,并且其中,所述元数据是与所存储的已压缩数字视频序列相关联

地存储的。

12. 一种用于形成数字视频序列的视频概要的方法,所述数字视频序列是使用数字视频捕获设备捕获的,所述方法包括:

- a) 使用所述数字视频捕获设备捕获数字视频序列,其中,所述数字视频序列包括多个视频帧;
- b) 使用处理器在所述数字视频序列被捕获时自动地分析所述数字视频序列中的至少一个所述视频帧,以确定一个或多个特征值;
- c) 压缩所述数字视频序列,其中,仅在已经确定了相应特征值之后压缩所分析的视频帧;
- d) 将已压缩数字视频序列存储在处理器可访问的存储器中;
- e) 在存储已压缩数字视频序列之后,使用处理器自动地分析所述特征值以识别一个或多个关键视频片段,而不对所存储的已压缩数字视频序列进行解压缩,每个关键视频片段与来自所述数字视频序列的一组视频帧相对应;
- f) 对所述关键视频片段进行组装以形成视频概要;
- g) 允许用户预览所述视频概要;
- h) 从所述用户接收调整的用户设置;
- i) 使用处理器来响应于所述调整的用户设置而自动地分析所述特征值以识别一个或多个更新的关键视频片段,而不对所存储的已压缩数字视频序列进行解压缩,每个更新的关键视频片段与来自所述数字视频序列的一组视频帧相对应;
- j) 对所述更新的关键视频片段进行组装,以形成更新的视频概要;以及
- k) 将所述更新的视频概要的表示存储在处理器可访问的存储器中。

13. 如权利要求 12 所述的方法,其中,重复步骤 g)-j),直到所述用户对所预览的视频概要满意为止。

视频概要方法

技术领域

[0001] 本发明涉及数字视频处理领域,并且更具体地涉及一种用于形成视频概要的方法。

背景技术

[0002] 许多数字捕获设备能够捕获视频以及静止图像。然而,管理数字视频内容可能是一项困难的任务。通常用视频的第一帧的缩略图来可视地表示视频。这可能不会提供很多的对于视频内容的洞察。确定在给定的视频中是否包含特定事件常常需要查看整个视频。对于较长的视频而言,用户可能更愿意能够获得视频的快速概要,而不必查看整个视频。

[0003] 从共享的角度来看,数字视频也可能呈现出实际问题。许多数字捕获设备以高达 1920×1080 个像素的空间分辨率、以每秒30或60帧的速度记录视频。即使在被压缩时,所生成的数据的量也可能使得共享甚至相对较短的视频不切实际。

[0004] 可以使用视频编辑软件来手工地将视频概括成能够更容易共享的较短的版本。然而,手工视频编辑可能是一个冗长的繁重的过程,并且许多用户对手工编辑并不感兴趣。也存在自动视频概括算法。这些方案开始于将捕获的视频作为输入,并分析视频以确定视频概要。然而,这样的算法是非常复杂的,这是因为需要对视频进行解码以执行确定视频概要所需要的分析。因而,不可能在数字捕获设备上立即查看与刚刚捕获的视频相对应的视频概要。这种缺点使得难以有助于对捕获的视频进行快速回顾和共享。

[0005] 因而,需要提供用于在数字捕获设备中计算视频概要的系统和方法。具体地,需要提供允许在完成视频捕获时在极小延迟的情况下在数字捕获设备上生成视频概要的方案。

发明内容

[0006] 本发明代表一种用于形成数字视频序列的视频概要的方法,其中所述数字视频序列是使用数字视频捕获设备捕获的,所述方法包括:

[0007] 使用所述数字视频捕获设备捕获数字视频序列,其中,所述数字视频序列包括多个视频帧;

[0008] 使用处理器来在捕获所述数字视频序列时自动地分析所述数字视频序列中的所述多个视频帧中的至少一个,以确定一个或多个特征值;

[0009] 压缩所述数字视频序列;

[0010] 将已压缩数字视频序列存储在处理器可访问的存储器中;

[0011] 使用处理器来自动地分析所述特征值,以识别一个或多个关键视频片段,而不需要对所存储的已压缩数字视频序列进行解压缩,每个关键视频片段与来自所述数字视频序列的一组视频帧相对应;

[0012] 组装所述关键视频片段以形成视频概要;以及

[0013] 将所述视频概要的表示存储在处理器可访问的存储器中。

[0014] 本发明具有以下优点:其在捕获时分析视频数据,以确定特征值,从而消除了

建视频概要时对解压缩视频数据的需要。

[0015] 其具有以下额外优点：其在完成捕获之后分析特征值，以识别一个或多个关键视频片段，而不需要对所存储的数字视频序列进行解压缩。对关键视频片段进行组装以形成视频概要。这允许在完成视频捕获时在极小的延迟的情况下在数字捕获设备上生成视频概要并查看该视频概要。

[0016] 在一些实施例中，使用元数据将视频概要编码在数字视频文件中，而不需要将视频概要编码成单独的文件。这具有以下优点：使用理解视频概要元数据的“智能”视频播放器通常可以查看这种视频概要，同时这种视频概要对于传统的播放器是透明的。

附图说明

[0017] 图 1 是示出了根据本发明实施例的用于形成视频概要的系统的组件的高级图；

[0018] 图 2 是根据本发明实施例的用于创建视频概要的方法的流程图；

[0019] 图 3 是根据本发明实施例的用于使用用户反馈来创建视频概要的方法的流程图；

[0020] 图 4 是根据本发明实施例的用于创建视频概要的方法的流程图，其中，将视频概要存储成元数据；以及

[0021] 图 5 是根据本发明实施例的用于查看视频概要的方法的流程图。

具体实施方式

[0022] 在以下描述中，将按照通常实现为软件程序的方式来描述本发明的优选实施例。本领域技术人员将容易地认识到，还可以用硬件来构造这种软件的等价形式。因为图像操作算法和系统是公知的，所以本描述将特别地针对形成根据本发明的系统和方法的一部分或更直接与该系统和方法合作的算法和系统。这种算法和系统的其它方面以及用于生成并以其它方式处理涉及于其中的图像信号的硬件或软件在本文中没有具体示出或描述，可以从本领域已知的这样的系统、算法、组件和元件中进行选择。给出在以下材料中根据本发明描述的系统，本文中没有具体示出、暗示或描述的用于实现本发明的软件是常规的，并且处于本领域普通技术知识范围内。

[0023] 此外，如本文中使用的，用于执行本发明的方法的计算机程序可以存储在计算机可读存储介质中，计算机可读存储介质可以包括例如：磁存储介质，例如磁盘（例如硬盘驱动器或软盘）或磁带；光存储介质，例如光盘、光带或机器可读条形码；固态电子存储设备，例如随机存取存储器（RAM）或只读存储器（ROM）；或者用于存储具有指令的计算机程序的任何其它物理设备或介质，其中所述指令用于控制一个或多个计算机实施根据本发明的方法。

[0024] 本发明包括本文描述的实施例的组合。提及“特定实施例”等是指在本发明的至少一个实施例中存在的特征。单独提及“实施例”或“特定实施例”等未必是指相同的一个实施例或相同的多个实施例；然而，这样的实施例并不是相互排斥的，除非这样指示或对于本领域技术人员是显而易见的。在提及“方法”或“多种方法”等时使用单数或复数并不是限制性的。应当注意的是，除非上下文明确地提到或需要，否则在本公开内容中使用的词语“或者”具有非排它性的意义。

[0025] 因为采用成像设备以及相关的用于信号捕获和处理以及显示的电路的数字照相

机是公知的,所以本描述将具体地针对形成根据本发明的方法和装置的一部分的元件或者与所述方法和装置更直接地协作的元件。在本文中未具体示出或描述的元件是从本领域已知的元件中选择的。将要描述的实施例的某些方面是以软件形式提供的。考虑到在以下材料中示出并描述的根据本发明的系统,在本文中未具体示出、描述或暗示的可用于实现本发明的软件是常规的,并在本领域普通技术的范围内。

[0026] 本领域技术人员将熟悉以下对数字照相机进行的描述。将明显的是,对于该实施例存在可能的供选择的许多变型,以降低成本,增加特征或改善照相机的性能。

[0027] 图 1 描述了根据本发明的包括能够捕获视频图像的数字照相机 10 的数字摄影系统的方框图。优选地,数字照相机 10 是便携式的用电池操作的设备,其足够小以至于用户在捕获和回顾图像时能够容易地手持。数字照相机 10 生成数字图像,使用图像存储器 30 将所述数字图像存储成数字图像文件。本文中所使用的短语“数字图像”或“数字图像文件”是指诸如数字静止图像或数字视频文件之类的任何数字图像文件。

[0028] 在一些实施例中,数字照相机 10 捕获运动视频图像和静止图像两者。在其它实施例中,数字照相机 10 是只捕获运动视频图像的数字视频照相机。数字照相机 10 还可以包括其它功能,包括但不限于数字音乐播放器(例如,MP3 播放器)、移动电话、GPS 接收机或可编程数字助理(PDA)的功能。

[0029] 数字照相机 10 包括镜头 4,其具有可调孔径和可调快门 6。在优选实施例中,镜头 4 是变焦镜头,并且由变焦和聚焦电机驱动器 8 来控制。镜头 4 将来自场景的光(未示出)聚焦到图像传感器 14 (例如,单芯片彩色 CCD 或 CMOS 图像传感器)上。镜头 4 是一种用于在图像传感器 14 上形成场景的图像的光学系统。在其它实施例中,该光学系统可以使用固定焦距的镜头,其具有可变或固定焦点。

[0030] 图像传感器 14 的输出由模拟信号处理器(ASP)和模拟-数字(A/D)转换器 16 转换成数字形式,并被临时存储在缓冲存储器 18 中。随后处理器 20 使用存储在固件存储器 28 中的嵌入式软件程序(例如,固件)来操作存储在缓冲存储器 18 中的图像数据。在一些实施例中,使用只读存储器(ROM)将软件程序永久地存储在固件存储器 28 中。在其它实施例中,可以通过使用例如 Flash EPROM 存储器来修改固件存储器 28。在这样的实施例中,外部设备可以使用有线接口 38 或无线调制解调器 50 来更新存储在固件存储器 28 中的软件程序。在这样的实施例中,固件存储器 28 还可以用于存储图像传感器校准数据、用户设置选择以及在关闭照相机时必须保留的其它数据。在一些实施例中,处理器 20 包括程序存储器(未示出),并且存储在固件存储器 28 中的软件程序在被处理器 20 执行之前被复制到程序存储器中。

[0031] 将理解的是,可以使用单个可编程处理器或通过使用包括一个或多个数字信号处理器(DSP)设备在内的多个可编程处理器来提供处理器 20 的功能。可替换地,可以通过定制电路(例如,通过被具体设计用在数字照相机中的一个或多个定制集成电路(IC))或者通过可编程处理器和定制电路的组合来提供处理器 20。将理解的是,可以使用公共数据总线来进行处理器 20 与图 1 中所示的各种组件中的一些或全部之间的连接。例如,在一些实施例中,可以使用公共数据总线在处理器 20、缓冲存储器 18、图像存储器 30 和固件存储器 28 之间进行连接。

[0032] 然后使用图像存储器 30 来存储被处理的图像。将理解的是,图像存储器 30 可以是

本领域技术人员已知的任何形式的存储器,包括但不限于可移除闪存卡、内部闪存芯片、磁存储器或光学存储器。在一些实施例中,图像存储器 30 可以包括内部闪存芯片以及与可移除闪存卡(例如安全数字(SD)卡)的标准接口。可替换地,可以使用不同的存储器卡形式,例如微 SD 卡、压缩闪存(CF)卡、多媒体卡(MMC)、xD 卡或存储棒。

[0033] 图像传感器 14 由时序生成器 12 控制,时序生成器 12 产生各种时钟信号,用以选择行和像素并同步 ASP 和 A/D 转换器 16 的操作。例如,图像传感器 14 可以具有 12.4 兆像素(4088×3040 个像素),以便提供大致 4000×3000 个像素的静止图像文件。为了提供彩色图像,图像传感器通常与滤色镜阵列重叠,滤色镜阵列提供的图像传感器具有包括不同颜色的像素的像素阵列。不同颜色的像素可以以许多不同的模式进行排列。作为一个例子,可以使用公知的 Bayer 滤色镜阵列来排列不同颜色的像素,如在共同转让给 Bayer 的美国专利 3971065“彩色成像阵列”中所描述的,该美国专利的公开内容通过引用方式并入本文。作为第二例子,可以对不同颜色的像素进行排列,如在共同转让给 Compton 和 Hamilton 的、于 2007 年 7 月 28 日提交的并且名称为“Image sensor with improved light sensitivity”的美国专利申请公开 2005/191729 中所描述的,该美国专利申请的公开内容通过引用方式并入本文。这些例子不是限制性的,并且可以使用许多其它的颜色模式。

[0034] 将理解的是,图像传感器 14、时序生成器 12 以及 ASP 和 A/D 转换器 16 可以是单独制造的集成电路,或者它们可以被制造成单个的集成电路,如通常对 CMOS 图像传感器所做的。在一些实施例中,这种单个的集成电路可以执行图 1 中所示的其它功能中的一些功能,包括处理器 20 所提供的功能中的一些。

[0035] 图像传感器 14 在第一模式中被时序生成器 12 驱动时是有效的,用以提供较低分辨率的传感器图像数据的运动序列以便组成图像,该运动序列在捕获视频图像时并且也在预览被捕获的静止图像时使用。可以作为例如具有 1280×720 个像素的 HD 分辨率图像数据,或者作为例如具有 640×480 个像素的 VGA 分辨率图像数据,或者使用其它分辨率(其与图像传感器的分辨率相比,具有显著更少列和行的数据),来提供这种预览模式的传感器图像数据。

[0036] 可以通过对具有相同颜色的邻近像素的值进行组合,或者通过消除一些像素值,或者通过对一些颜色的像素值进行组合同时消除其它颜色的像素值,来提供预览模式的传感器图像数据。可以如共同转让给 Parulski 等的、名称为“Electronic camera for initiating capture of still images while previewing motion images”的美国专利 6292218 中所描述的那样处理预览模式的图像数据,该美国专利申请通过引用方式并入本文。

[0037] 图像传感器 14 在第二模式中被时序生成器 12 驱动时也是有效的,以提供高分辨率的静止图像数据。提供这种最终模式的传感器图像数据,以作为高分辨率的输出图像数据,对于具有较高的光照水平的场景而言,所述高分辨率的输出图像数据包括图像传感器的全部像素,并且可以是例如具有 4000×3000 个像素的 12 兆像素的最终图像数据。在较低光照水平,可以通过在图像传感器上“放弃”(binning)一些数量的颜色相似的像素来提供最终的传感器图像数据,以便提高信号水平并从而提高传感器的“ISO 速度”。

[0038] 变焦和聚焦电机驱动器 8 由处理器 20 所供应的控制信号来控制,以便提供适当的焦距设置并将场景聚焦到图像传感器 14 上。通过控制可调孔径和可调快门 6 的 f/ 数量和

曝光时间、图像传感器 14 的经由时序生成器 12 的曝光时段以及 ASP 和 A/D 转换器 16 的增益(即,ISO 速度)设置,来控制图像传感器 14 的曝光水平。处理器 20 还对闪光灯 2 进行控制,闪光灯 2 可以对场景进行照明。

[0039] 可以通过使用共同转让给 Parulski 等的名称为“Electronic camera with Rapid Automatic Focus of an Image upon a Progressive Scan Image Sensor”的美国专利 5668597 中所描述的“通过镜头”的自动聚焦来在第一模式中对数字照相机 10 的镜头 4 进行聚焦。这是通过以下方式实现的:使用变焦和聚焦电机驱动器 8 将镜头 4 的焦点位置调节到在近焦点位置与无限远焦点位置之间变化的多个焦点位置,同时处理器 20 确定为图像传感器 14 所捕获的图像的中心部分提供峰值的最近的焦点位置。随后可以将与最近的焦点位置相对应的焦点距离用于若干目的,例如自动设置适当的场景模式,并且该焦点距离可以作为元数据与其它镜头和照相机设置一起存储在图像文件中。

[0040] 处理器 20 生成菜单和低分辨率的彩色图像,它们临时存储在显示存储器 36 中并在图像显示器 32 上进行显示。图像显示器 32 通常是有源矩阵彩色液晶显示器(LCD),虽然也可以使用其它类型的显示器,例如有机发光二极管(OLED)显示器。视频接口 44 向视频显示器 46 (例如平板 HDTV 显示器)提供来自数字照相机 10 的视频输出信号。在预览模式或视频模式中,处理器 20 对来自缓冲存储器 18 的数字图像数据进行操作,以形成一系列运动预览图像,所述一系列运动预览图像通常作为彩色图像在图像显示器 32 上进行显示。在预览模式中,使用来自存储在图像存储器 30 中的数字图像文件的图像数据来产生在图像显示器 32 上显示的图像。

[0041] 响应于通过用户控制 34 所提供的用户输入,来控制在图像显示器 32 上显示的图形用户界面。用户控制 34 用于选择各种照相机模式,例如视频捕获模式、静止捕获模式和预览模式,并且用于发起对静止图像的捕获以及对运动图像的记录。在一些实施例中,当用户部分地按压快门按钮时,发起上面描述的第一模式(即,静止预览模式),快门按钮是用户控制 34 中的一个,并且当用户完全按压快门按钮时,发起第二模式(即,静止图像捕获模式)。用户控制 34 还用于开启照相机,控制镜头 4,以及发起图片拍摄过程。用户控制 34 通常包括按钮、摇杆式开关、操纵杆或选转盘的某种组合。在一些实施例中,通过使用图像显示器 32 上的触摸屏幕覆盖来提供用户控制 34 中的一些。在其它实施例中,可以使用额外的状态显示器和图像显示器。

[0042] 可以使用用户控制 34 选择的照相机模式包括“定时器”模式。当选择“定时器”模式时,在用户完全按压快门按钮之后,在处理器 20 发起对静止图像的捕获之前出现较短的延迟(例如,10 秒)。

[0043] 连接到处理器 20 的音频编解码器 22 从麦克风 24 接收音频信号,并向扬声器 26 提供音频信号。这些组件可以用于与视频序列或静止图像一起记录以及回放音频轨迹。如果数字照相机 10 是多功能设备(例如,组合照相机和移动电话),那么麦克风 24 和扬声器 26 可以用于电话会话。

[0044] 在一些实施例中,扬声器 26 可以用作用户接口的一部分,例如以便提供指示已经按压了用户控制或者已经选择了特定模式的各种可听信号。在一些实施例中,可以使用麦克风 24、音频编解码器 22 和处理器 20 来提供语音识别,使得用户能够通过使用语音命令而不是用户控制 34 来向处理器 20 提供用户输入。还可以使用扬声器 26 将进入的电话呼叫

通知给用户。这可以通过使用存储在固件存储器 28 中的标准振铃音调,或者通过使用从无线网络 58 下载的并存储在图像存储器 30 中的定制振铃音调来完成。此外,可以使用振动设备(未示出)来提供进入电话呼叫的静音(例如,非可听的)通知。

[0045] 在一些实施例中,数字照相机 10 还包含加速计 27,其提供与照相机的运动有关的数据。优选地,加速计 27 检测三个正交方向中的每个方向上的线性和旋转加速(总共有 6 维的输入)。

[0046] 处理器 20 还提供对来自图像传感器 14 的图像数据的额外处理,以便产生被渲染的 sRGB 图像数据,其被压缩并被存储在图像存储器 30 中的“已完成”图像文件中,例如公知的 Exif-JPEG 图像文件。

[0047] 可以经由有线接口 38 将数字照相机 10 连接到接口/可再充电器 48,其连接到计算机 40,计算机 40 可以是位于家庭或办公室的台式计算机或便携式计算机。有线接口 38 可以遵循例如公知的 USB 2.0 接口规范。接口/可再充电器 48 可以经由有线接口 38 向数字照相机 10 中的一组可再充电电池(未示出)提供电力。

[0048] 数字照相机 10 可以包括无线调制解调器 50,其通过射频频带 52 与无线网络 58 对接。无线调制解调器 50 可以使用各种无线接口协议,例如公知的蓝牙无线接口或公知的 802.11 无线接口。计算机 40 可以经由互联网 70 将图像上载到照片服务提供商 72,例如 Kodak EasyShare Gallery。其它设备(未示出)可以访问照片服务提供商 72 所存储的图像。

[0049] 在可替换的实施例中,无线调制解调器 50 通过射频(例如,无线)链路与诸如 3GSM 网络之类的移动电话网络(未示出)通信,移动电话网络与互联网 70 连接,以上载来自数字照相机 10 的数字图像文件。可以将这些数字图像文件提供给计算机 40 或照片服务提供商 72。

[0050] 现在将参考图 2 来描述本发明。初始地,在捕获数字视频序列的步骤 210 中,使用诸如数字照相机 10 之类的数字视频捕获设备来捕获数字视频序列,其中,数字视频序列由多个视频帧组成。

[0051] 在捕获数字视频序列时,在确定特征值的步骤 220 中,分析视频帧的至少一个子集以确定一个或多个特征值。这些特征值可以表示与视频帧相关联的各种属性,例如颜色特性或在视频帧中出现任何人脸。特征值还可以表示基于运动的属性,例如连续视频帧之间的全局运动值,或者连续帧的相应元素之间的局部运动值。全局运动值通常与捕获设备的运动相对应,而局部运动值是指场景中的对象的移动。本领域技术人员将认识到前面描述的特征仅仅是例子,并且基于对视频帧的分析可以确定其它特征值。

[0052] 确定特征值的步骤 220 还可以包括在捕获数字视频序列时在通过麦克风 24 记录至少一个音频采样并用音频编解码器 22 处理所述至少一个音频采样时,分析所述至少一个音频采样。音频特征值可以包括时域中的以及关于特定频率范围的属性,例如信号强度。

[0053] 确定特征值的步骤 220 还可以包括在捕获数字视频序列时分析数字视频捕获设备设置。这些设备设置特征值可以包括用户控制 34 形式的用户输入,所述用户输入用于通过控制变焦和聚焦电机驱动器 8 来调整变焦镜头 4 的变焦位置。数字变焦还可以提高图像的放大率并且可以是另一特征值。

[0054] 确定特征值的步骤 220 还可以包括在捕获数字视频序列时分析来自加速计 27 的

加速计数据。加速计数据提供与数字照相机 10 的运动有关的度量,并且可以用于补充或替代基于视频帧数据的运动计算。

[0055] 确定特征值的步骤 220 还可以包括由处理器 20 分析视频编码器所生成的数据。这种数据包括作为在视频编码过程中使用的运动估计步骤的一部分而生成的运动向量信息。许多通用的视频编码过程通常生成这种运动向量信息,以作为其正常处理链的一部分。

[0056] 对于每个视频帧,在已经确定了特征值之后,压缩数字视频序列的步骤 230 对视频帧进行压缩。MPEG 和 H. 263 标准中规定的视频压缩算法对于本领域技术人员而言是公知的。可以将已压缩的视频帧包括在诸如 Apple QuickTime 之类的容器中,其为视频文件提供文件格式封装器(wrapper)。

[0057] 存储已压缩数字视频序列的步骤 240 将已压缩数字视频序列存储在处理器可访问的存储器中,例如图像存储器 30。所存储的已压缩数字视频序列可以包含视频和音频信息。

[0058] 在本发明的一个实施例中,将在确定特征值的步骤 220 中确定的特征值存储成与所存储的已压缩数字视频序列相关联的元数据。可以将元数据存储例如在用户数据原子中,其中用户数据原子是在 Apple QuickTime 文件格式的规范中定义的。

[0059] 在本发明的另一实施例中,可以在确定特征值的步骤 220 中确定的特征值存储成与所存储的已压缩数字视频序列相关联的单独文件。

[0060] 在本发明的又一实施例中,在确定特征值的步骤 220 中确定的特征值没有存储在任何形式的永久存储器中,而是在完成视频概要形成算法之后被丢弃。

[0061] 在已经完成了视频捕获过程并且已经存储了已压缩数字视频序列之后,执行识别关键视频片段的步骤 250,以识别数字视频序列的关键视频片段。识别关键视频片段的步骤 250 使用处理器自动地分析特征值以识别一个或多个关键视频片段,而不需要对所存储的已压缩数字视频序列进行解压缩,每个关键视频片段与来自数字视频序列的一组视频帧相对应。通常,组成关键视频片段的视频帧是数字视频序列中的连续帧,并且可以由开始帧号以及结束帧号或指示关键视频片段的长度的值来标识关键视频片段。

[0062] 在本领域中,用于基于与视频相关联的特征值来确定来自视频的关键视频片段的许多方法都是已知的。在这些方法中的一些方法中,首先识别视频关键帧,然后通过选择视频序列中的环绕每一个视频关键帧的一部分来形成关键视频片段。Luo 等的美国专利申请公开 2007/0183497 (其通过引用方式并入本文) 公开了基于数字运动估计技术所计算出来的在视频中的运动特性来选择视频关键帧的方法。Luo 等的美国专利申请公开 2007/0182861 (也通过引用方式包括在本文中) 公开了一种基于来自加速计的数据所提供的在视频中的运动特性来选择视频关键帧的方法,其中,加速计是与视频捕获设备相关联的。根据本发明,可以使用这些方法中的任何一个来根据所确定的特征值识别关键视频片段。

[0063] 对于将特征值存储成与所存储的已压缩数字视频序列相关联的单独文件的实施例而言,识别关键视频片段的步骤 250 可以从该单独文件读取特征值,以提供特征值信息以供分析。

[0064] 类似地,对于将特征值存储成与所存储的已压缩数字视频序列相关联的元数据的实施例而言,识别关键视频片段的步骤 250 可以从所存储的已压缩数字视频序列中提取特

征值,而不对所存储的数字视频序列的视频帧进行解压缩。应当注意到,从所存储的已压缩数字视频序列中提取特征值元数据不被视为对所存储的已压缩数字视频序列进行解压缩。相反,对所存储的已压缩数字视频序列进行解压缩是指对已压缩比特流中的数据(例如,用于重构已压缩数字视频序列的视频帧序列的视频数据或报头数据)进行解码。

[0065] 本发明的优点在于,其允许在不对所存储的已压缩数字视频序列进行解压缩的情况下确定关键视频片段。这允许在完成捕获过程之后以极小的延迟确定关键视频片段并接着形成视频概要。现有技术中的依赖于解压缩来从所存储的已压缩数字视频序列中提取各个视频帧的方法受到执行解压缩任务所需要的时间量的限制。

[0066] 在本发明的一些实施例中,基于在捕获数字视频序列时确定的特征值的分析,整个地识别关键视频片段。在本发明其它实施例中,基于对特征值以及从所存储的已压缩数字视频序列中提取的额外信息的分析,识别关键视频片段。在这些情况中,可能需要至少部分地对所存储的已压缩数字视频序列中的一些信息进行解压缩。

[0067] 例如,在本发明的一个实施例中,从所存储的已压缩数字视频序列中提取音频信息,并且该音频信息与特征值一起用于识别关键视频判断。音频数据通常包括整个已压缩数字视频文件的较小的百分比,并且相对于对视频帧的像素数据进行解压缩,可以快速地提取音频数据。在一些情况中,不可能在捕获时生成音频特征值。处理器可能不具有可用于确定音频特征值并且还确定其它特征值以及处理并解码所捕获的视频的计算周期。在这些情况中,通过从所存储的已压缩数字视频序列中提取音频信息,仍然可以使用音频信息来识别关键视频片段。这表示速度和性能之间的折中。音频信息可以用于改善识别关键视频片段的步骤 250 的性能,代价是识别关键视频片段所需要的整个处理时间更长。

[0068] 类似地,在本发明的另一实施例中,从所存储的已压缩数字视频序列中提取视频帧信息,并且视频帧信息与特征值一起用于识别关键视频片段。对整个视频序列进行解压缩消除了捕获时确定基于视频的特征值的许多优点。因而,优选地,最多只从所存储的已压缩数字视频序列中解码少量的视频帧。本领域技术人员将认识到,当各个帧是独立编码的帧时,可以高效地从已压缩数字视频序列中解码各个帧。而且,视频信息的使用表示速度和性能之间的折中。视频信息可以用于改善识别关键视频片段的步骤 250 的性能,代价是识别关键视频片段所需要的整个处理时间更长。

[0069] 在本发明的一些实施例中,识别关键视频片段的步骤 250 使用来自用户控制 34 的用户输入来控制视频概要的各种属性。用户输入可以包括但不限于指定视频概要的长度、各个关键视频片段的最小持续时间以及关键视频片段的总数。

[0070] 在优选实施例中,识别关键视频片段的步骤 250 使用共同转让给 Deever 的、名称为“Method For Determining Key Video Frames”的共同待审美国专利申请 12/786,471 中描述的方法。这种方法先分析数字视频序列,以确定依据时间的重要性值。重要性值可以是数字视频序列的包括全局运动和局部运动在内的特性的函数。优选地,重要性值是使用从确定特征值的步骤 220 确定的特征值计算的。使用重要性值来形成数字视频序列的扭曲时间表示。扭曲时间表示给数字视频序列中的一些视频帧赋予较大的时间权重,而给其它视频帧赋予较小的时间权重。扭曲时间数字视频序列表示被划分成一组相等的扭曲时间间隔。通过对每个扭曲时间间隔内的视频帧进行分析,来从每个扭曲时间间隔当中选择关键视频帧。优选地,每个扭曲时间间隔的关键视频帧是通过以下方式选择的:对从确定特征

值的步骤 220 确定的与相应扭曲时间间隔内的视频帧相关联的特征值进行分析。这种准则可以包括：选择在完成放大过程之后立即出现的关键视频帧，或者选择在视频帧的中心区域中具有中等水平的局部运动的关键视频帧。

[0071] 在一个实施例中，通过选择一组环绕每个关键视频帧的视频帧来形成关键视频片段。例如，可以通过选择关键视频帧前后的两秒的视频帧来形成关键视频片段，从而形成四秒长的关键视频片段。

[0072] 可替换地，可以对关键视频帧进行排名，并且可以只针对与排名最高的关键视频帧相对应的关键视频帧子集来形成关键视频片段。对关键视频帧进行排名可以包括分析数字视频序列以确定照相机固定模式，以及响应于照相机固定模式来对关键视频帧进行排名。优选地，从确定特征值的步骤 220 确定的与全局运动相对应的特征值可以提供全局运动轨迹，其指示在整个数字视频序列中照相机的固定区域。与高固定的区域（即，在视频捕获的较高百分比内照相机所保持固定的区域）相对应的视频帧被排名较高。排名过程可以迭代地执行，在每个步骤选择下一最高排名的关键视频帧，并且在每个步骤该过程可以将表示未被表示的固定区域的关键视频帧提升到表示已经被包括在排名过程中的固定区域的关键视频帧之上。一旦已经对关键视频帧进行了排名，就可以选择最高排名的关键视频帧以被包括在关键视频片段中。

[0073] 如之前所描述的，可以通过选择一组环绕每个关键视频帧的视频帧来形成关键视频片段。可替换地，可以通过设置诸如视频概要的总持续时间以及关键视频片段的最小允许持续时间之类的准则来形成关键视频片段。然后可以选择关键视频片段，以满足所设置的准则同时最优地响应于额外的准则。例如，可以选择关键视频片段以使得关键视频片段在讲话部分的中间开始或结束的可能性最小。

[0074] 在已经识别出关键视频片段之后，形成视频概要的步骤 260 形成视频概要。通过将各个关键视频片段组装成单个集合来形成视频概要。在优选实施例中，与关键视频片段在数字视频序列中的出现顺序相对应的时间顺序来组装这些关键视频片段。

[0075] 在存储视频概要的表示的步骤 270 中，将视频概要的表示存储在处理器可访问的存储器中。在本发明的一个实施例中，将视频概要的表示存储在处理器可访问的存储器中包括：存储帧标识元数据，其提供关于数字视频序列中的与视频概要相对应的视频帧的指示。可以将帧标识元数据与所存储的已压缩数字视频序列相关联地存储。例如，视频概要表示可以作为所存储的已压缩数字视频序列中的元数据进行存储，所述元数据指示与包含在视频概要中的关键视频片段相对应的一系列开始帧和结束帧。这允许除了存储帧标识元数据所需要的物理存储器以外，在不使用任何额外的物理存储器的情况下存储视频概要的表示。

[0076] 在本发明的另一个实施例中，将视频概要的表示存储在处理器可访问的存储器中包括：提取所存储的已压缩数字视频序列中的与所识别的关键视频片段相对应的视频帧，并对它们进行合并，已形成新合并的与视频概要相对应的视频序列。在一些情况中，这可能需要对已压缩数字视频序列的多个部分进行解压缩，并且然后对合并的视频序列进行压缩以形成已压缩视频概要。合并的视频序列通常将包括视频数据以及相应的音频数据，所述相应的音频数据也必须从已压缩数字视频序列中被提取出来。

[0077] 随后，可以将已压缩视频概要存储在处理器可访问的存储器中。在这个实施例中，

将已压缩视频概要存储成与已压缩数字视频序列不同的文件。在这种情况下,视频概要表示是可以独立于原始的已压缩数字视频序列而被查看或共享的数字视频文件。在优选实施例中,视频概要的所存储的表示是适合于使用标准视频播放器进行播放的格式的数字视频文件。

[0078] 在本发明的一些实施例中,对所提取出来的与视频概要相对应的视频帧进行再采样,以便生成不同空间分辨率的所提取的视频帧,之后对不同空间分辨率的所提取的视频帧进行压缩以形成已压缩视频概要。这种再采样对于共享以高分辨率捕获的视频而言可能特别有益。用这种方式,已压缩视频概要将比已压缩数字视频序列小,这不仅是因为其包含更少的视频帧,还是因为视频帧具有较低的空间分辨率。可以生成较低空间分辨率的视频概要,而不必对较高分辨率的视频帧的整个集合进行解压缩。只需要对生成视频概要所需要的那些视频帧进行解压缩。

[0079] 类似地,可以在时间维度中对所提取的与视频概要相对应的视频帧进行再采样,以生成不同时间分辨率的所提取的视频帧,之后对不同空间分辨率的所提取的视频帧进行压缩以形成已压缩视频概要。

[0080] 图 3 与本发明的另一实施例相对应,其中,用户预览所形成的视频概要,并可以请求使用调整的用户设置来形成更新的视频概要。在这个实施例中,执行之前参考图 2 所描述的捕获数字视频序列的步骤 210、确定特征值的步骤 220、压缩数字视频序列的步骤 230、存储已压缩数字视频序列的步骤 240、识别关键视频片段的步骤 250、形成视频概要的步骤 260 以及存储视频概要的表示的步骤 270。在存储由形成视频概要的步骤 260 形成的视频概要之前,允许用户使用预览视频概要的步骤 262 来预览视频概要。随后使用用户满意测试 264 提示用户来指示他或她对第一视频概要的满意。如果用于满意该视频概要,则执行到参考图 2 所描述的存储视频概要的步骤 270。然而,如果用于不满意该视频概要,则使用调整用户设置的步骤 226 来允许用户调整一个或多个设置,并且随后形成新的视频概要。这些用户的设置可以通过用户控制 34 来输入,并且可以包括对诸如视频概要的长度、关键视频片段的最小长度或视频概要中所包括的关键视频片段的数量之类的参数的设置。在用户已经调整了用户设置之后,使用识别关键视频片段的步骤 250 以及形成视频概要的步骤 260 来响应于用户设置形成新视频概要。本领域技术人员将认识到,用户可以预览视频概要并多次更新用户设置,从而形成相应的视频概要,直到用户对所生成的视频概要满意为止。

[0081] 图 4 示出了本发明的另一实施例,其中,将视频概要存储与所存储的数字视频序列相关联的元数据。捕获数字视频序列的步骤 410 使用数字视频捕获设备来捕获数字视频序列,其中,数字视频序列包括多个视频帧。存储数字视频序列的步骤 420 将数字视频序列存储在处理器可访问的存储器中。所存储的数字视频序列可以包括或不包括所确定的特征值,其是使用图 2 中的确定特征值的步骤 220 而确定的。

[0082] 接着,识别关键视频片段的步骤 430 识别与来自所存储的数字视频序列的一组或多组视频帧相对应的一个或多个关键视频片段。在一些实施例中,识别关键视频片段的步骤 430 可以使用参考图 2 中的识别关键视频片段的步骤 250 所描述的所存储的特征值。在其它实施例中,识别关键视频片段的步骤 430 使用视频分析算法来直接分析所存储的数字视频序列中的帧。在这种情况下,识别关键视频片段的步骤 430 通常将会需要对数字视频序列进行解压缩以执行其分析。可以使用本领域任何已知的方法来识别关键视频片段,包

括之前参考图 2 所描述的那些方法在内。

[0083] 形成视频概要的步骤 440 以与图 2 中的形成视频概要的步骤 260 类似的方式对所识别的关键视频片段进行组装,以形成视频概要。将视频概要存储成元数据的步骤 450 通过存储与所存储的数字视频序列相关联的元数据来将视频概要存储在处理器可访问的存储器中,其中,元数据提供关于与视频概要相对应的视频帧的指示。

[0084] 识别关键视频片段的步骤 430、形成视频概要的步骤 440 以及将视频概要存储成元数据的步骤 450 不必发生在与捕获数字视频序列的步骤 410 和存储数字视频序列的步骤 420 相同的数字视频捕获设备上,或使用与步骤 410 和 420 相同的处理器 20。例如,存储数字视频序列的步骤 420 可以导致将所存储的数字视频序列写入到图像存储器 30,图像存储器 30 例如可以是可移除存储卡。随后,可移除存储卡可以用于将所存储的数字视频序列移动到另一设备,在所述另一设备中可以应用识别关键视频片段的步骤 430、形成视频概要的步骤 440 以及将视频概要存储成元数据的步骤 450。在一些情况中,所述另一设备可以是另一数字视频捕获设备,诸如图 1 中所示出的。在其它实施例中,将所存储的数字视频序列加载到诸如计算机 40 (图 1)或视频编辑系统之类的另一设备上,在所述另一设备上可以处理所存储的数字视频序列以确定视频概要。

[0085] 在优选实施例中,识别关键视频片段的步骤 430 涉及从所存储的数字视频序列中提取至少一个视频帧,以及分析所述视频帧以确定与这些视频帧相关联的特征值。在优选实施例中,识别关键视频片段的步骤 430 还涉及从所存储的数字视频序列中提取至少一个音频采样,以及分析所述音频采样以确定与所述音频采样相关联的特征值。可以分析特征值以识别关键视频片段。

[0086] 在优选实施例中,将视频概要存储成元数据的步骤 450 包括存储帧标识元数据,其提供关于数字视频序列中的与视频概要相对应的视频帧的指示。以这种方式将视频概要存储成帧表示元数据具有以下优点:除了存储帧标识所需要的极小量的存储器以外,不需要使用任何额外的物理存储器来存储视频概要。优选地,将帧标识元数据与所存储的已压缩数字视频序列相关联地存储。例如,可以将视频概要表示存储成所存储的已压缩数字视频序列中的元数据,所述元数据指示与包含在视频概要中的关键视频片段相对应的一系列开始帧和结束帧。在其它实施例中,可以将帧标识元数据存储在与所存储的数字视频序列相关联的单独文件中。

[0087] 在本发明的一些实施例中,将视频概要存储成元数据的步骤 450 还包括存储元数据,所述元数据提供关于与组成视频概要的关键视频片段相对应的音频采样的指示。

[0088] 在本发明的一些实施例中,将视频概要存储成元数据的步骤 450 还包括存储元数据,所述元数据提供关于将在所识别的关键视频片段之间的边界处应用的视频过渡效果的指示。视频过渡效果可以包括:例如从一个关键视频片段衰落到另一个(即,“淡入淡出”)以及从关键视频片段衰落到白色(或黑色)然后衰落到下一视频片段。其它类型的过渡效果将包括特殊效果,例如交互转景效果、生长或收缩圆环效果、水平或垂直百叶窗效果以及棋盘过渡效果。本领域技术人员将认识到这些仅仅是几个例子,并且根据本发明可以使用许多不同类型的过渡效果。使用视频过渡效果可以通过消除从数字视频序列的一个段到另一个段的突然过渡,来改善视频概要的总体视觉质量。

[0089] 在本发明的一些实施例中,将视频概要存储成元数据的步骤 450 还包括存储元数

据,所述元数据提供关于在关键视频片段之间的边界处应用的音频过渡效果的指示。音频过渡效果可以包括:例如衰落到静音和背景音。这些音频过渡效果可以通过消除数字视频序列的一个音频段到另一个音频段的突然过渡来改善视频概要的总体音频质量。

[0090] 图 5 示出了根据本发明实施例的由用于显示视频概要的视频回放系统使用的过程的流程图。具体地,图 5 中的方法用于显示被存储成元数据而不是被存储成可直接播放的数字视频文件的视频概要。

[0091] 首先,读取数据的步骤 510 读取与所存储的数字视频概要相关联的数据,其中,所述数据提供关于所存储的数字视频序列中的与视频概要相对应的视频帧的指示。例如,可以从用于存储数字视频序列的数字视频文件中所存储的元数据当中提取提供关于视频帧的指示的数据,其中所述元数据指示哪些视频帧包括在视频概要中。可替换地,还可以从与所存储的数字视频序列相关联的单独文件中提取所述数据。

[0092] 接着,提取视频帧的步骤 520 从所存储的数字视频序列中提取与组成视频概要的关键视频片段相对应的视频帧。通常,所存储的数字视频序列是以压缩形式存储的。在这样的情况中,提取视频帧的步骤 520 包括对所存储的数字视频序列进行解压缩。

[0093] 形成视频概要的步骤 530 用于从所提取出来的针对关键视频片段的视频帧形成视频概要。在最简单的情况中,形成视频概要的步骤 530 简单地将所提取的视频帧汇编成连续的视频剪辑。

[0094] 一旦形成了视频概要,就使用显示视频概要的步骤 540 来在软拷贝显示器上显示视频概要。软拷贝显示器可以是例如数字视频捕获设备上的回顾屏幕、与计算机或电视屏幕相关联的显示器。

[0095] 读取数据的步骤 510 还可以提供关于与视频概要相对应的音频采样的指示。在这种情况下,提取视频帧的步骤 520 还可以包括从所存储的数字视频序列中提取与视频概要相对应的适当的音频采样。

[0096] 在本发明的一些实施例中,读取数据的步骤 510 还提供关于视频概要的视频过渡效果的指示。类似地,读取数据的步骤 510 可以提供关于视频概要的音频过渡效果的指示。在读取数据的步骤 510 提供关于视频过渡效果的指示的情况中,形成视频概要的步骤 530 根据指定的过渡效果来修改所提取的视频帧,使得汇编的视频概要具有期望的过渡。因而,例如,可以通过读取以下数据来应用衰落到黑色的视频过渡效果:所述数据指示在视频概要中所包含的关键视频片段的最后 15 个帧期间出现衰落。当从所存储的数字视频序列中提取出与衰落到黑色的视频过渡相对应的帧时,可以根据逐渐衰落到黑色的公式来修改视频帧数据。在下一关键视频片段的前 15 个帧期间可以保留衰落到黑色,在此期间,根据从黑色逐渐衰落回原始视频数据的公式来修改所提取的帧。

[0097] 在软拷贝显示器还包含一个或多个用于发送音频的扬声器的情况中,在经由扬声器发送音频之前可以类似地处理音频过渡效果。例如,可以通过读取以下数据来应用衰落到静音的音频过渡效果:所述数据指示在与关键视频片段相对应的最后 8000 个音频采样期间发生音频过渡效果。当从所存储的数字视频序列中提取与衰落到静音的过渡相对应的音频采样时,可以根据逐渐衰落到静音的公式来修改音频采样数据。可以在与下一关键视频片段相对应的前 8000 个音频采样期间保留衰落到静音,在此期间,根据从静音逐渐衰落回原始音频数据的公式来修改所提取的音频采样。

[0098] 在根据本发明的用于形成视频概要的一些方法中,与视频概要中所包含的视频帧或音频采样相对应的视频概要元数据以及视频或音频过渡效果存储在包含所存储的数字视频序列的数字视频文件中。并且,在根据本发明的用于显示视频概要的一些系统中,读取该视频概要元数据以提供关于需要从所存储的数字视频序列中提取哪些与视频概要相对应的视频帧和音频采样的指示以及关于在显示器上显示视频并向扬声器发送音频之前应当应用什么样的视频和音频过渡效果的指示。这些方法和系统具有以下优点:可以在不存储额外的与视频概要相对应的单独视频文件的情况下表示并显示视频概要。

[0099] 在一些实施例中,视频显示设备可以向用户呈现查看原始视频或视频概要的选项。例如,可以在数字视频摄像机上提供用于播放原始视频序列或视频概要的单独的播放按钮。在一些实施例中,可以响应于激活“快进”按钮而播放视频概要。这反映了以下事实:视频概要代表了以加速的节奏通过视频序列的替换方式,而不是简单地提供对视频序列中的帧的时间采样的传统快进方式。播放视频概要可能是一种供用户识别视频序列中的、他们将要查看的期望部分的更有用的方式。

[0100] 参考图 5 描述的视频回放系统在其能够从所存储的数字视频序列中提取视频概要元数据并处理视频概要元数据以理解如何显示数字视频序列的视频概要版本的意义上讲,是一种“智能”视频播放器。“智能”视频播放器可以向查看者呈现查看全部数字视频序列或视频概要版本的选项。常规的视频播放器不会识别与视频概要相关联的视频概要元数据。然而,常规的视频播放器仍然可以读取并显示原始的所存储的数字视频序列。在这种情况下,所存储的数字视频序列中所包含的元数据被简单地忽略,并播放全部的数字视频序列。

[0101] 再次参考图 1 和图 2,本发明的另一实施例是用于捕获数字视频序列并用于生成相应的视频概要的数字视频捕获系统(数字照相机 10),包括:图像传感器 14,用于捕获视频帧;光学系统(具有可调孔径的镜头 4 以及可调快门 6),用于将场景成像到图像传感器上;处理器 20,用于形成视频概要并将视频概要的表示存储在处理器可访问的存储器中;图像显示器 32,用于查看数字视频序列;以及诸如用户控制 34 之类的模块,用于向用户提供可在软拷贝显示器上查看所存储的数字视频序列或视频概要的所存储的表示的选项。数字视频捕获系统可以是例如数字视频照相机或具有静止和视频捕获模式的数字照相机。数字视频捕获系统还可以采用其它形式,例如具有网络照相机的膝上型计算机或台式计算机。

[0102] 向数字视频捕获系统中的处理器 20 提供用于执行图 2 中所示的方法步骤的软件。更具体地,响应于用户命令,处理器 20 可以发起捕获数字视频序列的步骤 210,该步骤捕获包括多个视频帧的数字视频序列。处理器 20 在数字视频序列被捕获时自动地分析数字视频序列中的视频帧子集,以使用确定特征值的步骤 220 来确定一个或多个特征值。随后,处理器 20 通过应用压缩数字视频序列的步骤 230 来压缩数字视频序列;并且通过应用存储已压缩数字视频序列的步骤 240 将所生成的已压缩数字视频序列存储在处理器可访问存储器中。随后,处理器 20 应用识别关键视频片段的步骤 250 来自动地分析特征值,而不对所存储的已压缩数字视频序列进行解压缩,从而识别出一个或多个关键视频片段,每个关键视频片段与来自数字视频序列的一组视频帧相对应。最后,处理器 20 应用形成视频概要的步骤 260 来组装关键视频片段,以形成视频概要,并使用存储视频概要的表示的步骤 270 将视频概要的表示存储在处理器可访问的存储器中。

[0103] 数字视频捕获系统中的处理器 20 还可以在数字视频序列被捕获时执行自动地分析数字视频序列中的至少一个音频采样的步骤,以确定一个或多个特征值。

[0104] 数字视频捕获系统中的处理器 20 还可以执行将所确定的特征值作为与所存储的已压缩数字视频序列相关联的元数据而存储在处理器可访问的存储器中的步骤。

[0105] 数字视频捕获系统中的处理器 20 还可以执行将所确定的特征值作为与所存储的已压缩数字视频序列相关联的单独文件而存储在处理器可访问的处理器中的步骤。

[0106] 数字视频捕获系统中的处理器 20 还可以执行自动地分析特征值以及用户输入以识别一个或多个关键视频片段的步骤。用户输入可以来自于用户控制 34,并且可以包括关于视频概要的长度、视频概要中所包含的任意关键视频片段的最小长度以及视频概要中所包括的关键视频片段的数量的约束。

[0107] 数字视频捕获系统中的处理器 20 还可以执行将视频概要的表示存储在处理器可访问的存储器中的步骤,其中,生成提供关于数字视频序列中的与视频概要相对应的视频帧的指示的元数据,并将该元数据与所存储的已压缩数字视频序列相关联地存储。

[0108] 用户可以激活数字视频捕获系统上的用户控制 34,以在软拷贝图像显示器 32 上或在外部视频显示器 46 上查看视频概要。在将视频概要存储成用于提供关于数字视频序列中的与视频概要相对应的视频帧的指示的元数据的情况中,可以提取并显示视频概要,如图 5 中所描述的。具体地,数字视频捕获系统还可以执行“智能”数字播放器的角色,并可以基于元数据提取合适的视频和音频数据。

[0109] 在一些实施例中,可以提供用户界面,以使得用户能够回顾视频概要并决定其是否是可接受的。如果用户确定视频概要不是所期望的,那么数字视频捕获系统可以通过用户控制 34 提供用于向用户呈现使用用户调节的用户设置来请求视频概要的选项的方式。这些调节的用户设置可以包括:例如,视频概要的长度、视频概要中的关键视频片段的最小长度或者视频概要中将包括的关键视频片段的数量。在一些实施例中,该系统可以使用不同的设置或者识别关键视频片段的步骤 250 的不同变形来自动形成一系列视频概要选项。因而,可以有效地使得用户能够在多个视频概要版本之间进行选择,从而减小了没有生成令人满意的视频概要的可能性,并提高了特征的总体成功。

[0110] 这种所描述的数字视频捕获系统提供了以下优点:可以快速地生成并回顾一系列视频概要,而不必对所存储的已压缩数字视频序列进行解压缩,并且不必将视频概要编码成新的文件。

[0111] 在将要在网络上共享视频概要或者与不配备有“智能”视频显示器的任意设备共享视频概要的情况中,使用单独的视频概要文件来存储视频概要是期望的,其中,“智能”视频显示器能够解释在已压缩数字视频序列中包含的元数据并从而能够提取视频概要。在这种情况下,数字视频捕获系统的处理器 20 所生成的数字视频概要的所存储的表示可以是以适合于使用标准播放器进行播放的格式进行存储的数字视频文件。在一些实施例中,视频概要总是被存储成单独的视频概要文件。可替换地,可以提供用户控制 34,以允许用户选择用于存储视频概要的优选方法。在一些实施例中,响应于用户启动用户界面上的“共享”按钮,视频概要可以被存储成单独的视频概要文件。

[0112] 在数字视频捕获系统将视频概要存储成单独文件的实施例中,处理器 20 还可以执行在处理器可访问的存储器中将视频概要的表示存储成数字视频文件的步骤。在这种情

况中,向处理器 20 提供用于对所存储的已压缩数字视频序列中的与视频概要相对应的至少一部分进行解压缩,然后对所提取的视频帧进行压缩以形成已压缩的视频概要的软件,其中,可以将已压缩的视频概要存储在处理器可访问的存储器中。此外,数字视频捕获系统中的处理器 20 可以执行对所存储的已压缩数字视频序列的至少一部分进行解压缩以提取与视频概要相对应的音频采样、压缩所提取的音频采样并将已压缩的所提取的音频采样包括在已压缩的视频概要中的步骤。

[0113] 还通过在共享之前降低原始视频序列的空间或时间分辨率,来有助于在计算机网络上共享视频文件。数字视频捕获系统中的处理器 20 还可以对所提取的视频帧进行再采样,以生成不同空间或时间分辨率的所提取的视频帧。本发明具有以下优点:其允许形成较低分辨率的视频概要文件,而不需要对整个原始的已压缩数字视频序列进行解压缩。只需要对与生成视频概要有关的已压缩数字视频数据进行解压缩。这允许从原始的已压缩数字视频序列到视频概要的更快速的编码转换。

[0114] 还通过提高用于减小已压缩数字视频序列的大小的压缩的积极性,来有助于在网络上共享视频文件。与所存储的已压缩数字视频序列相比,数字视频捕获系统中的处理器 20 还可以以更积极的压缩设置来压缩所提取的视频帧,以生成被压缩得更厉害的视频概要。本发明具有以下优点:其允许形成被压缩得更厉害的视频概要文件,而不必对整个原始的已压缩数字视频序列进行解压缩。

[0115] 数字视频捕获系统的图像传感器 14、光学系统(镜头 4)、处理器 20 以及软拷贝显示器(图像显示器 32)可以都是数字视频照相机的组件,并且数字视频照相机上的用户界面控制可以用于向用户提供在数字视频照相机的软拷贝显示器上查看所存储的已压缩数字视频序列或所存储的视频概要。在这种框架中,用户可以在数字视频照相机上捕获视频,立即查看该视频的概要版本,如果希望的话则请求修改,然后选择共享该视频概要,从而可以生成并共享单独的视频概要文件。

[0116] 数字视频照相机还可以连接到外部软拷贝显示器,使得用户可以在外部软拷贝显示器上查看所存储的已压缩数字视频序列或所存储的视频概要。

[0117] 软拷贝显示器还可以是能够访问所存储的数字视频序列和所存储的视频概要的独立视频查看系统的一个组件,并且数字照相机上的用户界面控制可以向用户提供在软拷贝显示器上查看所存储的已压缩数字视频序列或所存储的视频概要的选项。

[0118] 数字视频捕获系统的处理器 20 还可以在形成视频概要时提供在关键视频片段之间应用的视频过渡效果。可以在显示时计算视频过渡效果的结果。这对于被存储成与所存储的已压缩数字视频序列相关联的元数据的视频概要而言尤其有利。

[0119] 在本发明的一些实施例中,数字视频捕获系统的处理器 20 通过分析关键视频片段自动地从多个可用的过渡效果中选择视频过渡效果。对于一些视频内容,衰落到白色的过渡是优选的,而对于其它视频内容,衰落到黑色可能是优选的,并且对于额外的其它视频内容,额外的不同的视频过渡可能是优选的。对关键视频片段的自动分析可以提供指示哪个视频过渡效果在两个关键视频帧之间提供最好的视觉结果的信息。

[0120] 可以以类似的方式处理音频过渡效果。数字视频捕获系统的处理器 20 还可以在形成视频概要时包括在关键视频片段之间应用的音频过渡效果。可以在显示时计算音频过渡效果的结果。

[0121] 在根据本发明实施例的数字视频捕获系统的另一实施例中,处理器 20 应用识别关键视频片段的步骤 250 的变形,其涉及自动地分析特征值以及在所存储的已压缩数字视频序列中所包含的信息,以识别关键视频片段。在这个实施例中,牺牲了速度来换取确定组成视频概要的关键视频片段时可用的额外信息。在许多情况中,处理器 20 可能不具有可用于在捕获时计算所有的期望特征值的计算周期。在这种情况下,可以对所存储的已压缩数字视频序列进行部分解压缩,以提取要在计算关键视频片段中使用的信息。例如,在本发明的一些实施例中可以提取所存储的已压缩数字视频序列中所包含的音频信息。附加地,在本发明的一些实施例中,还可以提取所存储的已压缩数字视频序列中所包含的视频信息。

[0122] 本发明实现的另一特征是辅助的手工裁剪特征。许多数字视频捕获设备提供手工裁剪的编辑特征,其中,用户可以回顾所捕获的视频,并选择从视频的开始和结束进行裁剪。使用本发明,可以形成包括单个关键视频片段的视频概要,并且将视频概要的开始点和结束点提供成所建议的手工裁剪点。

[0123] 注意,除了形成视频概要以外,在捕获时确定特征值以及在捕获后使用所确定的特征值的过程对于其它应用可能也是有用的。其它选项,包括使用特征值来确定视频中的关键帧在内,被认为落入本发明的范围内。

[0124] 许多数字视频序列压缩算法在对视频进行编码时使用图片组。一个图片组包括独立编码的视频帧以及依赖于来自至少一个其它视频帧的信息以便进行解码的额外视频帧。图片组的一个流行的实例是具有独立编码的帧(“I”帧),之后是预测编码帧(“P”帧)序列,其中,每个“P”帧是从之前的帧预测的。图片组以下一“I”帧结束,并且新的图片组以下一“I”帧开始。在这样的压缩方案中,“I”帧提供对于已压缩数字视频序列的访问点,其中,可以提取以“I”帧开始的多个帧。具体地,可以从已压缩的数字视频序列中提取全部图片组,并将它们编码转换成视频概要,这仅仅需要对指示包括图片组的已压缩字节的位置和数量的报头信息进行解码。因而,在形成视频概要时施加以下约束可能是有利的:每个关键视频片段的开始帧是用于开始图片组的“I”帧,并且优选地,关键视频片段包含正好数量的图片组。这种约束允许在对原始的已压缩数字视频序列进行极小压缩的情况下生成单独的视频概要文件。

[0125] 计算机程序产品可以包括一个或多个存储介质,例如:诸如磁盘(例如软盘)或磁带之类的磁存储介质;诸如光盘、光带或机器可读条形码之类的光存储介质;诸如随机存取存储器(RAM)或只读存储器(ROM)之类的固态电子存储设备;或者用于存储计算机程序的任何其它物理设备或介质,所述计算机程序具有用于控制一个或多个计算机实施根据本发明的方法的指令。

[0126] 部件列表

[0127] 2 闪光灯

[0128] 4 镜头

[0129] 6 可调节孔径和可调节快门

[0130] 8 变焦和聚焦电机驱动器

[0131] 10 数字照相机

[0132] 12 时序生成器

[0133] 14 图像传感器

- [0134] 16 ASP 和 A/D 转换器
- [0135] 18 缓冲存储器
- [0136] 20 处理器
- [0137] 22 音频编解码器
- [0138] 24 麦克风
- [0139] 25 压力传感器
- [0140] 26 扬声器
- [0141] 27 加速计
- [0142] 28 固件存储器
- [0143] 30 图像存储器
- [0144] 32 图像显示器
- [0145] 34 用户控制
- [0146] 36 显示存储器
- [0147] 38 有线接口
- [0148] 40 计算机
- [0149] 42 倾斜传感器
- [0150] 44 视频接口
- [0151] 46 视频显示器
- [0152] 48 接口 / 再充电器
- [0153] 50 无线调制解调器
- [0154] 52 射频频带
- [0155] 58 无线网络
- [0156] 70 互联网
- [0157] 72 照片服务提供商
- [0158] 210 捕获数字视频序列的步骤
- [0159] 220 确定特征值的步骤
- [0160] 230 压缩数字视频序列的步骤
- [0161] 240 存储已压缩的数字视频序列的步骤
- [0162] 250 识别关键视频片段的步骤
- [0163] 260 形成视频概要的步骤
- [0164] 262 预览视频概要的步骤
- [0165] 264 用户满意测试
- [0166] 266 调整用户设置的步骤
- [0167] 270 存储视频概要的表示的步骤
- [0168] 410 捕获数字视频序列的步骤
- [0169] 420 存储已压缩的数字视频序列的步骤
- [0170] 430 识别关键视频片段的步骤
- [0171] 440 形成视频概要的步骤
- [0172] 450 将视频概要存储成元数据的步骤

- [0173] 510 读取数据的步骤
- [0174] 520 提取视频帧的步骤
- [0175] 530 形成视频概要的步骤
- [0176] 540 显示视频概要的步骤

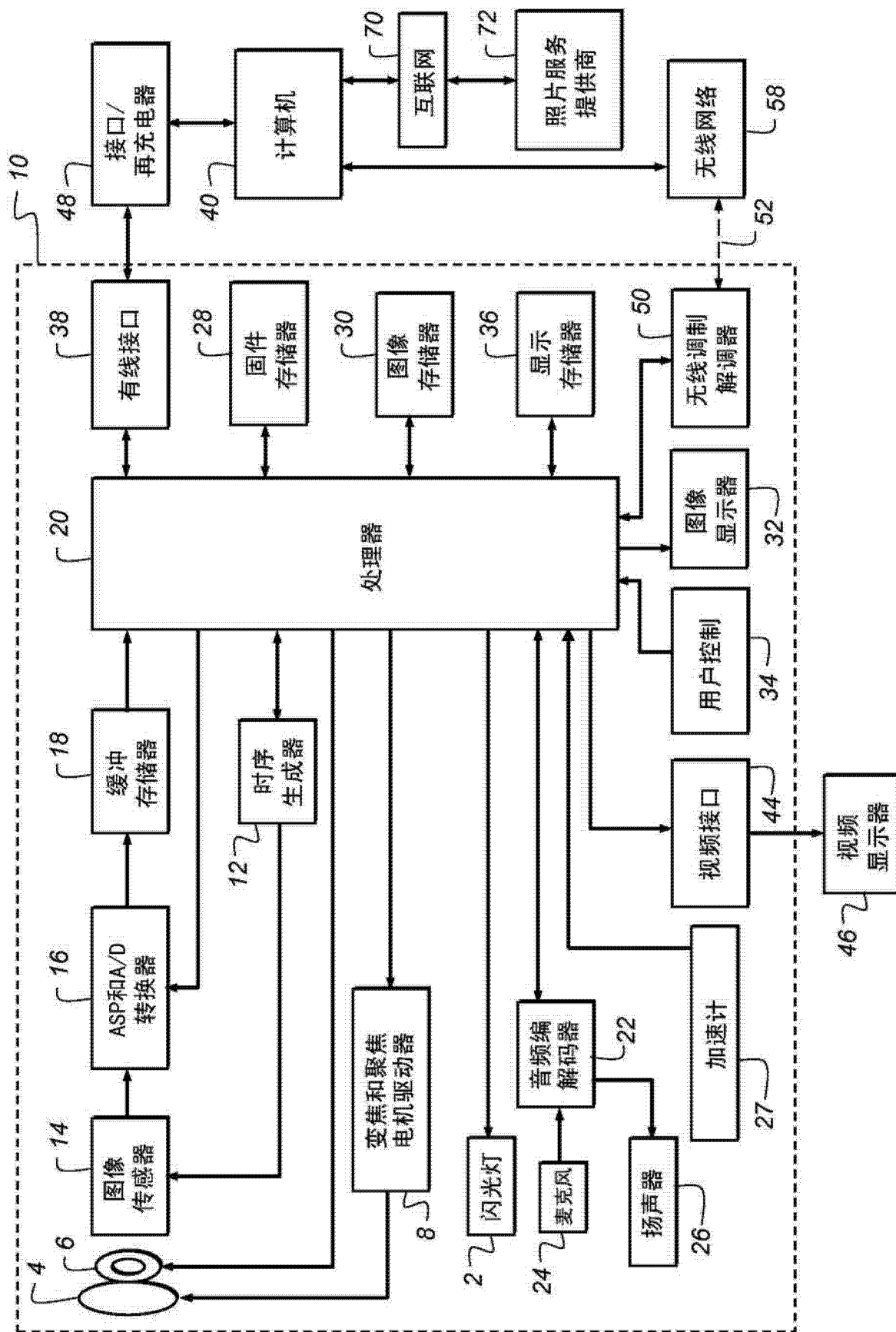


图 1

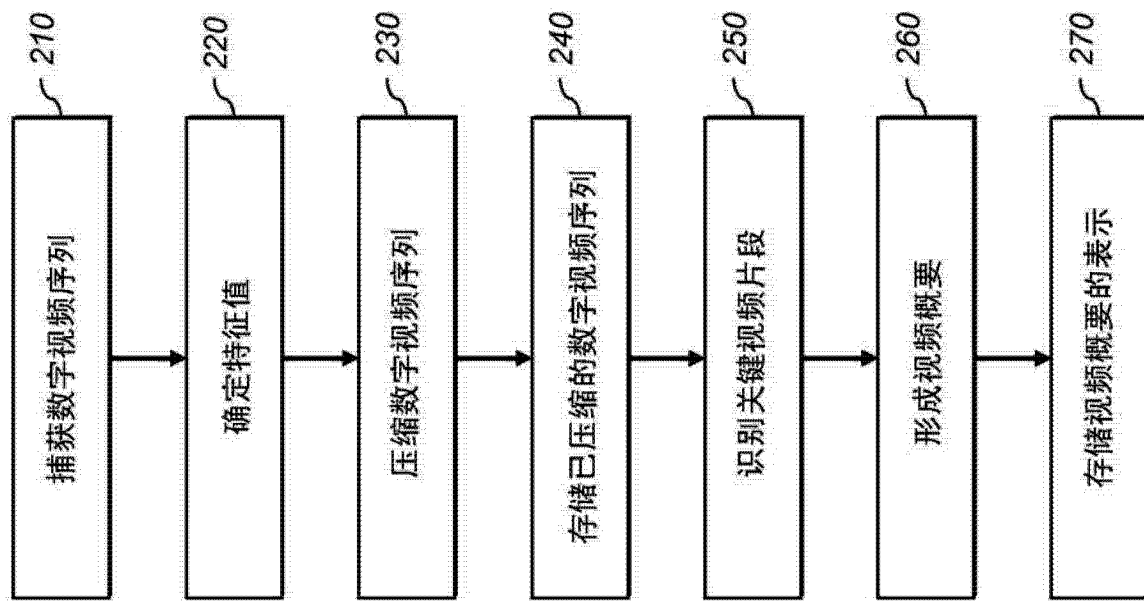


图 2

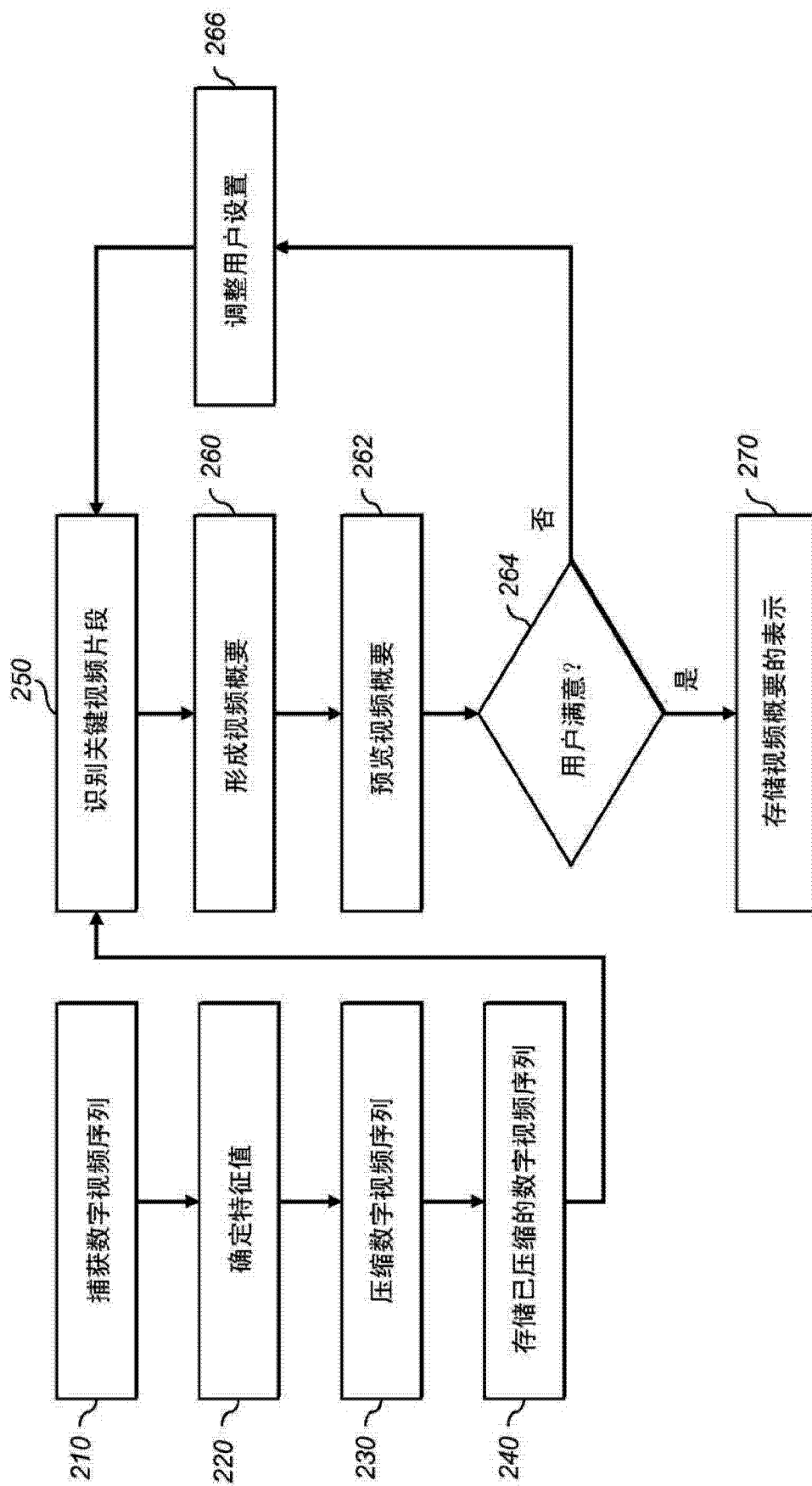


图 3

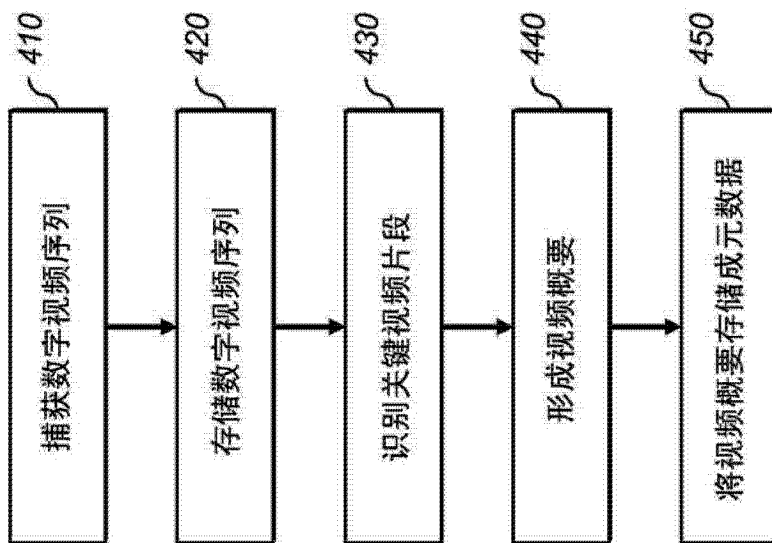


图 4

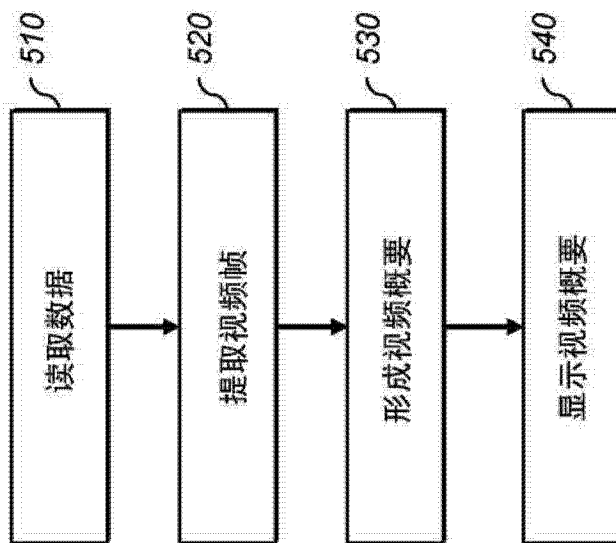


图 5