



(12) 发明专利

(10) 授权公告号 CN 101103364 B

(45) 授权公告日 2010.05.12

(21) 申请号 200580046762.4

(22) 申请日 2005.11.16

(30) 优先权数据

60/628,819 2004.11.17 US

60/628,861 2004.11.17 US

(85) PCT申请进入国家阶段日

2007.07.17

(86) PCT申请的申请数据

PCT/US2005/041253 2005.11.16

(87) PCT申请的公布数据

W02006/055512 EN 2006.05.26

(73) 专利权人 欧几里得发现有限责任公司

地址 美国马萨诸塞州

(72) 发明人 约翰·维斯 查尔斯·保罗·佩斯

(74) 专利代理机构 北京市金杜律师事务所
11256

代理人 王茂华 陈姗姗

(51) Int. Cl.

G06K 9/36 (2006.01)

G06K 9/46 (2006.01)

G06K 9/00 (2006.01)

(56) 对比文件

US 2004/0135788 A1, 2004.07.15, 说明书第
8 页 [0112]-第 9 页 [0113].US 6711278 B1, 2004.03.23, 说明书第 1 栏
第 13 行-第 6 栏第 42 行、附图 1-5.

US 2003/0194134 A1, 2003.10.16, 全文.

US 2004/0107079 A1, 2004.06.03, 说明书第
1 页 [0006], [0009].US 6661004 B2, 2003.12.09, 说明书第 5 栏
第 11-13 行.US 2003/0235341 A1, 2003.12.25, 说明书第
1 页 [0005], 第 5 页 [0064]-[0069].Toklu, Candemir, A. Murat Tekalp, A.
Tanju Erdem. Simultaneous alpha map
generation and 2-D mesh tracking for
multimedia application. Proceeding of
the international conference on image
processing: 1997. 1997, 1113-1116.Rene Vidal, Yi Ma, Shankar Sastry.
Generalized principal component
analysis (GPCA). Proceedings of the IEEE
conference on computer vision and pattern
recognition: 2003. 2003, 1621-1628.Paul Viola, Michael Jones. Rapid object
detection using a boosted cascade of
simple features. 2001 IEEE Computer Society
Conference on Computer Vision and Pattern
Recognition. 2001, 1511-1518.

审查员 孙艳

权利要求书 2 页 说明书 17 页 附图 10 页

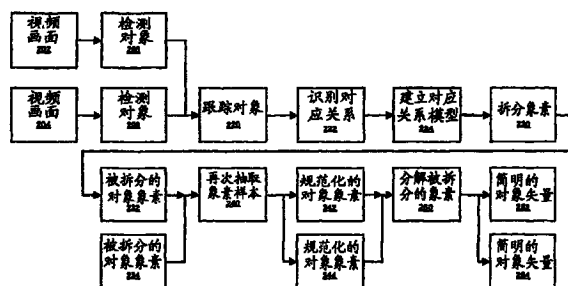
(54) 发明名称

用来处理视频数据的装置和方法

(57) 摘要

这项发明描述用来处理视频数据的装置和方法。本发明提供能用来评估数据和适合该数据的特定参数表达的拟合模型之间的一致性的视频数据表达。这允许比较不同的参数表达技术和选择最适合该特定数据的连续视频处理的技术。这种表达能以中间形式作为较大程序的铺垫或作为反馈机制被用于处理视频数据。在以它的中间形式被利用的时候, 本发明能被用于适合视频数据的储存、增强、细分、特征提取、压缩、编码和传输的

程序。本发明可用来以强健有效的方式提取显著的信息, 同时确定通常与视频数据来源相关联的问题的地址。



1. 一种用于从为数众多的视频画面中生成编码形式的视频信号数据的计算机装置,所述装置包括:

用于对主题视频信号数据的视频画面序列执行混合空间规范化压缩的装置,所述用于执行混合空间规范化压缩的装置包括:

检测装置,用于在主题视频信号数据的视频画面序列中检测至少一个对象;

跟踪装置,用于跟踪所述对象通过所述视频画面序列中的两幅或多幅画面;

拆分装置,用于在与所述对象相关联的视频画面中,将与所述对象相关联的像素数据与其它像素数据拆开;

识别装置,用于识别所述对象在与所述对象相关联的多幅视频画面之间的对应元素;

建模装置,用于建立对应关系的模型以产生模型化的对应关系;

再次抽样装置,用于在与所述对象相关联的所述视频画面中对像素数据再次抽样,所述再次抽样装置利用所述模型化的对应关系;

解除规范化装置,用于对所述再次抽样的像素数据的空间位置解除规范化,从而恢复所述再次抽样的像素数据的空间位置,所述解除规范化装置利用所述模型化的对应关系,以及

所述再次抽样的像素数据是所述主题视频信号数据的中间形式。

2. 根据权利要求1的装置,其中所述对象的检测装置和跟踪装置包括 Viola/Jones 面部检测装置。

3. 根据权利要求1的装置,其中所述对象是使用所述拆分装置从视频画面中拆分出来的,所述拆分装置包括时间整合装置。

4. 根据权利要求1的装置,其中所述用于执行混合空间规范化压缩的装置还包括因式分解装置,用于将所述对应关系的模型因式分解成整体模型,包括:

将所述对应关系整合成整体运动模型的装置;

所述对应关系的建模装置包括对二维仿射运动模型的解的强健的抽样共识装置;以及

所述对应关系的建模装置使用以基于区段的运动评估装置产生的有限差分为基础的抽样总体,所述基于区段的运动评估装置在所述序列之中的两幅或多幅视频画面之间执行运动评估。

5. 根据权利要求1的装置,其中所述用于执行混合空间规范化压缩的装置还包括用于对所述主题视频信号数据的中间形式进一步编码的装置,包括:

分解装置,用于将所述再次抽样的像素数据分解成编码表达;

改写装置,用于依据所述编码表达改写所述再次抽样的像素数据;

所述分解装置包括用于执行主要成份分析的装置,以及

所述改写装置包括用于执行主要成份分析的装置。

6. 根据权利要求5的装置,其中画面的非对象像素是以与所述对象的像素一样的方法建模的,包括:

当除去其它的对象的时候,所述对象是所述画面剩余的非对象。

7. 根据权利要求5的装置,其中所述用于执行混合空间规范化压缩的装置进一步包括组合装置,用于利用执行传统视频压缩/解压缩的装置来组合所述拆分的像素和再次抽样的像素,所述组合装置包括:

供应装置,用于将所述拆分的像素和所述再次抽样的像素作为标准视频数据供应给传统视频压缩处理;

储存和传输装置,用于将模型对应关系数据连同对应的编码视频数据一起进行储存和传输;以及

由此所述组合装置能使所述传统视频压缩装置提高压缩效率。

8. 根据权利要求 1 的装置,其中所述建模装置还包括用于将所述对应关系的模型因式分解成局部失真模型的装置,包括:

网孔定义装置,用于定义覆盖与所述对象相对应的像素的二维网孔,以及

将所述对应关系的测量结果建模到局部运动模型中的装置;

所述网孔定义装置以有顶点和边缘的规则网格为基础,以及

所述对应关系的测量结果包括以在所述序列中的两幅或多幅视频画面之间基于区段的运动评估而产生的有限差分为基础的顶点位移。

9. 根据权利要求 8 的装置,其中所述顶点对应于离散的图像特征,所述网孔定义装置包括:

识别与所述对象相对应的重要的图像特征的装置;

所述识别重要的图像特征的装置用于分析图像梯度 Harris 响应。

用来处理视频数据的装置和方法

[0001] 这份申请要求 2004 年 11 月 17 日以“System And Method For Video Compression Employing Principal Component Analysis”为题申请的美国专利临时申请第 60/628,861 号和 2004 年 11 月 17 日以“Apparatus and Methods for Processing and Coding Video Data”为题申请的美国专利临时申请第 60/628,819 号的优先权。这份申请是作为 2005 年 7 月 28 日申请的美国专利申请第 11/191,562 号的部分继续申请的于 2005 年 9 月 20 日申请的美国专利申请第 11/230,686 号的部分继续申请。上述的每份申请在此通过引证被并入本文作为参考。

发明领域

[0002] 本发明一般地涉及数字信号处理领域,更具体地说涉及用来有效地表达和处理信号或图像数据(最具体地说,视频数据)的计算机装置和用计算机实现的方法。

背景技术

[0003] 本发明能驻留在其中的现有技术的一般系统描述可以表示成图 1。在这里,方框图显示典型的现有技术视频处理系统。这样的系统通常包括下列几级:输入级 102、处理级 104、输出级 106 和一个或多个数据存储机制 108。

[0004] 输入级 102 可能包括诸如摄像机传感器、摄像机传感器阵列、距离检测传感器或从储存机制取回数据的装置之类的元素。输入级提供用来表达人造的和/或自然发生的现象的时间相关序列的视频数据。该数据的显著成份可能被噪音或其它不想要的信号掩盖或污染。

[0005] 呈数据流、数据阵列或数据包形式的视频数据可能被直接或通过中间储存元素 108 依照预先定义的传输协议送到处理级 104。处理级 104 可能采取模拟或数字专用装置或可编程装置(例如,中央处理器(CPU)、数字信号处理器(DSP)或现场可编程门阵列(FPGA))的形式来执行一组预期的视频数据处理操作。处理级 104 通常包括一个或多个 CODEC(编解码器)。

[0006] 输出级 106 产生能够影响使用者或外设的信号、显示或其它响应。通常,输出装置被用来产生指示器信号、显示、硬拷贝、处理过的数据在存储器中的表达,或初始化向远程站点的数据传输。它也可能被用来提供中间信号或控制参数供后续处理操作使用。

[0007] 存储器是作为储存这个系统中的非必选元素呈现的。在使用时,储存元素 108 可能是非易失性的(例如,只读储存媒体)或易失性的(例如,动态随机存取存储器(RAM))。为了包括几种类型的储存元素,它对于单一的视频处理系统并非是异乎寻常的,该元素有各种不同的对输入级、处理级和输出级的关系。这样的储存元素的例子包括输入缓冲器、输出缓冲器和处理高速缓冲存储器。

[0008] 图 1 所示视频处理系统的主要目标是处理输入数据产生对特定的应用具有深远意义的输出。为了实现这个目标,可以利用多种处理操作,包括噪声减少或消除、特征提取、对象拆分和/或规范化、数据分类、事件检测、编辑、数据选择、数据二次编码、和代码转换。

[0009] 许多产生缺乏限制的数据（尤其是声音和可视图像）的数据来源对人是重要的。在大多数情况下，这些来源信号的基本特征对有效地处理数据的目标产生不利的影响。来源数据固有的易变性是在不引进起因于在推演工程假设时使用的自然的、经验的和启发式的方法的误差的情况下以可靠和有效的方式处理数据的障碍。当输入数据被自然地或故意地限制在定义狭窄的特征集（例如，一组有限的符号值或狭窄的带宽）之中的时候，这种易变性将针对应用有所减弱。这些限制时常导致商业价值低的处理技术。

[0010] 信号处理系统的设计受该系统的倾向于用途和作为输入使用的来源信号的预期特性的影响。在大多数情况下，所需要的执行效率也将是一个重要的设计因素。执行效率依次受与可用的数据存储空间相比待处理的数据量以及与可用的计算能力相比应用的计算复杂性的影响。

[0011] 传统的视频处理方法有许多以数据通信速度慢、储存需求大和扰乱知觉人为现象的形式出现的无效性。这些可能因为人们希望使用和操纵视频数据的方法的多样性和因为人们对某些形式的视觉信息有先天的敏感性变成严重的问题。

[0012] “最佳的”视频处理系统在完成一组预期的处理操作方面是高效率的、可靠的和强健的。这样的操作可能包括数据的储存、传输、显示、压缩、编辑、加密、增强、分类、特征检测和确认。二次操作可能包括这样处理过的数据与其它数据来源的整合。在视频处理系统的情况下同等重要的是输出应该通过避免引进知觉人为现象与人类视觉相容。

[0013] 视频处理系统如果它的速度、效率和质量不强烈地取决于输入数据的任何特定特征的细节则可以被描述为“强健的”。强健也与在某些输入出现错误的时候完成操作的能力有关。许多视频处理系统未能强健到足以考虑到应用的一般类别，仅仅为在该系统的研发中使用的同样受狭窄限制的数据提供应用。

[0014] 显著信息由于输入元素的抽样速率与感知现象的信号特性不匹配可能在连续取值的数据来源的离散化中丢失。另外，当信号强度超过传感器极限导致饱和的时候也有遗失。同样，当输入数据的精度下降的时候，数据也会遗失，这在输入数据的完整的数值范围用一组离散数值表达，借此降低数据表达的精度的时候发生在任何量化程序中。

[0015] 总体易变性指的是一类数据或信息来源中的任何无法预测性。因为视觉信息通常不受限制，所以代表视觉信息特征的数据有非常大的总体易变性程度。视觉信息可以表达任何由于光线入射在传感器阵列上所形成的空间阵列序列或时间空间的序列。

[0016] 在仿制视觉现象时，视频处理器通常把一些限制组和 / 或结构强加在表达或解释数据方式上。结果，这样的方法可能引进将会影响输出质量、可能用来考虑输出的置信水平和能在该数据上可靠地完成的后续处理工作的类型的系统误差。

[0017] 一些量化方法降低视频画面中的数据精度同时试图保有那个数据的统计变化。通常，视频数据是这样分析的，以致数据值的分布被收集到概率分布之中。也有一些方法把数据映射到相空间之中，以便将数据的特色表示为空间频率的混合，借此允许精度下降以较少引起反对的方式扩散。这些量化方法在被大量地利用时往往导致知觉上难以相信的颜色和能在该视频画面原本平滑的区域中引起突然的怪异状态。

[0018] 差分编码通常也用来利用数据的局部空间相似性。在画面的一个部分中的数据倾向于聚集在那个画面中的相似数据周围和后续画面中的相似位置。然后，根据它的空间毗连数据表达该数据能与量化组合起来，而最终结果是对于给定的准确性表达差分比使用数

据的绝对值更精确。这个假定在原始视频数据的光谱分辨率有限的时候（例如，在黑白图像或颜色少的图像中）很好地工作。随着图像的光谱分辨率逐渐增加，相似性假定被严重破坏。这种破坏是由于没有能力有选择地保护视频数据准确性造成的。

[0019] 残差编码与差分编码类似，因为这种表达的误差被进一步差分编码，以便把原始数据的准确性恢复到预期的准确性水平。

[0020] 这些方法的变化尝试把视频数据变换成把数据相关关系暴露在空间相位和刻度之中的替代表达。一旦视频数据已经以这些方式变换，量化和差分编码的方法就能适用于被变换的数据，从而导致增加显著图像特征的保存。这些变换视频压缩技术中最普遍两种是离散余弦变换 (DCT) 和离散子波变换 (DWT)。DCT 变换的误差表明在视频数据数值方面有广泛的变化，因此，DCT 通常被用在视频数据的区段上，为的是使这些错误的相关关系定位。来自这种定位的假象往往沿着这些区段的边界出现。就 DWT 而言，更复杂的假象在基础函数和某些纹理之间有误配的时候发生，而且这引起模糊效应。为了抵消 DCT 和 DWT 的负面效应，提高表达的准确性以便以宝贵的带宽为代价减少失真。

发明内容

[0021] 本发明是一种在计算和分析方面均优于现有顶级技术的方法的计算机实现的视频处理方法。原则上本发明的方法是线性分解法，空间拆分法和空间规范化法的整合。从空间上限制视频数据大大提高线性分解法的强健性和适用性。此外，数据的空间拆分与空间规范化相对应能进一步用来增大单独的来自空间规范化的利益。

[0022] 具体地说，本发明提供一种装置，采用该装置能把信号数据有效地处理成一个或多个有益的表达。本发明在处理许多普遍发生的数据组时是有效的而且在处理视频和图像数据时是特别有效的。本发明的方法分析该数据并且提供那个数据的一种或多种简洁表达以使它的处理和编码变得容易。对于许多应用（包括但不限于：视频数据的编码、压缩、传输、分析、储存和显示），每种新的比较简洁的数据表达都允许减少计算处理、传输带宽和储存需求。本发明包括用来识别和提取视频数据的显著成份的方法，从而允许区分数据的处理和表达的优先次序。信号中的噪音和其它多余部分被看作是优先权比较低的，以致进一步的处理能集中在分析和表达视频信号中优先权比较高的部分上。结果，视频信号的表达比先前可能的表达更简洁。而且把准确性的损失集中在视频信号中知觉上不重要的部分。

附图说明

[0023] 图 1 是举例说明现有技术视频处理系统的方框图。

[0024] 图 2 是提供本发明的概观的方框图，它展示用来处理影像的主要组件。

[0025] 图 3 是举例说明本发明的运动评估方法的方框图。

[0026] 图 4 是举例说明本发明的整体配准方法的方框图。

[0027] 图 5 是举例说明本发明的规范化方法的方框图。

[0028] 图 6 是举例说明混合式空间规范化压缩方法的方框图。

[0029] 图 7 是举例说明本发明在局部规范化中使用的网孔生成方法的方框图。

[0030] 图 8 是举例说明本发明在局部规范化中使用的基于网孔的规范化方法的方框图。

[0031] 图 9 是举例说明本发明的整体和局部组合规范化方法的方框图。

[0032] 图 10 是举例说明本发明的 GPCA- 基本多项式拟合和求导方法的方框图。

[0033] 图 11 是举例说明本发明的 GPCA 递归细分法的方框图。

[0034] 具体实施方法

[0035] 在视频信号数据中, 视频画面被组装成通常描绘投影、成像到二维成像表面上的三维视场的图像序列。每个画面(或图像)都由代表响应抽样信号的成像敏感元件的象素组成。时常, 抽样信号对应于用二维敏感元件阵列抽样的一些反射的、折射的或发射的能量(例如, 电磁能、声能, 等等)。连续的顺序抽样导致时空数据流, 每个画面的两个空间维度和时间维度对应于该画面在视频序列中的次序。

[0036] 本发明如同图 2 举例说明的那样分析信号数据和识别显著成份。当信号由视频数据组成的时候, 时空流分析揭示时常作为特定对象(例如, 面部)的显著成份。识别程序限定显著成份的存在和重要性并且选择那些显著成份之中最重要的一个或多个显著成份。这不限在现在描述的处理之后或同时识别和处理其它较为不显著的成份。然后, 上述的显著成份被进一步分析, 以便识别易变的和不变的子成份。不变的子成份的识别是建立该成份的某个方面的模型的程序, 借此揭示该模型的参数表达法, 以允许将该成份合成到预期的准确性水平。

[0037] 在本发明的一个实施方案中, 检测和跟踪前景对象。该对象的象素被识别并且从每个视频画面中拆分出来。基于区段的运动评估被应用于从多个画面中拆分出来的对象。然后, 这些运动评估被整合成较高级的运动模型。该运动模型用来把该对象的例证隐藏到公用的空间配置中。对于特定的数据, 在这个配置中, 该对象更多的特征被对准。这种规范化允许将该对象的象素在多个画面上的数值的线性分解被紧凑地表达。属于该对象的外观的显著信息包含在这个紧凑的表达之中。

[0038] 本发明的优选实施方案详细描述前景视频对象的线性分解。该对象在空间上被规范化, 借此得出紧凑的线性外观模型。此外, 进一步的优选实施方案在空间规范化之前把前景对象从视频画面的背景中拆分出来。

[0039] 本发明的优选实施方案将本发明应用于人对着摄像机边说话边进行少量运动的影像。

[0040] 本发明的优选实施方案将本发明应用于影像中能通过空间转换被很好地表达的任何对象。

[0041] 本发明的优选实施方案明确地使用基于区段的运动评估来确定两个或多个视频画面之间的有限差分。为了要提供更有效的线性分解, 高级运动模型是依据那些有限差分因式分解的。

[0042] 检测 & 跟踪

[0043] 一旦已经确定信号的显著构成成份, 这些成份就可以被保留, 而所有其它的信号成份可以被减少或去除。检测显著成份的程序展示在图 2 中, 在那里视频画面(202)是用一个或多个检测对象(206)程序处理的, 从而导致一个或多个对象被识别并且随后被跟踪。保留的成份代表视频数据的中间形式。然后, 可以使用对于现有的视频处理方法通常不可得的技术给这个中间数据编码。因为该中间数据以几种形式存在, 所以标准的视频编码技术也能用来给这些中间形式中的一些编码。对于每个例证, 本发明都先确定然后使用最有效的编码技术。

[0044] 在一个优选实施方案中,特征分析程序完成显著信号模式的检测和分类。这个程序的一个实施方案使用专门为产生强度与在视频画面中检测到的对象特征有关的响应信号而设计的空间过滤器的组合。该分类程序是以不同的空间刻度应用于视频画面的不同位置的。来自分类程序的响应的强度指出显著信号模式出现的可能性。在把中心置于十分显著的对象上的时候,该程序用对应的强烈响应给它分类。显著信号模式的检测通过激活对视频序列中的显著信息的后续处理和分析来辨别本发明。

[0045] 给出显著信号模式在一个或多个视频画面中的检测位置,本发明分析显著信号模式的无变化特征。此外,对于无变化的特征,本发明分析该信号的残值,“较少显著的”信号模式。无变化特征的识别提供用来减少多余信息和拆分(即,分隔)信号模式的基础。

[0046] 特征点跟踪

[0047] 在本发明一个实施方案中,在一个或多个画面中的空间位置是通过空间强度场梯度分析确定的。这些特征对应于“线”的一些交点,这些交点能被宽松地描述为“拐角”。这样的实施方案进一步选择一组这样的拐角,这些拐角是强壮的而且在空间上是彼此完全不同的,在此称之为特征点。此外,使用光学流的分层次的多分辨率评估允许确定随着时间流逝特征点的平移位移。

[0048] 在图 2 中,跟踪对象(220)程序是为了把来自检测对象程序(208)的检测例证拉到一起和进一步识别一个或多个被检测对象的特征在许多视频画面(202 和 204)上的对应关系。

[0049] 特征跟踪的非限制性实施方案能被这样使用,以致这些特征被用来限定更规则的梯度分析法(例如,基于区段的运动评估)。

[0050] 另一个实施方案期待以特征跟踪为基础的运动评估的预测。基于对象的检测和跟踪

[0051] 在本发明的一个非限制性实施方案中,强健的对象分类程序被用来跟踪视频画面中的面部。这样的分类程序以已经在那些面部上训练过的定向边缘的级联响应为基础。在这个分类程序中,边缘被定义为一组基本的 Haar 特征和那些特征的 45 度旋转。级联分类程序是 AdaBoost 算法的变体。此外,响应计算能通过使用总面积表优化。

[0052] 局部配准

[0053] 配准包括在两个或多个视频画面中被识别对象的元素之间的对应关系的分配。这些对应关系变成建立视频数据中时间点截然不同的视频数据之间的空间关系模型的基础。

[0054] 为了根据广为人知的算法和那些算法的富有创造性的派生算法举例说明特定的实施方案和与它们相关联的实践缩减量,现在描述用于本发明的各种不同的非限制性的配准方法。

[0055] 在时空序列中建立明显的光学流模型的一种方法可以通过从视频数据的两个或多个画面产生有限差分域实现的。如果对应关系在空间和强度双重意义上符合特定的恒定不变的限制,光学流域能被稀疏地评估。

[0056] 菱形搜寻

[0057] 假定把视频画面分割成若干不重叠的区段,搜寻先前的与每个区段匹配的视频画面。以全面搜寻区段为基础(FSBB)的运动评估找出与当前画面中的区段相比较时在早先的视频画面中误差最小的位置。完成 FSBB 可能是计算费用十分浩大的,而且往往不产生比

以局域化运动假设为基础的其它评估方案更好的匹配。以菱形搜寻区段为基础 (DSBB) 的梯度下降运动评估是 FSBB 的常见的替代品,它使用各种不同尺寸的菱形搜寻图案朝着对于某个区段最好的匹配的方向反复地横越误差梯度。

[0058] 在本发明的一个实施方案中,为了产生数值稍后被因式分解成高阶运动模型的有限差分,DSBB 被用于一个或多个视频画面之间的图像梯度域分析。

[0059] 熟悉这项技术的人知道基于区段的运动评估能被视为规则网孔顶点分析的同价物。

[0060] 基于相位的运动评估

[0061] 在现有技术中,基于区段的运动评估通常是作为导致一个或多个空间匹配的空间搜寻实现的。基于相位的规范化的互相关 (PNCC) 如同图 3 举例说明的那样把来自当前画面和先前画面的区段变换到“相空间”中,并且寻找那两个区段的互相关。这种互相关被表达为位置与两个区段之间的边缘的“相移”相对应的数值域。这些位置通过定阈值被隔离,然后被逆变换成空间坐标。这些空间坐标是截然不同的边缘位移,而且对应于运动矢量。

[0062] PNCC 的优势包括反差掩蔽,该反差掩蔽在视频流中预留增益 / 曝光调节的容许偏差。另外,PNCC 允许来自单一步骤的结果,该单一步骤可能处理来自基于空间的运动评估程序的许多迭代。此外,该运动评估是子像素精确的。

[0063] 本发明的一个实施方案在一个或多个视频画面之间的图像梯度域的分析中利用 PNCC,为的是产生其数值稍后被因式分解成高阶运动模型的有限差分。

[0064] 整体配准

[0065] 在一个实施方案中,本发明将来自有限差分评估的域的一个或多个线性模型因式分解。发生这样的抽样的域在此被称为有限差分的一般总体。所描述的方法使用与 RANSAC 算法类似的强健的评估。

[0066] 如图 4 所示,在建立整体运动模型的情况下,有限差分是集中在通过那些运动评估的随机抽样 (410) 迭代处理的一般总体库 (404) 之中的平移运动评估 (402),而且线形模型被因式分解,提取那些样本的公因子 (420)。然后,那些结果被用来调节总体 (404) 以便通过排除该模型的异己样本更好地阐明该线性模型,如同通过随机处理发现的那样。

[0067] 在线性模型评估算法的一个实施方案中,运动模型评估程序以线性最小二乘解为基础。这种相关性使该评估程序摆脱异己样本数据。基于 RANSAC,所揭示的方法是一种通过反复评估数据子集抵消异己样本的效应探查将描述重要的数据子集的运动模型的强健方法。每个探头产生的模型都对它所代表的的数据百分比进行测试。如果有足够的迭代次数,则将发现与最大的数据子集拟合的模型。

[0068] 如同图 4 设想和举例说明的那样,本发明揭示一些在算法变更形式上超过 RANSAC 算法的改革,包括有限差分的初始抽样 (样本) 和线性模型的最小二乘评估。综合误差是使用已解的线性模型对一般总体中的所有样本评估的。根据残差符合预先设定的阈值的样本的数目把一个等级分配给该线性模型。这个等级被看作是“候选的共识”。

[0069] 初始抽样、求解和归类是通过迭代完成的,直到终止判据得到满足为止。一旦该判据得到满足,等级最高的线性模型被看作是总体的最后共识。

[0070] 非必选的改进步骤包括按照与候选模型拟合最好的次序反复地分析样本的子集和逐渐增加该子集的大小,直到再加一个样本将会超过整个子集的残留误差阈值。

[0071] 为了在将与某特定的线性模型相对应的另一个参数矢量空间中确定子空间的拓扑空间,所描述的本发明的非限制性实施方案可以作为对矢量空间(前面被描述为有限差分矢量的域)抽样的一般方法进一步推广。

[0072] 整体配准程序的进一步的结果是这个配准程序和局部配准程序之间的差异产生局部配准残差。这个残差是整体模型在近似局部模型时的误差。

[0073] 规范化

[0074] 规范化指的是朝着标准的或通常的空间配置方向再次抽取空间强度场样本。当这些相关的空间配置是这样的配置之间可逆的空间变换的时候,象素的再次抽样和附带插值也是直到拓扑极限可逆的。本发明的规范化方法是用图 5 举例说明的。

[0075] 当两个以上空间强度场被规范化的时候,提高的计算效率可以通过保存中间的规范化计算结果来实现。

[0076] 为了配准的目的,或等效地为了规范化,用来再次抽取图像样本的空间变换模型包括总体模型和局部模型。总体模型有从平移变换到影射变换逐渐增加的阶次。局部模型是有限差分,该有限差分暗示在基本上用区段或更复杂地用分段线性网孔确定的关于邻近象素的内插式。

[0077] 原始强度场向规范化强度场的插值增加基于强度场子集的 PCA 外观模型的直线性。

[0078] 如图 2 所示,对象象素(232 和 234)能被再次抽样(240)以便得到所述对象象素的规范化版本(242 和 244)。

[0079] 基于网孔的规范化

[0080] 本发明进一步的实施方案把特征点镶嵌到基于三角形的网孔中,跟踪该网孔的顶点,并且使用每个三角形的顶点的相对位置来评估与那三个顶点一致的平面的三维表面法线。当该表面法线与摄影机的投影轴相符的时候,成像象素能提供与该三角形相对应的对象的扭曲最小的透视图。创造倾向于支持正交表面法线的规范化图像能产生保存中间数据类型的象素,这将提高后来以外观为基础的 PCA 模型的直线性。

[0081] 另一个实施方案利用传统的以区段为基础的运动评估来含蓄地建立整体运动模型。在一个非限制性实施方案中,该方法将来自传统的以区段为基础的运动评估/预测所描述的运动矢量的整体仿射运动模型因式分解。

[0082] 图 9 举例说明整体和局部规范化的组合方法。

[0083] 渐进的几何规范化

[0084] 空间间断点的分类被用来对准镶嵌的网孔,以便在它们与网孔边缘一致的时候含蓄地建立间断点模型。

[0085] 同种区域的边界是用多角形轮廓近似的。为了确定每个多角形顶点的显著优先权,该轮廓是以逐次降低的精度逐次近似的。为了保护共享顶点的顶点优先权,顶点优先权在各个区域上传播。

[0086] 在这项发明的一个实施方案中,多角形分解方法允许与视场的同种分类相关联的边界的优先排序。象素是依照一些同种标准(例如,光谱相似性)分类的,然后把分类标签按空间连接到各个区域之中。在进一步优选的非限制性实施方案中,4-或 8-连通性判据被用来确定空间连通性。

[0087] 在优选的实施方案中,这些空间区域的边界随后被离散成多角形。所有多角形对所有同种区域的空间覆盖呈棋盘格状并且结合在一起形成初步的网孔。使用一些判据将这种网孔的顶点分解,以揭示保有最初网孔的大多数知觉特征的较简单的网孔表达。

[0088] 在优选的实施方案中,图像配准方法与这份说明书的另一部分揭示的一样用强壮的图像梯度向这些高优先权顶点偏置。由此产生的变形模型倾向于保护与成像对象的几何形状相关联的空间间断点。

[0089] 在优选的实施方案中,活跃的轮廓用来改善区域边界。每个多角形区域的活跃轮廓都允许增殖一次迭代。在不同的区域中每个活跃轮廓顶点的“变形”或移动是在计算平均值操作中结合的,以便考虑到隐式网孔受限制的增殖,对于该网孔它们有隶属关系。

[0090] 在优选的实施方案中,顶点被分配在适合也作为不同区域的轮廓部分的毗邻顶点的网孔中它有的毗邻顶点数的计数。这些其它的顶点被定义为处在对立状态。如果顶点计数为 1,则它有没有对立顶点,因此需要得到保护。如果两个毗邻的对立顶点的计数都为 1(意味着这两个顶点在不同的多角形中而且彼此相邻),那么一个顶点对另一个是可分辨的。当计数为 1 的顶点与数值为 2 的邻近的多角形顶点对立的时候,计数为 1 的顶点被转化为计数为 2 的顶点,而且那个顶点的计数等于 1。因此,如果出现另一个邻近的对立顶点,那么这个顶点能被再一次分辨。对于这种情况,保留最初的顶点计数是重要的,所以在分辨顶点的时候,我们能基于最初的顶点计数偏置求解方向。这是为了顶点 a 变得对顶点 b 清晰可见,那么顶点 b 对顶点 c 将不清晰可见,而顶点 c 应该对顶点 b 变得清晰可见,因为 b 已经被用于一种分辨率。

[0091] 在优选的实施方案中,T- 接合点被明确地处理。这些是在毗邻的多角形中没有点的多角形中的点。在这种情况下,每个多角形顶点都首先被画在图像点映射图上,这张映射图识别顶点的空间位置及其多角形标识符。然后横越和测试每个多角形的周长看看是否有任何来自另一个多角形的毗邻顶点。如果有来自另一个区域的邻近顶点,那么它们每个都被测试,看看它们是否已经有来自当前的多角形的邻近顶点。如果它们没有,那么当前的点作为当前的多角形的顶点被添加进去。这种额外的测试保证在另一个多角形中的孤立顶点被用来产生 T- 接合点。否则,这将在这个区域已经有匹配顶点的情况下仅仅添加新的顶点。所以,只有当邻近的顶点不与这个当前区域对立的时候才添加对立顶点。在进一步的实施方案中,通过使用掩模图像增加检测 T- 联接的效率。连续地访问多角形顶点,而且这样更新掩模,以致顶点的像素被确认为属于某个多角形顶点。然后多角形周长的像素被详细研究,如果它们与多角形顶点一致,那么它们被记录为在当前的多角形之内的顶点。

[0092] 在优选的实施方案中,当一个光谱的区域已经被一个或多个交叠的同种图像梯度区域再映射,而且另一个同种光谱区域也重叠的时候,先前被再映射的区域全被赋予与当前被再映射的那些区域相同的标签。因此基本上,如果光谱区域被两个同种区域遮住,那么所有被那两个同种区域遮住的光谱区域都将获得同样的标签,因此一个光谱区域真的被一个同种区域而不是两个同种区域覆盖是相似的。

[0093] 在本发明的一个实施方案中,为了找到邻接归并判据,处理区域映射图并非区域目录是有利的。在进一步的实施方案中,光谱拆分分类器能被修正以便训练该分类器使用非同种区域。这允许将处理集中在光谱区域的边缘。此外,增加以使用边缘(例如,稳定的边缘检测器)为基础的不同的拆分并且把那个馈送给活跃的轮廓识别最初的那组多角形

将考虑到同种区域的较大差别。

[0094] 局部规范化

[0095] 本发明提供能以“局部”方式完成像素在时空流中配准的方法。

[0096] 一种这样的局域化方法使用几何网孔的空间应用提供分析像素的方法,以致在成像现象中局域相干性在分辨与成像现象(或明确地说成像对象)的局部变形有关的表观图像亮度恒定性模棱两可的时候得到解释。

[0097] 这样的网孔被用来提供在像平面中表面变形的分段线性模型作为局部规范化的方法。当视像流的时间分辨率与视像中的运动相比高的时候,成像现象可能往往与这样的模型相对应。模型假设之例外是通过多种技术处理的,包括:像素和图像梯度区域的拓扑限制、邻近顶点限制和同种分析。

[0098] 在一个实施方案中,特征点用来产生由顶点与特征点对应的三角形元素构成的网孔。对应的特征点是其它画面暗示三角形及其对应像素的内插造成的“变形”产生局部变形模型。

[0099] 图 7 举例说明这样的对象网孔的产生。图 8 举例说明使用这样的对象网孔局部地规范化画面。

[0100] 在一个优选的实施方案中,产生一幅识别三角形的三角形映射图,其中所述映射图的每个像素都来自所述三角形。此外,与每个三角形相对应的仿射变换是作为优化步骤预先计算的。再者,在产生局部变形模型的时候,使用空间坐标在固定图像(先前的)上来回移动以确定来源像素的抽样坐标。这个被抽样的像素将代替当前像素位置。

[0101] 在另一个实施方案中,局部变形是在整体变形之后预先形成的。在先前揭示的说明中,整体规范化是作为使用整体配准方法从空间上规范化两幅或多幅视频画面中的像素的程序描述的。由此产生的整体规范化的视频画面能被进一步局部规范化。这两种方法的组合把局部规范化限制在整体上得到的解决办法的细分方面。这能大大减少求解所需要的局部方法的不明确性。

[0102] 在另一个非限制性实施方案中,特征点或“规则网孔”情况下的顶点是通过分析那些点邻近区域的图像梯度限定的。这个图像梯度能直接地或通过一些间接计算(例如,Harris 响应)被计算出来。此外,这些点能被用与图像梯度下降相关联的空间限制和运动评估结果误差过滤。合格的点能作为网孔的基础被许多棋盘格化技术之一使用,从而导致其元素是三角形的网孔。对于每个三角形,基于那些点和它们残留的运动矢量产生一个仿射模型。

[0103] 在优选的实施方案中,维持三角形仿射参数的目录。这个目录通过迭代构成当前的/早先的点目录(使用顶点查寻映射图)。当前的/早先的点目录被传送给用来评估为那个三角形计算仿射参数的变换的例行程序。然后,这些仿射参数或模型被保存在三角形仿射参数目录中。

[0104] 在进一步的实施方案中,该方法横移三角形标识符图像映射图,在这种情况下该映射图中的每个像素包含在该像素有隶属关系的网孔中的三角形的标识符。而且对于属于某个三角形的每个像素,计算适合那个像素的对应的整体变形和局部变形坐标。那些坐标依次用来完成对应像素的抽样并且把它的数值用在对应者的“规范化”位置。

[0105] 在进一步的实施方案中,以起因于图像梯度搜寻的密度和图像强度对应关系严格

性为基础把空间限制应用于那些点。在基于某种图像强度残差基准完成运动评估结果之后将那些点分类。然后,以空间密度限制为基础对这些点进行过滤。

[0106] 在进一步的实施方案中,使用空间光谱拆分,而且把小的同种光谱区域基于空间亲和力(它们的强度和/或颜色与邻近区域的相似性)合并。然后,使用同种合并把光谱区域以它们与同种质地(图像梯度)区域的重叠为基础组合在一起。进一步的实施方案然后使用中心周围点(那些点是被较大的区域包围的小区域)作为合格的感兴趣的点来支持网孔的顶点。在进一步的非限制性实施方案中,中心周围点的定义为其边界框在尺寸为 3×3 或 5×5 或 7×7 像素的一个像素之内而且对于那个边界框空间图像梯度是角落形状的区域。该区域的中心能被归类为角落,从而进一步限定那个位置为有利的顶点位置。

[0107] 在进一步的实施方案中,水平和垂直的像素有限差分图像被用来分类每个网孔边缘的强度。如果边缘有许多与它的空间位置一致的有限差分,那么该边缘和那个边缘的顶点被认为是对于成像现象的局部变形非常重要的。如果在边缘的有限差分之和的平均值之间有大的派生差异,那么该区域边缘通常很可能对应于质地变化边缘,而不是量化步骤。

[0108] 在进一步的实施方案中,空间密度模型终止条件被用来优化网孔顶点的处理。当检查过数目足以覆盖大部分检测矩形始端的空间区域的点的时候,于是可以结束该处理。终止产生得分。进入处理的顶点和特征点用这个得分来分类。如果那个点与现有的点在空间上挨得太近,或者那个点不与图像梯度的边缘相对应,则将它丢弃。否则,在那个点的邻近地区中的图像梯度下降,而且如果梯度的残差超过某个界限,那么那个点也被丢弃。

[0109] 规则网孔规范化

[0110] 本发明利用规则网孔扩展上述的局部规范化方法。这种网孔是不考虑潜在像素构成的,然而它的位置和尺寸与被检测对象相对应。

[0111] 给定被检测对象区域,空间画面位置和指出面部大小的刻度在面部区域的始端上产生规则网孔。在优选的实施方案中,使用一组不重叠的瓦片描绘矩形网孔,然后完成瓦片的对角线分割产生有三角形网孔元素的规则网孔。在进一步的优选实施方案中,瓦片与用于传统的视频压缩算法(例如,MPEG-4AVC)的那些成比例。

[0112] 在优选的实施方案中,与上述网孔相关联的顶点通过分析在用于训练的特定的视频画面中包围这些顶点的像素区域区分优先次序。分析这样的区域的梯度提供关于与每个顶点相关的将依靠局部图像梯度的处理(例如,基于区段的运动评估结果)的置信度。

[0113] 顶点位置在多个画面的对应关系是通过简单的逐步降低图像梯度找到的。在优选实施方案中,这是通过基于区段的运动评估实现的。在目前的实施方案中,高置信度的顶点考虑到高置信度的对应关系。置信度较低的顶点对应关系是通过推理经过求解不明确的图像梯度从置信度较高的顶点对应关系获得的。

[0114] 在一个优选实施方案中,规则网孔是在最初的跟踪矩形上制作的。产生 16×16 的瓦片,并且沿着对角线切割,形成三角形网孔。对这些三角形的顶点进行运动评估。运动评估结果取决于每个点的质地类型。质地被分为三类:角落、边缘和同种,它们也定义顶点的处理次序。角落顶点使用邻近顶点的评估结果,即,邻近点(如果可得)的运动评估被用于预言性运动矢量,而运动评估结果适用于每一个。提供最低的疯狂误差的运动矢量是作为这个顶点运动矢量使用的。用于角落的搜寻策略是所有的(宽的、小的和原点)。对于边缘,再一次使用最近的相邻运动矢量作为预言性运动矢量,而且使用误差最小的那一个。边

缘的搜寻策略是小的和原点。对于同种区域,搜寻邻近的顶点并且使用误差最小的运动评估。

[0115] 在一个优选实施方案中,每个三角形顶点的图像梯度被计算出来,而且基于类别和大小被分类。所以,角落先于边缘,边缘先于同种区域。对于角落,强的角落先于弱的角落,对于边缘,强的边缘先于弱的边缘。

[0116] 在一个优选实施方案中,每个三角形的局部变形以与那个三角形相关联的运动评估为基础。每个三角形都有对它评估的仿射。如果三角形不作拓扑逆转,或变成退化的,那么作为三角形部分的像素被用来以获得的评估仿射为基础抽取当前图像的样本。

[0117] 拆分

[0118] 通过进一步描述的拆分程序识别的空间间断点是通过它们各自边界的几何参数表达法(被称为空间间断点模型)有效地编码的。这些空间间断点模型可以不断地考虑到与编码子集相对应的更简洁的边界描述的渐进方式编码。渐进式编码提供一种在保留空间间断点的许多显著方面的同时区分空间几何学优先次序的强健方法。

[0119] 本发明的优选实施方案把多分辨率拆分分析和空间强度梯度域分析结合起来并且进一步使用时间稳定性限制,为的是实现强健的拆分。

[0120] 如图 2 所示,一旦已经随着时间的流逝跟踪对象的特征的对对应关系(220)并且建立了模型(224),遵守这个运动/变形模型能用来拆分与那个对象相对应的像素(230)。可以对画面(202 和 204)中已探测到的许多对象(206 和 208)重复这个程序。

[0121] 本发明使用的无变化特征分析的一种形式被集中在空间间断点的识别上。这些间断点是作为边缘、阴影、遮蔽、线、拐角或任何其它的在一个或多个视频成像画面中引起像素之间突然的可辨认的分离的可见特征出现的。此外,在颜色和/或纹理类似的对象之间的细微的空间间断点可能仅仅出现在视频画面中各个对象的像素相对于那些对象本身正在经历粘附运动而相对于其它对象正在经历不同的运动之时。本发明利用频谱拆分、纹理拆分和运动拆分的组合强健地识别与显著信号模式有关的空间间断点。

[0122] 时间拆分

[0123] 把平移运动矢量或在空间强度场中等价的有限差分测量结果按时间整合成高阶运动模型是现有技术描述的一种运动拆分形式。

[0124] 在本发明的一个实施方案中,产生运动矢量的稠密域,表现视频画面中对象运动的有限差分。这些导数是通过规则地分割瓦片或借助某种初始化程序(例如,空间拆分)按空间集合的。每个集合的“导数”使用线性最小二乘评估程序整合成一个高阶运动模型。然后,由此产生的运动模型作为矢量在运动模型空间中使用 k-means 群集技术群集。这些导数是基于与它们拟合最好的群分类的。然后,群标是作为空间分割的演化按空间群集的。该程序一直继续到空间分割稳定为止。

[0125] 在本发明的进一步的实施方案中,适合给定的孔径的运动矢量被内插到一组与该孔径相对应的像素位置。当用这种内插定义的区段横越与对象边界相对应的像素时候,由此产生的分类是该区段的某种不规则的对角线分割。

[0126] 在现有技术中,用来整合导数的最小二乘评估程序对离群值是非常敏感的。这种敏感性能产生使运动模型的群集方法严重地向迭代结果大大发散的点倾斜的运动模型。

[0127] 在本发明中,运动拆分方法通过分析两个以上视频画面上明显的像素运动识别空

间间断点。明显的运动是针对这些视频画面上的一致性分析的并且被整合成参数运动模型。与这种一致的运动相关联的空间间断点被识别出来。运动拆分也可以被称为时间拆分，因为时间变化可能是由运动引起的。然而，时间变化也可能是由一些其它的现象（例如，局部变形、照明变化，等等）引起的。

[0128] 通过所描述的方法，与规范化方法相对应的显著信号模式能被识别而且能通过几种背景减法之一与环境信号模式（背景或非对象）分开。时常，这些方法从统计上建立背景模型，因为像素在每个时间样本都呈现最小的变化量。变化能被视为像素数值差异。作为替代，运动拆分能在给出显著图像模式的探测位置和数值范围的情况下实现。距离变换能用来确定每个像素距探测位置的距离。如果与最大距离相关联的像素数值被保留，合理的背景模型能被求解。换句话说，环境信号能使用信号差异度量标准按时间再次抽样。

[0129] 给出环境信号的模型，就能按每个时间样本使完全的显著信号模式有差别。这些差别每个都能通过再次抽样变成空间规范化的信号差异（绝对差异）。然后，这些差异相互对准和累积。由于这些差异相对于显著信号模式已按空间规范化，所以，差异的峰通常将对应于与显著信号模式相关联的像素位置。

[0130] 非对象的分辨率

[0131] 给出清晰的背景图像，这个图像和当前画面之间的误差可以按空间规范化和按时间累积。这样的清晰背景图像是在“分辨率”部分中描述的。

[0132] 然后，由此产生的累积误差通过阈值检验提供初始轮廓。然后，该轮廓在空间上被扩展以使残留误差与轮廓变形平衡。

[0133] 梯度拆分

[0134] 纹理拆分方法或同义的强度梯度拆分分析像素在一个或多个视频画面中的局部梯度。梯度响应是一种表征空间间断点的统计尺度，其中所述空间间断点对于该视频画面中的像素位置是局部的。然后，使用几种空间群集技术之一把这些梯度响应组合成一些空间区域。这些区域的边界在识别一个或多个视频画面中的空间间断点方面是有用的。

[0135] 在本发明的一个实施方案中，来自计算机图形纹理生成的总面积表概念被用于加快强度场梯度计算的目的。累加值域的产生使通过与四次加法运算结合的四次查询计算任何长方形原始域的总和变得容易。

[0136] 进一步的实施方案使用对图像产生的 Harris 响应，而每个像素的邻近区域被归类为同种的、边缘或拐角。响应数值是依据这个信息产生的并且指出画面中每种元素的边缘化或拐角化的程度。

[0137] 多刻度梯度分析

[0138] 本发明的实施方案通过以几种空间刻度产生图像梯度值进一步约束图像梯度支持。这个方法能帮助限定图像梯度的资格，以致在不同刻度下的空间间断点能用来彼此相互支持，只要“边缘”在几种不同的空间刻度下能被分辨，该边缘应该是“显著的”。更有资格的图像梯度将倾向于与更显著的特征相对应。

[0139] 在优选实施方案中，纹理响应区域是首先产生的，然后，这个区域的数值以 k-means 分区间 / 分割为基础被量化成若干区间。然后，使用每个区间作为单一迭代能把分水岭拆分能应用于它的数值间隔渐进地处理最初的图像梯度数值。这种方法的好处是同种是在相对意义上用强烈的空间偏置定义的。

[0140] 光谱拆分

[0141] 光谱拆分方法分析视频信号中黑白像素、灰度像素或彩色像素的统计概率分布。频谱分类程序是通过完成关于那些像素的概率分布的群集操作构成的。然后,使用该分类程序把一个或多个像素分类,使之属于某个概率类别。然后,由此产生的概率类别和它的像素被赋予类别标签。然后,使这些类别标签在空间上合并成有截然不同的边界的像素区域。这些边界识别在一个或多个视频画面中的空间间断点。

[0142] 本发明可以利用基于光谱分类的空间拆分来拆分视频画面中的像素。此外,各个区域之间的对应关系可以是基于各个光谱区域与先前拆分的区域的重叠确定的。

[0143] 业已观察到当视频画面大体上由空间上被连接成与视频画面中的对象相对应的较大区域的连续彩色区域组成的时候,彩色(或光谱)区域的识别和跟踪能促进图像序列中对象的后续拆分。

[0144] 背景拆分

[0145] 本发明包括以每幅视频画面中的探测对象和每个个别像素之间的空间距离测量结果的瞬时最大值为基础建立视频画面背景模型的方法。给定探测到的对象位置,应用距离变换,产生适合画面中每个像素的标量距离数值。在所有的视频画面上每个像素的最大距离的映射图被保留。当最初分配最大数值的时候,或后来用不同的新数值更新该最大数值的时候,适合于那幅视频画面的对应的像素被保留在“清晰的背景”画面中。

[0146] 建立外观模型

[0147] 视频处理的共同目标往往是建立模型和保存视频画面序列的外观。本发明以允许通过预处理的运用以强健的和广泛适用的方式应用强制性外观建模技术为目标。先前描述的配准、拆分和规范化明显地适合这个目的。

[0148] 本发明揭示建立外观变化模型的方法。建立外观变化模型的主要基础在线性模型的情况下是分析特征矢量,以揭示开发利用线性相关关系的坚实基础。表达空间强度场像素的特征矢量能被组装成外观变化模型。

[0149] 在替代实施方案中,外观变化模型是依据被拆分的像素子集计算的。此外,该特征矢量能被分成若干空间上不重叠的特征矢量。这样的空间分解可以用空间铺瓦来实现。计算效率可以通过处理这些临时总体来实现,而不牺牲更普遍的 PCA 方法的维数减少。

[0150] 在产生外观变化模型时,空间强度场规范化能用来减少空间变换的 PCA 建模。

[0151] PCA

[0152] 产生外观变化模型的优选方法是通过把视频画面作为图案矢量组装成一个训练矩阵或总体然后把主要成份分析(PCA)应用在该训练矩阵上。当这样的展开式被截取的时候,由此产生的 PCA 变换矩阵被用来分析和合成后面的视频画面。基于截取水平,改变像素的初始外观质量水平能实现。

[0153] 图案矢量的特定的构成和分解方法对于熟悉这项技术的人是广为人知的。

[0154] 给出来自环境信号的显著信号模式的空间拆分和这个模式的空间规范化,像素本身或同义的由此产生的规范化信号的外观能被因式分解成线性相关的成份,其中低级参数表达考虑到适合表达像素外观的近似值误差和比特率之间的直接交换。

[0155] 如图 2 所示,为了得到量纲上简明的数据版本(252 和 254),规范化的对象像素(242 和 244)能投射到矢量空间中而且线性对应关系能使用分解程序(250)建立模型。

[0156] 连续的 PCA

[0157] PCA 使用 PCA 变换把图案编码成 PCA 系数。用 PCA 变换表达的图案越好,给该图案编码所需要的系数就越少。承认图案矢量可能随着时间在获得训练图案和待编码图案之间流逝降级,更新变换能帮助抵消这种降级。作为产生新变换的替代品,现有图案的连续更新在特定的情况下是计算上更有效的。

[0158] 许多最新技术的视频压缩算法依据一个或多个其它画面预测某视频画面。预测模型通常基于把每个预测画面分割成与在另一画面中对应的补丁相匹配的不重叠的瓦片和相关联的用偏移运动矢量参数化的平移位移。这个非必选地与画面索引耦合的空间位移提供瓦片的“运动预测”版本。如果预测的误差在特定的阈值以下,则瓦片的像素适合残差编码;而且在压缩效率方面有对应的增益。否则,瓦片的像素被直接编码。这种基于瓦片的换句话说是基于区段的运动预测方法通过平移包含像素的瓦片建立影像模型。当影像中的成像现象坚持这种建模的时候,对应的编码效率增加。为了与在基于区段的预测中固有的平移假定一致,这个建模限制为了与在基于区段的预测中固有的平移假设一致假定特定的时间分辨率水平(或帧频)对于正在运动的成像对象是存在的。这种平移模型的另一个必要条件是对于特定的时间分辨率空间位移必须受到限制;换言之,用来推导预测结果的画面和被预测的画面之间的时间差必须是比较短的绝对时间。这些时间分辨率和运动限制使存在于视频流中的某些多余的视频信号成份的识别和建模变得容易。

[0159] 基于残差的分解

[0160] 在 MPEG 视频压缩中,当前的画面是通过先使用运动矢量对先前的画面进行运动补偿,然后把残差更新应用于那些补偿区段,最后将任何没有充份匹配的区段作为新区段完成编码构成的。

[0161] 对应于残留区段的像素通过运动矢量映射到先前画面的像素上。结果是像素通过能通过连续应用残值合成的影像的瞬时路径。这些像素被确认为能使用 PCA 最明确地表达的像素。

[0162] 基于遮挡的分解

[0163] 本发明的进一步提高确定适用于多个区段的运动矢量是否将导致来自先前画面的任何像素被移动像素遮挡(覆盖)。对于每个遮挡事件,都把遮挡像素劈成新层。没有历史的像素也将暴露出来。暴露出来的像素被放到任何将在当前画面中与它们拟合而且历史拟合也能在那层上完成的层上。

[0164] 像素的时间连续性是通过像素对不同层的接合和移植得到支持的。一旦获得稳定的层模型,每层中的像素就能基于对条理分明的运动模型的隶属关系编组。

[0165] 分波段时间量化

[0166] 本发明的替代实施方案使用离散余弦变换(DCT)或离散子波变换(DWT)把每个画面分解成分波段图像。然后,将主要成份分析(PCA)应用于这些“分波段”影像之中的每幅影像。概念是视频画面的分波段分解与原始视频画面相比较减少任何一个分波段中的空间变化。

[0167] 就移动对象(人)的影像而言,空间变化倾向于支配用 PCA 建模的变化。分波段分解减少任何一个分解影像中的空间变化。

[0168] 就 DCT 而言,任何一个分波段的分解系数都按空间安排在分波段影像之中。举例

来说,DC 系数是从每个区段获取的并且被安排在看起来像原始影像的邮票版本一样的分波段影像之中。这将对所有其它的分波段重复,而且使用 PCA 处理每个由此产生的分波段影像。

[0169] 就 DWT 而言,分波段已经按针对 DCT 描述的方式排列好。

[0170] 在非限制性实施方案中,PCA 系数的截取是变化的。

[0171] 子波

[0172] 当使用离散子波变换(DWT)分解数据的时候,多个带通数据组以较低的空间分辨率为结果。变换程序能被递归地应用于导出数据直到仅仅产生单一的标量数值为止。在已分解的结构中标量元素通常以分等级的父母/孩子方式相关。由此产生的数据包含多分辨率的分等级结构以及有限差分。

[0173] 当 DWT 被应用于空间强度场的时候,许多自然发生的图像现象由于空间频率低是用第一或第二低带通导出数据结构以微不足道的知觉损失表达的。截短该分等级结构在高频率空间数据不是不存在就是被视为噪音的时候提供简明的表达。

[0174] 尽管 PCA 可以用来以为数不多的系数实现精确的重建,但是这种变换本身可能是相当大的。为了减少这个“初始”变换的规模,可以使用子波分解的嵌零树(EZT)结构来建立变换矩阵的越来越精确的版本。

[0175] 子空间分类

[0176] 如同实践这项技术的人充分理解的那样,离散抽样的现象数据和导出数据能被表达成一组与代数矢量空间相对应的数据矢量。这些数据矢量以非限制性方式包括拆分后对象的规范化外表中的像素、运动参数和特征或顶点的任何二或三维结构位置。这些矢量都存在于矢量空间之中,而且该空间的几何分析能用来产生样本或参数矢量的简洁表达。有益的几何条件是借助形成紧凑子空间的参数矢量代表的。当一个或多个子空间混合,形成表面上更复杂的单一子空间的时候,那些要素子空间可能难以辨别。有几种拆分方法考虑到通过检查通过原始矢量的一些交互作用(例如,内积)产生的高维矢量空间中的数据分离这样的子空间。

[0177] 一种差分矢量空间的方法包括把矢量投射到表达多项式的 Veronese 矢量空间之中。这种方法在现有技术中是作为通用的 PCA 或 GPCA 技术广为人知的。通过这样的投射,多项式的法线被找到、聚集,而且与原始矢量相关联的那些法线能聚集在一起。这种技术的实用性的例子是把随着时间推移跟踪的二维空间点对应关系因式分解成三维结构模型和那个三维模型的运动。

[0178] GPCA 技术在作为明确定义的仅仅在以少许噪音产生数据矢量的时候易受影响的结果应用的时候是不完全的。现有技术假定管理程序使用者介入对 GPCA 算法的管理。这个限制大大限制该技术的潜能。

[0179] 本发明扩展了 GPCA 方法的概念基础,以便在有噪音和混合余维数存在时强健地处理多个子空间的识别和拆分。这种改革在技术状态上为该技术提供无人监督的改进。

[0180] 在现有技术中,GPCA 在 Veronese 映射图的多项式的法向矢量上操作,不考虑那些法向矢量的正切空间。本发明的方法扩充 GPCA,以便找到与通常在 Veronese 映射图中找到的法向矢量的空间正交的正切空间。然后使用这个“正切空间”或 Veronese 映射图的子空间把该 Veronese 映射图因式分解。

[0181] 正切空间是通过平面波膨胀和揭示几何对象（明确地说，Veronese 映射图的多项式的法线的切线）的表达的二元性的 Legendre 变换在位置坐标和正切平面坐标之间的应用识别的。离散的 Legendre 变换是通过凸分析应用于定义与法向矢量相对应的导数的约束形式。这种方法用来在有噪音存在的情况下通过算法向矢量拆分数据矢量。这个凸分析与 GPCA 合并提供一种比较强健的算法。

[0182] 本发明在应用 GPCA 的时候利用迭代的因子分解法。具体地说，在现有技术中发现的基于导数的落实被延伸到通过在此描述的同一 GPCA 方法细分分类数据矢量的总体。被重复应用，这项技术能用来强健地找出 Veronese 映射中的候选法向矢量，然后使用这种扩展的 GPCA 技术进一步限定那些矢量。就因子分解步骤而言，从原始数据组中除去与那组细分的矢量相关联的原始数据。剩余的数据组能用这种改进的 GPCA 技术分析。这种改进对于以无人监督的方式使用 GPCA 算法是至关重要的。图 11 举例说明数据矢量的递归细分。

[0183] 人们将进一步确认，本发明对 GPCA 技术的改进在 Veronese 多项式矢量空间中有多根的情况下有较大的优势。此外，当 Veronese 映射图的法线平行于矢量空间轴线之时，现有技术在遇到退化情形的时候，本发明的方法不会退化。

[0184] 图 10 举例说明基本的多项式拟合和求微分的方法。

[0185] 混合空间规范化压缩

[0186] 本发明通过把拆分视频流添加到“规范化”的视频流之中充分发挥以区段为基础的运动预测编码方案的效率。然后，这些视频流分开编码以允许传统的编码解码器的平移运动假设是有效的。在完成规范化视频流的解码之时，视频流解除规范化，进入它们适当的位置并且被组合在一起产生原始的视频序列。

[0187] 在一个实施方案中，一个或多个对象是在视频流中探测到的，而与探测到的每个个别对象有关的象素随后被拆分，离开非对象象素。接下来，针对对象象素和非对象象素产生整体空间运动模型。这个整体模型用来完成对象象素和非对象象素的空间规范化。这样的规范化已经有效地把非平移的运动从视频流中除去并且已经提供一组影像，这组影像的相互遮挡经被减到最少。这些是本发明的方法的两个有益的特征。

[0188] 象素已按空间规范化的对象和非对象的新影像是作为给传统的以区段为基础的压缩算法的输入提供的。在这些影像解码时，整体运动模型的参数被用来还原规范化的解码画面，对象象素一起合成到非对象象素之上，产生最初的视频流的近似。

[0189] 如图 6 所示，对于一个或多个对象（630 和 650）先前探测到的对象例证（206 和 208）每个都用传统视频压缩方法（632）的分开例证处理。此外，起因于对象的拆分（230）的非对象（602）也使用传统的视频压缩（632）压缩。这些分开的压缩编码（632）之中的每一个的结果是分开的传统编码流，每个编码流（634）分开地对应于每个视频流。在某个点，可能在传输之后，这些中间编码流（234）能被解压缩（636）成规范化的非对象（610）和许多对象（638 和 658）的合成物。这些合成后的象素能解除规范化（640），变成它们的已解除规范化的版本（622、642 和 662），把这些象素按空间相对于其它象素放置在正确的位置，以致合成程序（670）能把对象象素和非对象象素结合成完整的合成画面（672）。

[0190] 混合编码解码的整合

[0191] 在把传统的基于区段的压缩算法和本发明描述的规范化 - 拆分方案结合起来时，有一些已经产生结果的本发明的方法。首先，有专门的数据结构和必要的通信协议。

[0192] 主要的数据结构包括整体空间变形参数和对象拆分规范掩模。主要的通信协议是包括传输整体空间变形参数和对象拆分规范掩模的各个层面。

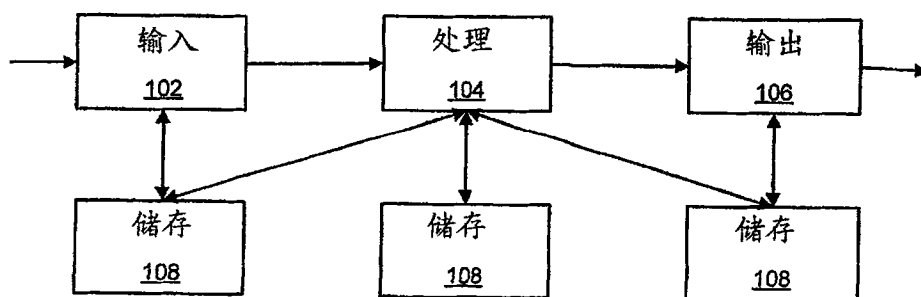


图 1 现有技术

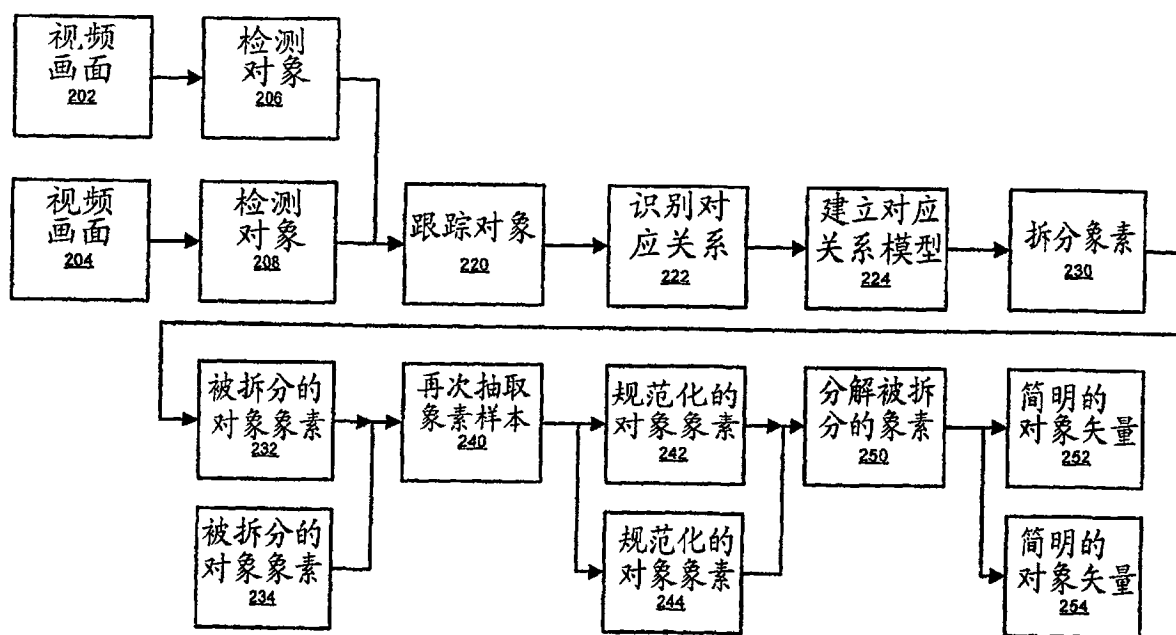


图 2

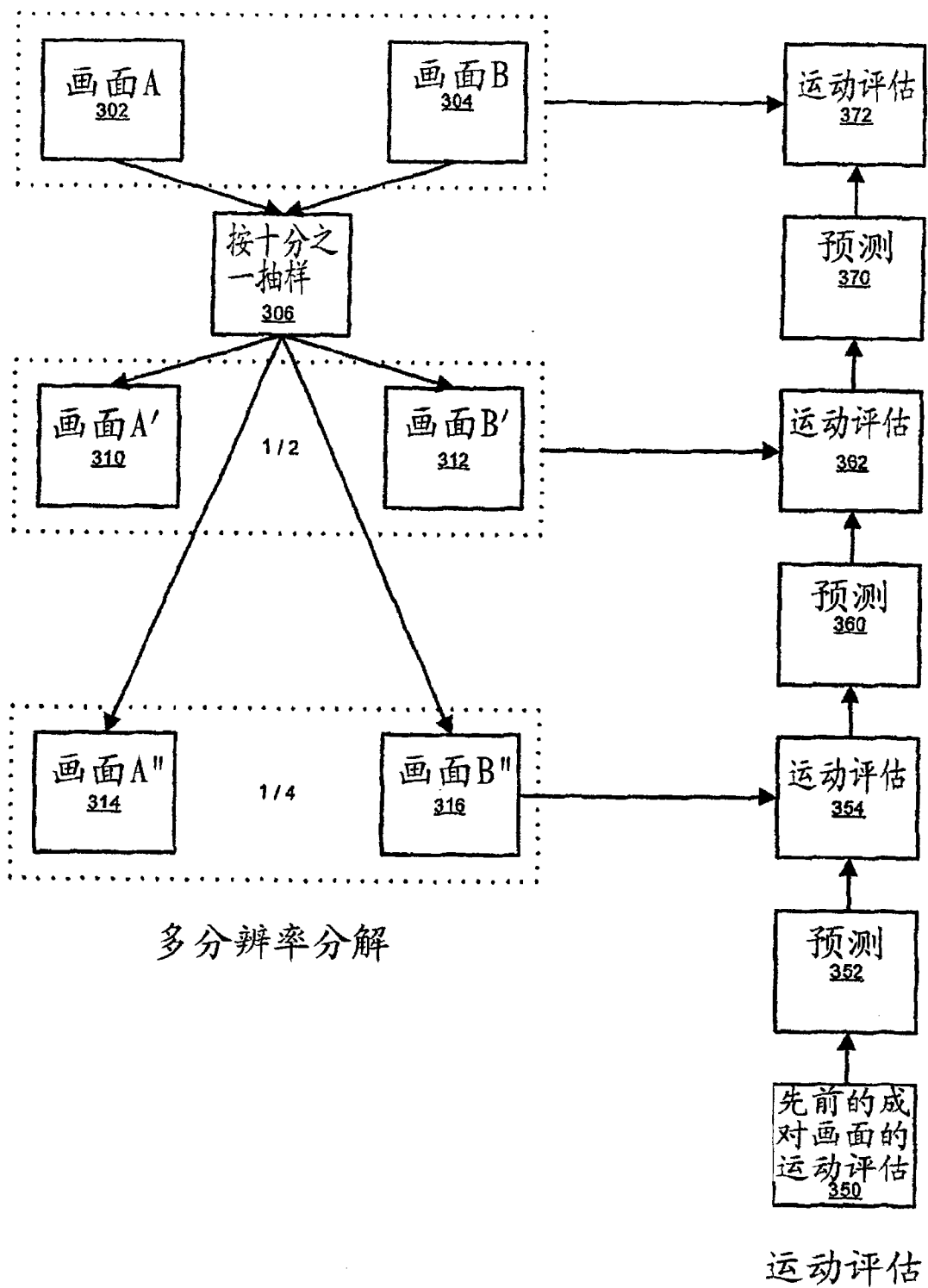
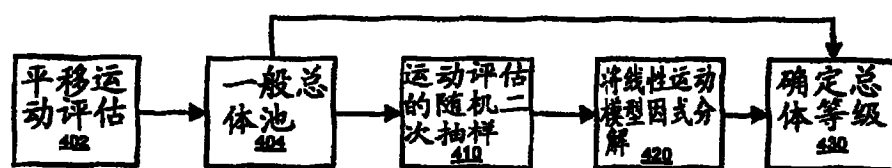
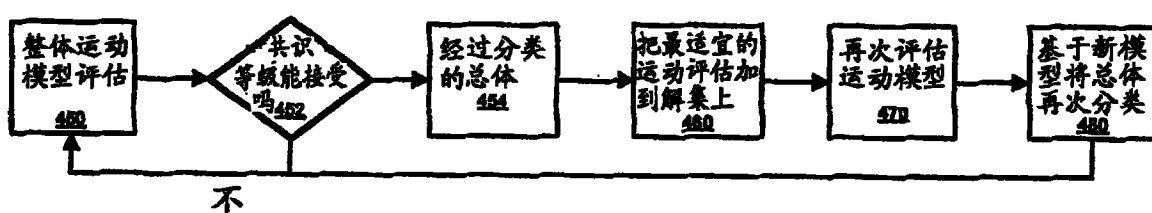


图 3



整体运动模型评估



共识等级验收测试和处理

图 4

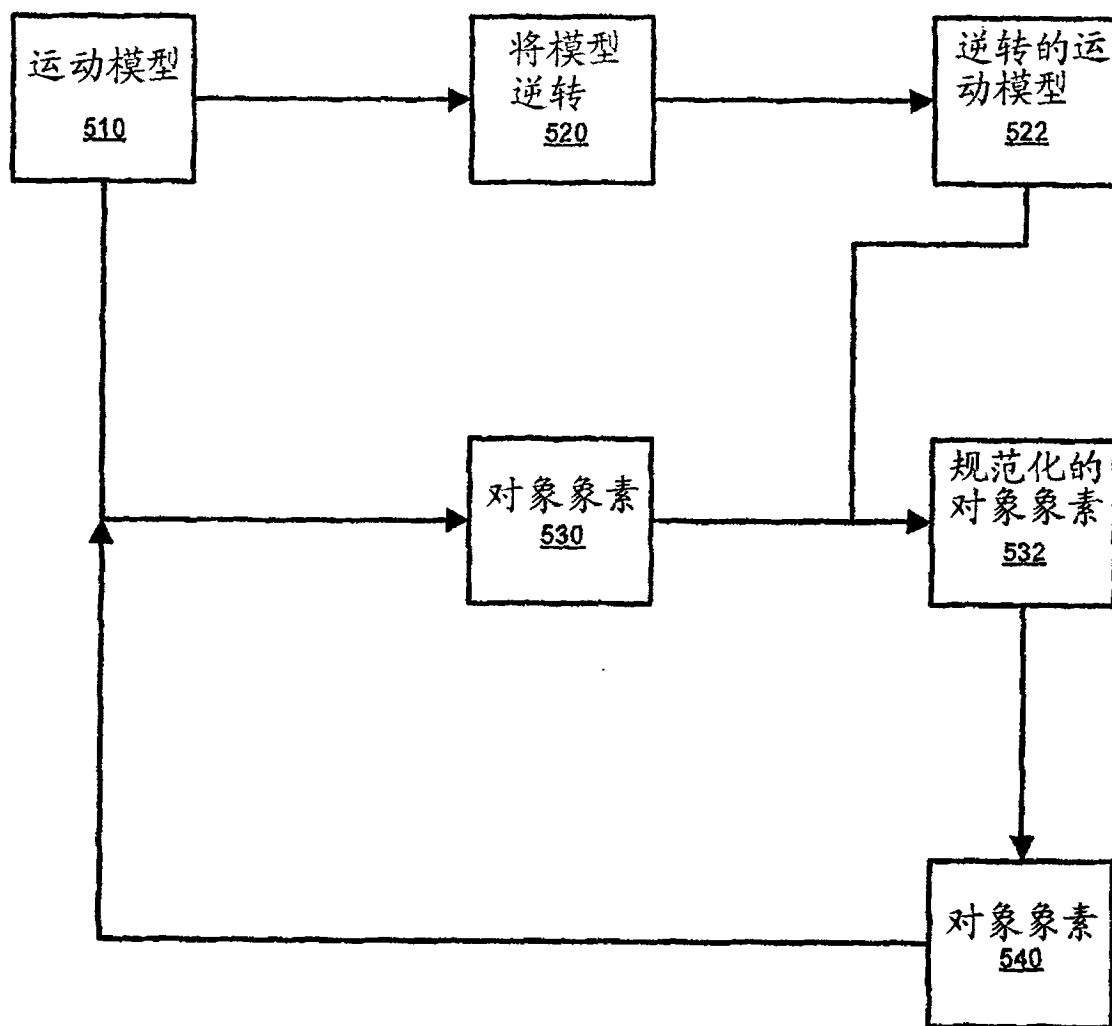


图 5

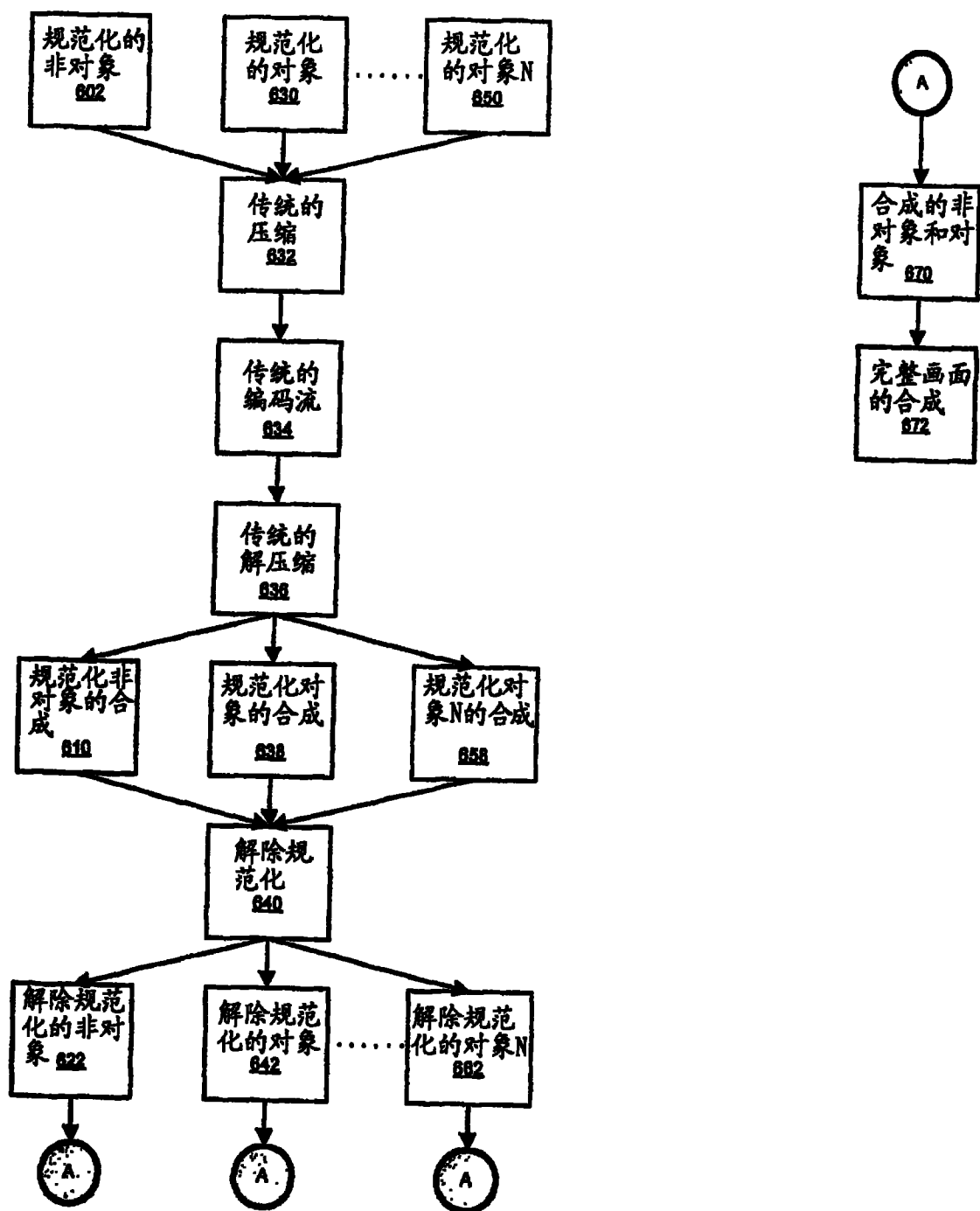


图6 混合式

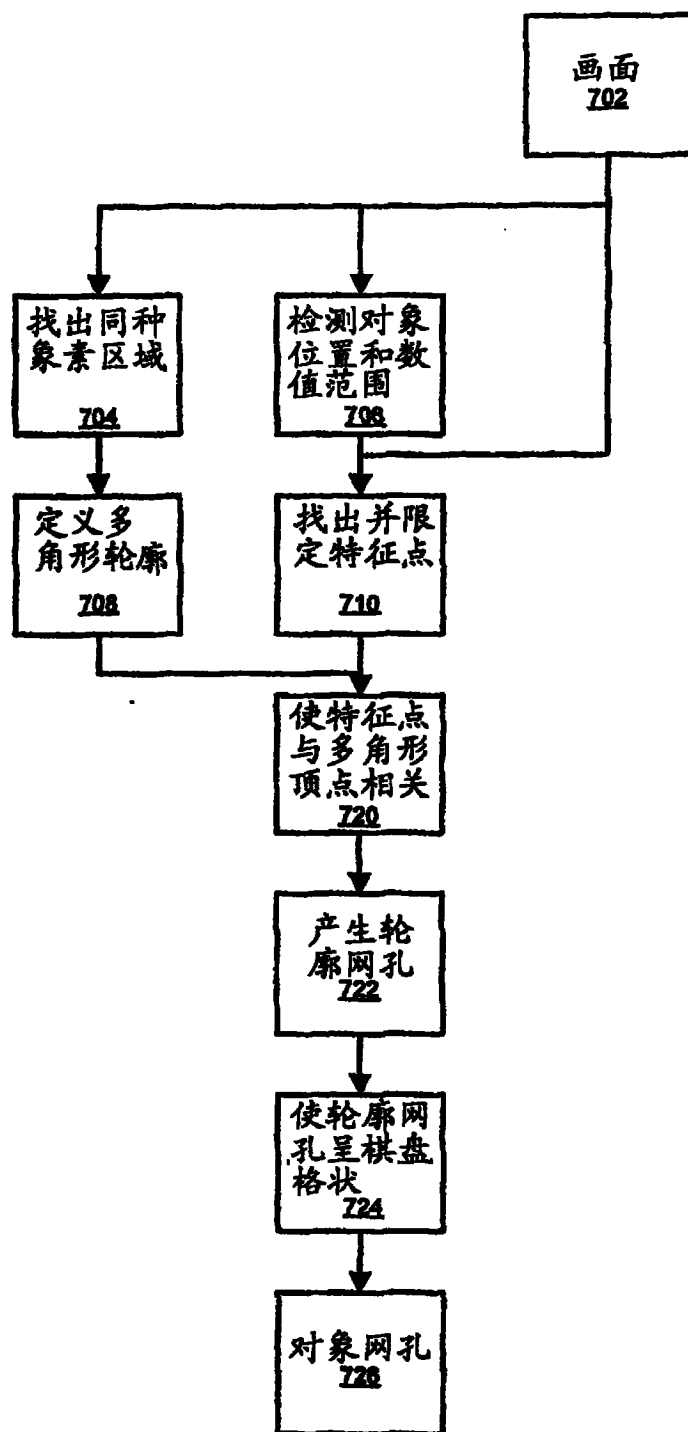


图 7 局部规范化 - 网孔的生成

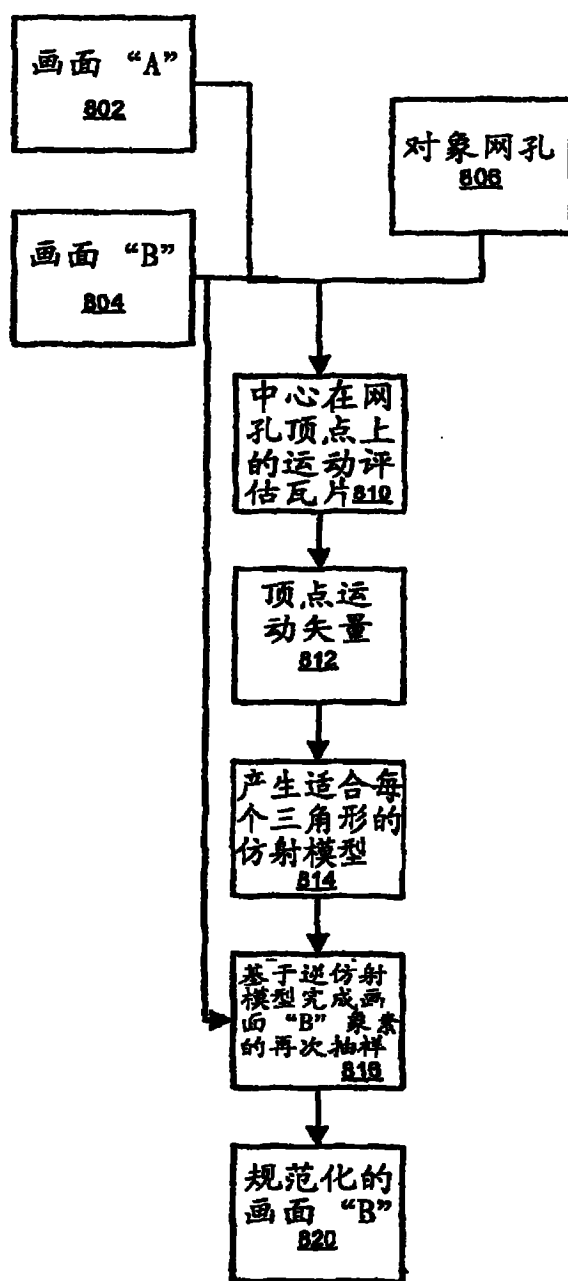


图 8

局部规范化 - 基于网孔的规范化

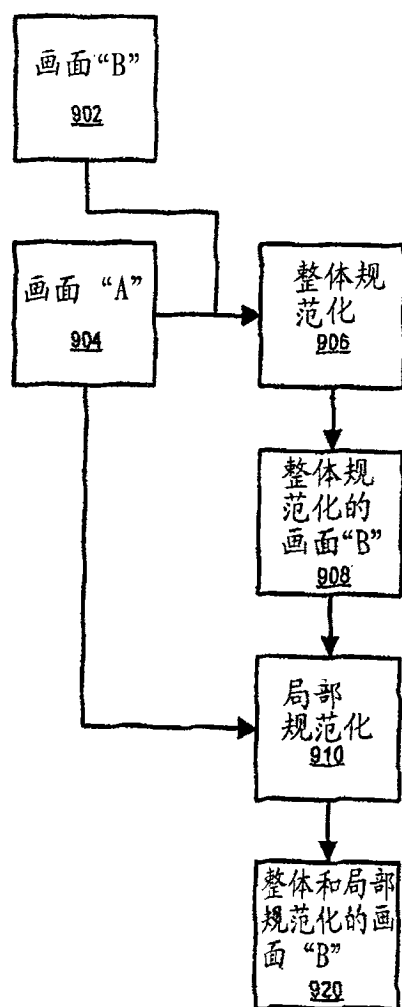


图9 整体和局部组合规范化

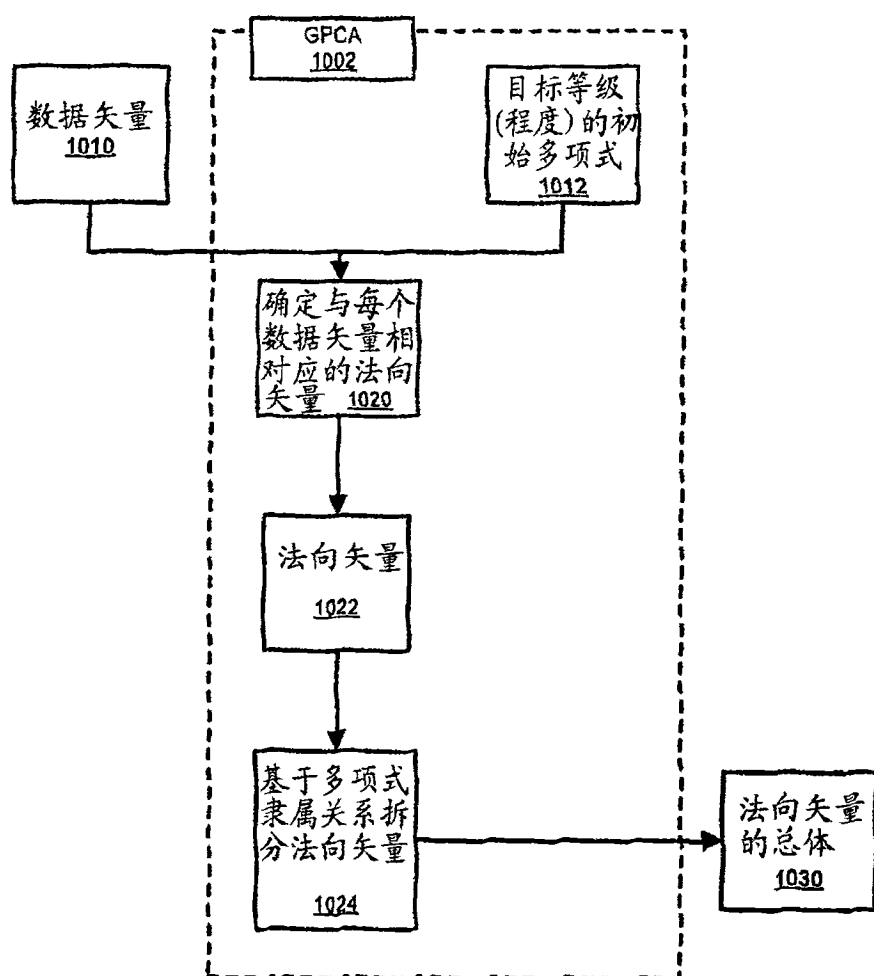


图 10GPCA- 基本多项式拟合和微分

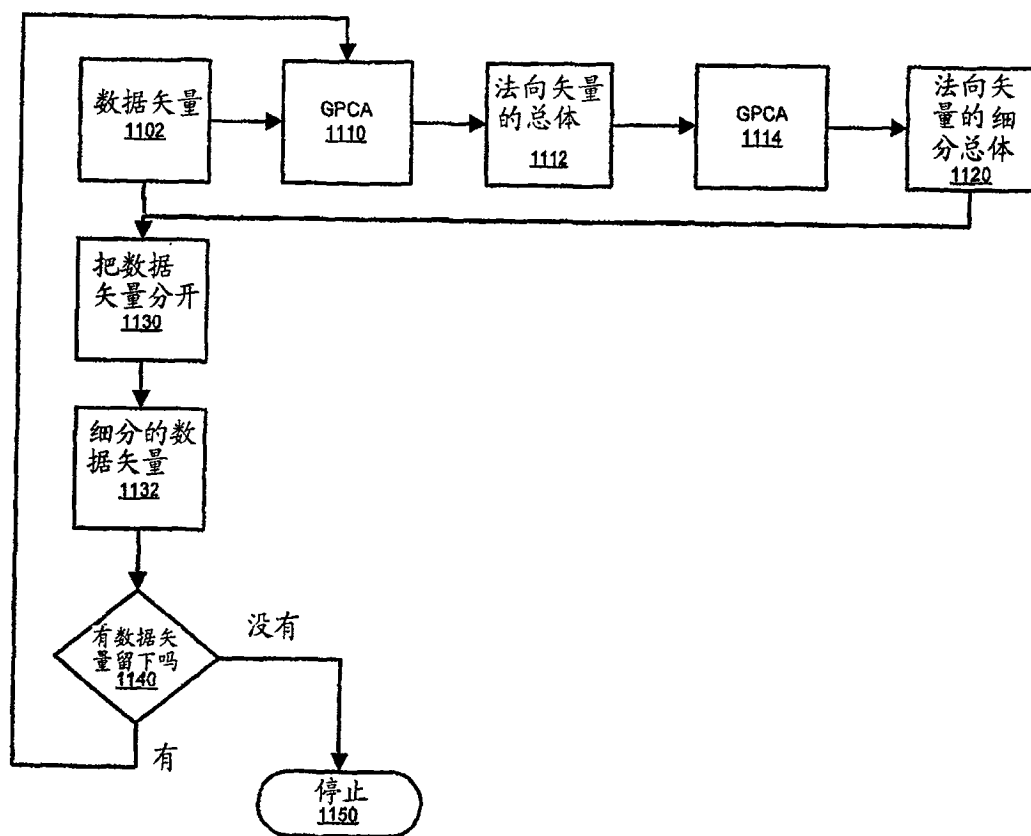


图 11GPCA- 递归的 GPCA 细分