



(12) 发明专利

(10) 授权公告号 CN 101151640 B

(45) 授权公告日 2010.12.08

(21) 申请号 200680010469.7

(22) 申请日 2006.01.20

(30) 优先权数据

60/648,094 2005.01.28 US

60/653,810 2005.02.17 US

(85) PCT申请进入国家阶段日

2007.09.28

(86) PCT申请的申请数据

PCT/US2006/001907 2006.01.20

(87) PCT申请的公布数据

W02006/083567 EN 2006.08.10

(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 查尔斯·保罗·佩斯 约翰·维斯

(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.

G06T 9/00 (2006.01)

(56) 对比文件

JP 特开 2002-74375 A, 2002.03.15, 全文.

CN 1333976 A, 2002.01.30, 全文.

US 6711278 B1, 2004.03.23, 全文.

Viola, P. Jones, M. Rapid

object detection using a boosted cascade of simple features. Computer Vision and Pattern Recognition, 2001. 2001, 11-9.

BILGE GUNSEL et. al. Content-based access to video objects: Temporal Segmentation, visual summarization, and feature extraction. Signal Processing 66 2. 1998, 66 (2), 261-280.

审查员 张小丽

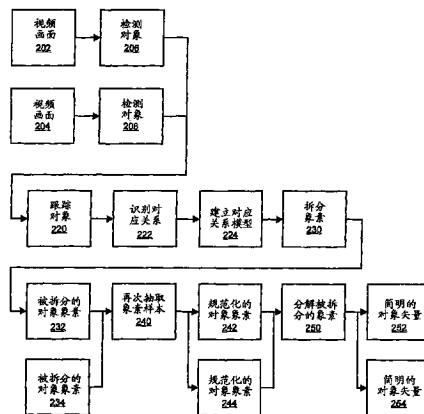
权利要求书 2 页 说明书 19 页 附图 13 页

(54) 发明名称

用来处理视频数据的装置和方法

(57) 摘要

本发明描述用来处理视频数据的装置和方法。本发明提供一种视频数据的表达,该数据表达能用来评定所述数据和用于所述数据的特定参数表达的拟合模型之间的一致性。这允许比较不同的参数化技术和选择继续处理该特定视频数据的最佳技术。所述表达能以中间形式作为较大程序的一部分或作为反馈机制被用于处理视频数据。当以它的中间形式被利用的时候,本发明能被用于视频数据的储存、增强、提炼、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。



CN 101151640 B

1. 一种用来从为数众多的视频画面中生成视频信号数据的编码形式的计算机装置,该装置包括:

用于在目标视频信号数据的视频画面序列中识别在两幅画面或多幅画面之间的对象的对应元素的装置;

用于建立这种对应关系的模型以产生轮廓模型的装置,所述轮廓模型被用来拆分对象和视频画面序列中非对象背景,其中轮廓模型通过确定轮廓顶点在与对象相一致的图像亮度的累积误差区域中的位置来建立;

用于对与所述对象相关联的所述视频画面中的象素数据进行拆分和再次抽样的装置,所述拆分和再次抽样的装置利用所述轮廓模型;

用于恢复再次抽样象素数据的空间位置的装置,所述恢复装置利用轮廓模型;

所述对象是一个或多个对象;以及

所述再次抽样象素数据是目标视频信号数据的中间形式。

2. 根据权利要求 1 的装置,其中对象是用跟踪方法跟踪的,包括:

在视频画面序列中探测对象的装置;

追踪通过视频画面序列的两幅画面或多幅画面的所述对象的装置;

所述对象的探测和跟踪装置包括 Viola/Jones 面部探测算法。

3. 根据权利要求 1 的装置,其中:

拆分和再次抽样的装置包括分离器;以及

所述对象是使用空间拆分方法从视频画面中拆分出来的,所述空间拆分方法通过以下步骤来实现:

分离器把与所述对象相关联的所述象素数据与所述视频画面序列中的其它象素数据拆开;

把所述恢复的再次抽样的象素数据与相关联的拆分数据组装在一起形成原始视频画面的装置;

所述分离器使用时间整合。

4. 根据权利要求 1 的装置,其中轮廓模型被因式分解成若干整体模型,其中包括:

把轮廓测量结果整合成整体运动模型的轮廓建模装置;

所述轮廓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;以及

所述轮廓建模装置包括以在视频画面序列中的两幅或多幅视频画面之间依据基于区段的运动评估产生的有限差分为基础的抽样总体。

5. 根据权利要求 1 的装置,其中目标视频信号数据的中间形式被进一步编码,其中包括:

把所述对象象素数据分解成编码表达的装置;

从编码表达中重组所述对象象素数据的装置;

所述分解装置包括主要成份分析,以及

所述重组装置包括主要成份分析。

6. 根据权利要求 5 的装置,其中画面的非对象象素如对象象素被建模,当其它对象被除去的时候,所述非对象象素与画面的剩余对象相对应。

7. 根据权利要求 5 的装置,其中被拆分和再次抽样的象素数据是用传统的视频压缩/

解压缩程序组合的,其中包括:

把所述再次抽样像素数据作为标准视频数据供应给传统的视频压缩/解压缩程序的装置;

连同对应的编码视频数据一起储存和传输轮廓模型对应关系数据的装置;

借此所述的压缩/解压缩程序能提高压缩效率。

8. 根据权利要求 1 的装置,其中轮廓模型被因式分解成局部变形模型,其中局部变形模型由以下来执行:

定义覆盖与所述对象相对应的像素的二维网孔的装置;

将轮廓测量结果生成局部运动模型的装置;

所述网孔定义装置以顶点和边缘的规则格子为基础,以及

所述轮廓测量结果包括以在视频画面序列中两幅或多幅视频画面之间依据基于区段的运动判断所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图像特征,其中包括:

识别与所述对象相对应的重要的图像特征的装置;以及

所述识别装置是分析图像梯度的 Harris 响应。

10. 根据权利要求 1 的装置,其中通过确定轮廓顶点在与累积误差相关联的画面中的位置,轮廓模型被用来将对象与非对象背景拆分开。

11. 根据权利要求 1 的装置,其中建立对应关系模型从而生成轮廓模型进一步包括使来自轮廓模型中一个的轮廓位置适合网孔顶点从而提供相符合的空间轮廓的装置。

12. 一种从多幅视频画面中生成视频信号数据的编码形式的方法,该方法包括以下步骤:

识别在目标视频信号数据的视频画面序列中的两幅或多幅画面之间的对象的对应元素;

建立对应关系模型以生成轮廓模型,所述轮廓模型被用来拆分对象和视频画面序列中的非对象背景,其中通过确定轮廓模型顶点在与对象相一致的图像强度的累积误差区域中的位置来建立轮廓模型;

使用轮廓模型,对与对象相关联的视频画面中的像素数据进行拆分和再次抽样,其中再次抽样的数据是目标视频信号数据的中间形式;以及

使用轮廓模型,恢复再次抽样的像素数据的空间位置。

13. 根据权利要求 12 的方法,其中轮廓模型被用来通过确定轮廓顶点在与累积误差相关联的画面中的位置来拆分对象和非对象背景。

14. 根据权利要求 12 的方法,其中建立对应关系模型以生成轮廓模型进一步包括使来自轮廓模型中一个的轮廓位置适合网孔顶点从而提供相符合的空间轮廓的步骤。

用来处理视频数据的装置和方法

[0001] 这份申请要求 2005 年 2 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 60/653,810 号和 2005 年 1 月 28 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 60/648,094 号的优先权。这份申请是作为 2005 年 9 月 20 日申请的美国专利申请第 11/230,686 号的部分继续申请的于 2005 年 11 月 16 日申请的美国专利申请第 11/280,625 号的部分继续申请,而其中所述的美国专利申请第 11/230,686 号又是 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。上述的每份申请在此通过引证被全部并入。

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)的计算机装置和计算机实现的方法。

背景技术

[0003] 本发明存在于其中的现有技术的一般系统描述能用图 1 表示。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 和一个或多个数据储存机制 108。

[0004] 输入级 102 可能包括若干元素,例如,照相机敏感元件,照相机敏感元件阵列、测距敏感元件、或从储存机制取回数据的方法。输入级提供表达人造的和/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] 视频数据可以依照预先定义的转移协议以数据流、阵列或数据包的形式直接地或通过中间的储存元素 108 提交给处理级 104。处理级 104 可以采用专用模拟或数字器件或可编程器件(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))的形式来执行所需要的一组视频数据处理操作。处理级 104 通常包括一个或多个 CODEC(编码/解码器)。

[0006] 输出级 106 产生能够影响使用者或外部装置的信号、显示或其它响应。通常,输出器件被用来产生指示信号、显示、硬拷贝、处理过的数据在存储器中的表达,或开始向远程站点传输数据。它也可以用来提供在后面的处理操作中使用的中间信号或控制参数。

[0007] 存储器在这个系统中是作为非必选元素出现的。在使用时,储存元素 108 可以是非易失的,例如,只读储存媒体,或易失的,例如,动态随机存取存储器(RAM)。单一的视频处理系统包括若干种储存元素并非罕见的,这些元素对输入级、处理级和输出级有各种不同的关系。这样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的是处理输入数据,产生对特定应用意义深长的输出。为了实现这个目标,可以利用多种处理操作,包括减少或消除噪音、特征提取、对象拆分和/或规范化、数据分类、事件探测、编辑、数据选择、数据重新编码和代码变换。

[0009] 产生受不良约束的数据的许多数据来源（尤其是声音和可视图像）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数据处理的目标有不利的影响。来源数据固有的易变性是在不引进起因于在推导工程假定中使用的未试过的经验的和探索性的方法的误差的情况下以可靠且有效的方式处理数据的障碍。这种易变性对于某些应用当输入数据被自然地或故意地限制在定义狭窄的特征组（例如，一组有限的符号值或狭窄的带宽）之中的时候被减轻。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预期用途和作为输入使用的来源信号的预期特征的影响。在大多数情况下，所需的完成效率也将是重要的设计因素。完成效率反过来受待处理的数据的数量与可用的数据储存器相比较的影响以及受该应用程序的计算复杂性与可得的计算能力相比较的影响。

[0011] 传统的视频处理方法因具有许多低效率性质而蒙受损害，这些低效率性质是以数据通信速度慢、存储需求大和干扰感性假象的形式出现的。这些可能是严重的问题，因为人们希望有多种方法使用和操纵视频数据和人们对某些形式的可视信息有先天的敏感性。

[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。在画面的一个部份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两个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，与空间规范化相对应的数据空间拆分能进一步用来增加单独来源于空间规范化的利益。

[0022] 具体地说，本发明提供一种能有效地把信号数据处理成一个或多个有益的表达的手段。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，所以进一步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模块。

[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组合式整体和局部规范化方法的方框图。

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

[0033] 图 11 是举例说明本发明的回归 GPCA 提炼方法的方框图。

[0034] 图 12 是举例说明背景分辨方法的方框图。

[0035] 图 13 是举例说明本发明的对象拆分方法的方框图。

[0036] 具体实施方法

[0037] 在视频信号数据中, 视频画面被组装成通常描绘在二维成像表面上投影(成像)的三维现场的图像序列。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像素组成。时常, 抽样信号对应于被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)。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一个时间维度对应于该画面在视频序列中的次序。

[0038] 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。识别程序限定该显著成份的存在和重要性并且选择在那些被限定的显著成份之中最重要的一个或多个显著成份。这不限制在现在描述的处理之后或同时识别和处理其它显著性较低的成份。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。不变子成份的识别是该成份某个方面的建模程序, 借此揭示该模型的参数表达法, 以允许将该成份被合成到预期的准确性水平。

[0039] 在本发明的一个实施方案中, 探测和跟踪前景对象。识别该对象的像素并且把这些像素从每个视频画面中拆分出来。把基于区段的运动评估应用于从多个画面中拆分出来的对象。然后, 把这些运动评估结果整合成一个高级的运动模型。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。这种规范化允许紧凑地表达多个画面上的对象像素的数值的线性分解。属于对象外观的显著信息被包含在这个紧凑表达之中。

[0040] 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。该对象是按空间规范化的, 借此得出紧凑的线性外观模型。此外, 进一步优选的实施方案在空间规范化之前先把前景对象从视频画面的背景中拆分出来。

[0041] 本发明的优选实施方案将本发明应用于一个人对着摄像机边说话边进行少量运动的影像。

[0042] 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很好地表达的任何对象。

[0043] 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。

[0044] 探测和跟踪

[0045] 一旦已经确定信号的显著构成成份, 就可以保留这些成份, 而可以减少或去除所有其它的信号成份。探测显著成份的程序展示在图 2, 其中视频画面(202)是用一个或多个探测对象(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几种形式编码。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。

[0046] 在一个优选实施方案中,显著特征分析程序完成显著信号模式的探测和分类。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探测到的对象显著特征有关的响应信号而设计的空间过滤器的组合。以不同的空间刻度在视频画面的不同位置应用该分类程序。来自该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。显著信号模式的探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。

[0047] 给出显著信号模式在一个或多个视频画面中的探测位置,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冗余信息和拆分(即,分离)信号模式的基础。

[0048] 特征点跟踪

[0049] 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“一些线”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异类的,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。

[0050] 在图 2 中,跟踪对象(220)程序是为了把来自探测对象程序(208)的探测例证和在许多视频画面(202 和 204)上一个或多个被探测对象的特征的进一步的识别对应关系(222)拉到一起而展示的。

[0051] 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)。

[0052] 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。

[0053] 基于对象的探测和跟踪

[0054] 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。这样的分类程序以对已在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。该级联分类程序是 AdaBoost 算法的变体。此外,响应计算能通过使用总面积表优化。

[0055] 局部配准

[0056] 配准包括在两个或多个视频画面中被识别对象的诸元素之间的对应关系的分配。这些对应关系变成建立该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。

[0057] 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它们与实践相关联的缩减量,现在描述用于本发明的各种不同的非限制性的配准方法。

[0058] 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来实现的。如果该对应关系在空间和强度双重意义上符合特定的恒定性限制,则光学流场能被稀疏地评估。

[0059] 如图 3 所示,画面(302 或 304)有可能通过十取一程序(306)或一些其它的二次抽

样程序（举例来说，低通过滤器）按空间被二次抽样。这些在空间上减少的图像（310&312）也可能被进一步二次抽样。

[0060] 菱形搜寻

[0061] 假定把一个视频画面分割成若干不重叠的区段，搜寻与每个区段匹配的先前的视频画面。以全面搜寻区段为基础（FSBB）的运动评估找出在先前的视频画面中与当前画面中的区段相比较时误差最小的位置。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。

[0062] 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。

[0063] 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等同物。

[0064] 基于网孔的运动评估

[0065] 基于网孔的预测使用顶点被边缘连接起来的几何网孔描绘视频画面的不连续区域，随后通过受网孔顶点位置控制的变形模型预测后续画面中那些区域的变形和运动。为了预测当前画面，因为顶点是移动的，所以在用顶点定义的区域之内的象素也是移动的。原始象素数值的相对运动和由此产生的近似是通过一些插值方法完成的，这些插值方法将象素位置与邻近那个象素的顶点的位置联系起来。当这样的运动存在于视频信号中的时候，缩放和旋转的附加建模与纯粹的平移相比较能产生更精确的画面象素预测。

[0066] 通常，网孔模型能被定义为是规则的或自适应的。规则的网孔模型是在不考虑基础信号特性的情况下设计的，而自适应的方法尝试按空间安排与基础视频信号的特征相关的顶点和边缘。

[0067] 规则网孔表示法提供一种方法，倘若影像中的成像对象有较多的空间间断点与网孔边缘相对应，运动或运动中固有的变形就能用该方法预测或建模。

[0068] 自适应网孔是在实质上比规则网孔更多地考虑基础视频信号的特征的情况下形成的。此外，这种网孔的自适应性质可以随着时间逝去考虑到网孔的各种不同的提炼。

[0069] 为了实现网孔和象素配准，本发明使用同种判据调整顶点搜寻。空间上与异种强度梯度相关联的顶点是先于那些有比较同种的梯度的顶点完成运动评估的。

[0070] 在优选的实施方案中，网孔的顶点运动评估是通过针对同等或近乎同等的同种顶点的运动评估的空间填注另外区分优先次序的。

[0071] 在优选的实施方案中，最初的网孔空间配置和最后的网孔配置是通过使用标准的图解式填充例行程序用小平面标识符填充映射图像在小平面水平上相互映射的。与每个三角形相关联的仿射变换能很快地从变换表中查出，而在一个网孔中与小平面对应的象素位置能很快地转换成在另一个网孔中的位置。

[0072] 在优选的实施方案中，为了评定与每个运动评估匹配相关联的残留误差针对顶点进行初步的运动评估。这个初步的评估被另外用来区分顶点运动评估次序的优先次序。这样的残差分析的好处是与比较少的失真相关联的运动评估将导致维持更似乎真实的网孔拓扑。

[0073] 在优选的实施方案中,网孔顶点运动评估被依比例缩减到某个有限范围,而且多种运动评估是通过一些迭代完成的,为的是允许网孔接近更全面优化的和拓扑正确的解。

[0074] 在优选的实施方案中,利用中心在每个顶点上矩形瓦块邻域的基于区段的运动评估被用来确定考虑到内插多角形邻域的顶点位移。除了针对误差梯度由来避免象素的空间插值和变形之外,这项技术也允许运动评估的平行计算。

[0075] 基于相位的运动评估

[0076] 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交叉相关 (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交叉相关。这种交叉相关被表达为位置与两个区段之间边缘的“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。

[0077] PNCC 的优势包括对比度遮掩,该对比度遮掩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,该单一步骤或许处理来自基于空间的运动评估程序的许多迭代。此外,该运动评估是子象素精确的。

[0078] 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。

[0079] 整体配准

[0080] 在一个实施方案中,本发明将来自有限差分评估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。

[0081] 如图 4 所示,在建立整体运动模型的情况下,有限差分是收集到借助那些运动评估的随机抽样 (410) 被迭代处理的一般总体库 (404) 中的平移运动评估 (402),而且线性模型被因式分解,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) 以便通过排除通过随机处理发现的该模型的异己样本更好地阐明该线性模型。

[0082] 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从而探查将描述重要的数据子集的运动模型的最强健方法。每个探头产生的模型都对它所代表的数据的百分比进行测试。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。

[0083] 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。根据残差与预先设定的阈值一致的样本的数目给该线性模型分配一个等级。这个等级被看作是“候选的共识”。

[0084] 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该总体的最后共识。

[0085] 最初的抽样、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该总体的最后共识。

[0086] 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分析该样本子集并

且逐渐增加子集规模,直到再多加一个样本将超过整个子集的残留误差阈值。

[0087] 如图 4 所示,整体模型评估程序 (450) 一直重复到共识等级可接受性测试令人满意 (452) 为止。当该等级尚未实现的时候,在设法揭示线性模型把与已发现的模型相关的有限差分的总体 (404) 分类。最好的 (最高等级的) 运动模型被添加到程序 460 的解集中。然后,在程序 470 中再次评估该模型。完成后,该总体 (404) 被再次分类。

[0088] 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簇,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域) 抽样的一般方法被进一步推广。

[0089] 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。

[0090] 规范化

[0091] 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在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。

[0092] 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。

[0093] 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。

[0094] 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。

[0095] 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)。

[0096] 基于网孔的规范化

[0097] 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。

[0098] 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。

[0099] 图 9 举例说明整体和局部规范化的组合方法。

[0100] 渐进的几何规范化

[0101] 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。

[0102] 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。

[0103] 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。

[0104] 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。

[0105] 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。

[0106] 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”或移动被结合在计算平均值操作中,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。

[0107] 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。

[0108] 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,以致顶点的象素被确认为属于某个多角形顶点。然后多角形周长的象素被详细研究,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。

[0109] 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。

[0110] 在本发明的一个实施方案中,为了找到邻接归并判据,处理区域映射图而不是处理区域目录是有利的。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。

[0111] 局部规范化

[0112] 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。

[0113] 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现图像亮度恒定性模棱两可的时候得到解释。

[0114] 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。

[0115] 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。

[0116] 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。

[0117] 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。

[0118] 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。在先前揭示的说明书中,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。

[0119] 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。

[0120] 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。

[0121] 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象素包含在该象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象素,计算适合那个象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象素的抽样并且把它的数值用在对应者的“规范化”位置。

[0122] 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。

[0123] 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 象素的一个象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。

[0124] 在进一步的实施方案中,水平和垂直的象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。

[0125] 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。

[0126] 规则网孔规范化

[0127] 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象素构成的,然而它的位置和尺寸与被检测对象相对应。

[0128] 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-4 AVC)的那些成比例。

[0129] 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象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。

[0130] 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。

[0131] 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。

[0132] 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。

[0133] 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。

[0134] 在替代实施方案中，基于训练周期内低的空间变形能限定像素区域。较少的变形根据局部变形模型指出较高的优先权。基于网孔的运动评估和规范化能被偏置以支持将倾向于包含最大数目的一致地发生的像素的网孔小平面。在这项技术的进一步的实施方案中，边缘探测响应被用来进一步限定小平面的。

[0135] 拆分

[0136] 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。

[0137] 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。

[0138] 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。

[0139] 时间拆分

[0140] 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。

[0141] 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。

[0142] 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。

[0143] 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。

[0144] 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。

[0145] 通过所述的方法,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例证都呈现最小的变化量。变化能被视为像素数值差异。

[0146] 基于拆分周界的整体变形模型是通过先创造围绕对象的周界达成,然后使该周界向探测到的对象中心倒塌直到该周界的顶点已经实现位置与异种图像梯度一致。运动评估是针对这些新的顶点位置推断的,而强健的仿射评估被用来发现整体变形模型。

[0147] 基于拆分网孔顶点图像世系的有限差分被整合成整体变形模型。

[0148] 对象拆分

[0149] 图 13 所示方框图展示对象拆分的一个优选实施方案。该程序由规范化图像(1302)的系综开始,然后所述规范化图像在该系综当中被逐对地计算差分(1304)。然后,这些差分被逐元素地积聚到积聚缓冲区之中(1306)。为了识别比较重要的误差区域,给该积聚缓冲区设定阈值(1310)。然后,为了确定累积误差区域的空间支持(1310),对通过阈值的元素掩模进行形态学分析(1312)。然后,由此产生的形态学分析(1312)的抽出物(1314)与探测到的对象位置进行比较(1320),以便将后来的处理集中在与该对象一致的累积误差区域。然后,用形成其凸壳(1324)的多角形近似孤立空间区域(1320)的边界(1322)。然后,调整壳的轮廓(1330),以便更好地初始化用于活跃轮廓分析(1332)的顶点位置。一旦活跃轮廓分析(1332)已经在累积误差空间中会聚在低能解上,该轮廓被用作最后轮廓(1334),而且被限制在该轮廓之中的像素被视为最有可能是对象像素的那些像素,而该轮廓之外的那些像素被视为非对象像素。

[0150] 在优选的实施方案中,运动拆分能在给定显著图像模态的被探测位置和规模的情况下实现。距离变换能用来确定每个像素离开该被探测位置的距离。如果与最大距离相关像素数值被保留,合理的背景模型能被求解。换句话说,环境信号是使用信号差衡量标准按时再次抽样的。

[0151] 进一步的实施方案包括使用与当前探测位置相关的距离变换来把距离分配给每个像素。如果到某个像素的距离大于在某个最大像素距离表中的距离,那么该像素数值被记录下来。在适当的训练周期之后,如果适合于那个像素的最大距离是大的,则假定该像素有最高的作为背景像素的可能性。

[0152] 给出环境信号模型,完全的显著信号模态能按每个时间例证计算差分。这些差分

每个都能被再次抽样变成空间规范化的信号差（绝对差）。然后，这些差分彼此对准并且被累积。由于这些差分相对于显著信号模式是空间规范化的，所以差分的峰值将主要对应于与显著信号模式有关的象素位置。

[0153] 在本发明的一个实施方案中，训练周期被定义，其中一些对象探测位置被确定而且那些位置的质心被用来与远离这个位置将考虑到画面差分化产生将有最高的作为非对象象素的概率的背景象素的探测位置一起确定最佳的画面数。

[0154] 在本发明的一个实施方案中，活跃轮廓模型被用来通过在累积误差“图像”中确定轮廓顶点位置把前景对象从非对象背景中拆分出来。在优选的实施方案中，活跃轮廓的边缘被细分得与被探测对象的比例相称，以便得到较大的自由度。在优选的实施方案中，最后轮廓的位置能快速灵活地移动到最接近的规则网孔顶点，以便得到有规律地隔开的轮廓。

[0155] 在对象拆分的一个非限制性实施方案中，使用一个导向核来针对时间上成对的图像产生误差图像过滤器响应。对取向与总运动方向正交的过滤器的响应倾向于在相对于背景的运动从遮蔽背景到显露背景的时候增强误差表面。

[0156] 规范化图像系综的规范化图像画面强度矢量是依据一个或多个创造残差矢量的参考画面计算差分的。这些残差矢量是逐元素地累积的，以形成累积残差矢量。然后，这个累积残差矢量被按空间地调查，以便定义适合对象象素和非对象象素的空间拆分的空间对象边界。

[0157] 在一个优选实施方案中，为了得出能用来给累积残差矢量设定阈值的统计阈值，完成了最初的累积残差矢量统计分析。通过先腐蚀后膨胀的形态学操作，形成初步的对象区域掩模。然后，分析该区域的轮廓多角形点以揭示那些点的凸壳。然后，把该凸壳作为初始轮廓用于活跃轮廓分析法。该活跃轮廓在它会在该对象的累积残差空间边界上之前是一直增殖的。在进一步优选的实施方案中，初步的轮廓边缘通过添加中点顶点被进一步细分到实现适合于所有的边缘长度的最小边缘长度。这个进一步的实施方案意味着逐渐增加活跃轮廓模型的自由度以便更精确地适合该对象的轮廓。

[0158] 在优选的实施方案中，细化的轮廓用来产生通过覆盖该轮廓暗示的多角形和覆盖规范化图像中的多角形指出该对象的象素的象素掩模。

[0159] 非对象的分辨

[0160] 图 12 所示的方框图揭示非对象拆分或同义地背景分辨的一个优选实施方案。通过背景缓冲区 (1206) 和初始最大距离值缓冲区 (1204) 的初始化，该程序是为通过把“稳定性”与距被探测对象位置 (1202) 的最大距离联系起来确定最稳定的非对象象素而工作的。给出新探测的对象位置 (1202)，该程序检查每个象素位置 (1210)。对于每个象素位置 (1210)，使用距离变换计算离开被探测对象位置 (1202) 的距离。如果那个象素的距离大于在最大距离缓冲区 (1204) 中先前储存的位置 (1216)，那么先前的数值被当前的数值代替 (1218)，而且该象素数值被记录在象素缓冲区中 (1220)。

[0161] 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“背景分辨”部分中描述的。

[0162] 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，该轮廓在空间上扩展以使残留误差与轮廓变形平衡。

[0163] 梯度拆分

[0164] 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。

[0165] 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。

[0166] 进一步的实施方案使用对一幅图像产生的 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。

[0167] 多刻度梯度分析

[0168] 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。

[0169] 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。

[0170] 光谱拆分

[0171] 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。

[0172] 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。

[0173] 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。

[0174] 背景拆分

[0175] 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。

[0176] 建立外观模型

[0177] 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。

[0178] 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。表达空间强度场象素的特征矢量能被组装成外观变化模型。

[0179] 在替代实施方案中,外观变化模型是依据被拆分的象素子集计算的。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。

[0180] 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。

[0181] PCA

[0182] 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,改变象素的初始外观质量水平能实现。

[0183] 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。

[0184] 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,象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象素外观的近似值误差和比特率之间的直接交换。

[0185] 如图 2 所示,为了得到量纲上简明的数据版本 (252 和 254),规范化的对象象素 (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。

[0186] 连续的 PCA

[0187] 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。

[0188] 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,则瓦片的象素适合残差编码;而且在压缩效率方面有对应的增益。否则,瓦片的象素被直接编码。这种基于瓦片的换句话说基于区段的运动预测方法通过平移包含象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。

[0189] 基于残差的分解

[0190] 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域作为新区段

完成编码构成的。

[0191] 对应于残留区段的象素通过运动矢量映射到先前画面的象素上。结果是象素通过能通过连续应用残值合成的影像的瞬时路径。这些象素被确认为能使用 PCA 最明确地表达的象素。

[0192] 基于遮挡的分解

[0193] 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象素被移动象素遮挡（覆盖）。对于每个遮挡事件，都把遮挡象素劈成新层。没有历史的象素也将暴露出来。暴露出来的象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。

[0194] 象素的时间连续性是通过象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，每层中的象素就能基于对条理分明的运动模型的隶属关系编组。

[0195] 分波段时间量化

[0196] 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。

[0197] 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。

[0198] 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。

[0199] 就 DWT 而言，分波段已经按针对 DCT 描述的方式排列好。

[0200] 在非限制性实施方案中，PCA 系数的截取是变化的。

[0201] 子波

[0202] 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。

[0203] 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。

[0204] 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。

[0205] 子空间分类

[0206] 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象素、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。

[0207] 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。

[0208] 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。

[0209] 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。

[0210] 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。

[0211] 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)的表达的多元性的 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受约束形式。这种方法用来在有噪音存在的情况下通过计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。

[0212] 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另一 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。

[0213] 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个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遇到退化情形的时候,本发明的方法不会退化。

[0214] 图 10 举例说明基本的多项式拟合和求微分的方法。

[0215] 混合空间规范化压缩

[0216] 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。

[0217] 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。

[0218] 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。

[0219] 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)。

[0220] 混合编码解码的整合

[0221] 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。

[0222] 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。

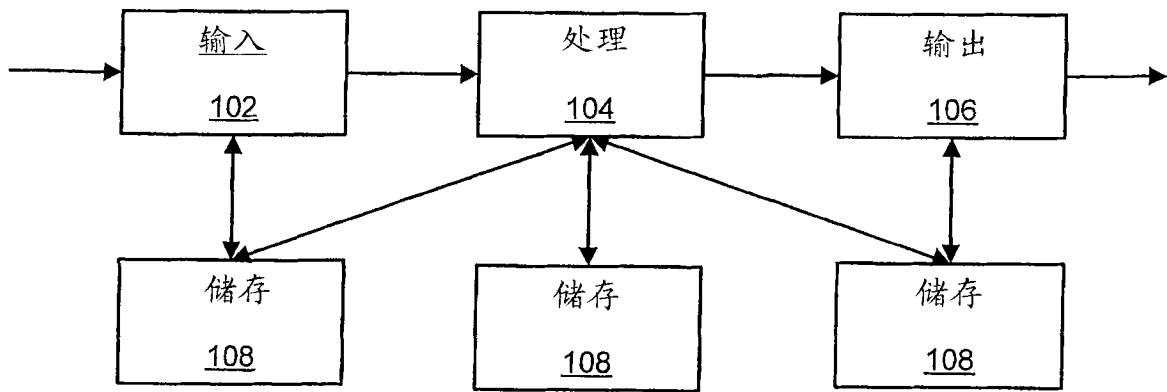


图 1
(现有技术)

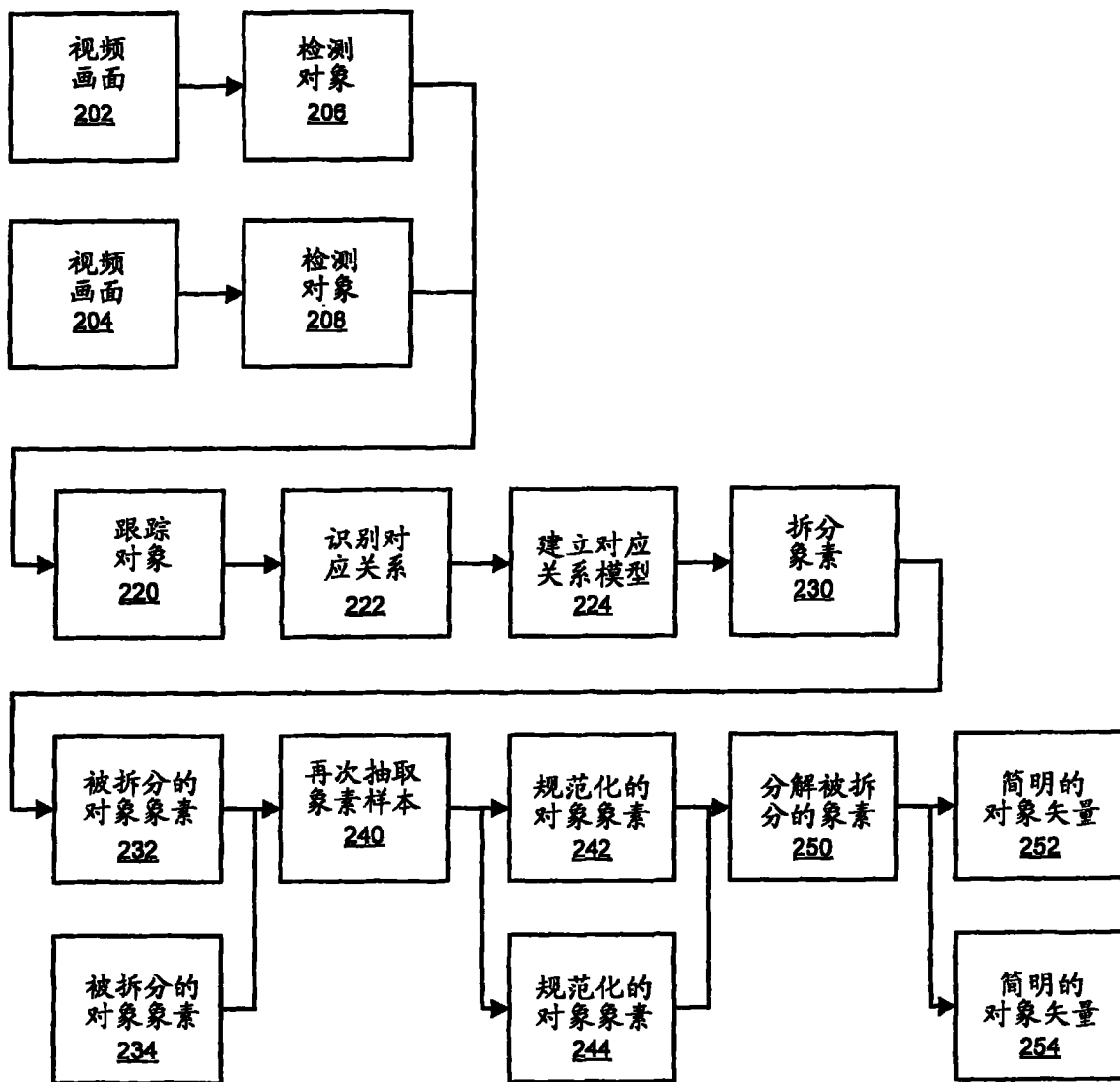


图 2

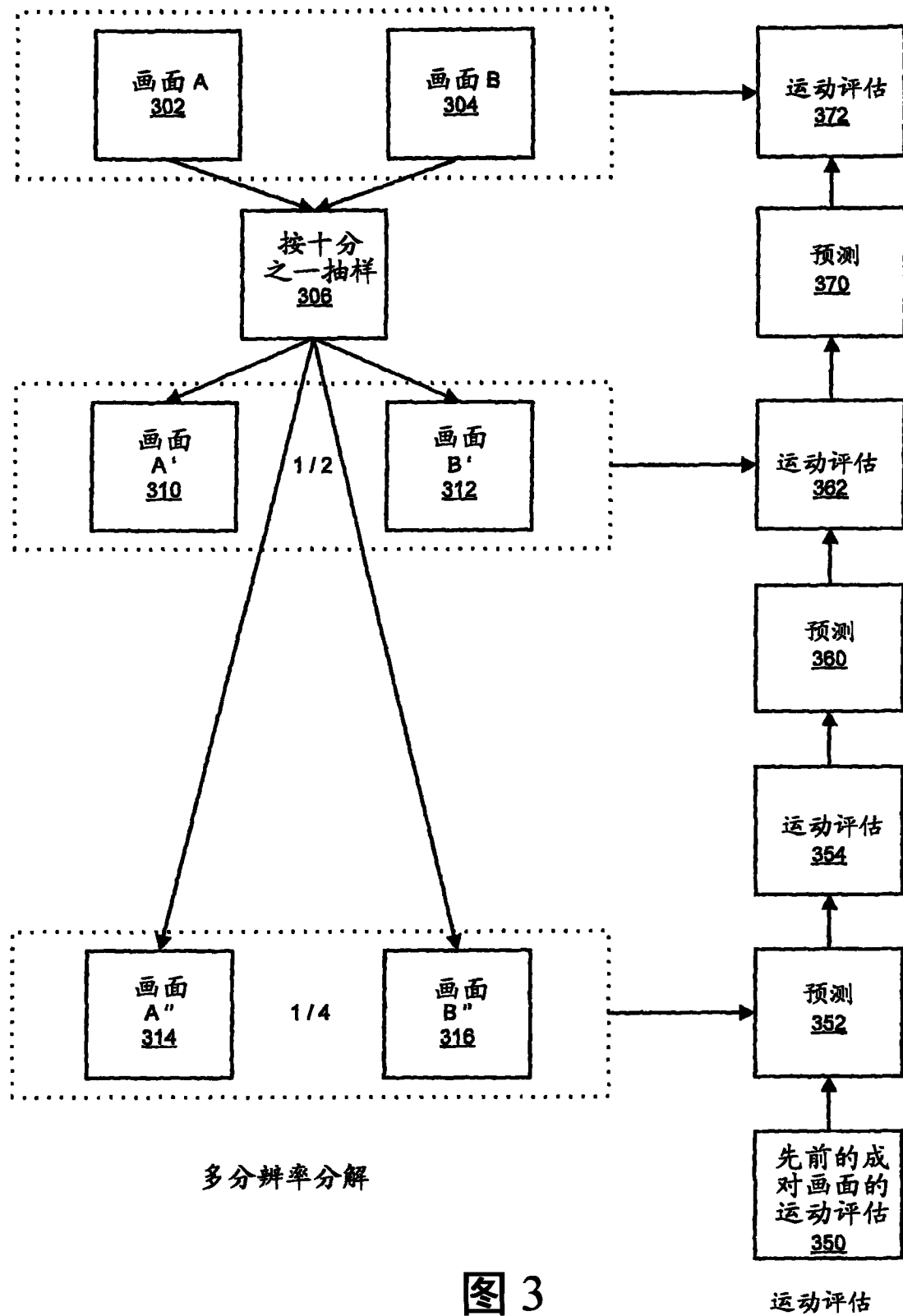
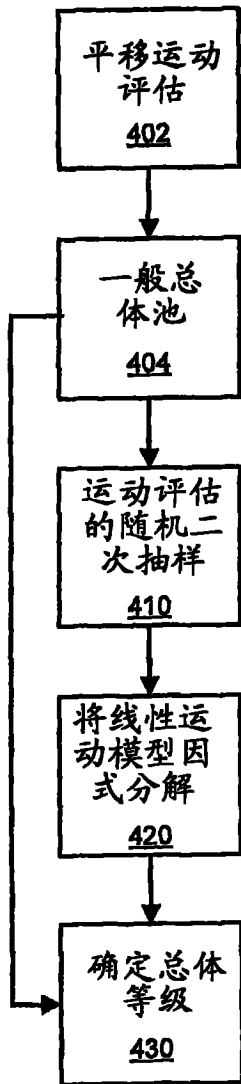
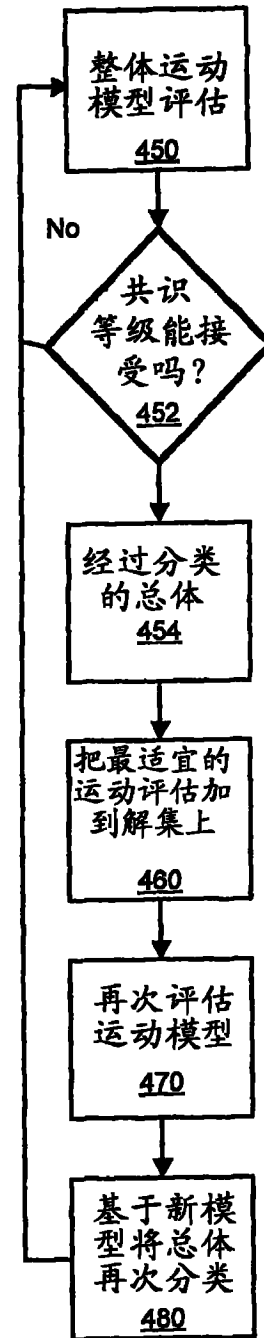


图 3



整体运动模型评估



共识等级验收测试和处理

图 4

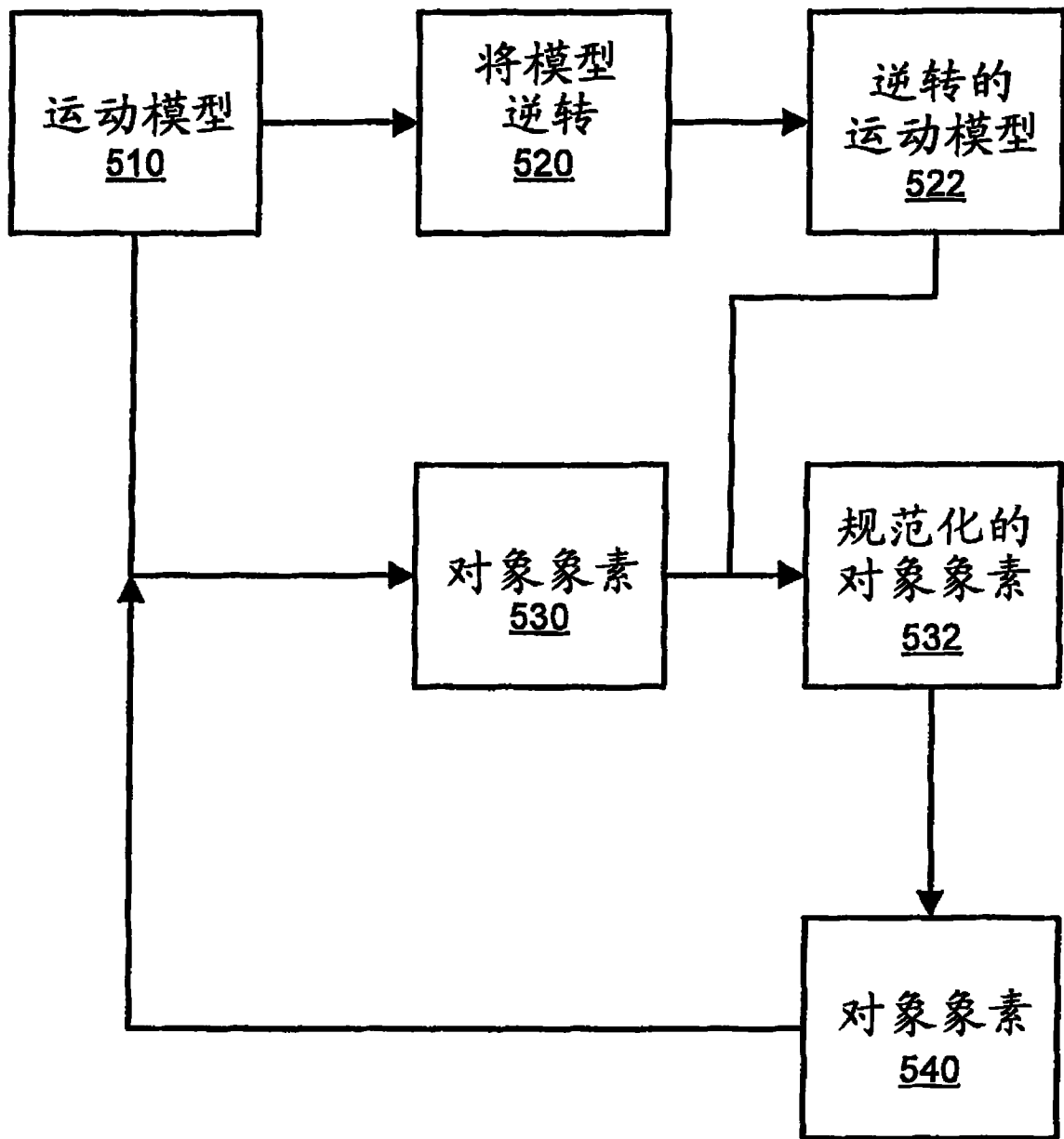


图 5

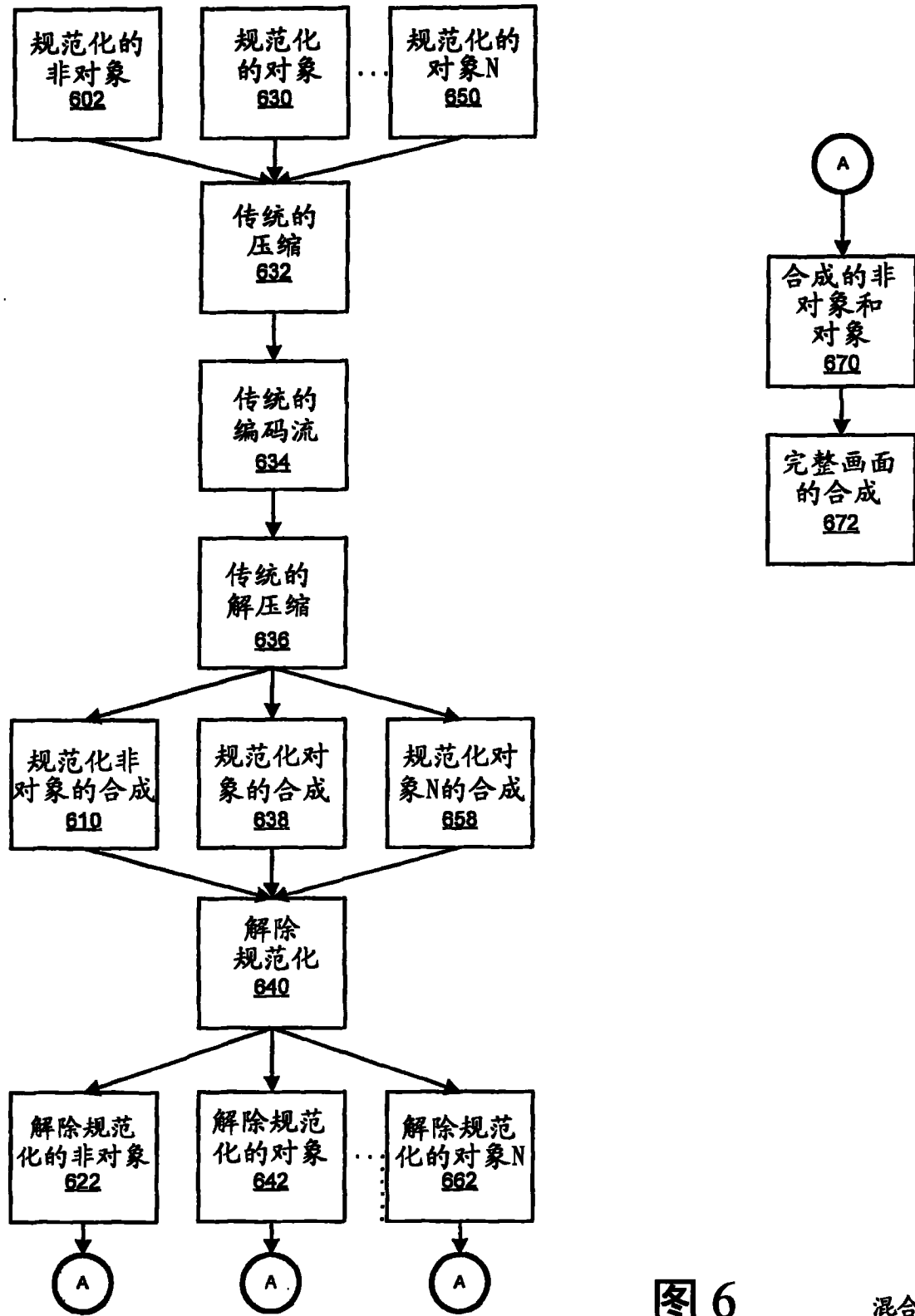


图 6

混合

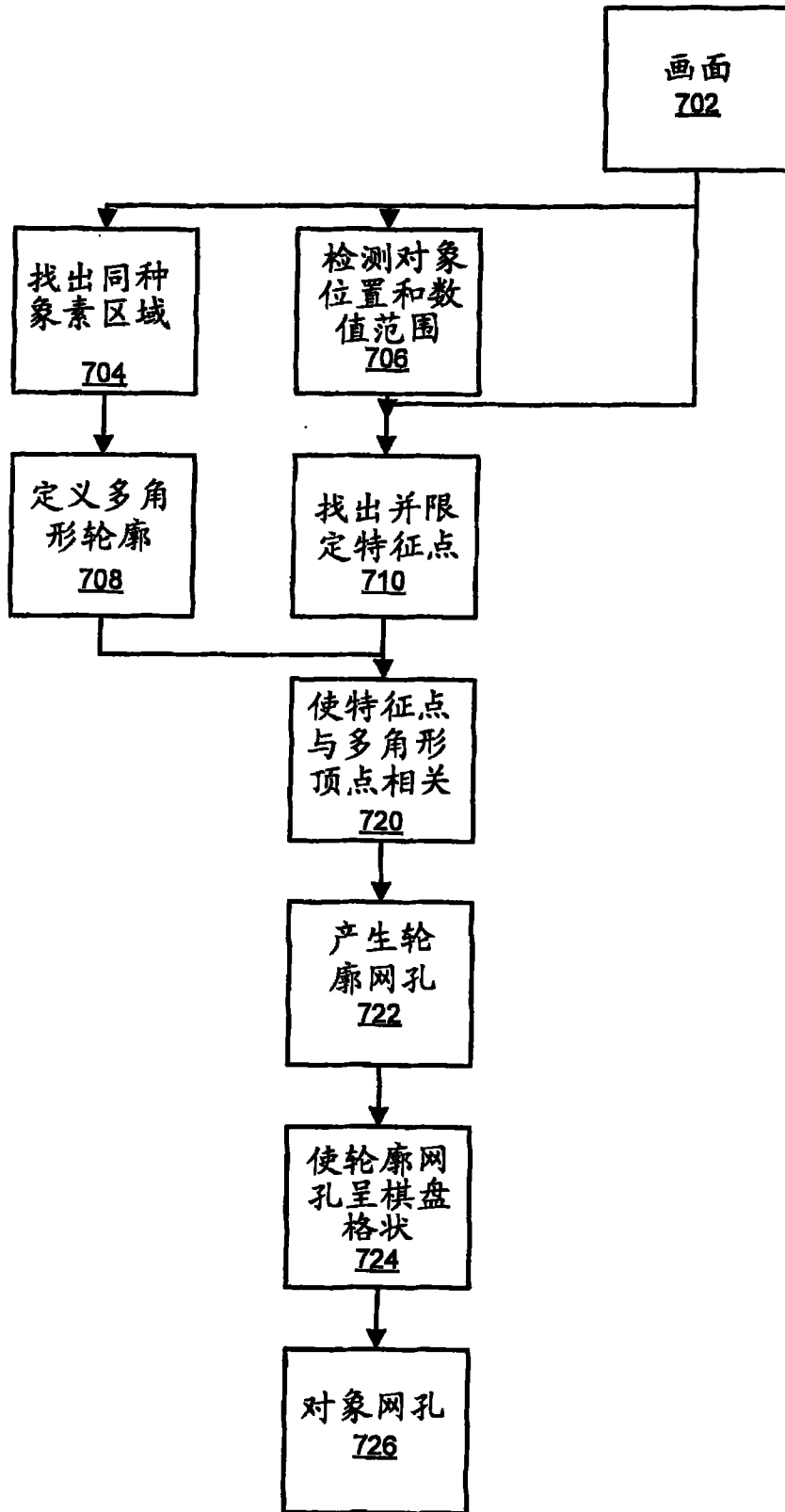


图 7

局部规范化 -
网孔的生成

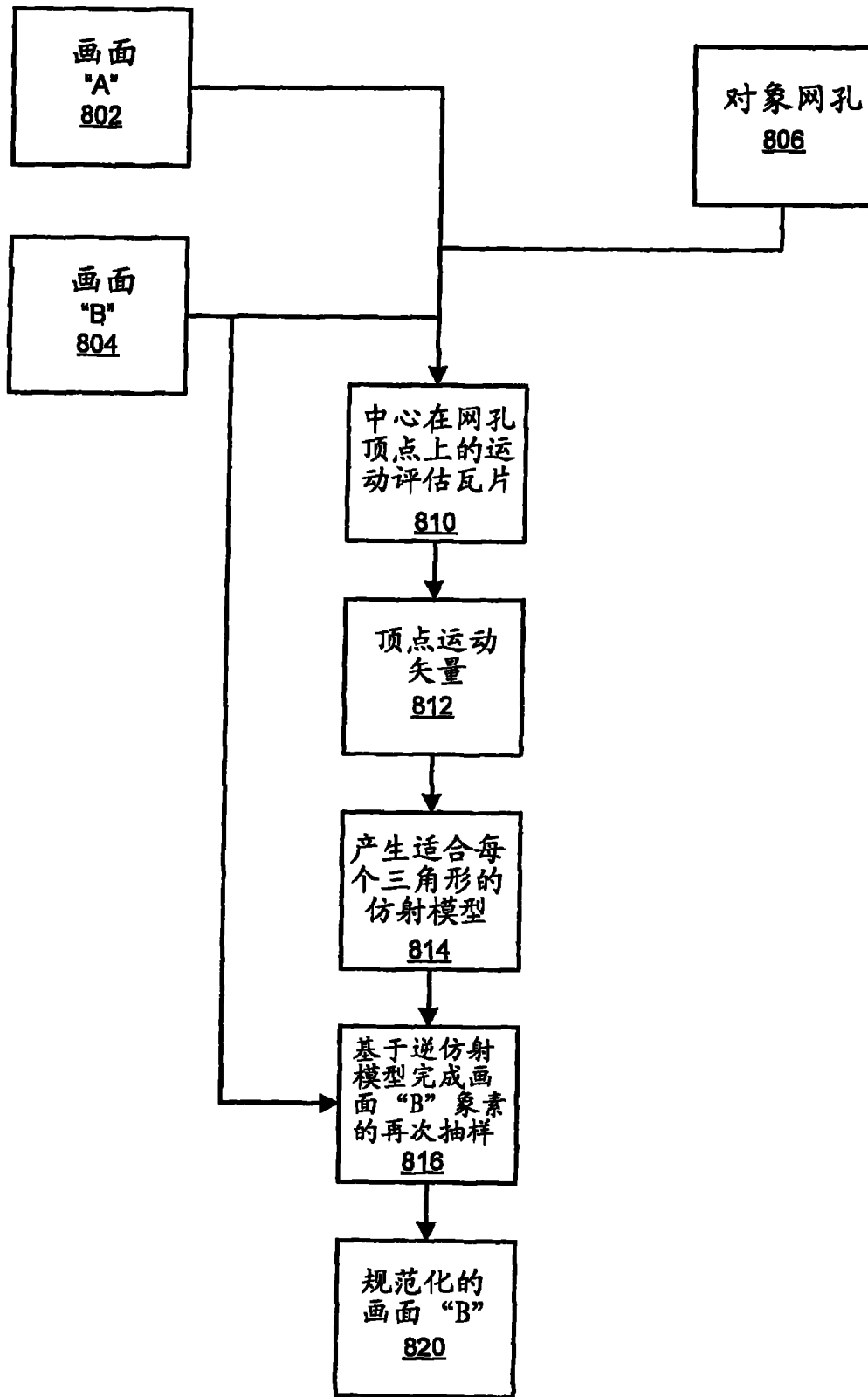


图 8

局部规范化 -
基于网孔的规范化

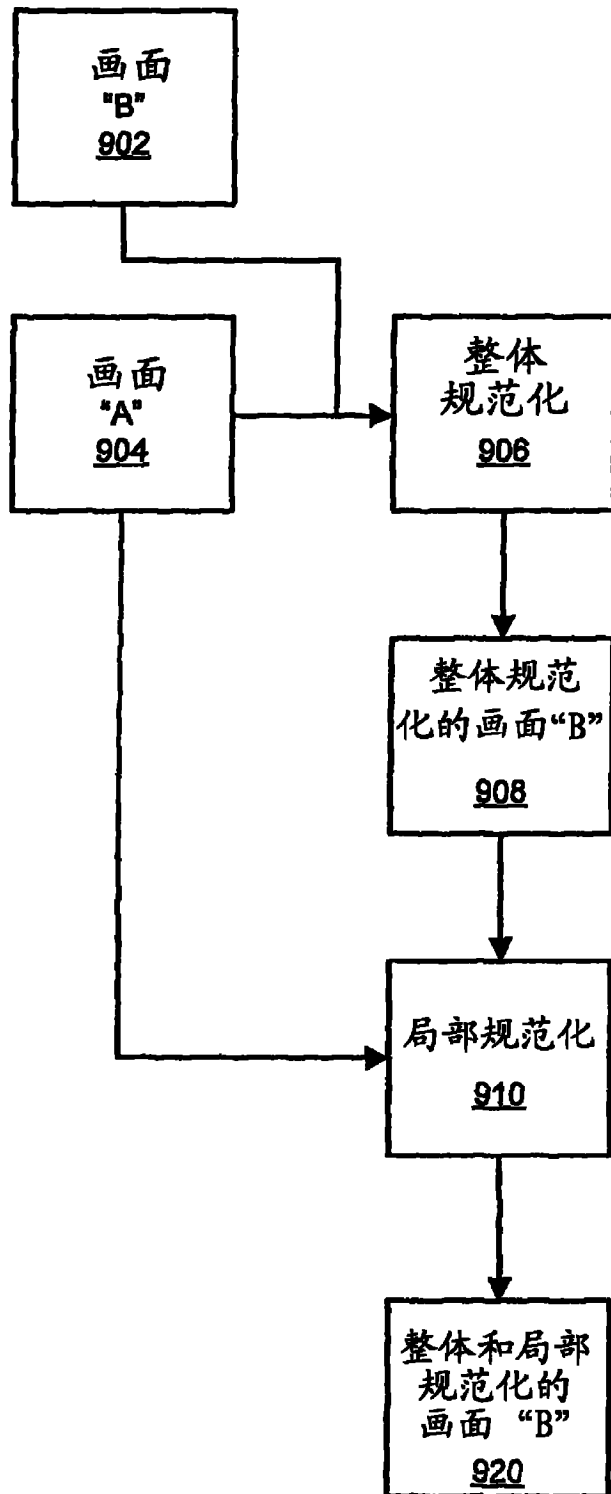


图 9

整体和局部规范化的组合

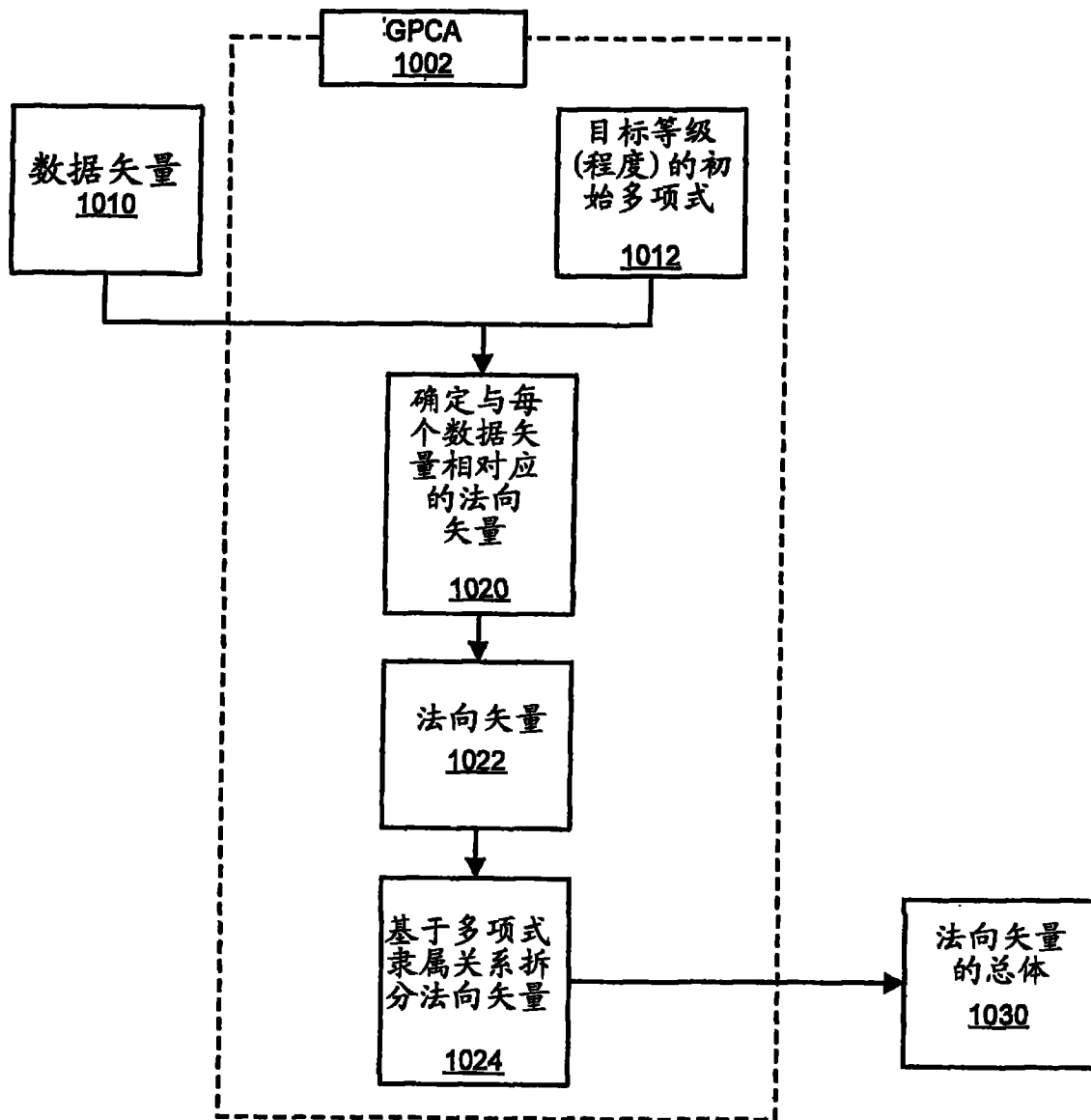


图 10

GPCA - 基本多项式拟合和微分

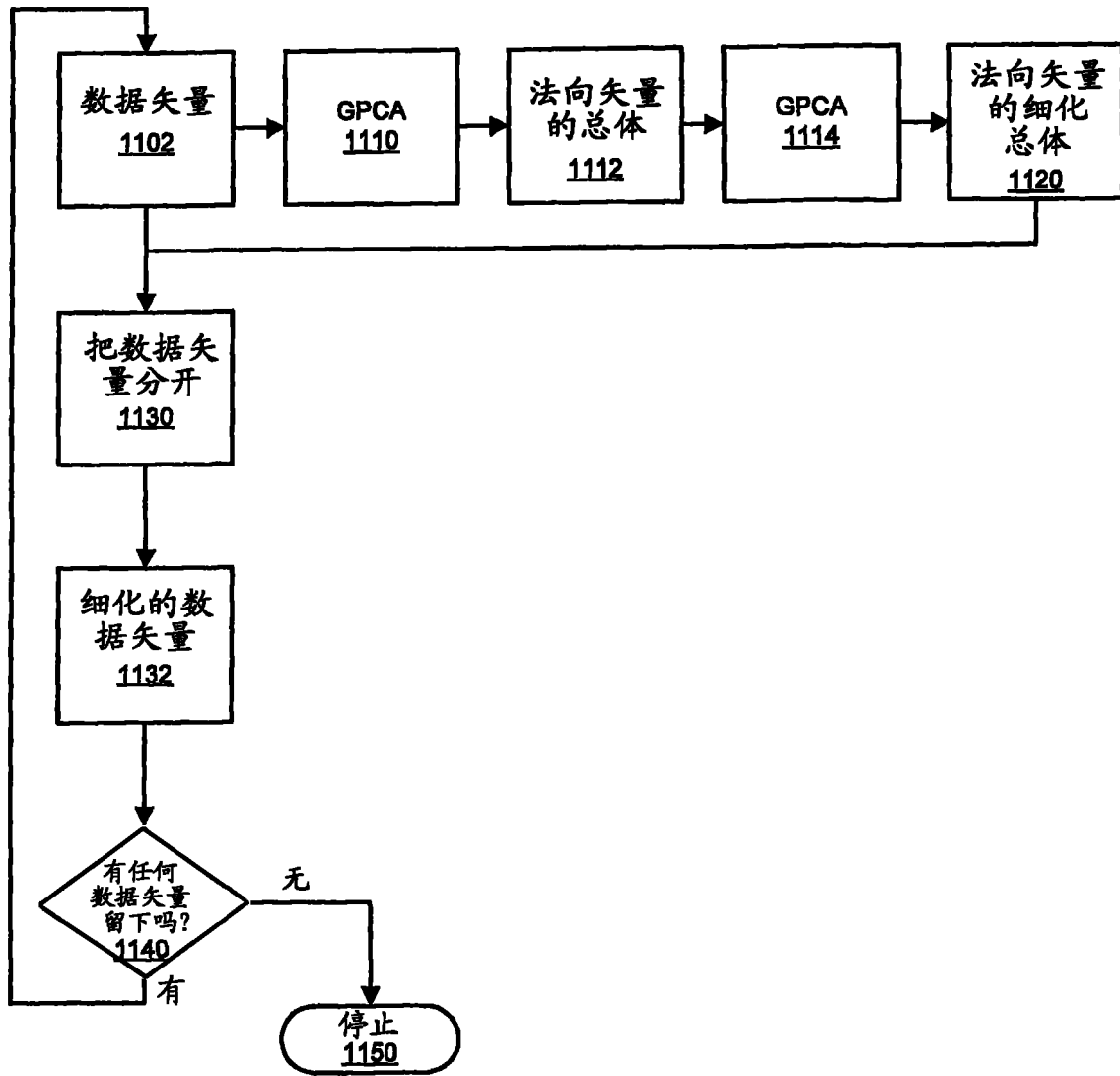
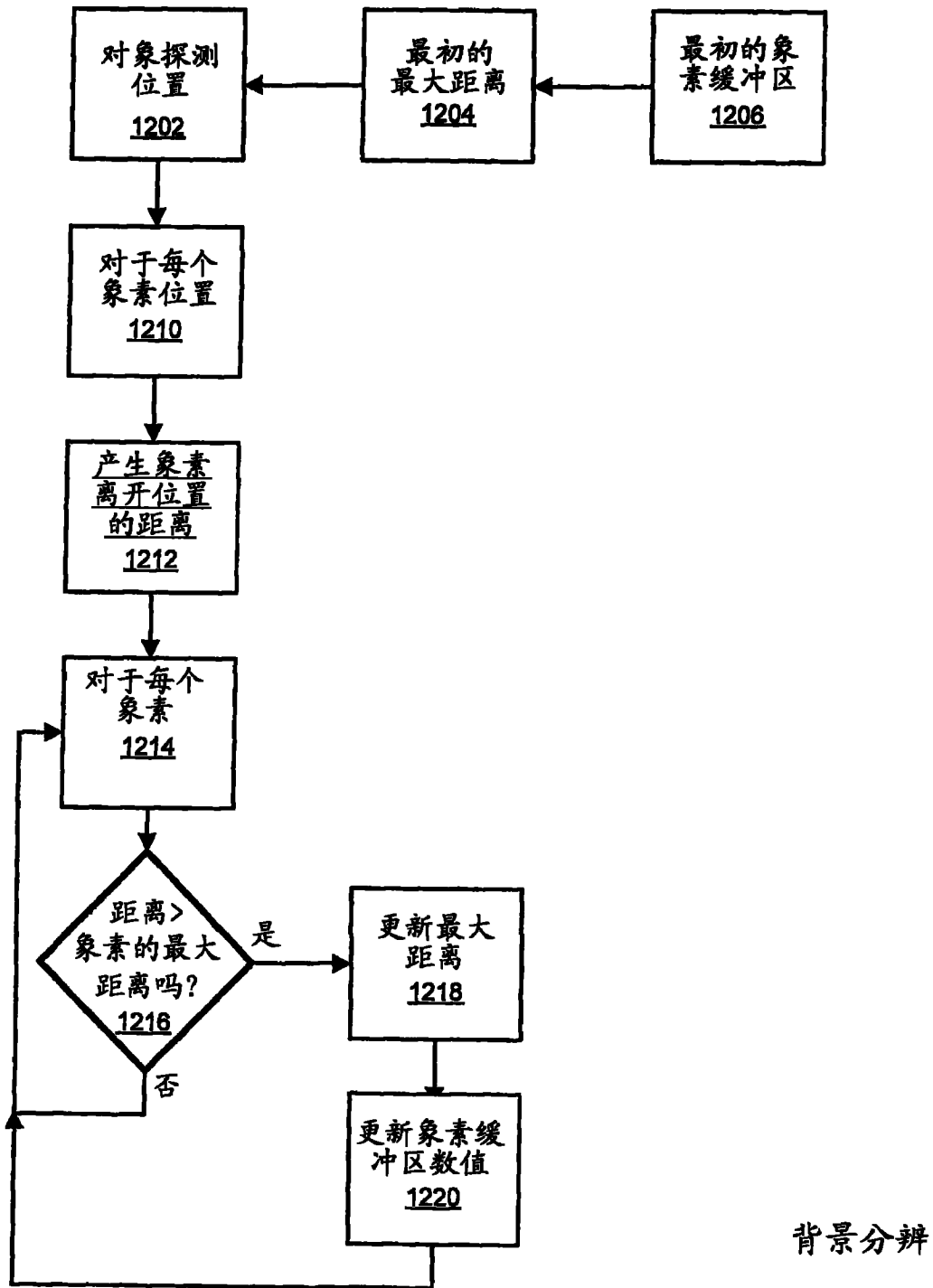


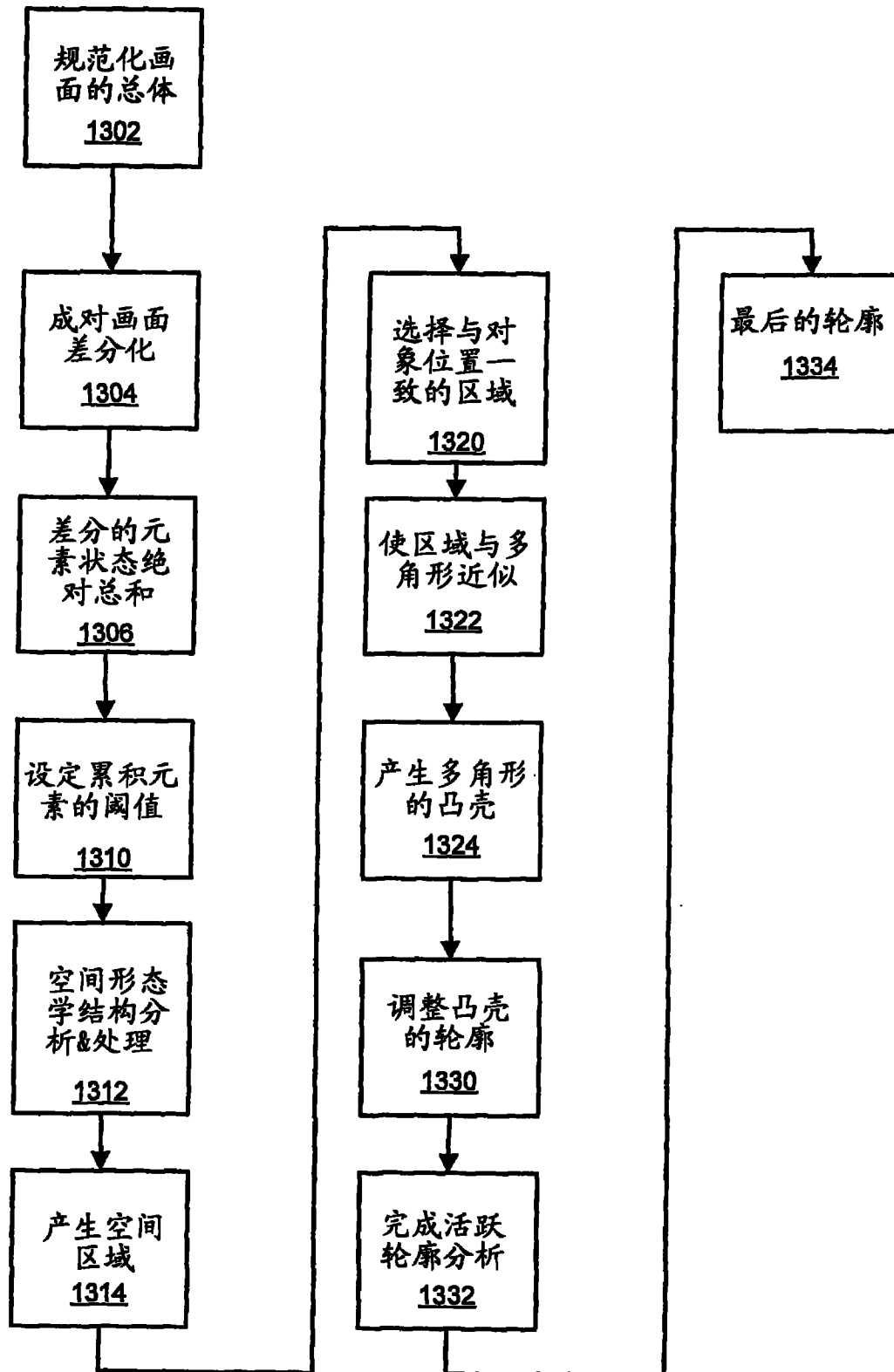
图 11

GPCA - 递归的GPCA
细化



背景分辨

图 12



对象拆分

图 13