



(12) 发明专利

(10) 授权公告号 CN 102804178 B

(45) 授权公告日 2016. 05. 04

(21) 申请号 201080026591. X

(22) 申请日 2010. 06. 04

(30) 优先权数据

12/487, 686 2009. 06. 19 US

(85) PCT国际申请进入国家阶段日

2011. 12. 15

(86) PCT国际申请的申请数据

PCT/US2010/001637 2010. 06. 04

(87) PCT国际申请的公布数据

W02010/147623 EN 2010. 12. 23

(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 D·马迪拉克施 A·C·路易

(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.

G06F 17/30(2006. 01)

G06K 9/62(2006. 01)

(56) 对比文件

US 2003/0198390 A1, 2003. 10. 23,

CN 1959705 A, 2007. 05. 09,

US 2008/0033991 A1, 2008. 02. 07,

审查员 胡雅娟

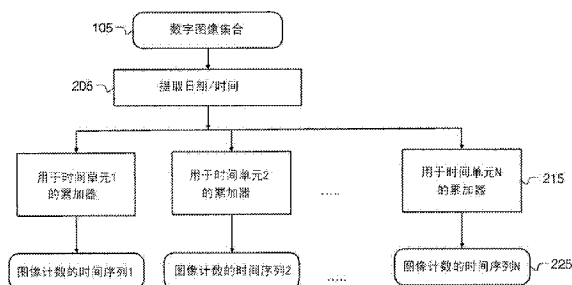
权利要求书1页 说明书6页 附图8页

(54) 发明名称

检测用户的图像集合中的重要事件

(57) 摘要

一种用于确定数字图像集合中的重要事件的方法,包括:使用处理器以执行以下操作:根据所述图像集合生成图像计数的时间序列;计算所述图像计数的时间序列的模型;以及使用所述图像计数的时间序列和所述模型来确定重要事件。



1. 一种用于确定数字图像集合中的重要事件的方法,包括:
使用处理器来执行以下操作:
 - a) 根据所述图像集合生成图像计数的时间序列;
 - b) 计算所述图像计数的时间序列的模型,所述模型包括捕获与先前的数据值的关系的自回归组件,并包括捕获图片拍摄行为中的逐渐改变的移动平均组件;以及
 - c) 使用所述图像计数的时间序列和所述模型来确定重要事件,所述重要事件是由所述图像计数的时间序列与所述模型的输出的偏差来识别的。
2. 根据权利要求1所述的方法,其中,所述模型是ARIMA(p,d,q)模型,其中p是自回归分量的阶,q是移动平均分量的阶,d是处理与稳态的偏差所需的差分的阶。
3. 根据权利要求2所述的方法,其中,所述模型是使用适合度分数从多个模型中选择出来的。
4. 根据权利要求1所述的方法,其中,步骤c)包括:
 - i) 计算残差,其中,所述残差是所述图像计数的时间序列在每个时间单元的图像计数与所述模型在每个时间单元所预测的图像计数之间的差;
 - ii) 识别所关注的时间步长,其具有高于计算出的阈值的残差;以及
 - iii) 合并连续的所关注的时间步长,并且将图像计数阈值应用于所合并的连续的所关注的时间步长以确定重要事件。
5. 根据权利要求1所述的方法,其中,步骤a)中的所述图像计数的时间序列是以不同的时间粒度生成的。
6. 一种用于确定数字图像集合中的重要事件的方法,包括:
使用处理器来执行以下操作:
 - a) 根据所述图像集合以不同的时间粒度生成图像计数的时间序列;
 - b) 计算在每个时间粒度的所述图像计数的时间序列的模型,所述模型包括捕获与先前的数据值的关系的自回归组件,并包括捕获图片拍摄行为中的逐渐改变的移动平均组件;
 - c) 使用所述图像计数的时间序列和所述模型来确定在每个时间粒度的重要事件,所述重要事件是由所述图像计数的时间序列与所述模型的输出的偏差来识别的;以及
 - d) 基于用户操作或系统需求来确定所述重要事件。
7. 根据权利要求6所述的方法,其中,所述模型是ARIMA(p,d,q)模型,其中p是自回归分量的阶,q是移动平均分量的阶,d是处理与稳态的偏差所需的差分的阶。
8. 根据权利要求6所述的方法,其中,所述模型是使用适合度分数从多个模型中选择出来的。
9. 根据权利要求6所述的方法,其中,步骤c)包括:
 - i) 计算残差,其中,所述残差是所述图像计数的时间序列在每个时间单元的图像计数与所述模型在每个时间单元所预测的图像计数之间的差;
 - ii) 识别所关注的时间步长,其具有高于计算出的阈值的残差;以及
 - iii) 合并连续的所关注的时间步长,并将图像计数阈值应用于所合并的连续的所关注的时间步长以确定重要事件。

检测用户的图像集合中的重要事件

技术领域

[0001] 概括地说,本发明涉及数字图像处理领域,具体地说,涉及一种用于识别对用户的图像集合中的重要事件进行描绘的数字图像组的方法。

背景技术

[0002] 数字照相机和扫描仪的普及导致数字图像的激增,创建了大型个人图像数据库。因为拍摄数字图片很容易并且实际上是免费的,因此用户不再将图片拍摄局限于重要的事件和特殊的场合。图像正在被频繁地捕获,并且正在用户的生活之中日复一日地发生。因为一般的用户已经积累了多年的数字图像,因此浏览集合以找出在重要事件期间拍摄的图像对于用户而言已经成为耗时的过程。

[0003] 已经在将图像分组为事件方面进行了研究。于2003年8月12日公布的、转让给A.Loui和E.Pavie的、题目为“A method for automatically classifying images into events”的美国专利No.6,606,411以及于2002年2月26日公布的、转让给A.Loui和E.Pavie的、题目为“A method for automatically comparing content of images for classification into events”的美国专利No.6,351,556公开了用于根据瞬时事件和子事件来对图像内容进行聚类的算法。根据美国专利No.6,606,411,事件具有一致的颜色分布,因此,这些图片可能是在相同的背景下拍摄的。对于每一个子事件,针对一起拍摄的所有背景区域计算单个颜色和纹理表示。上面两个专利教导了如何将数字图像集合中的图像和视频聚类为瞬时事件和子事件。在客观意义上,使用术语“事件”和“子事件”来指示计算机调节的过程的结果,其中计算机调节的过程尝试与用户对特定事件发生(与事件相对应)和这些事件发生的分割(与子事件相对应)的主观认知相匹配。在于2005年7月5日公布的、转让给A.Loui、M.Jeanson和Z.Sun的、题目为“Event clustering of images using foreground and background segmentation”中公开了另一种自动地将图像组织为事件的方法。检测到的事件按照从最早到最晚的年代顺序排序在时间轴上。

[0004] 通过使用上述方法,可以通过沿着时间轴查看事件的表示而不是查看每个图像缩略图,来减少用户定位特殊事件所需的浏览量。然而,一般的用户仍然可以在历年期间生成100个以上的此类事件,并且更加多产的图片拍摄者可能很容易就超出几百个检测的事件。这些事件中的很多事件描绘了日复一日的活动,而不是用户识别的重要或特殊的场合。需要创建较小的重要或特殊事件(在本文中,表示为重要事件)集合,这使得用户易于浏览其集合的概况。此外,需要针对特定用户的图片拍摄行为定制重要事件。例如,除了在特殊场合期间很少拍摄任何图片的人应当将其大多数图像包括在重要事件中;而习惯每天拍摄很多图片的人可以将所有捕获的图像中的一小部分包括在重要事件中。

发明内容

[0005] 根据本发明,提供了一种用于确定数字图像集合中的重要事件的方法,包括:使用处理器来执行以下操作:

[0006] (a)根据所述图像集合生成图像计数的时间序列；

[0007] (b)计算所述图像计数的时间序列的模型；以及

[0008] (c)使用所述图像计数的时间序列和所述模型来确定重要事件。

[0009] 每个用户具有典型的图片拍摄行为；并且重要事件通常是不符合该典型行为的事件。在本发明中，将用户的图片拍摄行为转换为时间序列，其中，针对每一个时间步长存在一个数据点。然后，使用时间序列建模技术来以不同的时间粒度对该数据进行建模。该模型描述了用户的典型图片拍摄行为，并且随着用户而改变。重要事件是由实际数据与适合用户数据的模型的输出的偏差来识别的。

[0010] 优点

[0011] 图像和视频的组织 and 检索对于一般用户而言是一个问题。对于用户而言，能够浏览其集合中的重要事件的概况是有用的。现有技术中公开的技术允许将集合中的图像分类为事件，而不具有确定这些事件的重要意义或重要性的能力。因此，这些包括无趣的或常见的日复一日的事件，这使事件的数量膨胀到即使当浏览事件列表时也难以找到重要事件的程度。本发明对集合中检测到的、具有较高概率成为重要事件的事件的较小子集进行识别。还针对用户的图片拍摄行为定制对重要事件的检测。

附图说明

[0012] 图1是实现本发明的系统的框图；

[0013] 图2是由图1实现的方法的整个流程图；

[0014] 图3是图2的方框110中所示的时间序列发生器的更详细的流程图；

[0015] 图4是图2的时间序列建模方框115的更详细的流程图；

[0016] 图5(a)、(b)和(c)示出了根据本发明产生的图像计数的时间序列和分析的具体例子；

[0017] 图6是图2的方框120中所示的重要事件检测器的更详细的流程图；以及

[0018] 图7是根据本发明的时间粒度选择方法的流程图。

具体实施方式

[0019] 如本领域技术人员众所周知的，本发明可以在计算机系统中实现。在下面的描述中，本发明的一些实施例将被描述为软件程序。本领域技术人员将容易认识到，这种方法的等同形式也可以被理解为本发明的范围内的硬件或软件。

[0020] 因为图像处理算法和系统是众所周知的，因此该描述将特别针对形成根据本发明的方法的一部分或者与所述方法更直接合作的算法和系统。可以从本领域已知的这些系统、算法、组件和元件中选择在本文中未明确示出或描述的这些算法和系统的其它方面以及用于产生或者另外处理与之相关的图像信号的硬件或软件。考虑到在下面的说明书中阐述的描述，其所有软件实现都是常规的，并且属于这些领域中的普通技能。在其余的描述中，集合中的视频包括在术语“图像”中。

[0021] 本发明可以在计算机硬件和计算机化的装备中实现。例如，可以在数码照相机、多媒体智能电话、数字打印机中、在互联网服务器上、在公用电话亭上以及在个人计算机上执行该方法。参照图1，示出了用于实现本发明的计算机系统。虽然为了说明优选实施例的目

的而示出该计算机系统,但是本发明不限于所示的计算机系统,而是可以在例如数码照相机、家用计算机、公用电话亭或任何其它系统中找到的任何电子处理系统上使用,以处理数字图像。计算机10包括基于微处理器的单元20(其在本文中也称作处理器),用于接收并处理软件程序并且用于执行其它处理功能。存储器单元30存储用户提供的以及计算机生成的数据,该数据可以由处理器20在运行计算机程序时访问。显示设备(例如,监视器)70电连接到计算机10,用于例如通过图形用户界面显示与软件相关联的信息和数据。键盘60也连接到该计算机。如本领域中众所周知的,作为使用键盘60进行输入的替换,鼠标可以用于在显示设备70上移动选择器并且用于选择该选择器所覆盖的项目。诸如压缩光盘(CD)和DVD的输入设备50可以插入到计算机10中,以用于向计算机10和处理器20输入软件程序和其它信息。进一步地,如本领域中众所周知的,可以对计算机10进行编程,以用于在内部存储软件程序。此外,可以通过诸如存储器卡、拇指驱动器、CD和DVD的输入设备50或者通过将捕获设备(例如,照相机、手机、录像机)作为输入设备直接连接到计算机10,来将媒体文件(例如,图像、音乐和视频)传送到计算机10的存储器30。计算机10可以具有到诸如局域网或因特网的外部网络的网络连接,例如,电话线或无线连接80。可以通过网络连接从其它计算机或因特网向该计算机传送软件程序和媒体文件。

[0022] 还应当注意的是,本发明可以在软件或硬件的组合中实现,而不局限于物理连接的或位于相同的物理位置内的设备。图1中所示的设备中的一个或多个可以位于远端并且可以经由网络连接。所述设备中的一个或多个可以例如通过射频链路直接地或经由网络无线连接。

[0023] 参照图2,用户的数字图像集合105位于计算机10的存储器30中。附图中的其它方框是通过软件程序来实现的并且由计算机10的处理器20来执行。将数字图像集合105提供给时间序列发生器110。图3示出了时间序列发生过程中的步骤。从EXIF元数据中提取205图像的捕获日期和时间信息,其中所述EXIF元数据是由捕获设备(例如照相机)存储在数字图像集合105的图像文件中的。确定一组时间单元,这些时间单元可以是几个月(捕获一个季度)、一个月、一周、一天或几小时,单元的大小被称作粒度。因为针对每个时间单元生成一个图像计数,因此所选择的时间单元的范围受到图像集合的大小的限制。因为需要至少40至50个数据点来产生可靠的结果,因此将跨越一年的集合限于最大粒度为一周;并且季度级别的粒度将需要跨越10年或更长时间的集合。在优选的实施例中,所使用的最精细的粒度是一天的几个部分(早上、下午、晚上),其中,最大粒度是由图像集合的大小来确定的。推荐使用至少一年的用户集合。累加器215对每个时间单元中的图像的数量进行计数,其中集合的开始位于单元0,从而产生一组图像计数的时间序列225。

[0024] 参照图2,对图像计数的时间序列进行分析以找出适合图像计数的模型115。图4示出了用于生成适当模型的时间序列建模中的步骤。存在很多用于时间序列建模的行之有效的方法(参照“Introduction to Time Series and Forecasting”、Brockwell和Davis、Springer-Verlag 2002)。图像计数的时间序列通常是不固定的(即,其均值和标准偏差可能随着时间而改变)。因为图片通常是以组为单位拍摄的,例如,在假期和家庭聚会期间可能存在连续几天进行图片拍摄,并且在工作日期间可能存在连续几天的零图像计数,因此,模型需要包括捕获与先前的数据值的关系的自回归组件。该模型还需要包括可以捕获图片拍摄行为中的逐渐改变的移动平均组件。适合于这种情形的模型是ARIMA(自回归综合移动

平均)模型(上面提及的Brockwell和Davis的文献的第179至187页)。模型ARIMA(p,d,q)具有三个主要参数-p是自回归分量的阶,q是移动平均分量的阶,d是处理与稳态的偏差所需的差分的阶。ARIMA(p,d,q)模型由下式给出:

[0025]

$$\left(1 - \sum_{i=1}^p \phi_i L^i\right) (1-L)^d X_t = \left(1 + \sum_{j=1}^q \theta_j L^j\right) \varepsilon_t \quad (1)$$

[0026] 其中,L是滞后算子, ϕ 是模型的自回归部分的参数, θ 是移动平均部分的参数,而 ε 是误差项。通常假设误差项是从具有零均值的正态分布中采样的独立的、同分布的(iid)变量。

[0027] 在优选的实施例中,对于每个图像计数的时间序列305,第一步骤是估计时间序列模型310的阶。根据数据的自相关图(ACF)和偏自相关图(PACF)来确定p和q的初始值(上面提及的Brockwell和Davis的文献的第94至95页)。如果ACF指数地衰减至零或者在衰减至零的同时显示出阻尼振荡,则将p的初始值选择为比滞后小的值,在该值处PACF下降至零,其中,通常考虑在零附近的95%的置信区间带而不是刚好零值。如果ACF显示出一个或多个尖峰,则将q的初始值选择为比滞后小的值,在该值处ACF变为零(通常,在零附近的95%的置信区间带内)。根据对大量用户的集合的实验,在该域中未发现指示需要更高阶的差分的严重非平稳的情况,即ACF未衰减至零或者衰减得非常缓慢。因此,d的初始估计被设置为1。图5中示出了例子。图5(a)示出了覆盖一年的图像计数的时间序列,其中将历日作为时间单元。图5(b)示出了ACF图,图5(c)示出了PACF图。实的垂直线示出了95%的置信区间。这两个图都示出了直到滞后为2的重要值,并且在此之后都衰减为零。根据该信息,p和q的值被设置为1。d的值处于默认水平1。

[0028] 参照图4,针对图像计数的时间序列来拟合315一组ARIMA模型,以包括阶参数的初始估计周围的变化。生成具有所有组合p+1/p/p-1、d+1/d/d-1、q+1/q/q-1的模型。注意,所有这些组合可以不是严格的ARIMA过程,如果一些阶为零,则它们可以退化为AR、MA或ARMA,例如,d=0对应于ARMA(p,q)模型。模型拟合过程确定方程(1)中的 ϕ 和 θ 的值。注意,因为存在p个 ϕ 参数、q个 θ 参数并且额外的参数是误差项 ε 的标准偏差,因此需要被确定的参数的数量等于p+q+1。模型拟合过程是在大多数商业可用的统计分析软件包(例如,SAS公司的JMP)中实现的。

[0029] 实际上,诸如ARIMA的混合模型(其包括自回归分量和移动平均分量)难以识别并包括许多试验和误差。可能必须丢弃在315中根据较差的参数估计或对数据的较差拟合而拟合的一些模型。这是在步骤320中基于以下两个核查完成的:(1)在315中拟合的ARIMA模型的模型参数中的一些可以具有小于2的t比率,这指示参数值为零的概率大于5%。这指示由拟合过程生成的参数值的相对较低的置信,并且丢弃这些ARIMA模型;(2)计算在315中拟合的模型与图像计数的时间序列之间残差。残差被定义为由模型预测的值与该时刻的实际值之间的差。在理想情况下,在模型很适合数据的情况下,残差是近似iid(独立的且同分布的)。如果对于所有的滞后,残差都是iid,则残差的ACF图将具有处于95%置信水平内的值。如果对于40个滞后中的3个以上的滞后不是这样的,或者如果一个处于95%边界之外的远端,则可以拒绝iid假设(上面提及的Brockwell和Davis的文献的第166至167页)。在该情况下,ARIMA模型没有很好地适合数据,并且丢弃该ARIMA模型。

[0030] 其余的模型均适合于数据并且是可行的。可以基于用于确定适合度的多个准则来识别最佳模型(上面提及的Brockwell和Davis的文献的第171至174页)。在优选的实施例中,将Akaike信息准则(AICC)用作适合度测量,如上面提及的Brockwell和Davis的文献的第171页中所定义的。为步骤320以后剩余的ARIMA模型来计算325该准则。选择330显示出最佳(最低)AICC值的模型。

[0031] ARIMA模型广泛用于预测经济和金融市场,并且对于本领域技术人员而言公知的是,大多数商业可用的统计分析软件包包括用于通过指定p、d和q的值来拟合ARIMA模型的工具(例如,SAS公司的JMP、业务预测系统公司的自动预测系统和预测专家的Autobox)以及用于分析ACF图和PACF图并且执行上面提到的测试的工具。

[0032] 参照图2,将步骤330中选择的ARIMA模型提供给重要事件检测器120,该重要事件检测器120用于确定集合中的重要事件。图6示出了重要事件检测器120中遵循的步骤。参照图6,将所选择的ARIMA模型410的预测输出与用于生成模型的图像计数的时间序列405进行比较。在每个时间步长,将残差计算415模型的预测输出与图像计数的时间序列之间的差。计算残差的方差,并且根据该方差来确定420阈值。在优选的实施例中,使用的阈值为 $\sigma/3$ 。识别430残差的绝对值大于阈值的时间步长,作为“所关注的时间步长”。通过合并相邻的所关注的时间步长并且通过仅保留图像计数大于最小阈值的时间步长或者合并的一组时间步长,来识别440重要事件。在优选的实施例中,该最小阈值是图像计数的时间序列的平均图像计数。

[0033] 参照图7,步骤225中生成的图像计数的时间序列中的每一个通过时间序列建模115和重要事件检测120,以产生不同时间粒度的重要事件510。使这些重要事件510对于时间粒度选择器530是可用的,该时间粒度选择器530基于额外的输入520来选择要使用的一组重要事件。额外的输入可以包括用户操作、系统需求或用户偏好。在浏览应用中,可以按用户选择观看集合的时间粒度来选择重要事件。例如,如果用户正在观看较短时间跨度的单日,则显示出最精细粒度的重要事件;而如果用户正在观看超过五年的集合,则周时间粒度的重要事件是合适的。在显示能力方面的系统需求也可以指定重要事件的数量,并且因此指定所选择的粒度。例如,如果约10个事件将适合显示,则粒度被选择为使得重要事件的数量接近于这个数量。用户也可以设置以某个粒度来观看重要事件的偏好。

[0034] 部件列表

[0035] 10 计算机

[0036] 20 处理器

[0037] 30 存储器

[0038] 50 输入设备

[0039] 60 键盘

[0040] 70 显示设备

[0041] 80 网络连接

[0042] 105 数字图像集合

[0043] 110 时间序列发生器

[0044] 115 时间序列建模步骤

[0045] 120 重要事件检测器

[0046]	205	提取日期/时间步骤
[0047]	215	用于不同时间单元的累加器
[0048]	225	图像计数的时间序列组
[0049]	305	图像计数的时间序列
[0050]	310	估计初始参数的步骤
[0051]	315	拟合ARIMA模型的步骤
[0052]	320	选择可行模型的步骤
[0053]	325	计算合适度测量的步骤
[0054]	330	选择最佳ARIMA模型的步骤
[0055]	405	图像计数的时间序列
[0056]	410	ARIMA模型
[0057]	415	计算残差的步骤
[0058]	420	确定阈值的步骤
[0059]	430	识别所关注的时间步长的步骤
[0060]	440	识别重要事件的步骤
[0061]	510	重要事件
[0062]	520	额外的输入
[0063]	530	时间粒度选择器

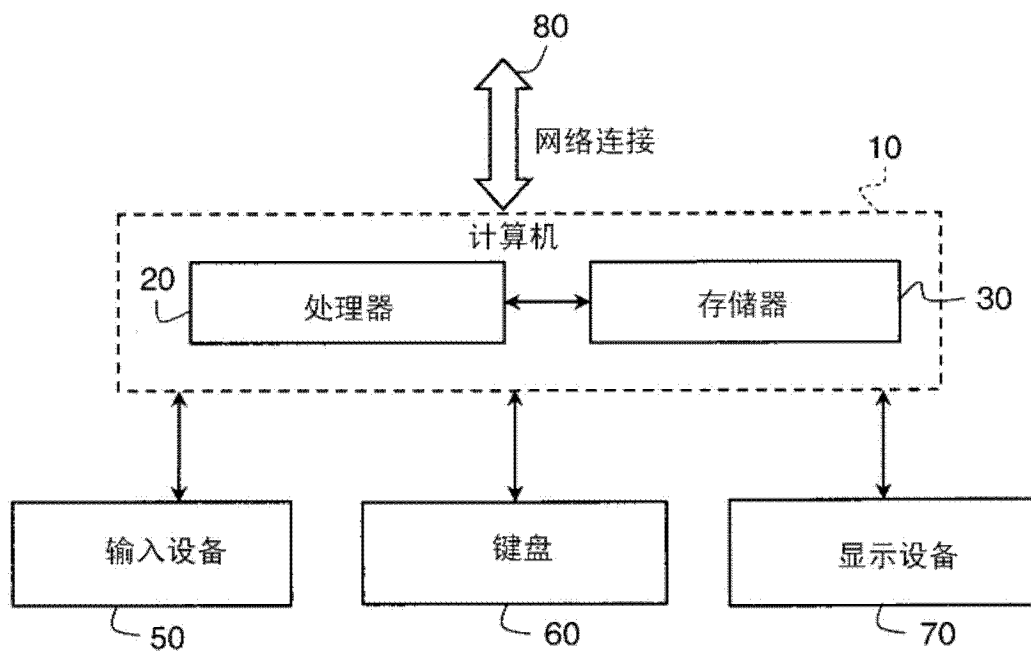


图1

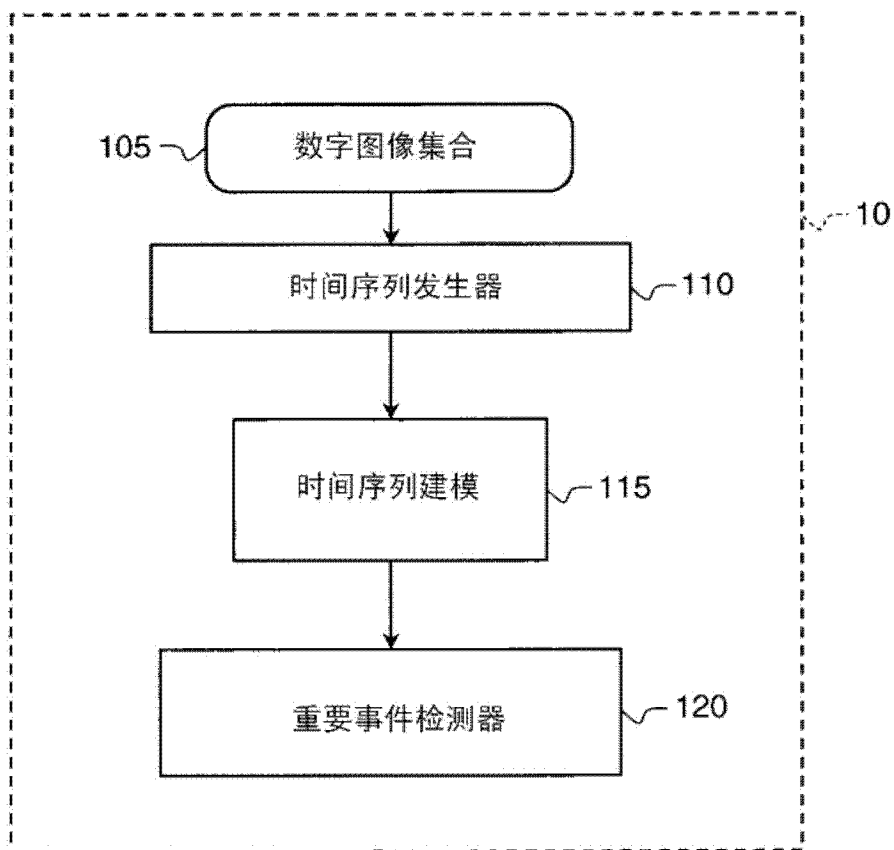


图2

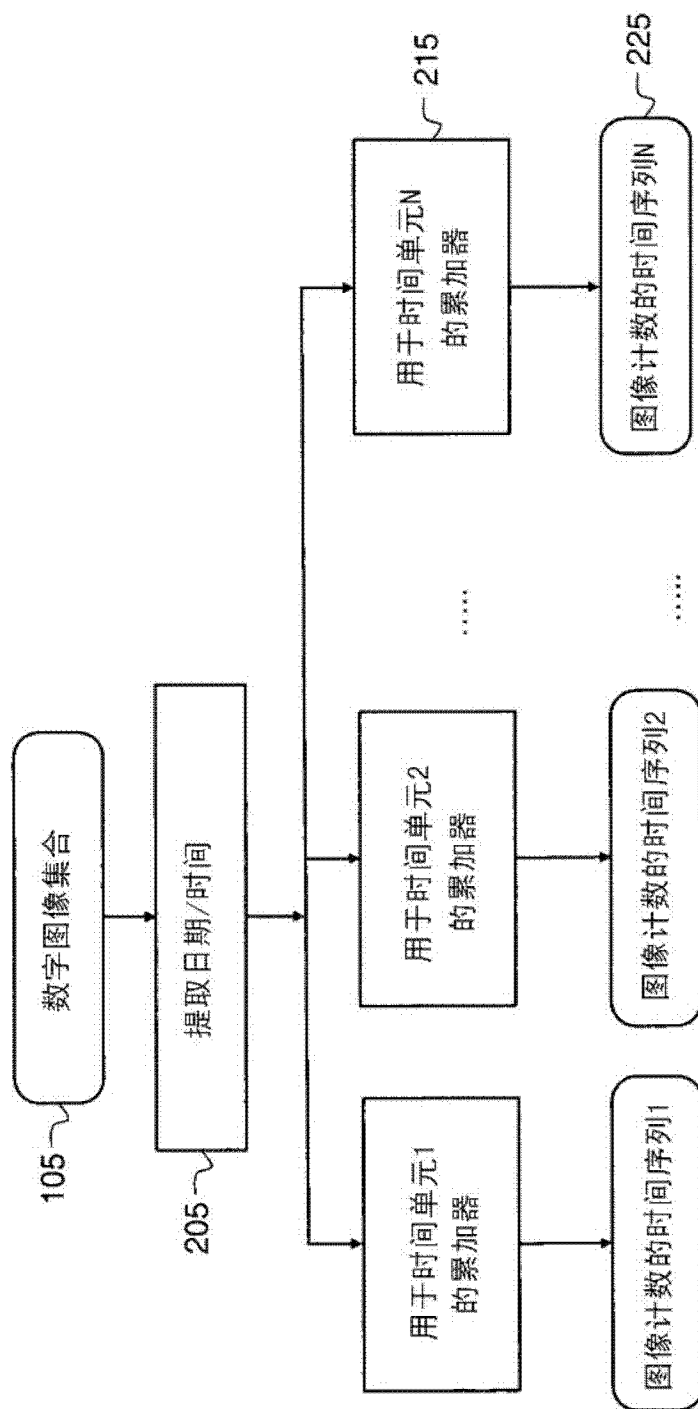


图3

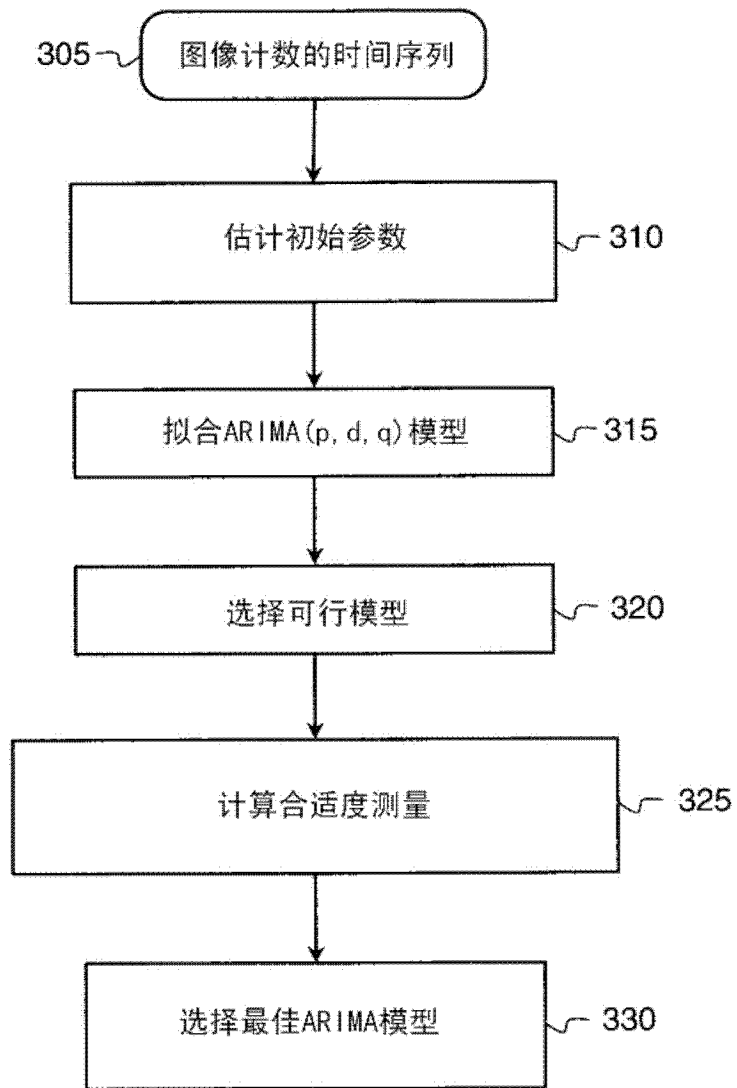


图4

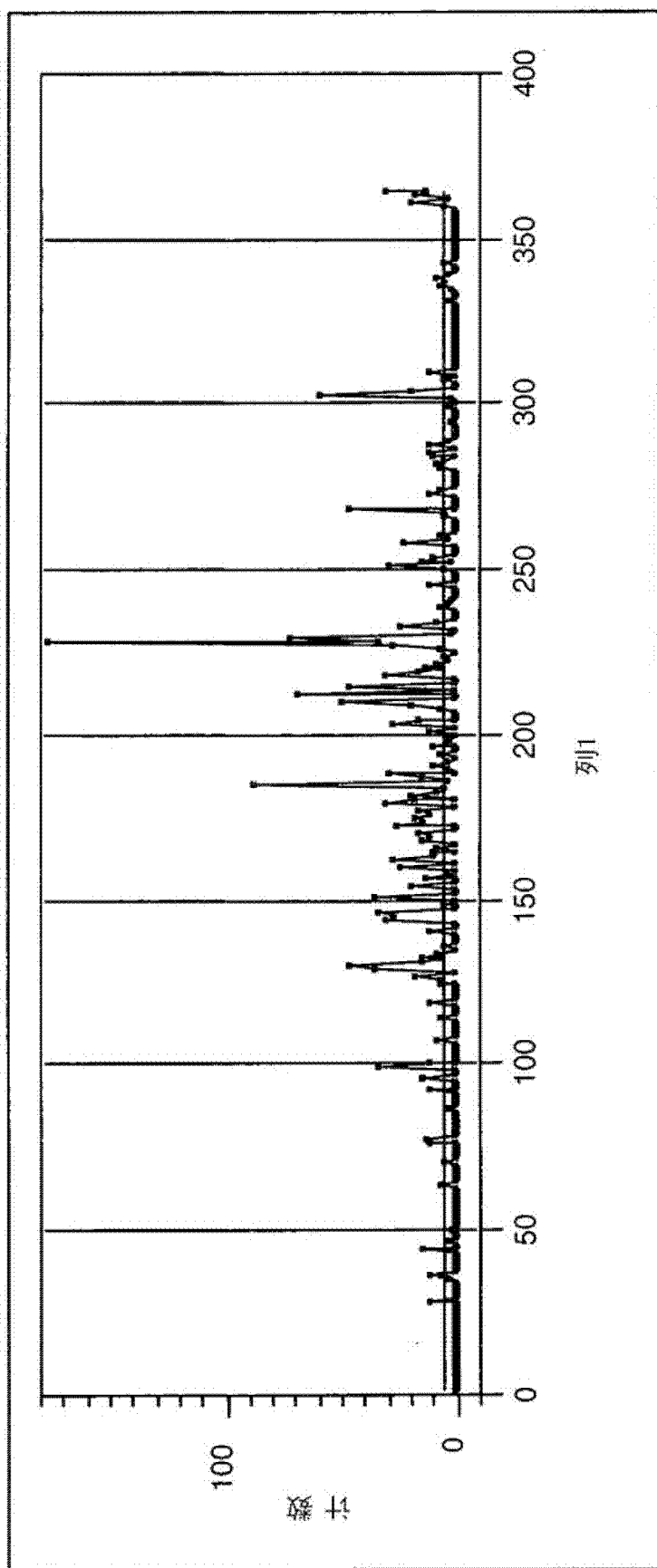


图5(a)

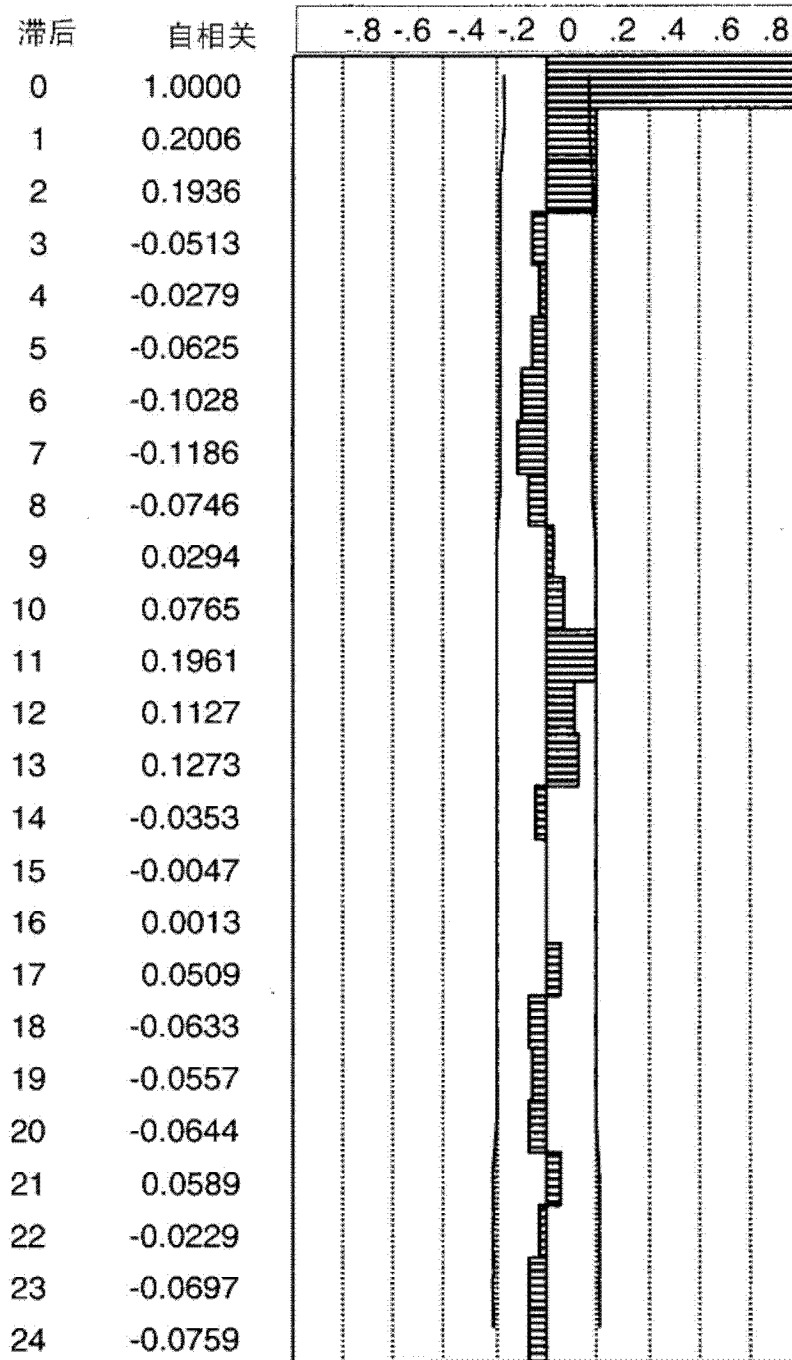


图5(b)

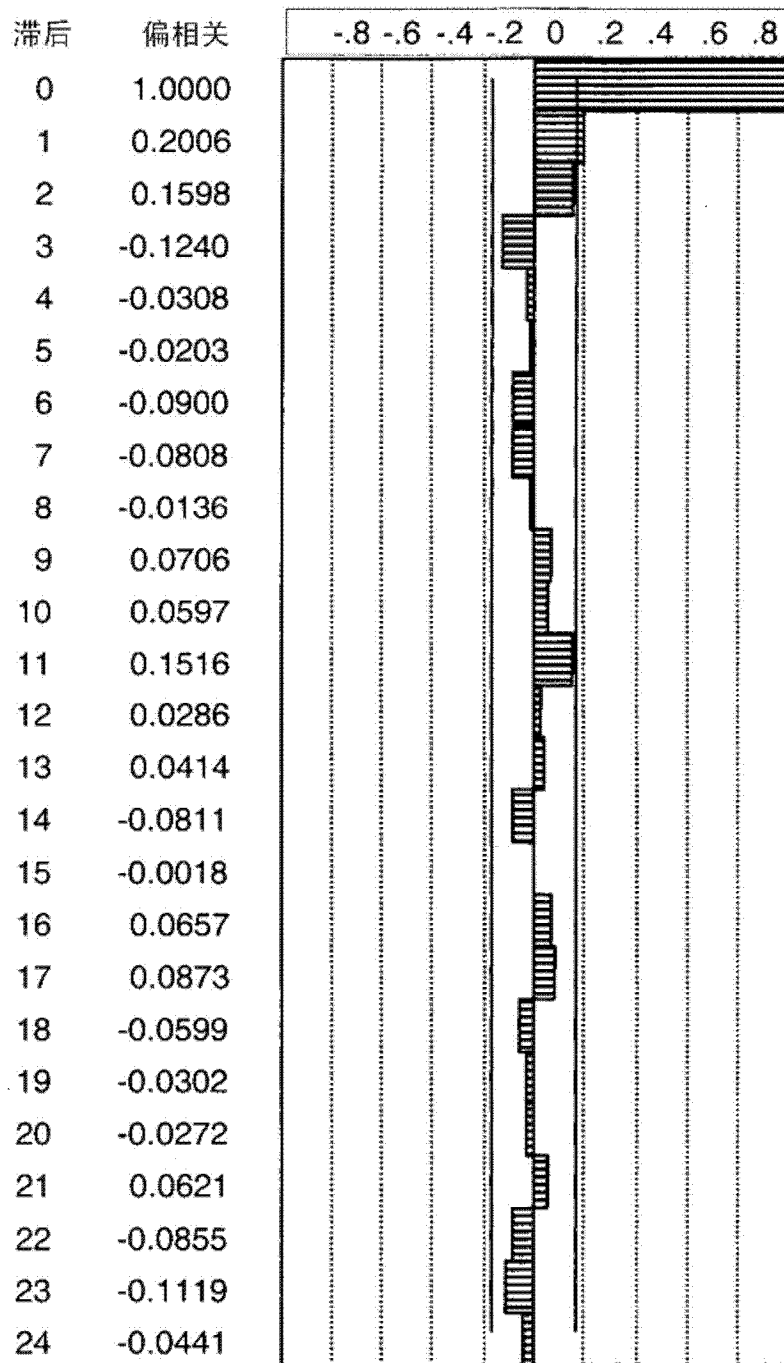


图5(c)

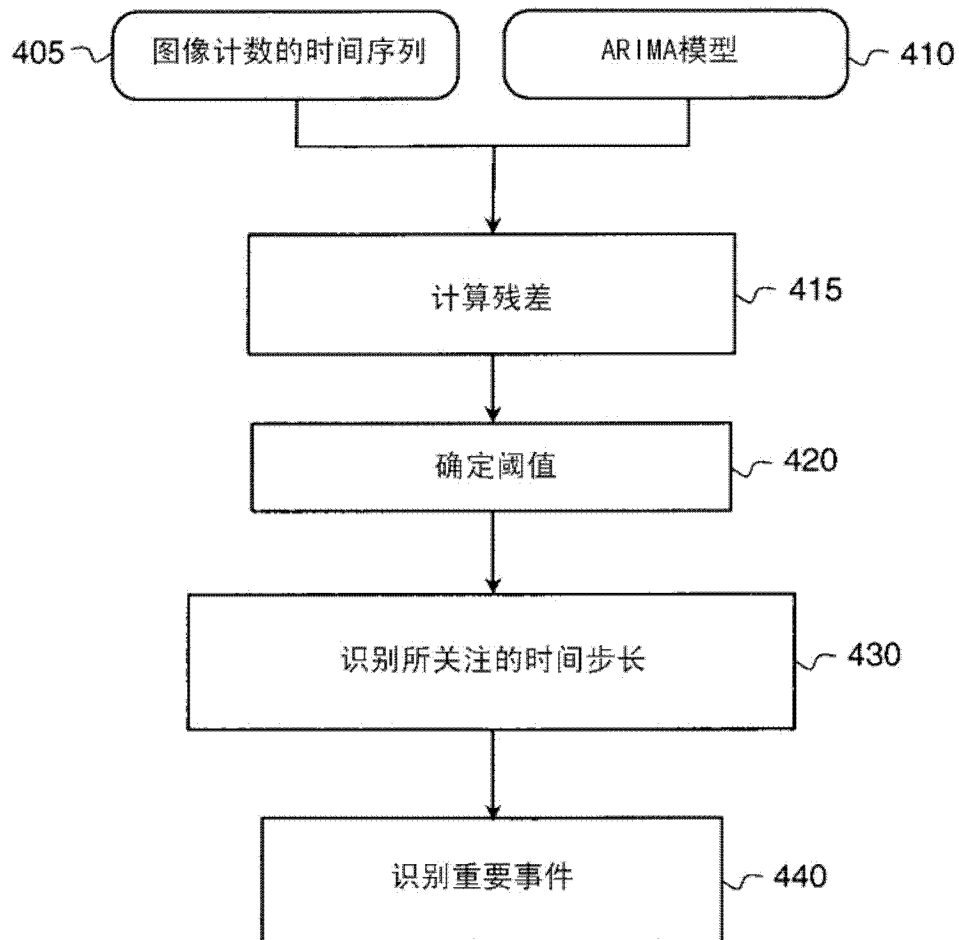


图6

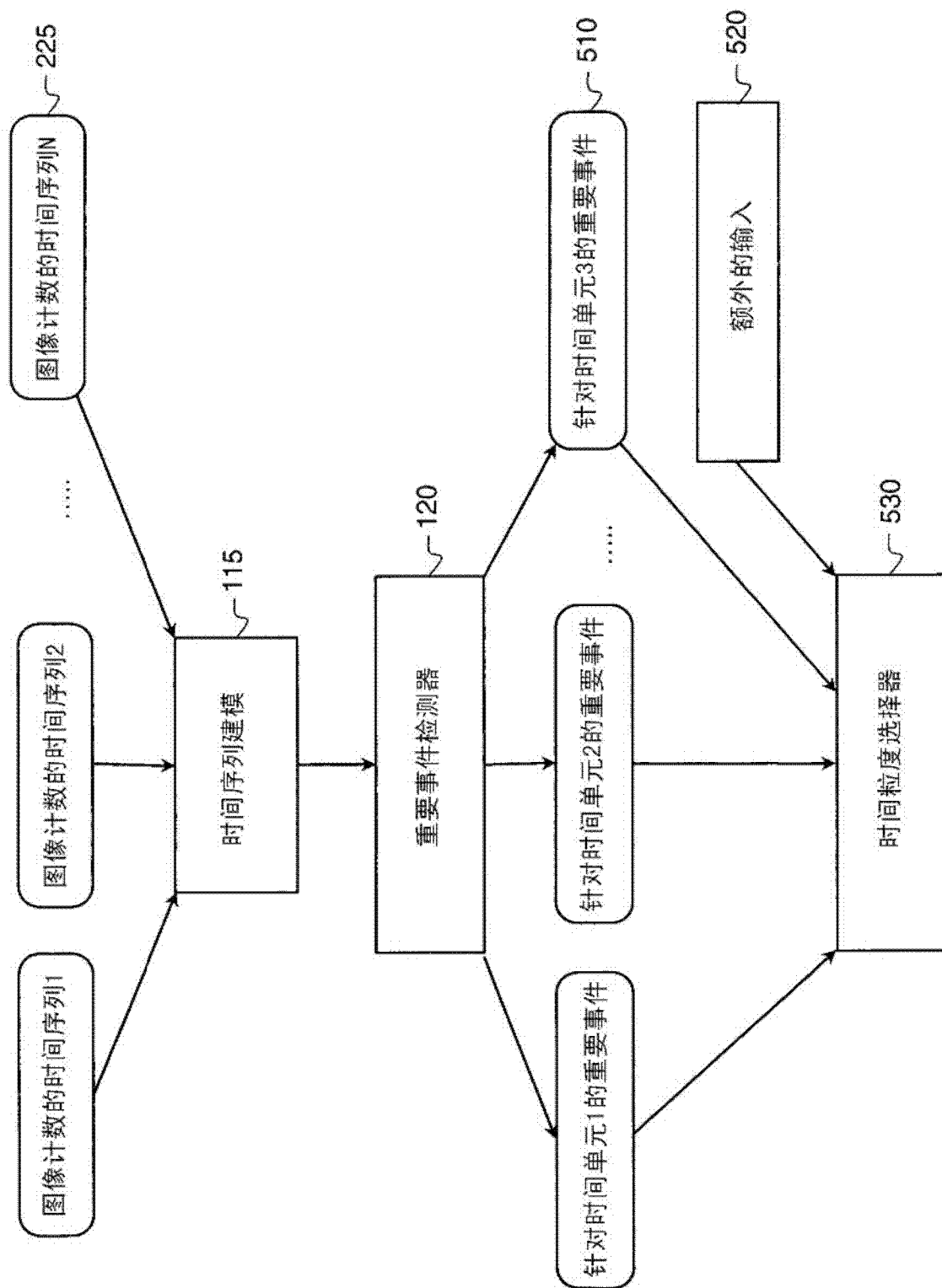


图7