[19] 中华人民共和国国家知识产权局

[51] Int. Cl[7]
H04N 5/445

[12] 发明专利申请公开说明书

[21] 申请号 03814700.9

[43] 公开日 2005 年 8 月 31 日

[11] 公开号 CN 1663245A

[22] 申请日 2003.6.10 [21] 申请号 03814700.9

[30] 优先权

[32] 2002. 6.25 [33] US [31] 10/179,351

[86] 国际申请 PCT/IB2003/002558 2003.6.10

[87] 国际公布 WO2004/002137 英 2003.12.31

[85] 进入国家阶段日期 2004.12.23

[71] 申请人 皇家飞利浦电子股份有限公司

地址 荷兰艾恩德霍芬

[72] 发明人 S·古塔

[74] 专利代理机构 中国专利代理(香港)有限公司

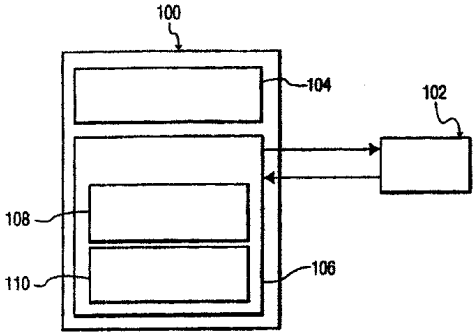
代理人 程天正 张志醒

权利要求书 4 页 说明书 7 页 附图 3 页

[54] 发明名称 通过投票进行成块的观看历史/档案的趋势分析

[57] 摘要

一种用于识别观看者历史中的趋势的方法。所述方法包括：生成表示观看者所访问的内容中的至少一个特征的出现次数的观看历史；将观看历史分成两个或多个观看历史部分，每一个观看历史部分都对应于第一预定时间间隔；将两个或多个观看历史部分中的每一部分都分成两个或多个观看历史子部分，其中两个或多个观看历史子部分中的每一部分都对应于第二预定时间间隔；估计两个或多个观看历史子部分中的每一部分的趋势；以及根据估计出的两个或多个观看历史部分中的每一部分的两个或多个观看历史子部分的趋势，来估计这两个或多个观看历史部分中的每一部分的趋势。



ISSN1008-4274

1．一种用于识别观看者历史中的趋势的方法，所述方法包括：

生成表示观看者所访问的内容中的至少一个特征的出现次数的观看历史；

将观看历史分成两个或多个观看历史部分，每一个观看历史部分都对应于第一预定时间间隔；

将两个或多个观看历史部分中的每一部分都分成两个或多个观看历史子部分，其中两个或多个观看历史子部分中的每一部分都对应于第二预定时间间隔；

估计两个或多个观看历史子部分中的每一部分的趋势；以及

根据估计出的两个或多个观看历史部分中的每一部分的两个或多个观看历史子部分的趋势，来估计这两个或多个观看历史部分中的每一部分的趋势。

2．根据权利要求1所述的方法，其中所访问的内容是电视节目。

3．根据权利要求1所述的方法，其中所访问的内容是Web站点。

4．根据权利要求1所述的方法，其中至少两个第二预定时间间隔重叠。

5．根据权利要求1所述的方法，进一步包括为两个或多个观看历史部分中的每一部分中的所述至少一个特征计算条件概率，其中估计两个或多个观看历史子部分中的每一部分的趋势是根据计算出的条件概率来进行的。

6．根据权利要求1所述的方法，其中估计出的两个或多个观看历史部分中的每一部分的趋势都基于大多数估计出的其两个或多个观看历史子部分的趋势。

7．一种用于管理推荐装置中的用户档案的存储的方法，所述方法包括：

随着时间观察用户的行为；以及

生成用户档案，作为两个或多个观看历史部分，每一部分都对应于第一预定时间间隔，其中两个或多个观看历史部分中的每一部分都被分成两个或多个观看历史子部分，每一个子部分都对应于第二预定时间间隔。

8．根据权利要求7所述的方法，其中所述用户档案与电视节目推

荐器相关联。

9. 根据权利要求 7 所述的方法，其中所述用户档案与 Web 站点推荐器相关联。

10. 根据权利要求 7 所述的方法，其中所述两个或多个观看历史子部分是连续的。

11. 根据权利要求 7 所述的方法，其中所述两个或多个观看历史子部分按第三预定时间间隔重叠。

12. 根据权利要求 7 所述的方法，进一步包括根据估计出的两个或多个观看历史部分中的每一部分的两个或多个观看历史子部分的趋势，来估计这两个或多个观看历史部分中的每一部分的趋势。

13. 一种嵌入在计算机可读介质中的计算机程序产品，用于识别观看者历史中的趋势，所述计算机程序产品包括:

用于生成表示观看者所访问的内容中的至少一个特征的出现次数的观看历史的计算机可读程序代码装置;

用于将观看历史分成两个或多个观看历史部分的计算机可读程序代码装置，每一部分都对应于第一预定时间间隔;

用于将两个或多个观看历史部分中的每一部分分成两个或多个观看历史子部分的计算机可读程序代码装置，其中两个或多个观看历史子部分中的每一部分都对应于第二预定时间间隔;

用于估计两个或多个观看历史子部分中的每一部分的趋势的计算机可读程序代码装置; 和

用于根据估计出的两个或多个观看历史部分中的每一部分的两个或多个观看历史子部分的趋势来估计这两个或多个观看历史部分中的每一部分的趋势的计算机可读程序代码装置.

14. 根据权利要求 13 所述的计算机程序产品，进一步包括用于计算两个或多个观看历史部分中的所述至少一个特征的条件概率的计算机可读程序代码装置，其中估计两个或多个观看历史子部分中的每一部分的趋势都是根据计算出的条件概率来进行的。

15. 根据权利要求 13 所述的计算机程序产品，其中估计出的两个或多个观看历史部分中的每一部分的趋势都基于大多数估计出的其两个或多个观看历史子部分的趋势。

16. 一种机器可读的程序存储装置，确实地具体化由该机器可执

行的指令程序以执行用于识别观看者历史中的趋势的方法步骤，所述方法包括:

生成表示观看者所访问的内容中的至少一个特征的出现次数的观看历史;

将观看历史分成两个或多个观看历史部分，每一部分都对应于第一预定时间间隔;

将两个或多个观看历史部分中的每一部分都分成两个或多个观看历史子部分，其中两个或多个观看历史子部分中的每一部分都对应于第二预定时间间隔;

估计两个或多个观看历史子部分中的每一部分的趋势; 以及

根据估计出的两个或多个观看历史部分中的每一部分的两个或多个观看历史子部分的趋势，来估计这两个或多个观看历史部分中的每一部分的趋势。

17. 根据权利要求 16 所述的计算机存储装置，其中所述方法进一步包括为两个或多个观看历史部分中的每一部分中的所述至少一个特征计算条件概率，其中估计两个或多个观看历史子部分中的每一部分的趋势是根据计算出的条件概率来进行的。

18. 根据权利要求 16 所述的计算机存储装置，其中估计出的两个或多个观看历史部分中的每一部分的趋势都基于大多数估计出的其两个或多个观看历史子部分的趋势。

19. 一种用于识别观看者历史中的趋势的装置，所述装置包括:

用于生成表示观看者所访问的内容中的至少一个特征的出现次数的观看历史的装置;

用于将观看历史分成两个或多个观看历史部分的装置，所述每一观看历史部分都对应于第一预定时间间隔;

用于将两个或多个观看历史部分中的每一部分都分成两个或多个观看历史子部分的装置，其中两个或多个观看历史子部分中的每一部分都对应于第二预定时间间隔;

用于估计两个或多个观看历史子部分中的每一部分的趋势的装置; 和

用于根据估计出的两个或多个观看历史部分中的每一部分的两个或多个观看历史子部分的趋势来估计这两个或多个观看历史部分中的

每一部分的趋势的装置。

20. 根据权利要求 19 所述的装置，其中所访问的内容是电视节目。

21. 根据权利要求 19 所述的装置，其中所访问的内容是 Web 站点。

22. 根据权利要求 19 所述的装置，其中至少两个第二预定时间间隔重叠。

23. 根据权利要求 19 所述的装置，进一步包括用于为两个或多个观看历史部分中的每一部分中的所述至少一个特征计算条件概率的装置，其中估计两个或多个观看历史子部分中的每一部分的趋势是根据计算出的条件概率来进行的。

24. 根据权利要求 19 所述的装置，估计出的两个或多个观看历史部分中的每一部分的趋势都基于大多数估计出的其两个或多个观看历史子部分的趋势。

通过投票进行成块的观看历史/档案的趋势分析

本发明总体上涉及趋势分析，更具体而言，涉及用于产生电视节目推荐的电视推荐器。

目前，用于构造观看档案的方法是公知的，所述观看档案从观看的和未观看的电视节目历史中反映出观看者对电视节目的偏好。例如，可从加利福尼亚州桑尼维尔市的Tivo公司购得的Tivo™系统允许观看者通过"拇指向上（thumbs up）和拇指向下（thumbs down）"特征来对表演分等级，并借此来表示观看者喜欢或不喜欢的节目。这些系统的有利特征就是：逐渐增加地更新，在所述更新中并不永久地存储观看历史。更确切地讲，不论何时将少量的新节目添加到观看历史中，都逐渐增加地更新档案。然而，使用这类系统的一个困难就是：难以实现用于忘记观看历史中的旧节目的方案，因为往往在观看者档案中丢失了观看历史的临时性质。忘记对于追踪观看者的偏好方面的变化而言是有利的。在存储器中保留全部的观看历史对存储空间来说通常成本太高了。

这类方案的一个示例就是如2000年11月15日提交的名为"Television Program Recommender With Interval-Based Profiles for Determining Time-Varying Conditional Probabilities （带有基于时间间隔的档案的用于判断随时间变化的条件概率的电视节目推荐器）"的共同待审美国专利申请09/718,256中所描述的临时性的大量构思，将该篇申请的全部内容引入于此以供参考。这篇共同待审申请利用特殊观看历史部分（或者在此称为"观看历史窗口"）中的全部观看历史来评估趋势分析。

因此，本发明的一个目的是，提供一种用于识别观看者历史中的趋势的方法，来克服与现有技术相关的缺陷。

相应地，提供了一种用于识别观看者历史中的趋势的方法。所述方法包括：生成表示观看者所访问的内容中的至少一个特征的出现次数的观看历史；将观看历史分成两个或多个观看历史部分，每一个观看历史部分都对应于第一预定时间间隔；将两个或多个观看历史部分中的每一部分都分成两个或多个观看历史子部分，其中两个或多个观

看历史子部分中的每一部分都对应于第二预定时间间隔；估计两个或多个观看历史子部分中的每一部分的趋势；以及根据估计出的两个或多个观看历史部分中的每一部分的两个或多个观看历史子部分的趋势，来估计这两个或多个观看历史部分中的每一部分的趋势。优选地，所访问的内容是电视节目或 Web 站点。优选地，至少两个第二预定时间间隔重叠。

优选地，所述方法进一步包括：为两个或多个观看历史部分中的每一部分中的所述至少一个特征计算条件概率，其中估计两个或多个观看历史子部分中的每一部分的趋势是根据计算出的条件概率来进行的。

优选地，估计出的两个或多个观看历史部分中的每一部分的趋势都基于大多数估计出的其两个或多个观看历史子部分的趋势。

还提供了一种用于管理推荐装置中的用户档案的存储的方法。所述方法包括：随着时间观察用户的行为；以及生成用户档案，作为两个或多个观看历史部分，每一部分都对应于第一预定时间间隔，其中两个或多个观看历史部分中的每一部分进一步被分成两个或多个观看历史子部分，每一个子部分都对应于第二预定时间间隔。优选地，所述用户档案与电视节目推荐器或 Web 站点推荐器相关联。

两个或多个观看历史子部分都可以是连续的或按第三预定时间间隔重叠。

优选地，所述方法进一步包括：根据估计出的两个或多个观看历史部分中的每一部分的两个或多个观看历史子部分的趋势来估计这两个或多个观看历史部分中的每一部分的趋势。

还提供了一种嵌入在计算机可读介质中的计算机程序产品，用于识别观看者历史中的趋势。所述计算机程序产品包括：用于生成表示观看者所访问的内容中的至少一个特征的出现次数的观看历史的计算机可读程序代码装置；用于将观看历史分成两个或多个观看历史部分的计算机可读程序代码装置，每一部分都对应于第一预定时间间隔；用于将两个或多个观看历史部分中的每一部分分成两个或多个观看历史子部分的计算机可读程序代码装置，其中两个或多个观看历史子部分中的每一部分都对应于第二预定时间间隔；用于估计两个或多个观看历史子部分中的每一部分的趋势的计算机可读程序代码装置；和用

于根据估计出的两个或多个观看历史部分中的每一部分的两个或多个观看历史子部分的趋势来估计这两个或多个观看历史部分中的每一部分的趋势的计算机可读程序代码装置。

优选地，所述计算机程序产品进一步包括：用于计算两个或多个观看历史部分中的至少一个特征的条件概率的计算机可读程序代码装置，其中估计两个或多个观看历史子部分中的每一部分的趋势都是根据计算出的条件概率来进行的。

优选地，估计出的两个或多个观看历史部分中的每一部分的趋势都基于大多数估计出的其两个或多个观看历史子部分的趋势。

又提供了一种机器可读的程序存储器装置，确实地具体化由机器可执行的指令程序以便执行用于识别观看者历史中的趋势的方法步骤。所述方法包括：生成表示观看者所访问的内容中的至少一个特征的出现次数的观看历史；将观看历史分成两个或多个观看历史部分，每一部分都对应于第一预定时间间隔；将两个或多个观看历史部分中的每一部分都分成两个或多个观看历史子部分，其中两个或多个观看历史子部分中的每一部分都对应于第二预定时间间隔；估计两个或多个观看历史子部分中的每一部分的趋势；以及根据估计出的两个或多个观看历史部分中的每一部分的两个或多个观看历史子部分的趋势，来估计这两个或多个观看历史部分中的每一部分的趋势。

优选地，所述方法进一步包括：为两个或多个观看历史部分中的每一部分中的所述至少一个特征计算条件概率，其中估计两个或多个观看历史子部分中的每一部分的趋势是根据计算出的条件概率来进行的。

优选地，估计出的两个或多个观看历史部分中的每一部分的趋势都基于大多数估计出的其两个或多个观看历史子部分的趋势。

又还提供了一种用于识别观看者历史中的趋势的装置。所述装置包括：用于生成表示观看者所访问的内容中的至少一个特征的出现次数的观看历史的装置；用于将观看历史分成两个或多个观看历史部分的装置，所述每一部分都对应于第一预定时间间隔；用于将两个或多个观看历史部分中的每一部分都分成两个或多个观看历史子部分的装置，其中两个或多个观看历史子部分中的每一部分都对应于第二预定时间间隔；用于估计两个或多个观看历史子部分中的每一部分的趋势

的装置；和用于根据估计出的两个或多个观看历史部分中的每一部分的两个或多个观看历史子部分的趋势来估计这两个或多个观看历史部分中的每一部分的趋势的装置。优选地，所访问的内容是电视节目或Web站点。优选地，至少两个第二预定时间间隔重叠。

优选地，所述装置进一步包括：用于为两个或多个观看历史部分中的每一部分中的所述至少一个特征计算条件概率的装置，其中估计两个或多个观看历史子部分中的每一部分的趋势是根据计算出的条件概率来进行的。

优选地，估计出的两个或多个观看历史部分中的每一部分的趋势基于大多数估计出的其两个或多个观看历史子部分的趋势。

通过论及下列说明书、所附的权利要求以及附图，将更好地理解本发明的设备和方法的这些及其它特征、方面和优点，在附图中：

图1举例说明了用于识别观看者历史中的趋势的装置的优选实施方式。

图2举例说明了图1的观看历史，所述观看历史具有至少两个观看历史部分。

图3a举例说明了图2的特定观看历史部分，所述观看历史部分具有连续的观看历史子部分。

图3b举例说明了图2的特定观看历史部分，所述观看历史部分具有重叠的观看历史子部分。

图4举例说明了用于识别观看者历史中的趋势的方法的优选实施方式。

尽管本发明适用于由观看者访问的许多的且各种类型的内容，但是人们发现它在电视节目环境中尤为有用。因此，在不将本发明的可用性局限在电视节目的情况下，将在这样的环境中描述本发明。然而，本领域的技术人员将会认识到的是，本发明的方法和装置能被用在产生诸如web档案之类的观看者档案的任何应用中。为了本公开起见，在此"观看者"和"用户"可互换使用，并且认为它们具有相同的一般含义。不应将"观看者"解释成限于在显示器上观看的内容，而是应该将其解释成访问任何类型的内容的任何人。

现在，参照图1，举例说明了一种用于识别观看者历史中的趋势的装置，总体上用参考标记100来表示。装置100可操作地连接于诸如

电视、个人电脑或PDA（个人数字助理）之类的观看终端102。优选地，将装置100配置成机顶盒，并且可操作地通过适当的电缆而连接于观看终端102，但是也可以与之集成形成所述装置。所述装置100除了其它特征以外还具有处理器104和存储器106。存储器106除了其它特征以外还包括观看历史108和观看者档案110。正像在共同待审申请09/718,256号中论述的那样，观看者档案110包含多条记录，每一条记录都与不同的节目特征相关联。

现在参照图2，观看历史108包含用户在观看终端102上所访问的内容的历史列表。所访问的内容是用户在观看终端102上所访问的任意内容，比如用户所观看（或未观看）的像电视节目或者用户访问的Web站点。正如共同待审申请09/718,256号中论述的那样，优选地将观看历史108作为一系列观看历史部分112到114加以维护，其使得不固定的观看偏好能够被识别出来。所述装置100能够将观看历史108作为一系列连续的定期（比如每隔一个月）观看历史部分112...114加以维护。无论在哪种情况下，观看历史部分112...114都相当于观看者在时间跨度$T_1...T_n$期间所访问的内容，这每一个时间间隔都小于观看历史108所涵盖的整个时间周期。因此，每个观看历史部分112...114都与第一预定时间间隔$T_1...T_n$相关联，其中n是观看历史部分的数目。尽管这里把与观看历史部分相关联的时间称为第一预定时间间隔，但是这并不意味着要将时间间隔限制于相同的时间长度，而是只用来区别于其它的时间间隔。实际上，符合观看历史部分112...114的第一预定时间间隔$T_1...T_n$可以具有不同的时间间隔或者所有都具有相同的时间间隔。正如下面将要描述的那样，最好为每个观看历史部分112...114确定条件概率，由此只需要在存储器中维护观看历史部分112...114直到计算出这个条件概率。

现在参照图3a，在该图中示出了观看历史部分112。优选地进一步将至少其中一个观看历史部分（举例来说示出的是部分112）维护到两个或多个观看历史子部分116...118中。类似于观看历史部分112...114，两个或多个观看历史子部分中的每一部分都对应于第二预定时间间隔$T_{1/n}...T_{k/n}$，其中k是观看历史子部分的数目，并且其中每个第二预定时间间隔都小于观看历史部分112的整个时间间隔。例如，如果观看历史部分112是月份，那么观看历史子部分就可以是该月中

的可选的某些天，该月中的头 20 天，甚至是该月的每一天的某些时刻。像观看历史部分 112...114 一样，如图 3a 所示，观看历史子部分 116...118 可以是连续的，或者如图 3b 所示可以是重叠的。图 3b 举例说明了具有四个观看历史子部分 116、120、122 和 118（即，k 为 4）的观看历史部分 112 的简单示例。每个观看历史子部分 116、120、122 和 118 分别符合时间间隔 $T_{1/n}$、$T_{2/n}$、$T_{3/n}$和 $T_{4/n}$。正如上面所论述的那样，对于观看历史部分 112...114 而言，与观看历史子部分相关联的时间在此称为第二预定时间间隔。然而，这并不意味着要将第二预定时间间隔限制于相同的时间长度，而是只用来区别于其它的时间间隔，比如观看历史部分 112...114 的时间间隔。实际上，就像能够在时间间隔 T_{01}、T_{02}和 T_{03}中重叠那样，符合观看历史子部分 116...118 的第二预定时间间隔 $T_{1/k}...T_{n/k}$能够具有不同的时间间隔或者所有都具有相同的时间间隔，作为选择在此称为第三预定时间间隔。

现在参照图 4，现在将论述一种用于识别观看者历史中的趋势的方法，所述方法总体上指代参考标记 200。所述方法 100 包括在步骤 202 生成观看历史。所述观看历史表示观看者所访问的内容的至少一个特征的出现次数。在本领域中，根据观看者的行为来构造观看历史是为大家所熟知的，比如像在共同待审申请 09/718,256 号中公开的那样。正如先前论述的那样，所访问的内容的示例是观看电视节目或访问 Web 站点。

在步骤 204，将观看历史分成两个或多个观看历史部分，每一部分都对应于第一预定时间间隔。例如，观看历史部分可以是一个月长或 20 个离散的内容块，比如 20 个电视节目。在步骤 206，优选地计算每个观看历史部分中的每个特征的条件概率（CP）。例如，可以根据等式 1 来计算 CP:

$$CP = \frac{\text{具有特征的时间显示的数目}}{\text{在分类中显示的数目}} \qquad (1)$$

在步骤 208，将至少一个观看历史部分和（优选地）所有观看历史部分都分成两个或多个观看历史子部分。每个观看历史子部分都对应于第二预定时间间隔。正如上面所论述的那样，观看历史子部分可以是连续的或重叠的。然后，在步骤 210 利用已知的方法来为每个观看

历史子部分估计观看趋势，比如在共同待审申请 09/718,256 号中公开把这种趋势基于先前计算出的 CP 的方法。

在步骤 212，根据估计出的两个或多个观看历史部分中的每一部分的两个或多个观看历史子部分的趋势来估计这两个或多个观看历史部分中的每一部分的趋势。例如，如果在所有观看历史子部分当中特定趋势获得最大数目的投票，那么就把所述趋势选作为观看历史部分的最有可能的趋势。投票的一个简单示例就是某个观看历史部分对应于有 31 天的月份，这个观看历史部分被分成 31 个观看历史子部分，每一个子部分都对应于该月中的一天。如果估计出 16 个观看历史子部分具有线性趋势、而估计出 15 个观看历史子部分具有循环趋势，那么就为相应的观看历史部分选择线性趋势。

正如上面所论述的那样，用于识别观看者历史中的趋势的装置 100 除了被具体化成电视的"机顶盒"以外，还可以被具体化成任何计算装置，比如像具有处理器 104 和存储器 106 的个人电脑或工作站。此外，在提供 web 档案的情况下，也可以将它具体化成 web 浏览器，由具有处理器 104 和存储器 106 的服务器来为所述 web 浏览器提供服务。

本发明的方法尤其适合于通过计算机软件程序来实施，这类计算机软件程序优选地包含与所述方法的各个步骤相对应的模块。当然，也可以将这类软件嵌入在诸如集成芯片或外围设备之类的计算机可读介质中。

尽管已经示出并描述了作为本发明的优选实施例的那些实施例，但是当然将要理解的是，在不背离本发明的精神的情况下能够容易地作出形式或细节上的各种修改和改变。因此，意味着本发明不限于所描述和举例说明的确切形式，而应该视为涵盖了可能落入所附权利要求的范围内的所有修改。

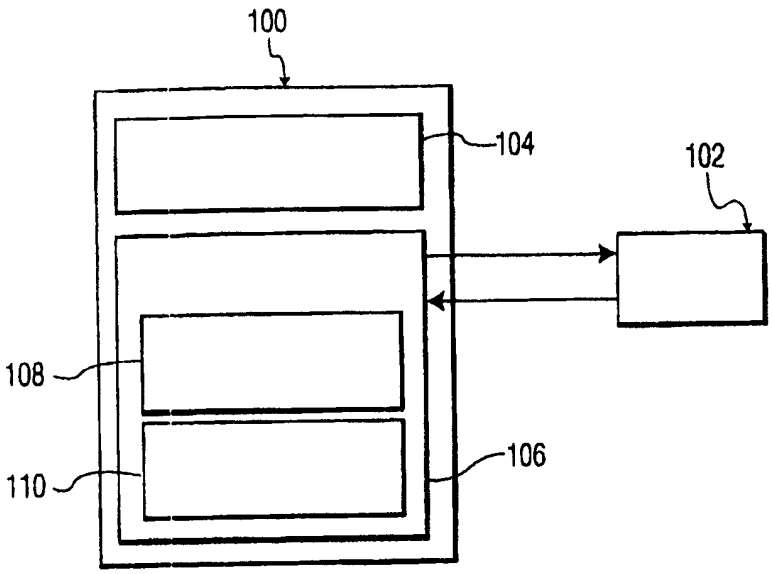


图 1

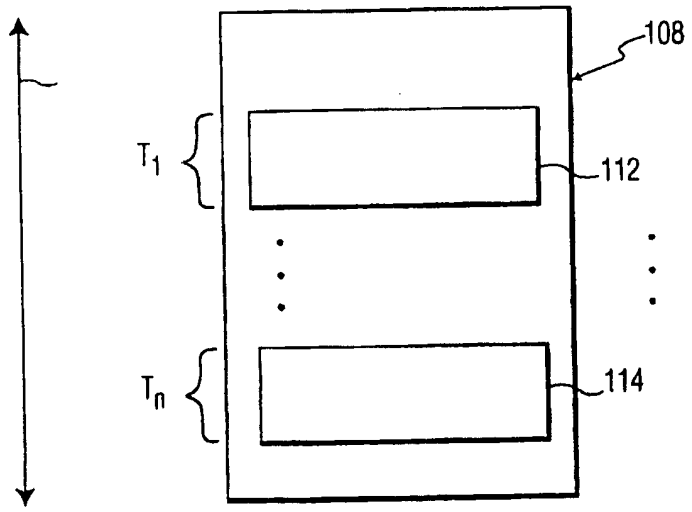


图 2

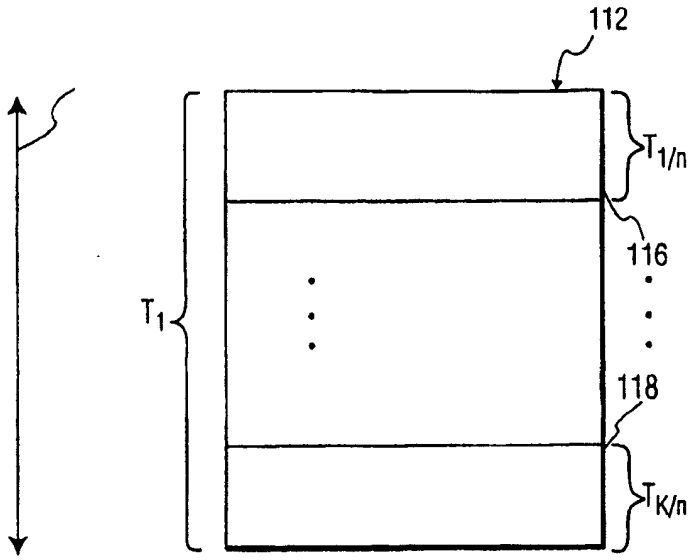

112
$T_{1/n}$
116
T_1
118
$T_{K/n}$

图 3A

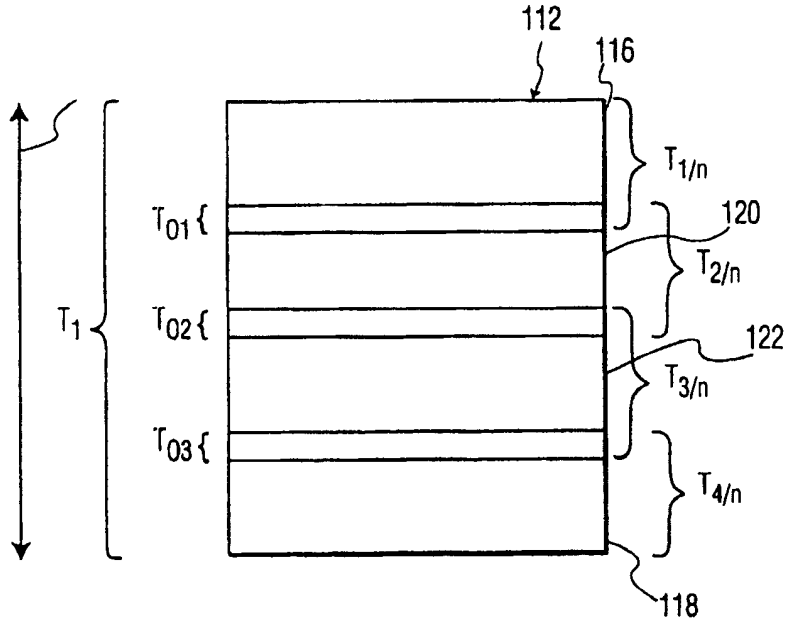

112
116
$T_{1/n}$
T_{01}
120
$T_{2/n}$
T_1
T_{02}
122
$T_{3/n}$
T_{03}
$T_{4/n}$
118

图 3B

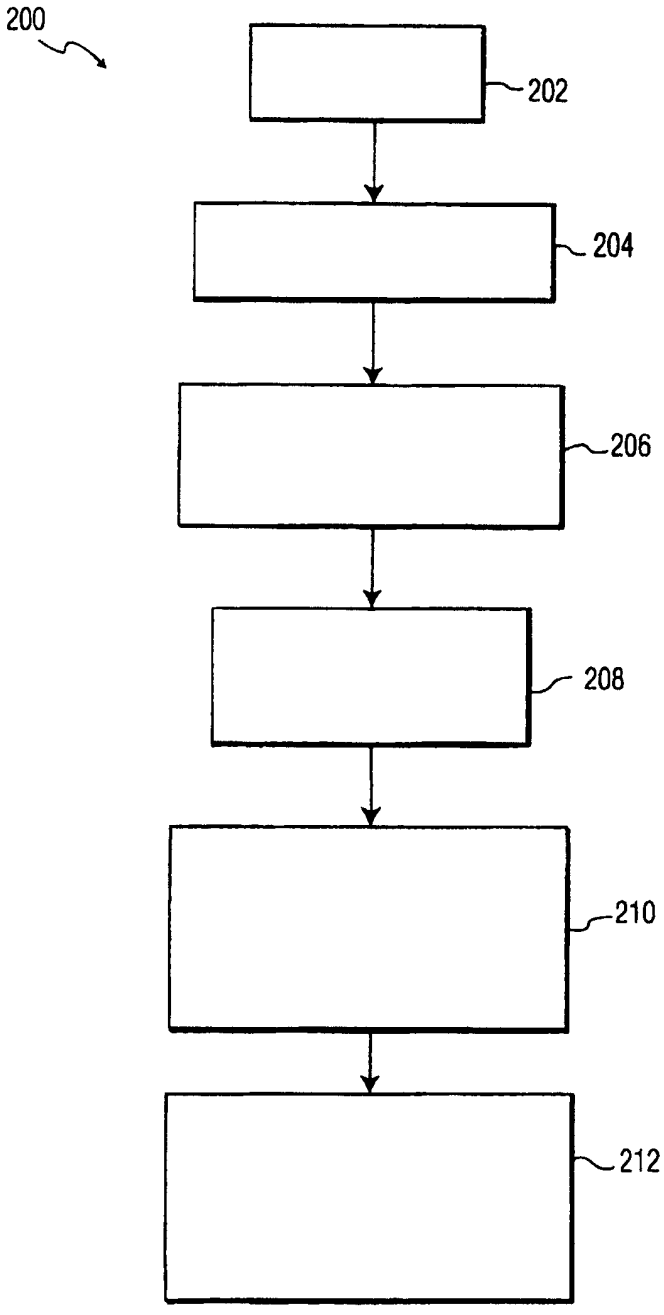


图 4