



[12] 发明专利说明书

[21] ZL 专利号 94116476.4

[43] 授权公告日 2003 年 4 月 16 日

[11] 授权公告号 CN 1105978C

[22] 申请日 1994.9.22 [21] 申请号 94116476.4

[30] 优先权

[32] 1993.9.24 [33] US [31] 126846

[71] 专利权人 美国电报电话公司

地址 美国纽约

[72] 发明人 斯蒂芬·G·艾克

保罗·J·卢卡斯

杰弗里·D·施米特

[56] 参考文献

US 5062147A 1991.10.29 G06F11/32

US 5121470A 1992.06.09 G06F15/20

审查员 黄毅斐

[74] 专利代理机构 中国国际贸易促进委员会专利

商标事务所

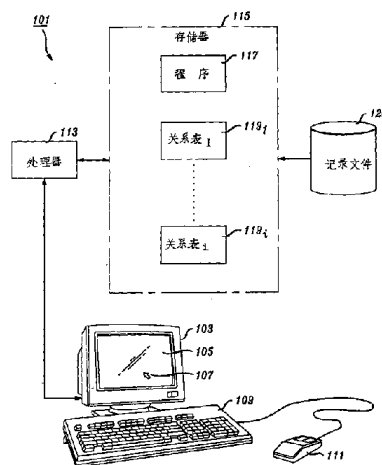
代理人 蒋世迅

权利要求书 2 页 说明书 17 页 附图 23 页

[54] 发明名称 用于显示数据的可视表示的设备

[57] 摘要

用于交互分析系统记录文件的一种装置(101)和技术。被技术人员和系统专家所监控,用以确定系统的执行、状态和软件故障的系统记录文件通常产生于各种硬件和软件的监控运行期间。每个记录文件(120)都包含时标报告。这个技术对于分析大的记录文件特别有用。一个新发表的软件可能包含很多递增版本,它需要进行测试。每个新版软件的测试可能产生一个包含数千报告的记录文件。使用这种装置(101)和技术,报告被联系起来,故障被隔离,并且临时模式比使用常规、非图形技术能更快、更有效地被识别。



1. 一种具有用于产生多个时标消息的处理器装置以及可视显示装置的设备，每个消息有一组特征，该设备的特征在于：

用于在所述可视显示装置上可视地显示多个符号的装置，其中每个符号对应于所述多个消息中的一个相应消息；

该可视显示装置显示所述符号，以便每个符号具有一种形状，该形状根据其相应消息特征组中的第一特征而变化；以及

该可视显示装置还显示所述符号，以便每个符号在一个位置显示，该位置由时标以及其相应消息特征组中的第二特征确定。

2. 根据权利要求1的设备，还包括处理装置，用于根据所述特征的值定制所述特征。

3. 根据权利要求1的设备，其中所述可视显示装置显示每个符号为可变倾斜度和可变颜色的线段。

4. 根据权利要求3的设备，其中所述可视显示装置显示每个符号为可变倾斜度、可变颜色和可视结构的线段。

5. 根据权利要求1的设备，其中所述可视显示装置显示每个符号为颜色可变的几何形状。

6. 根据权利要求1的设备，其中所述可视显示装置显示每个符号为可视结构可变的几何形状。

7. 根据权利要求1的设备，其中所述可视显示装置显示每个符号为颜色和可视结构可变的几何形状。

8. 根据权利要求1的设备，其特征在于，还包括显示对应于任何符号的所述消息的消息显示装置。

9. 根据权利要求8的设备，其特征在于，还包括用于搜寻用于特定文本模式的多个消息的装置。

10. 根据权利要求1的设备，其特征在于，还包括用于显示一个指定消息类型的消息的总发生次数的装置。

11. 根据权利要求10的设备，其中所述用于显示一个指定消息类型的

消息的总发生次数的装置显示在指定的时间间隔内发生的所述总数。

12. 根据权利要求1的设备，其特征在于，还包括用于在一个时间间隔内显示所有时标消息总数的装置。

13. 根据权利要求1的设备，其特征在于，还包括用于关掉除来自所选模块的符号外的所有符号的显示的装置。

14. 根据权利要求13的设备，其中所述选择模块是处理器模块。

15. 根据权利要求1的设备，其特征在于，还包括用于显示选择器窗口的装置，由该选择器窗口选择的与符号相关的属性值被禁止显示。

用于显示数据的可视表示的设备

技术领域

本发明一般涉及在有一个或多个处理器的系统，例如分布式系统的图形显示，特别是系统记录文件报告的图形显示。

背景技术

很多系统都生成记录文件作为其正常操作的一部分。这种文件典型地包括系统性能、系统状态和软件故障报告。这些报告通常是自由格式的文件。每个报告都是具有单独的时间标记的，用以指示报告产生的时间。通过检查记录文件，系统操作者可以发现和纠正系统和软件问题，以免这些问题影响系统的操作。

记录文件的一个共同特点就是很多不重要的报告和重要的报告一起产生。这些“噪声”报告混杂在记录文件中，使得重要的报告含糊不清。例如，一个包括很多升级版本的软件程序新版在 15 个小时的测试期间所产生的记录文件，可能包含 55,000 个报告，它由 100,000 行文字组成，相当于 1600 页。但这些报告中实际上只有几百个很重要。这种记录文件中的“噪声”报告可能遮蔽了一个或多个重要的报告，而使操作者忽视了它。

发明内容

本发明的第一个目的是提供一种用图形来分析记录文件的装置。

本发明的第二个目的是提供一种用图形显示记录文件的装置，使得分析者能够发现记录文件内的重要报告。

本发明的第三个目的是提供一种装置，它可以依照类别和发生时间显示记录文件中的报告，以使用户看到每一个报告的详情。

本发明的第四个目的是提供一种装置，它可以依照类别和发生时间显示记录文件中的报告，使得操作者能交互地浏览报告以分析报告的起因。

本发明公开了一种具有用于产生多个时标消息的处理器装置以及可

视显示装置的设备，每个消息有一组特征，该设备的特征在于：

用于在所述可视显示装置上可视地显示多个符号的装置，其中每个符号对应于所述多个消息中的一个相应消息；

该可视显示装置显示所述符号，以便每个符号具有一种形状，该形状根据其相应消息特征组中的第一特征而变化；以及

该可视显示装置还显示所述符号，以便每个符号在一个位置显示，该位置由时标以及其相应消息特征组中的第二特征确定。

本发明还公开了一种具有一个处理器以及一个随机存取存储器的设备，其特征在于，它还包括一个大容量存储装置，在该存储装置中存储有多个未处理的记录文件消息，该处理器将记录文件消息存入所述随机存取存储器中的一组关系表中；并且还包含一个显示装置，用于可视地显示从所述多个记录文件消息中得到的所述关系的非文本几何图形。

而且，本发明还公开了一种可视地提供一个具有多个时标消息的记录文件的设备，每个消息具有一组特征，其特征在于，该设备包括：用于显示消息的装置，以便根据该特征组，可视地将每个消息区分出来，从而使观看所显示消息的人很容易地识别出具有不同特征的消息。

此外，本发明公开了一种由微处理设备实现的、用于显示由系统产生的多个时标消息的方法，每个消息具有一个特征组，其特征在于，该方法包括步骤：根据时标以及特征，存储该多个消息；分配一个符号以可视地表示至少具有类似时标或类似特征的消息；以及在显示器上可视地显示每个位置上的消息，该位置是由至少其中一个时标以及消息的特征所确定。

依照对本发明的一个方面的简述，可以看到上述的目的是通过提供一种用于显示许多时间标记及具有一组特征的消息的装置和方法来实现的。这个装置包括多种符号，其中每种符号对应于一条消息。每种符号有一种显示形状，它按照其相应的消息特征以及由该消息的时间和该消息特征确定的位置而变化。

附图说明

尽管该说明书以特别指出和清楚地要求保护主题的附加权利要求书来结束，但我们相信，通过下面连同附图的描述，本发明能更容易被理解。

图 1 是记录文件分析系统例子的框图。

图 2 是在优选实施例中记录文件分析系统操作者所观察到的典型可视显示图。

图 3 与图 2 相似，区别在于时间分区更窄。

图 4 与图 3 相似，其中只给出所选择的系统测试特性的报告。

图 5 与图 3 相似，其中只给出了与系统数据库完整性检测器和依照问题计数的校正器有关的报告。

图 6 与图 3 相似，其中只给出了与系统数据库完整性检测器和依照问题编码的校正器有关的报告。

图 7 与图 6 相同，只是在显示的一部分上叠加了一个选择器窗口。

图 8 与图 7 相同，只是在显示的一部分叠加了一个浏览窗口。

图 9 是存储器中被处理的记录文件的详情。

图 10 是存储器中数据结构的详情。

图 11 是存储器中关系表的详情。

图 12 是存储器中属性名的全局区详情。

图 13 是存储器中的多元组详情。

图 14 是存储器中的选择器详情。

图 15 是本发明产生显示的过程的概况。

图 16 是生成关系表的详细流程图。

图 17 是处理多元组的详细流程图。

图 18 是依照本发明生成显示的详细流程图。

图 19 是创建图表的详细流程图。

图 20 是产生图表条线的详细流程图。

图 21 是产生时间-条的详细流程图。

图 22 是读记录文件的详细流程图。

图 23 是给记录文件上色的详细流程图。

图 24 是挑选相关符号的详细流程图。

图 25 是一个绘制多元组于屏幕上的详细流程图。

具体实施方式

现在参照图 1, 图 1 给出了记录文件分析系统例子 101 的框图。这个 101 系统包括终端 103, 它用以提供输出给系统操作者和从该处接收输入; 处理器 113, 它执行实际的分析功能; 存储系统 115, 它包括被处理器 113 执行的程序 117 和关系表 119₁ 至 119_i, 每个关系表都包括一个相应的多元组。系统 101 还有一个海量存储系统 120, 它用以存储还没有处理的记录文件, 即已建立的时标信息组。

更详细地说, 终端 103 包括显示屏 105, 处理器 113 在其上显示信息给操作者。显示屏 105 还包括指针 107, 它用以在显示屏 105 上指定一个位置, 还可以在键盘 109 或鼠标 111 的控制下移动。操作者通过从键盘 109 和/或鼠标 111 的输入, 控制系统 101 的操作。处理器 113 可以是从个人计算机到工作站或甚至巨型计算机的任何种类的处理器。存储系统 115, 包括任何可以传到系统 101 的数据, 因此它包括随机存取存储器或只读存储器。连接到存储系统 115 上的是一个海量存储系统 120, 它把数据读到存储系统 115, 使得这些数据更容易存取, 或者存储这些作为长期使用的数据。海量存储系统 120 可以是磁盘或是光盘。

当用系统 101 来分析记录文件中的信息时, 操作者可以用键盘 109 或鼠标 111 作为输入设备。处理器 113 执行程序 117, 用来完成分析关系表 119₁-119_i, 把它们的结果显示在屏幕 105 上。操作者可以使用键盘 109 和/或鼠标 111 更细致地交互检查这些结果。

本系统 101 的优选实施例在 UNIX 操作系统下运行(UNIX 是 UNIX 系统实验室的注册商标), 它使用有 X 窗口系统的工作站。

参看图 2, 显示 201 呈现在屏幕 105 上。显示 201 是在开发期间测试 5ESS 分布系统 15 个小时所得到的记录文件(5ESS 是 AT&T 公司的注册商标)。这个记录文件包括超过 55,000 个报告, 它由 100,000 行文字组成。显示 201 有四种报告形式: 判断, 即对软件/数据不一致的检测; 审查, 即系统数据库完整性的检测器和校正器; 运行和维护报告, 即硬件组件的调换、诊断、修复和清除处理的过程; 以及在通信建立期间中继错误的报告。

为使分析记录文件的显示 201 更为容易, 第一步是只选择那些有“兴

趣”的报告。这里我们所说的“兴趣”是指那些能够表明影响服务或可能影响服务的事件和软件错误的报告。记录文件中的“噪声”报告仅仅通过处理或显示过程中不选择它们而把它们滤掉。考虑到记录文件中，很多报告是由跟踪报告(如栈帧，栈迹和寄存器转储)而产生的。从分析出发，这些“噪声”报告既无益于系统操作者对模式和相互关系的发现，又势必使那些有益于系统的报告模糊不清。

为使分析记录文件的显示 201 更容易的第二步是利用它的临时变量。换句话说，就是用每个报告的时标作为它在显示 201 上的位置的一个坐标。先前的基于文本的分析技术，像那些使用可视文本编辑器的技术，由于有兴趣的报告的空间分离，使时标记录文件的固有性质模糊不清，即使用文本编辑器时遇到的那些分离它们的行的数量与产生这些行所需的时间没有什么关系。例如，一个五分钟周期的记录文件可由几百行的记录来表示，而另外一个五分钟周期的记录文件可由几千行来表示。

按照本发明，显示 201 的显示技术有置放在栅格里的有角度的勾号标记。举一个例子，勾号标记 202 指明一个报告的发生，这个报告是有关在纵轴 204 处的判据 21101。勾号标记 202 的倾斜度和颜色被译成码，这将在下面解释。沿着栅格的纵轴 204，报告按种类和类型被分别分成 206、208、210 和 201 带及行。206~212 的每个带都是由不同种类的报告组成的。带 206 由判断报告行组成，带 208 由审查报告行组成，带 210 由运行和维护报告行组成，而带 212 由中继错误报告行组成。在一个给定的类内，每行都是由一种类型的报告组成。在纵轴 204 的每行的左侧印有类型名称，而发生的总数是以条状图的形式显示在显示屏右侧的其对应行的结束处。用滑动光标 226 可以调节条线图，最长的指示条已用标度尺截短，以浅灰色显示。显示 201 可以同时显示屏 105 上显示近 70 行，这包括类型各行和用于带之间的分离线的行。横轴 220 表示时间。时间向显示 201 的右边增长。一个给定类型报告的发生，通过在水平位置上沿着对应于这种类型的行画勾号来标记，它们对应于用时标表示的各自出现时间。单位时间内所有报告类型所发生的总数，以从第二横轴 221 方向上延伸的叠加矩形图的形式表示在显示屏底部，第二横轴也表示时间。滑动光标(未示

出) 还调整矩形图的尺寸。

关系数据模型用于显示 201 的显示数据。带 206~212 的每个类对应于单个关系表 119_1-119_i (示于图 1)。关系表的主要特性的唯一值对应于一个类中的不同类型。在显示 201 的左侧, 有一个相互影响的彩色标尺 224。相互影响的彩色标尺 224 用于通过所有关系表 119_1-119_i 的从属特性 226 的数值来产生颜色编码报告。一组从属特性 226 包括装置、误差计数、误差编码、事件, 从属特性 226 的信令组包括类型、处理器识别, 以及主干组类型, 这些在显示 201 的左下角以缩写方式显示。图 2 中, 按照它们各自处理器模型, 灰度标尺阴影的不同代表着不同报告的颜色编码, 分为与报告有关的 AM(未在图 2 中标记), SM21, SM23, SM25, SM47 和 SM48。报告的颜色编码是: 浅兰色用于 SM21, 深绿色用于 SM23, 浅绿色用于 SM25, 黄色用于 SM47, 红色用于 SM48。除了使用颜色的方式外, 还借助勾号标记对报告发生的特征进行编码, 勾号标记的倾斜程度也可用于编码可能的不同特征。采用颜色和倾斜度是互相补充的。如果有很多数值, 相邻的色调可能太近使得操作者难以分辨。编码显示 201 以便颜色和倾斜度编码每个独立的特征, 使得相近的数值, 即使在颜色上相近, 但角度上的不同也能使操作者区分它们。显示 201 使用预先定义的有 6 个倾斜角度的设定, 由于有足够多的差别, 使得操作者能区分它们。如果有多于 6 个特征必须进行编码, 那么可再用相同倾斜度。这是因为当倾斜度编码方式“环绕一周”时, 颜色编码方式已经发展到有足够多的不同色彩去分辨它们之间的差别。

从横轴 221 伸出的叠加矩形条也是用颜色编码的, 以指示哪个特征对应于在矩形条的时间周期内生成的报告。这是用叠加矩形图的不同灰度标尺阴影来表示的。这样, 通过观察显示 201 的叠加矩形图, 操作者可以看到系统 101 的执行开始在第 11 小时时退化。同样通过观察在叠加矩形图中深灰的比例(在相关的颜色图中为深绿色), 操作者可以得出结论, 很多的报告是在处理器模块 SM23 中发生的。

参照图 3, 显示 301 除了横轴 321 被分成 5 分钟的间隔而不是横轴 201 的 1 小时的间隔外, 和显示 201 是一样的。这表明, 从横轴 321 伸出的矩

形图将表示 5 分钟的时间周期，而不是 1 个小时的周期。使用这种更细的时间分隔，操作者现在就可揭示在测试 1.5 小时以后开始的报告活动的峰值，并且大约每隔 20 分钟重复一次。

观察行末的条线图，可以看出多数报告活动发生在带 308 的审查类中，而运行和维护报告发生在带 310 中。按照条线图的长度，出现最频繁的审查类名称为：PORTLA, CKTDATA, CDBCOM 和 ISANBUS。按照同一标准，出现频繁的判断类名称为 39999。重新观察显示 301 中的 PORTLA, CKTDATA, CDBCOM, ISANBUS 和 39999 行，得出了处理器模块 SM23 中的带有倾斜度和颜色特征的勾号标记的实际数量，证实了条线图所显示的一切。

如果处理机模块 SM21, SM25, SM47 和 SM48 所遇到的问题与模块 SM23 的问题相似，只不过是在程度上低一些的话，将指示一个可能的系统范围问题。然而，如果所发生的问题与 SM23 无关，将指示一个局部问题。使用显示 301 左边的相互影响的彩色标尺 224，一个或多个处理器模块 SM21 ~ SM48 的勾号标记可以被关掉，也就是不显示，用以显示在记录文件中其它的重要报告模式。

图 4 中的显示 401 显示了处理器模块 SM25 和 SM48 单独的执行过程。处理器模块 SM25 有较浅灰度的勾号标记，在彩屏上显示浅绿色，更多垂直方向排列。另一方面，处理器模块 SM48 有较深的勾号标记，在彩屏上显示红色，更多按平行方向排列。处理器模块 SM25 具有大约 2.5 小时实验中所得到的各种类型的报告活动突发和经过 15 个小时实验所产生的 ISANBUS 审查报告的相当稳定的数据流。另一方面，处理器模块 SM48 有一报告模式，它与 SM23 处理器模块呈现的报告模式非常相似。这种近似可指示处理器模块 SM23 和 SM48 之间的模块间的错误。

关掉大部分的报告时，显示 401 显示另一个相互关系。判断行 39999 报告和与它同时出现的审查“波”之间有明确的关系。一个“波”是用报告的几个类型的几乎垂直的序列来表示的。另外，一个过程也被清除，带 410 行打出 PURGED，而当很多中继错误出现时，带 412 打出 CRR。处理器模块 SM23 也有一个相似的报告模式，因此进一步指出在 SM23 和

SM48 之间有一个模块间错误发生。

参照图 5, 错误计数的第二特征从特征组 526 中选取, 它和图 2 中的 226 基本上一致, 只不过是用 ERRORCNT 高亮代替 SM。由于有大量的审查报告, 所以错误计数第二特征被选择用来分析这些报告。每个审查过程检查全局数据和纠正不一致的东西。如果有不一致的报告被发现, 那些本过程的错误计数就增加。图 5 中, 按照它们各自的 ERRORCNT 特性, 给出了审查报告的倾斜度和灰度阴影编码(用以代表它们各自的颜色)。由于只有审查有 ERRORCNT 特征, 所以只有审查勾号标记按照它们各自的倾斜度与灰度阴影得以显示。判断带 506, 操作和维护带 501 和中继错误带 512 的勾号标记都变成黑色阴影并且都按垂直方向排列, 用以指出它们没有以这种操作模式定义。

显示 501 表明, 带 508 的多数审查仅发现少量的错误, 这些错误用蓝色的勾号标记表示。例外的是 PORTLA 审查, 它是出现最频繁的一种。PORTLA 一直在记录文件中发现大量错误。它用绿色的勾号标记表示在相应的颜色图中。

现在参看图 6, 显示 601 展示记录文件报告的 ERRORCODE 特征。显示 601 按照它们各自的 ERRORCODE 特征, 表示出审查报告的倾斜度和灰度阴影编码。从显示 601 中可以很容易地看出, 由于勾号标记都有一样的倾斜角度, 所以在整整 15 个小时的测试中大多数审查的错误编码都一致。这里, 每行内的倾斜角度可能比灰度阴影或颜色更有效, 因为虽然相邻的错误编码有非常接近的阴影或编码, 但它们没有如此靠近的倾斜角度。

在显示 601 中, 和先前的显示 501 一样, PORTLA 审查是个例外。可视交错阴影线是由很多有各种倾斜度的勾号标记所造成的。这种交错阴影模式表示有很多个问题触发这个审查。这也就是为什么 RORTLA 是最常出现的审查, 长条图 640 表示了这一点。

通常就像这种记录文件的情况一样, 有很多问题。因此它将使得我们把注意力集中到那些引起大多数错误的问题上。对于软件错误, 5ESS 指定一个事件号一个与错误报告有关的序列。例如, 判断 39999 的每次发生

和它的审查波共享同样的事件号。这个事实使得我们只选择那些具有最大相关报告数的事件。图7和图6一致，只不过是使用选择器。选择器是一个弹出窗口，允许一个特征值以使用彩色标尺的相同方式被关掉。选择器也有条形图，用以表示所有数值发生的总次数。

图7中，事件号选择器701把它的数值按降序排列。当然，只有那些最常发生的数值放在左边。这是通过在707点击“none”和使用键盘109或鼠标111来选择最顶部事件实现的。这只显示那些具有最大相关报告数的错误。然而，我们宁愿把我们的注意力集中到SM23和SM25上，因为它们是其上错误最多的两个SM。（尽管这个信息在先前的图中已经表示，这个信息还是由在SM选择器的弹出窗口705上的长条707的长度表示）。使用选择器701和705的作用是改变显示，使得对应被选中报告的勾号标记被显示出来。

现在参看图8，除了浏览窗口850外，显示801与显示701完全一致。如果需要很多信息来证实与一个已被可视分析过的特殊问题有关的假设，也许我们必须退回来并参看原始的记录文件；浏览报告的文字以便寻找那些没有显示的附加细节。举一个例子，在大约 $2\frac{2}{3}$ 小时内对第一个ISANBUS审查的发生的测试中，为了那些感兴趣的东西，操作者可能想查看原始报告。点击鼠标111到对应ISANBUS审查的勾号标记803上，从集中于该报告记录文件中弹出该报告的浏览窗口850，记录文件中的文字行的颜色编码和颜色标尺的颜色相符合。在浏览窗口850中的滚动条允许操作者在记录文件中浏览附近的报告。浏览窗口850还有一个搜寻文字的模式854区，在这里操作者打一个或多个字母数字字符，诸如关键字INT，在记录文件中向前或向后进行搜寻使之相符。

如图2到图8所示的各个显示，最好分别用颜色来帮助表示不同方面。

现在参看图1和图9，从记录文件中得到显示201-801的方法将在下面描述。最初，系统的记录文件被存储在海量存储120中。记录文件被顺序读入存储区115，在此它按程序117被处理器113处理，被读到用以组成记录文件的时标报告所选择部分的相关数据库中。由于记录文件被读取和被处理，关系表 119_1-119_i 被存储在存储区115。例如判断、审查、运

行维护报告和中断错误报告是存储的四个关系表。

每个关系表 119_1-119_i 都有一个描述文件，这个文件包括它的特征的名称。所以，这些特征描述了一个关系表。当这些特征被读入存储区 1702 时，关系表和特征被联系起来。这些关系表用来辨别它们使用的特征，并把特征指针存储到被关系表所使用的每个特征内。如果两个关系表使用同一个特征，在 1702 区只有一个入口，每个关系表都有一个指针指向那个入口。同样，当记录文件被读入时，多元组被存储在存储区 1704，并且关系表辨别这些多元组，它们每一个都使用和存储多元组指针。

图 12 给出了存储在特性区 1702 的数据结构的详情。首先，有一个特征名为 1802 的全局区，在此每个特征的名称都与用 AVL 树 1804 指向命名特征的指针有关。每个特征入口都有一个名称，特征 AVL 树 1804 的相关整数索引，指向数值 1808 的 AVL 树的指针，和指针—数值—节点—TO—节点数值信息 AVL 树 1806。存储所述特征的所有数值的数值 1808 的 AVL 树可以是数字的或文字的。数值 1808 的 AVL 树内的每个值，其数值本身被存储，每个数值的顺序索引被存储，而且每个值的发生次数或频率被存储。

特征 AVL 树 1806 还为每个特征存储一个指向“特征—数值—节点”的指针，一个屏蔽/不屏蔽的标志，以及一个具有所述特征的多元组指针的动态矢量。

现在参看图 13，描述存储在多元组区域 1704 的数据结构的详情。定义中的多元组是一组相关特征的数值(IEEE Std.100—1992)。每个多元组被一个，而且仅仅是一个关系表 119_1-119_i 拥有。每个多元组包含指向它自己的关系表的指针；从记录文件开始的以秒表示的时间；记录文件中的行号范围，它对应多元组所表示的信息；屏蔽的计数器（这个计数器很必要，因为多元组可能被多个选择器所屏蔽，然而只需要一个屏蔽用以阻止数值的显示）；以及指针—特征—数值—节点矢量，而不是过多地存储数值本身。

参看图 10 和图 14，存储器 115 的选择器区 1706 存储选择器窗口所使用的数据结构，如图 7 所示。选择器区 1706 在窗口内存储特征值的虚

坐标到显示 701 内数据的实坐标的映射图。当数据以降序排列时，这一点很必要。每个值有一个屏蔽/不屏蔽位，和一个用以指示当前是否按字母顺序排列的标记。虚坐标是按升序排列的数值的坐标，也就是说，如果在选择器的窗口内没有进行排序，则为数值的位置坐标。实际的坐标是数值排序以后，显示在选择器窗口内的数值坐标。当使用鼠标 111 在选择器窗口内选择排序的数值时，为了定位正确的值，使用映射图是有必要的。

图 15 的框图 2101 给出了为获得刚刚描述的数据结构和使用这些数据结构而产生像显示 201 或 701 那样的显示所执行的过程。关系表生成过程 2102 生成关系表 119_1-119_i 。它已经处理成为在关系数据库模式的记录文件上进行处理。

现在参考图 16，关系生成过程 2102 从关系表 119_1-119_i 中的第一个关系表 119_1 开始处理。动作 2202 读取第一关系表 119_1 的描述文件。然后，动作 2204 把描述文件的特征名加到特征名 1802 的全局区上，并标明在字符中哪些特征是数字的。在动作 2204 以后，动作 2206 从记录文件数据库中读取一个多元组。然后，动作 2208 处理当前关系表的多元组，这将在下面的图 17 中描述。在动作 2208 后，动作 2210 查看是否还有更多的当前关系表的多元组需要处理。如果回答是，过程 2102 返回到动作 2206 去读取当前关系表另一个多元组。如果回答不，那就意味着当前关系表的所有多元组都已被处理，过程 2102 转到动作 2212。动作 2212 查看是否要建立关系表 119_1-119_i 中的其他关系表，如果是，过程 2102 返回动作 2202，并读取要建立的关系表 119_1-119_i 中下一个关系表的描述文件。如果回答不，这意味着过程 2102 已经建立了所有的关系表 119_1-119_i 。过程 2102 可以转到建立显示过程 2104。

现在参看图 17，在描述建立显示过程 2104 之前，描述上述的处理多元组过程 2208 的详情。过程 2208 以当前关系表的多元组开始并转到动作 2302。在动作 2302，相关特征的当前多元组的数值被加到所述特征的数值 1808 的全局区上，而这个多元组的指针被加到多元组表 1810 上。在动作 2302 后，动作 2304 加一个指针到当前的多元组，该指针指向 AVL 树中的当前特征值节点。下一个动作 2306 查看是否要处理多元组的另一个

值。如果有另一个值要处理，过程 2208 返回到动作 2302，像处理当前多元组的先前值一样处理下一个数值。如果没有另一个数值要处理，过程 2208 转到动作 2308。动作 2308 把多元组的指针指向当前的关系表作为它自己的关系表，此时这个多元组的处理过程就全部完成。在过程 2208 把对应于记录文件的关系数据库每个关系表 119_1-119_i 的每个多元组处理完以后，所得结果即为图 13 所示的多元组数据结构。

现在参看图 18，描述生成显示过程 2104 的详情。这个过程产生一个象在显示屏 105 上的显示 201 那样的显示。如果使用颜色显示，动作 2402 沿显示 201 的左边产生一个颜色标尺选择器，如图 2 所示。基于一个给定特征的唯一数值上的分布，在着色过程中，从兰到红的颜色都可以被指定。这些颜色对应于勾号标记颜色，作为图像相关技术的一部分。下一步，动作 2404 产生标号，也就是在显示中所用的“名称”或符号，用于对应所述关系表原始键的唯一值的当前关系。这些标号是显示中的“固定”部分，剩余部分涉及到进一步处理对应记录文件的关系表和先前提到的它们各自的数据结构。

动作 2406 是一个产生图形的过程，在正面的图 19 中描述。动作 2406 之后的动作 2408 创建用于图表的条线。动作 2408 在下面的图 20 中描述。动作 2408 之后的动作 2401，查看是否有另一个关系表需要显示。如果有另一关系表需要显示，过程 2104 返回到动作 2404，并转到下一关系表的过程 2406 和 2408。如果没有另一个关系表需要显示，过程 2104 转到动作 2412。动作 2412 是一个创建时间条的过程，它将在下面的图 21 中描述。这些过程完成了在显示屏 105 生成显示，例如显示 201 的过程。

现在参看图 19，描述创建图表过程 2406。创建图表过程 2406 始于创建显示过程 2104 所选择的关系表。过程 2406 的动作 2502 访问当前关系表的一个多元组。然后，动作 2504 把多元组指针加到四叉树中多元组表的点(x, y)处，该点由多元组的时间以及沿 y 轴的属性的多元组值的索引确定。动作 2504 使用来自数据库的数据和如图 13 所示的多元组数据结构。然后，动作 2506 查看，为了创建绘图过程 2406，是否有另一多元组需要访问。如果有，过程 2406 返回到动作 2502 去访问当前关系表的另一个多

元组。如果没有，创建图表过程 2406 的所有当前关系表的多元组都已经被处理，则过程 2406 转到创建图表条线过程 2408。

现在参看图 20，描述创建图表条线过程 2408。在图 2 中，显示 201 上从 214 到 220 的这些条线都是横向的，从显示的右侧纵轴伸展。同先前的过程一样，创建图表条线过程 2408 始于已经选择的关系表 119_1-119_i 。动作 2602 将用于显示的每个条线的总数初始化为零。这以后，过程 2408 转到动作 2604。

动作 2604 查看当前关系表是否有选择用于颜色编码的特征。如果当前关系表没有所述的特征，那就意味着当前显示有关系表不具备的特性，而且在这种情况下，没有条被显示。例如，观察图 5 中显示 501 的带 506。在这种情况下，过程 2408 向前跳到创建显示过程的动作 2410 处，如图 18 所示。

如果当前关系确定具备选择用于颜色编码的特征，那就意味着当前显示具有关系表所具备的种特征，而且可以画出一个或多个长条，在这种情况下，过程 2408 进入到动作 2606。动作 2606 访问当前关系表的一个多元组，并且进入动作 2608。动作 2608 查看已经被访问的多元组是否被屏蔽。如果在图 2 中，在显示 201 的显示 224 上，作为颜色编码特征的多元组的值没有被鼠标 111 所选中，或者是如图 13 所示，多元组有一个非零的屏蔽计数，那么就认为多元组被屏蔽。如果被访问的多元组没有被屏蔽，过程 2408 进入动作 2610，它用以增加相应于被访问的多元组的条总数，而且过程 2408 进入动作 2612。如果被访问的多元组已经被屏蔽，那么过程 2408 向前跳至动作 2612，而且相应条的总数不增加。动作 2610 后，动作 2612 查看是否有当前关系表的另一个多元组要访问。如果有另一个多元组要访问，过程 2408 返回动作 2606 去访问另一个多元组。如果没有另一个多元组要访问，那么当前关系表的所有条已经创建，过程 2408 进入动作 2410。

动作 2410，像在图 18 所描述的那样，回到过程 2104 用以为图表标号，图表和条线处理另一个关系表。在所有的关系表 119_1-119_i 通过动作 2402 至 2410 后，创建显示过程 2104 进入创建时间条过程 2412。

现在参看图 21, 描述创建时间条过程 2412。创建时间条过程 2412 与每个关系表互相独立, 而且取代了所有的多元组。在过程 2412 开始时的动作 2702, 把所有的时间条总数初始化为零。然后过程 2412 进入动作 2704, 在这里, 用颜色编码的特征值被访问。接着, 动作 2706 查看被访问的数值是否被颜色选择器屏蔽。如果被访问的值被屏蔽, 那么这个值将不影响时间条, 过程向前跳至动作 2716, 这一点将在下面解释。如果被访问的值没被屏蔽, 那么这个值将影响时间条, 过程 2412 进入动作 2708。

动作 2708 访问具有被访问值的多元组。然后, 动作 2710 查看被访问的多元组是否被屏蔽。如果这个多元组被屏蔽, 那么过程 2412 跳至动作 2714, 而且这个多元组并不影响当前的时间条。如果这个多元组没有被屏蔽, 过程 2412 进入动作 2712, 在此, 在进入动作 2714 之前, 当前时间条总数增加。相关的时间条是通过用当前时间间隔的长度把多元组的时间划分(以秒为单位)而得到的。

动作 2714 查看是否有另一个多元组具有被访问的数值。如果有, 那么过程 2412 返回到动作 2708 去处理所述多元组。如果没有, 那么过程 2412 进入动作 2716。动作 2716 查看是否有另外的值需要处理。如果有, 那么 2412 返回到动作 2704 去处理这个值。如果没有, 那么过程 2412 就完成了所有的时间条, 并且创建显示过程 2104 也全部结束。这里除了勾号标记(在图 25 中描述)以外用于显示 201 的每一部分的数据全部被创建。但是, 为了创建部分特殊显示 701, 还需要过程 2106 和 2108。

现在参看图 22, 描述读记录文件过程 2106。这个过程用浏览窗口。动作 2802 把记录文件读入存储器 115。然后, 动作 2804 初始化位于存储器内的记录文件开始处的缓冲器指针。接着, 动作 2806 检查缓冲器指针所指的字符。这个检查用以确定该字符是一个新行字符还是别的字符, 然后过程 2106 进入动作 2808。动作 2808 查看缓冲器指针指向的字符是否为一个新行字符。如果不是新行字符, 过程 2106 跳至动作 2814。如果缓冲器指针是指向一个新行字符, 那么过程 2106 进入动作 2810。

由于新文本行以新行字符表示开始, 所以过程 2106 到达动作 2810。动作 2810 把新行字符变为空字符, 从而结束了存储器中的文本行。然后,

动作 2812 把用于行 1610 的指针放到已过去的新行字符处,也就是文本新行的开端,并且进入动作 2814。动作 2814 查看是否有另一字符需要检查,如果有,过程 2106 返回到动作 2806 去检查另一个字符。过程 2106 将以这种方式循环反复,直至每个新行的开端都被存储到存储器 115 中,并且在记录文件中不再有字符需要检查。此时,总的过程 2101 进入过程 2108。

参照图 23,描述给记录文件上色过程 2108。这个过程把记录文件文本的颜色置成与作为可视装置的显示屏 105 上显示的多元组的颜色一样,向操作者证明,用指针 107 指定和按下鼠标 III 的键的方法,带进浏览窗口 850 的记录文件报告,正如在图 8 中看到的那样,与在显示中选定的勾号标记是相关连的。

动作 2902 是第一个动作,这个动作访问一个特征。接着,动作 2904 访问一个关系表。然后动作 2906 访问受访关系表的多元组。在动作 2906 以后,动作 2908 查看已经访问的关系表是否有被访问的特征。如果没有,过程 2108 进入动作 2912,把相应于当前多元组行号范围的颜色目录 1620 置成白色,并且进入动作 2914。如果有,过程 2108 进入动作 2910,把相应于符号范围当前多元组颜色目录 1620 置成在当前特征中的多元组数值的索引,并且进入动作 2914。

动作 2914 查看是否需要访问关系表的另一个多元组。如果是,那么过程 2108 返回到动作 2906,访问这个多元组用以检查当前特征。如果不是,过程 2108 进入动作 2916,用以查看是否有另一个关系表。如果有另外一个关系表,过程 2108 返回到动作 2904 访问的这个关系表用以检查用于该特征的多元组。如果没有另外的多元组要检查,过程 2108 进入动作 2918 去查看是否有另外的特征要访问。如果有另外的特征要访问,过程 2108 返回到动作 2902 去访问这个特征,用以查看当前关系表中的关系是否有这种特征。如果没有另外的特征要访问,那就意味着所有的特征都已被访问,并且,所有的记录文件行都已设置颜色编码并且存储为它们各自多元组值的颜色或白色。

现在参看图 24,描述挑选相关性过程 3001。符号的挑选相关性过程 3001 用以使操作者通过指针 107 和鼠标 111 与显示交互作用。动作 3002

把鼠标物理坐标(x, y)转换为以秒为单位的时间,也就是横轴,以及沿 y 轴属性的值的索引。接着,动作 3004 使用经转换后的坐标(x', y'),在所述坐标上查找用于多元组表的四叉树表。在动作 3004 之后,动作 3006 查看是否在(x', y')处发现了多元组表。如果没有发现,过程 3301 结束,并且,关系处理过程结束,也就是没有形成相关性。

如果发现多元组的表,动作 3008 通过指针访问表上的多元组,而且过程 3001 进入动作 3010。动作 3010 查看被访问的多元组是否被屏蔽。如果被屏蔽,过程 3001 进入动作 3012 去查看表上是否有另一个多元组,如果有,过程 3001 返回到动作 3008 去访问并查问这个多元组的屏蔽情况。如果没有另外的多元组,过程 3001 结束。如果被动作 3008 访问的多元组没有被屏蔽,过程 3001 就从动作 3010 转到动作 3014。从操作者的观点来看,动作 3014 就是希望的结果。因为动作 3014 交互地滚动浏览窗 850 以显示在记录文件中对应于鼠标 111 选定的勾号标记 803 的报告。

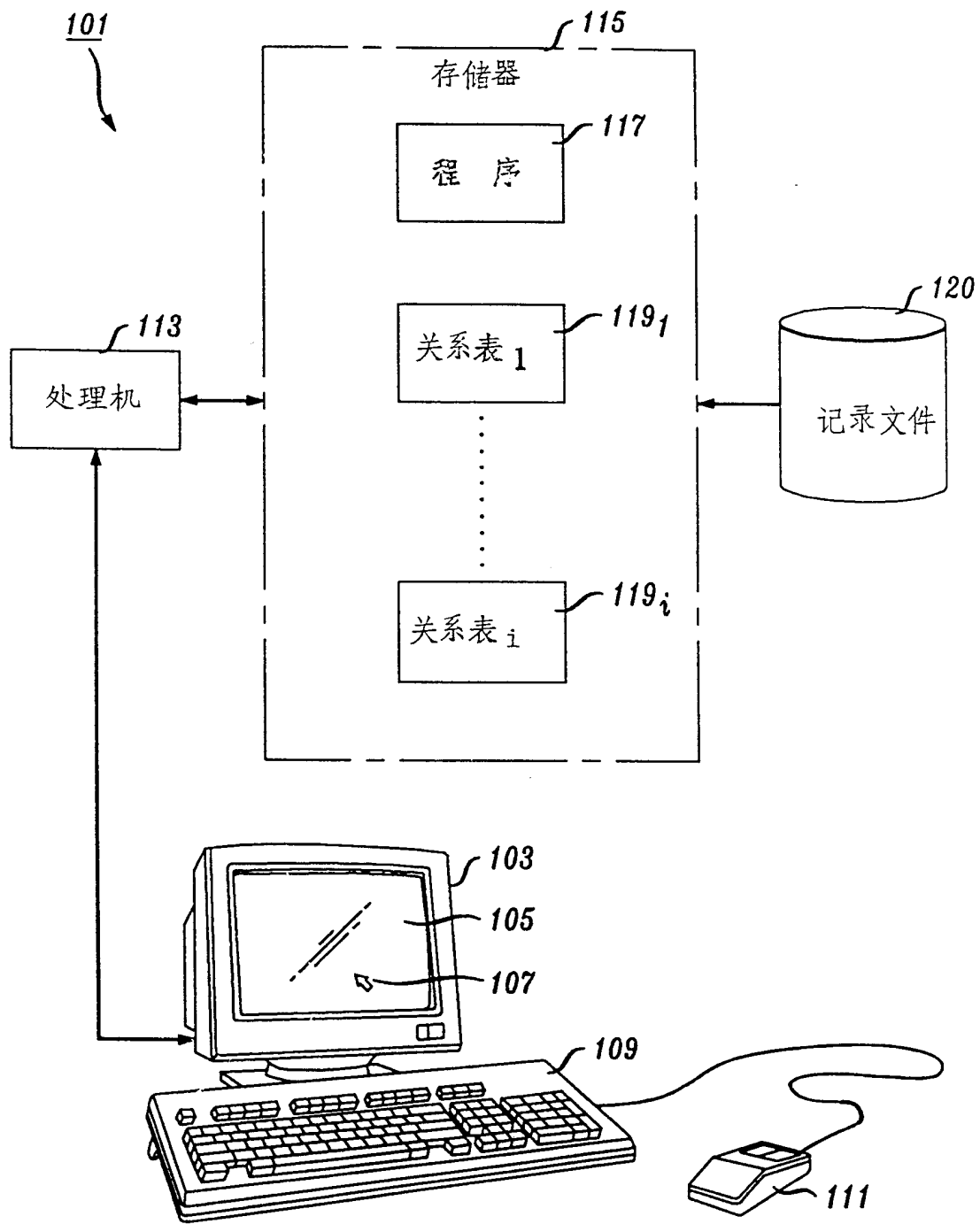
现在参看图 25,描述显示屏 105 上的多元组绘图过程 3101。动作 3102 访问要被描绘的显示上的多元组。接着,动作 3104 查看这个多元组是否被屏蔽。如果这个多元组被屏蔽,过程 3101 跳到动作 3114 去寻找另一个多元组。如果这个多元组没有被屏蔽,过程 3101 进入动作 3106。动作 3106 按多元组值的索引把颜色置成用颜色编码的特征值。然后,动作 3108 根据多元组的时间和在沿 Y 轴使用的特征值的索引来决定(x, y)的位置。然后,动作 3110 按照角度编码的特征值的多元组值索引来设置要画出的勾号标记的倾斜角度。接着,动作 3112 在显示屏 105 上以(x, y)为中心,画出倾斜的以颜色编码的线。在动作 3112 后,过程 3101 进入动作 3114,去查看是否有另一个多元组要绘制。如有另外的多元组,过程 3101 返回到动作 3102,描绘这个多元组。如果没有另一个多元组,则所有的多元组已绘在显示屏 105 上,过程 3101 执行完毕。

现在回头参看图 15,过程 2110 是运行程序过程,该过程从由过程 2102、2104、2106 和 2108 建立和存储的所有图形数据形成各种显示。另外,运行程序过程 2110 使用过程 3001 和 3101 提供给操作者一个交互的图形显示,用以分析记录文件。

根据本发明的装置能使操作者在处理器系统甚至是分布处理器系统中很快地发现和分离出感兴趣的信息。还有，由于感兴趣的信息被如此快地发现，本发明能使操作者发现二级信息，而使用以前的文本基础技术是不可能发现某些二级信息的。

尽管通过参考优选实施例来说明和描述本发明，但应理解，本发明在格式、细节和应用方面都可以进行多种改变。比如，本发明可以用于分布式计算机系统而不是分布式交换处理器系统。

图. 1



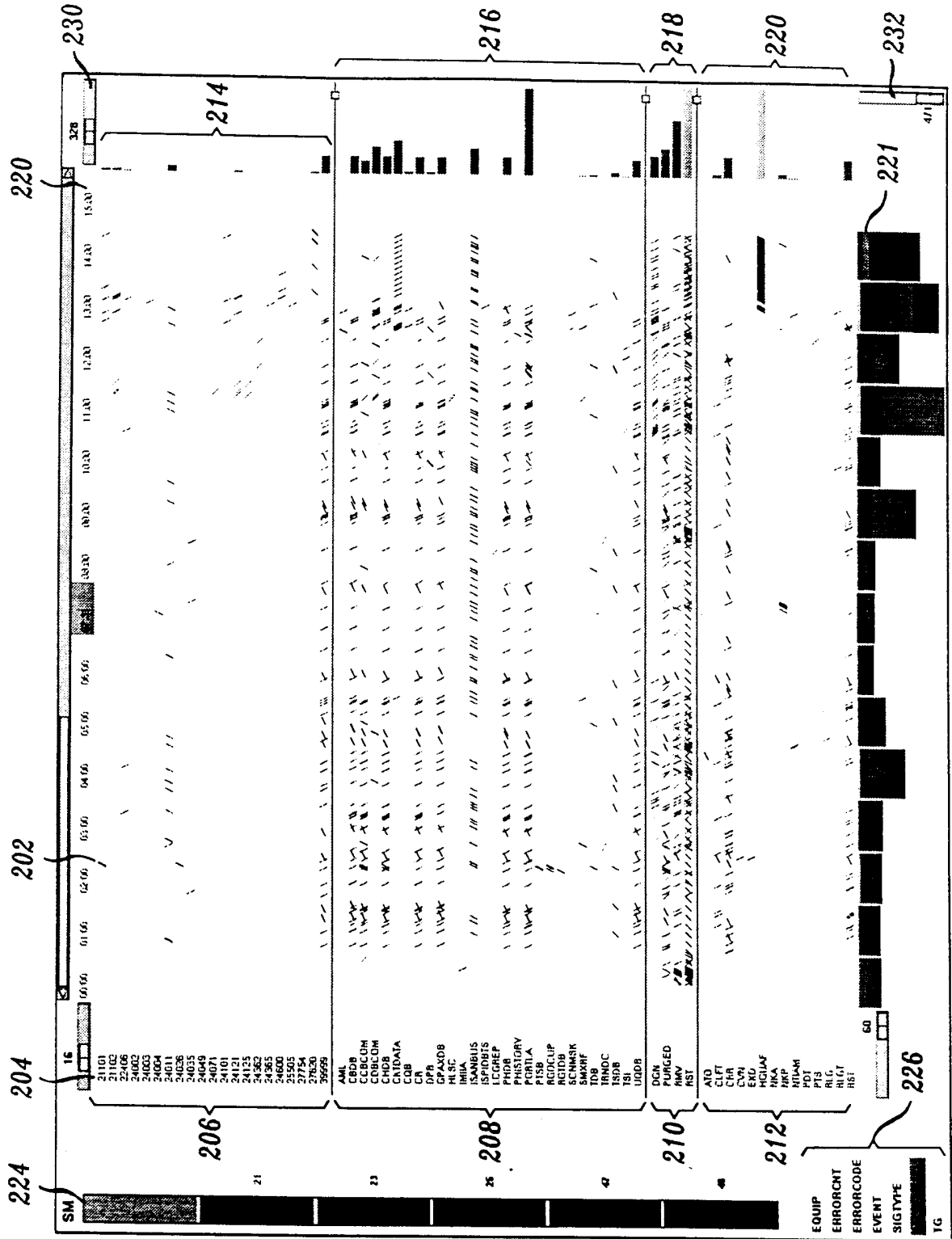
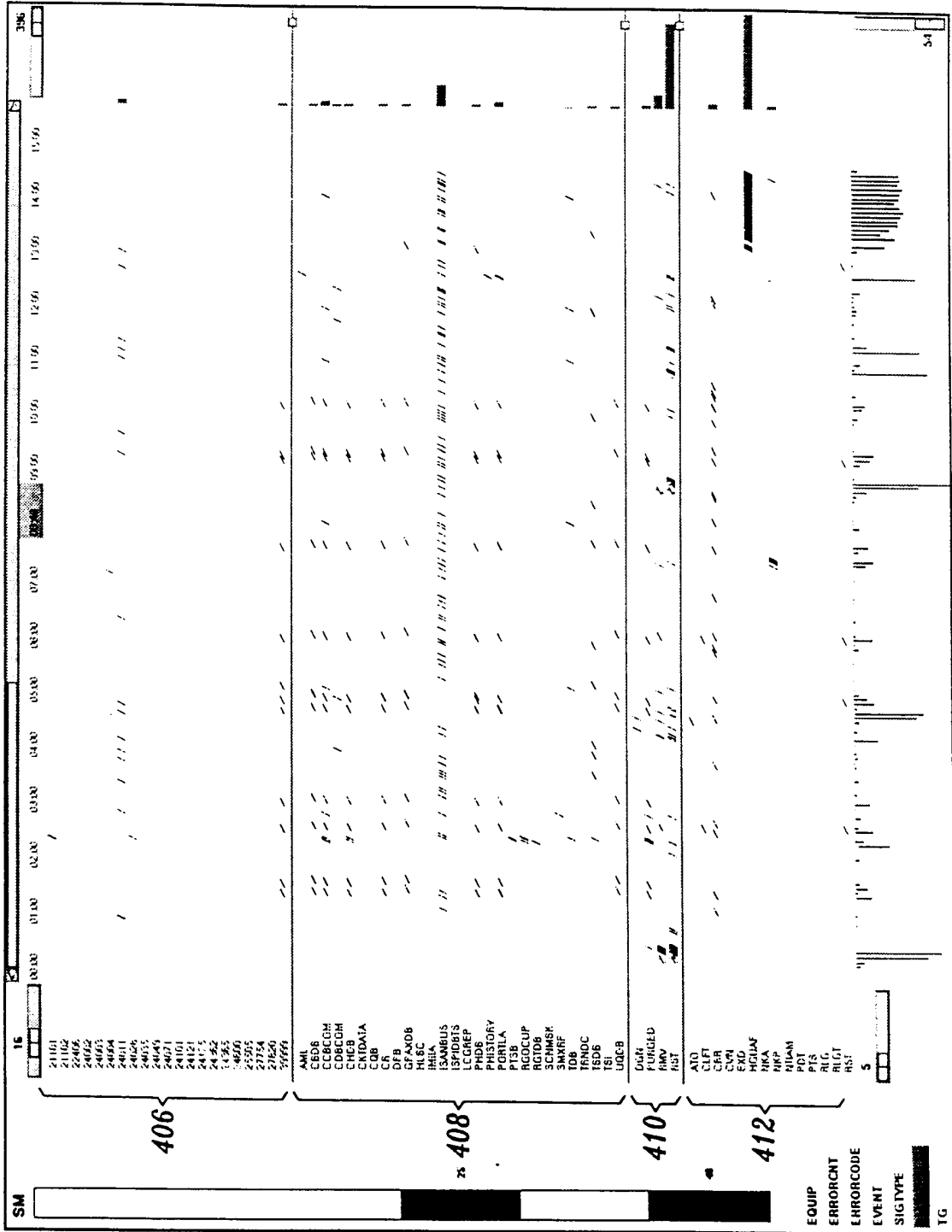


图 2
201



4
401

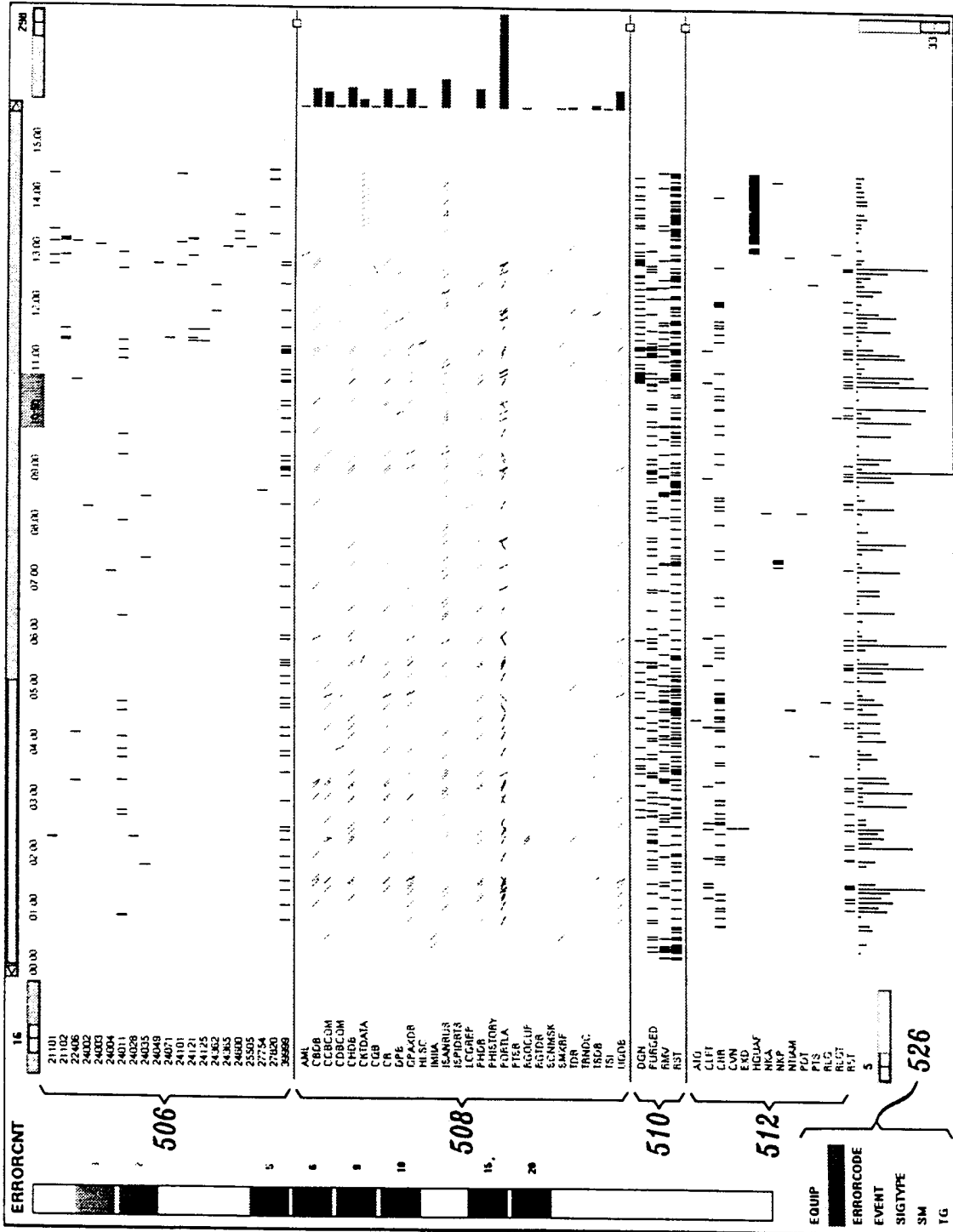


图 5
501

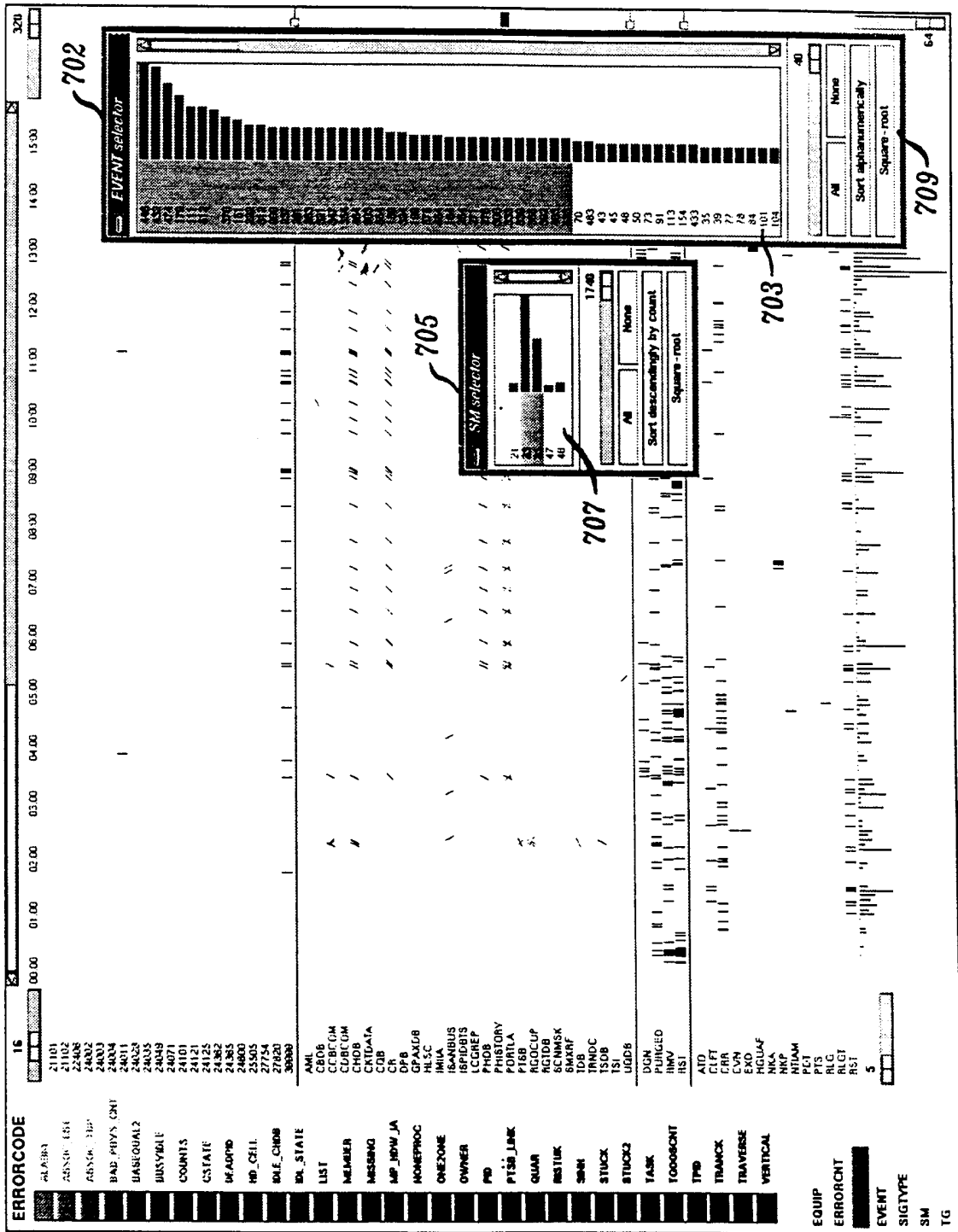


图 7
701

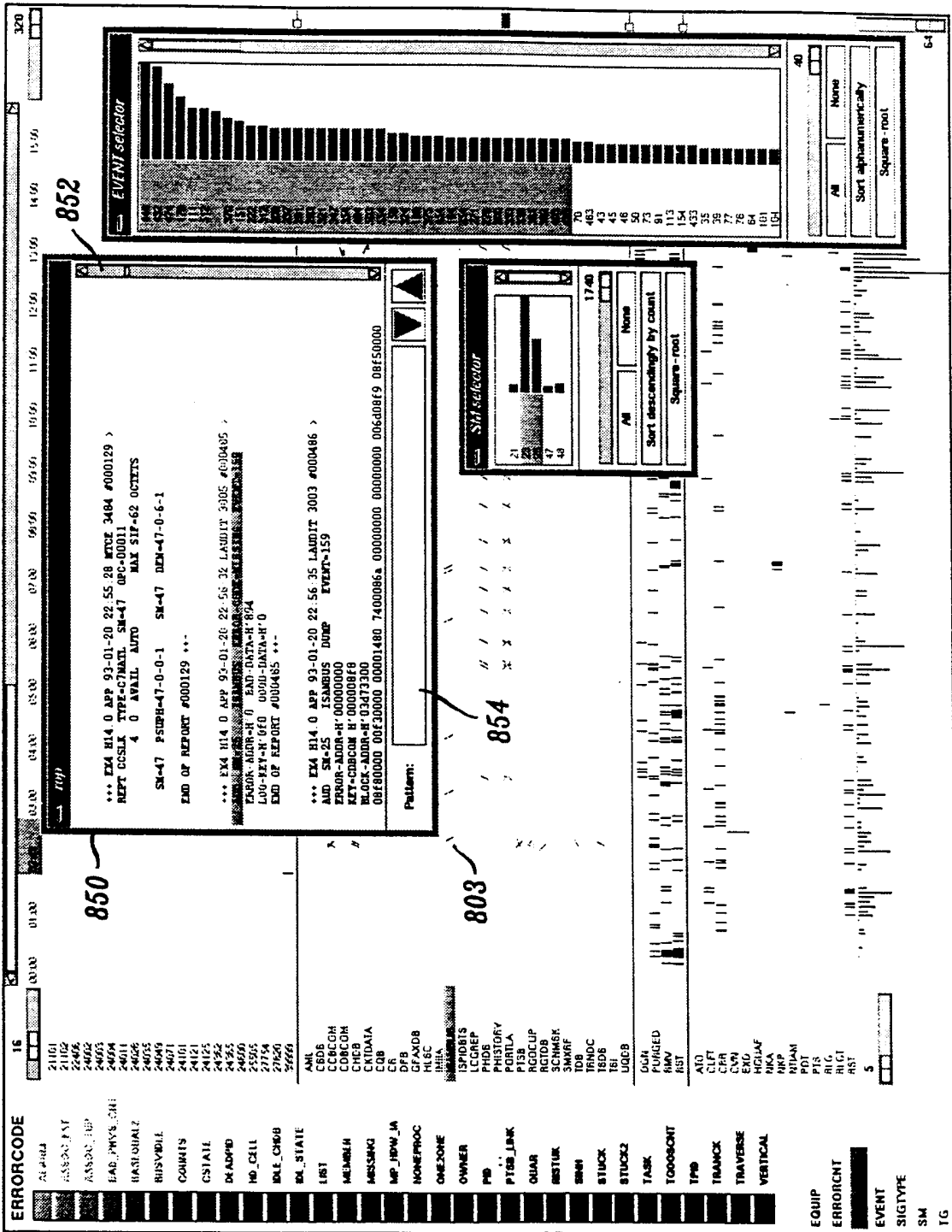


图 8
801

图9

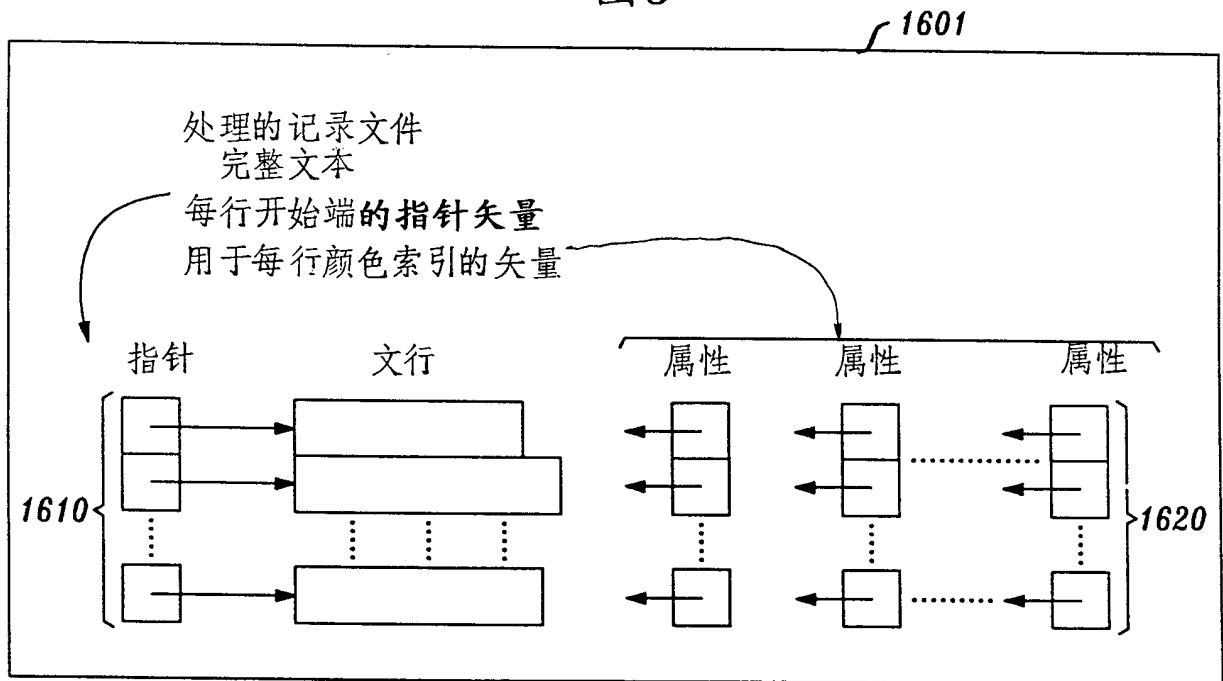


图 10

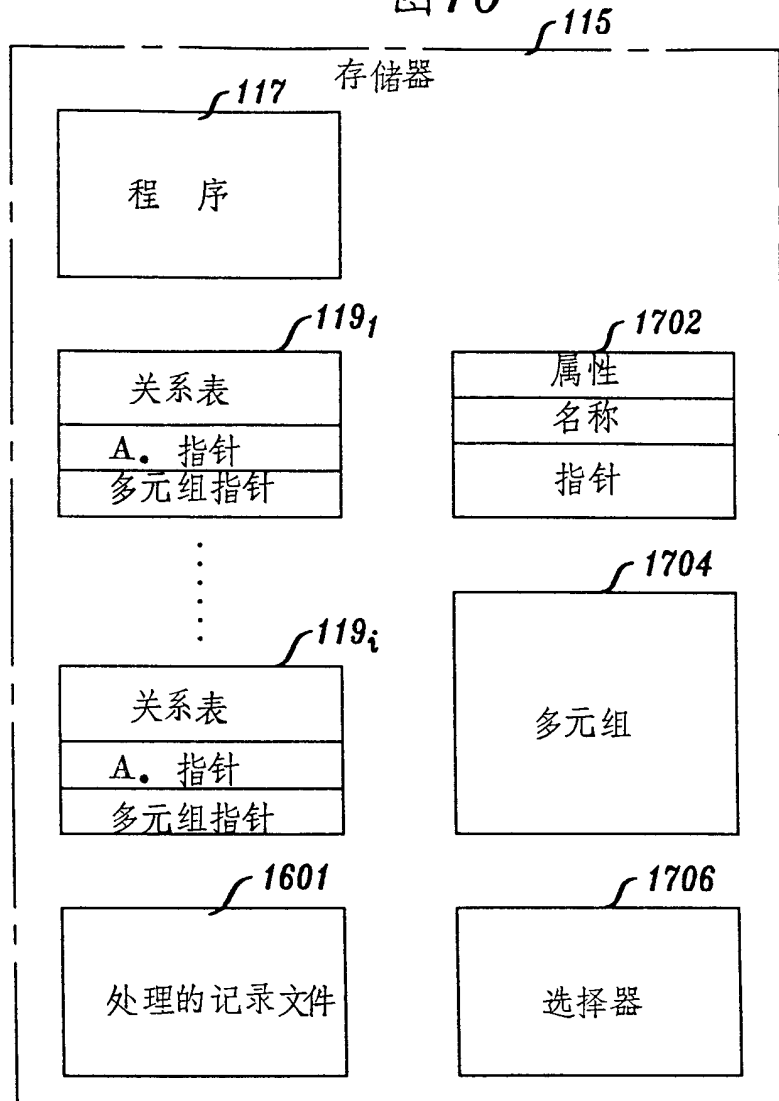


图 11

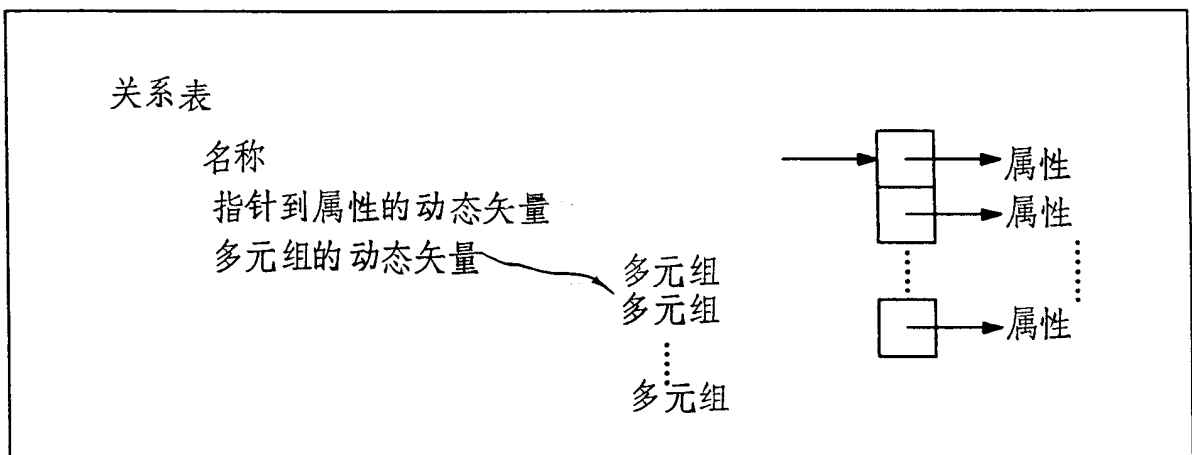


图12

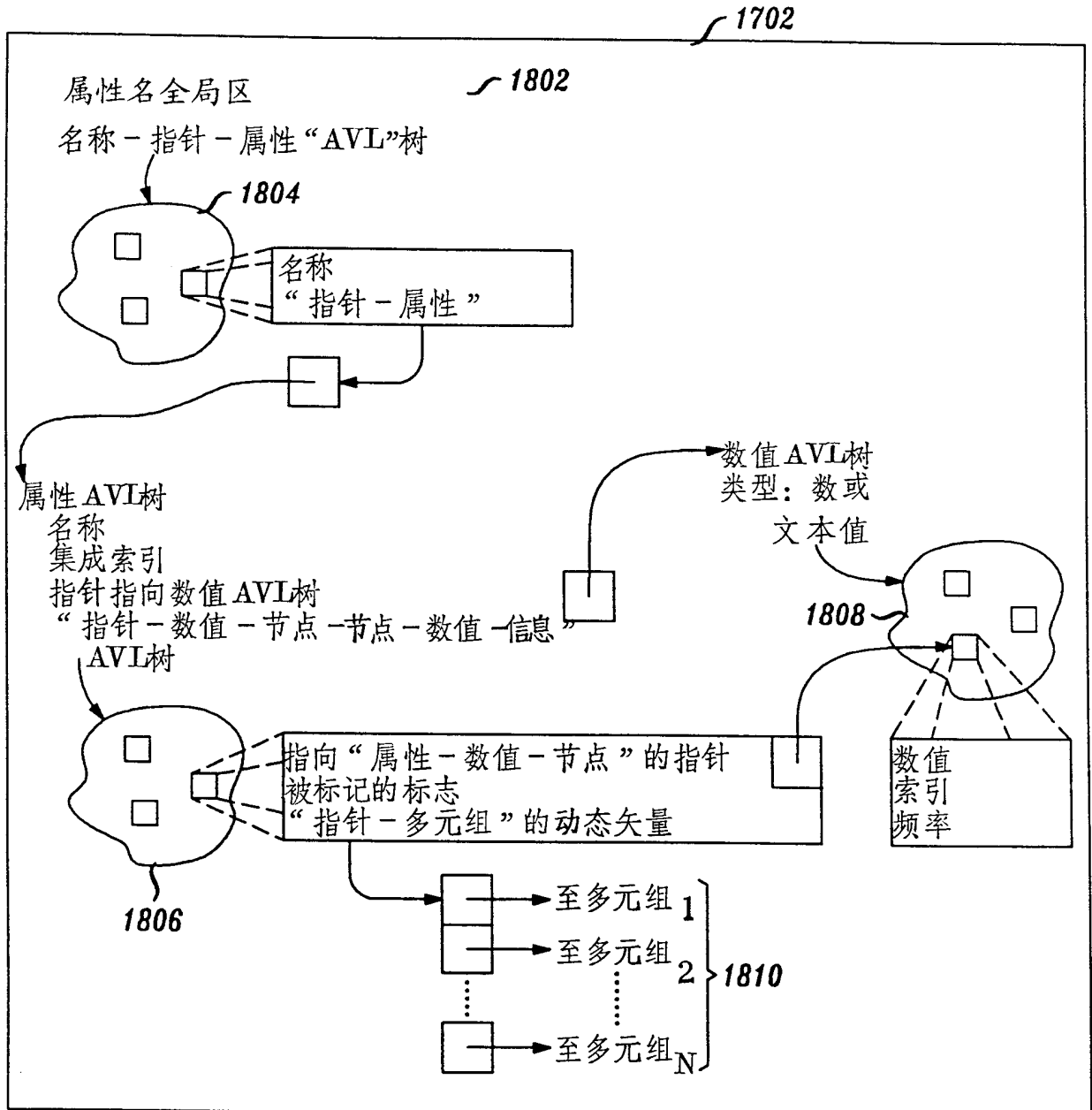


图13

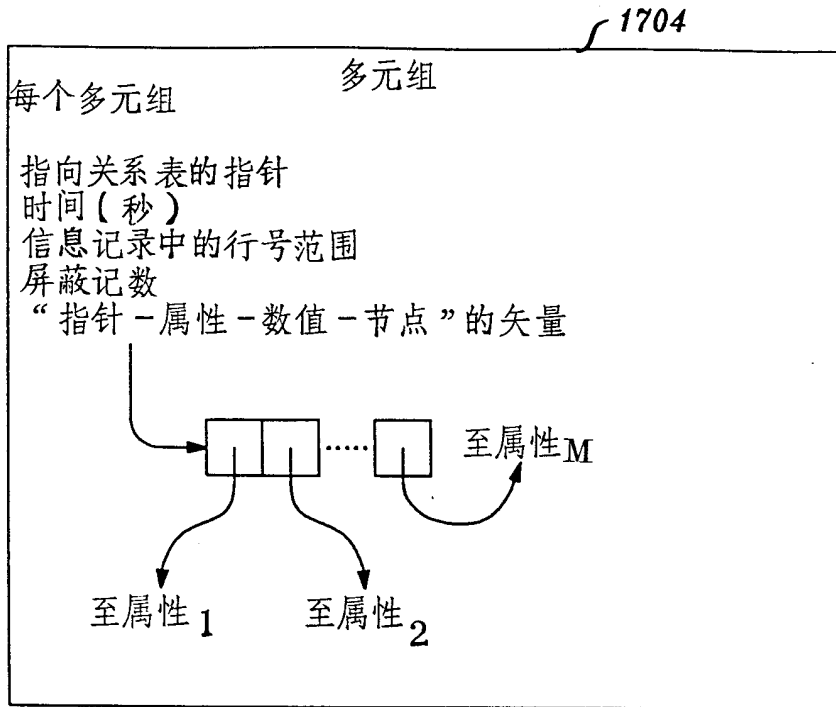
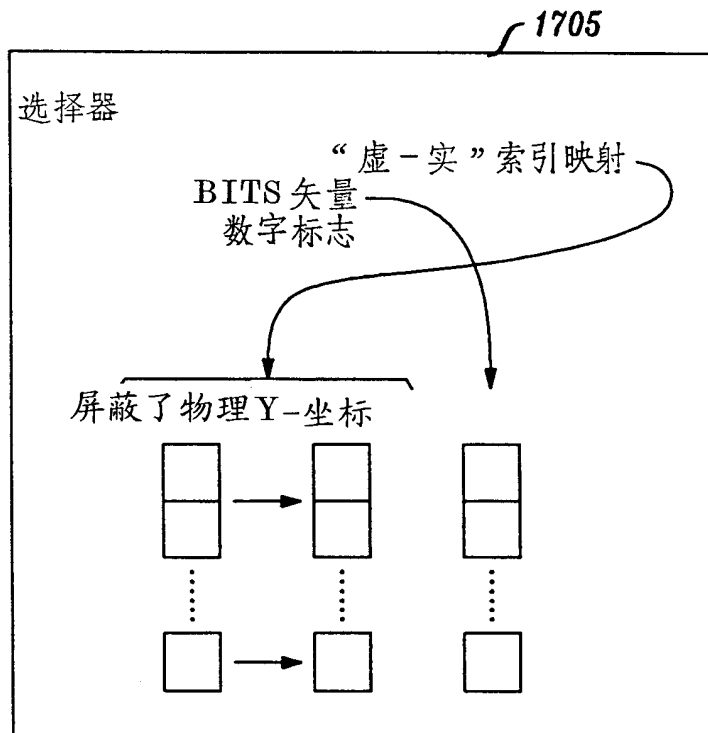


图14



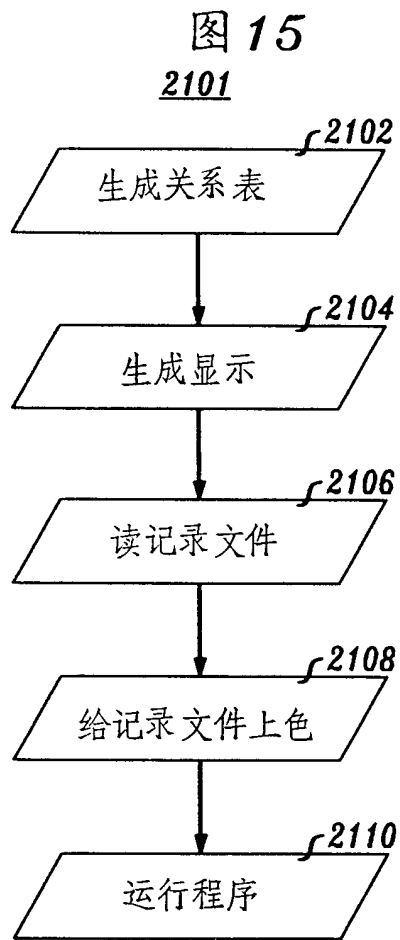


图 16
2102
生成关系表

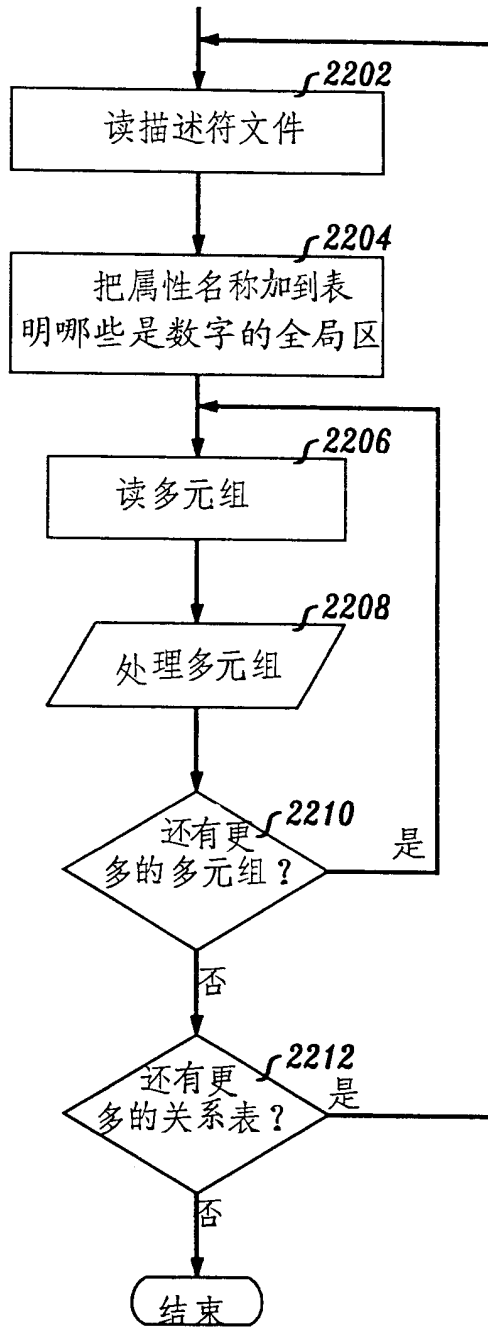


图 17

2208

处理多元组

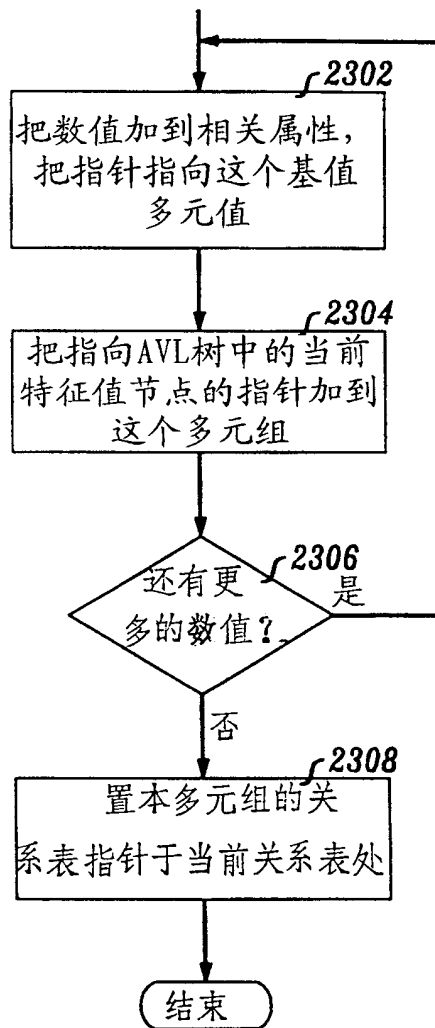


图 18

2104
生成显示

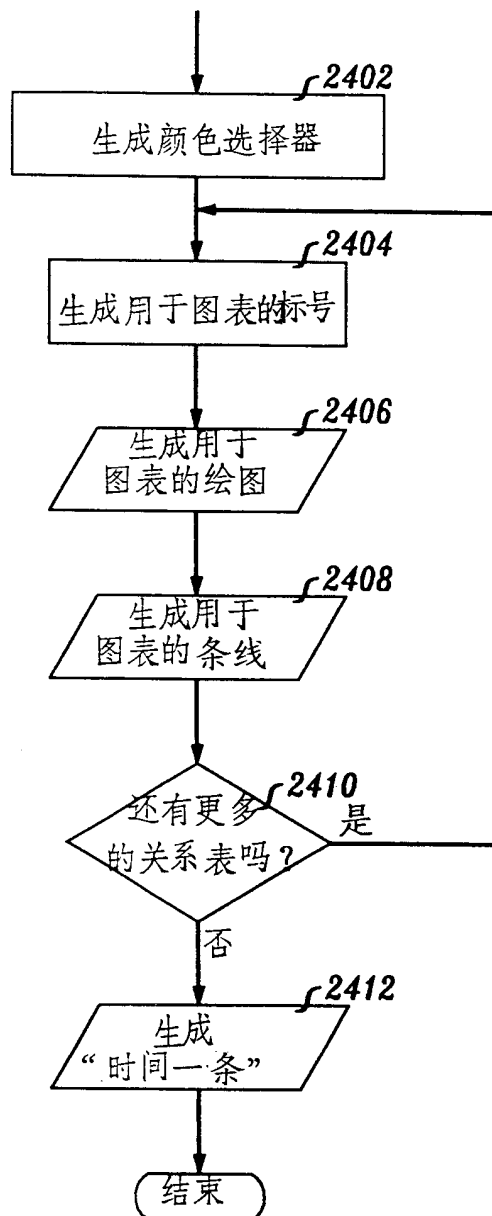


图19

2406
创建图表

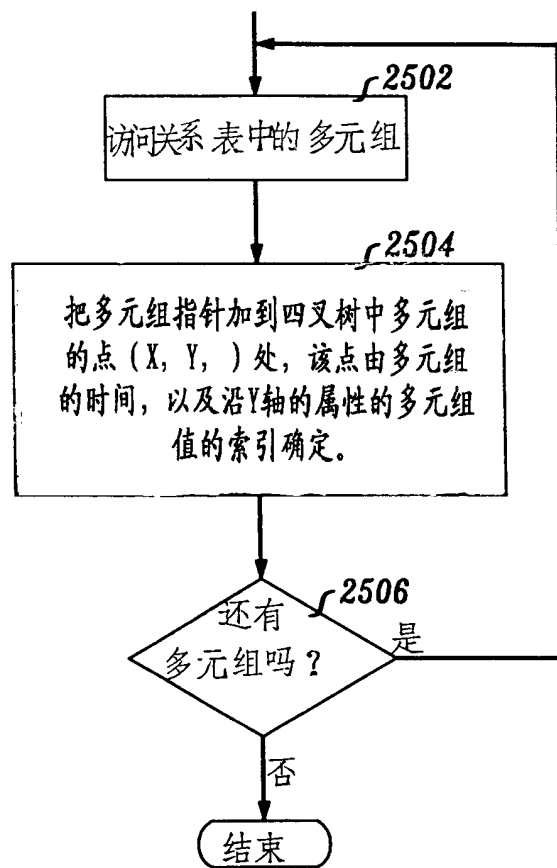


图20
2408
创建用于图表的条线

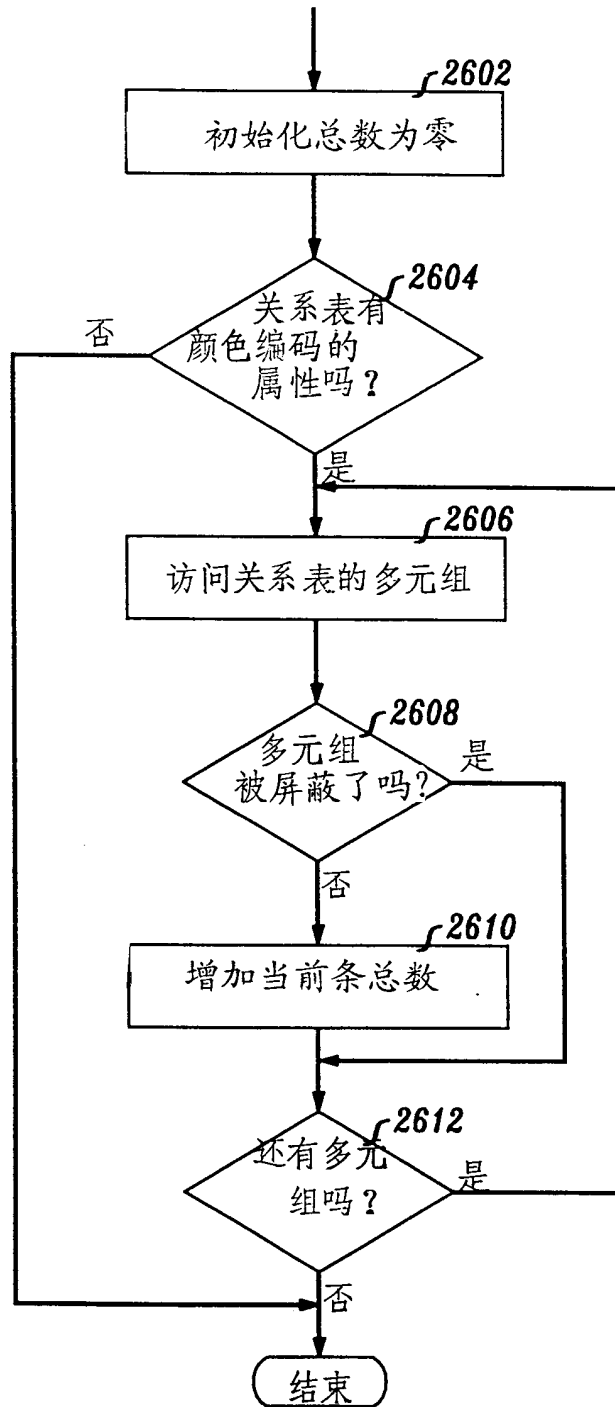


图 21
2412
创建时间 - 条

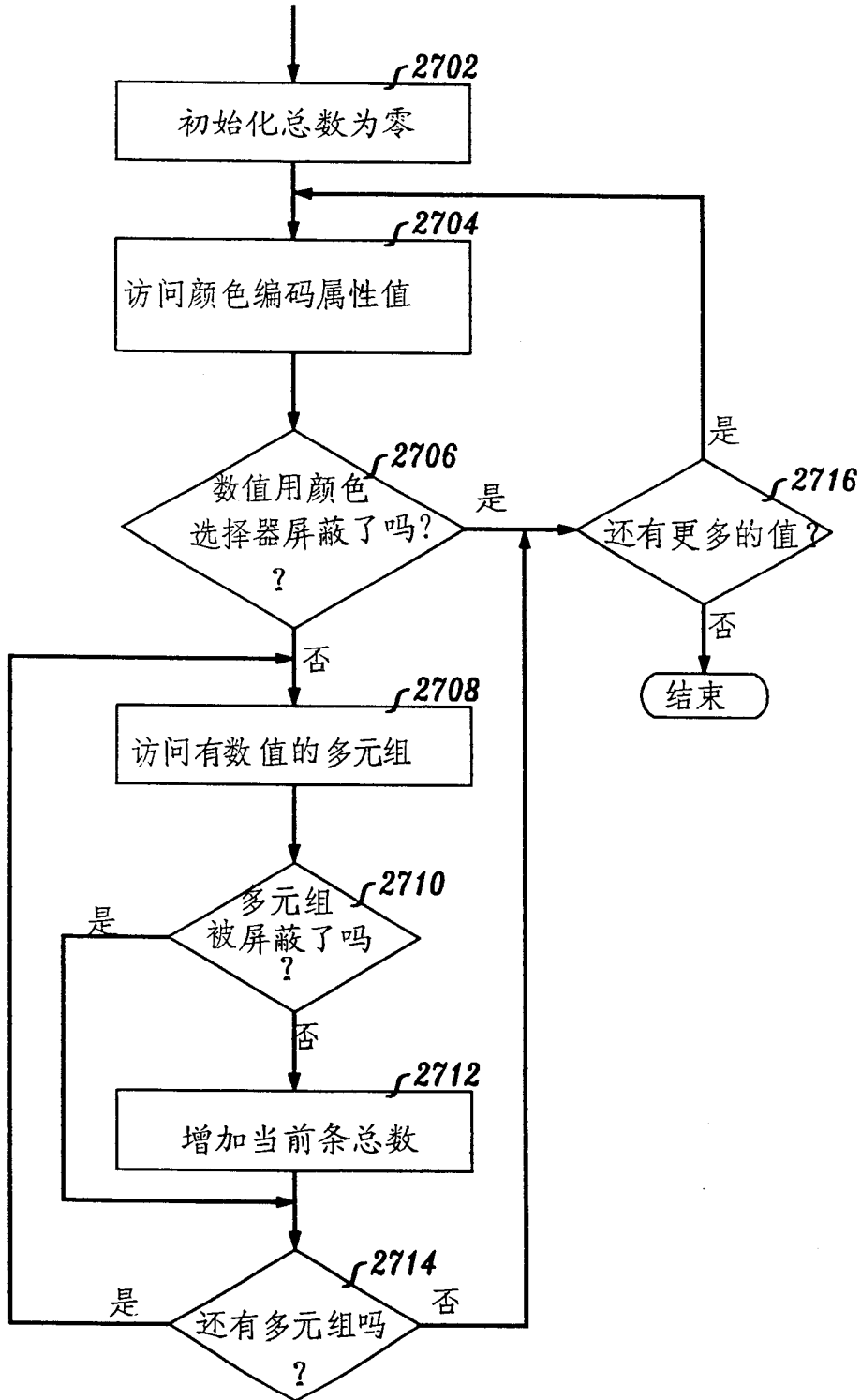


图 22

2106

读记录文件

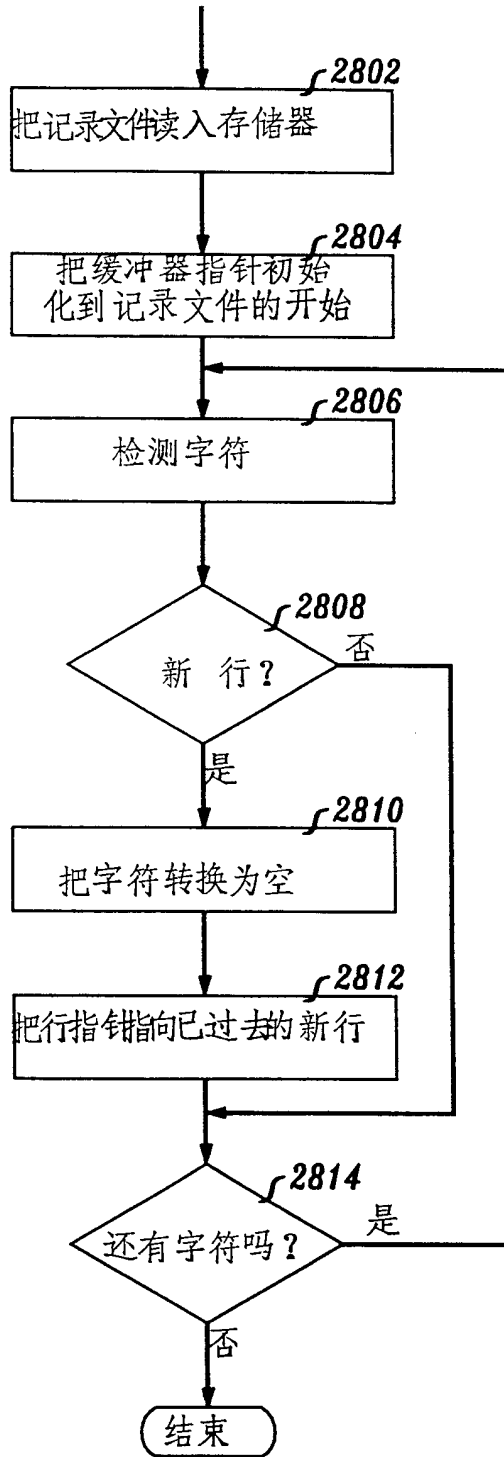


图 23

2108

给记录文件上色

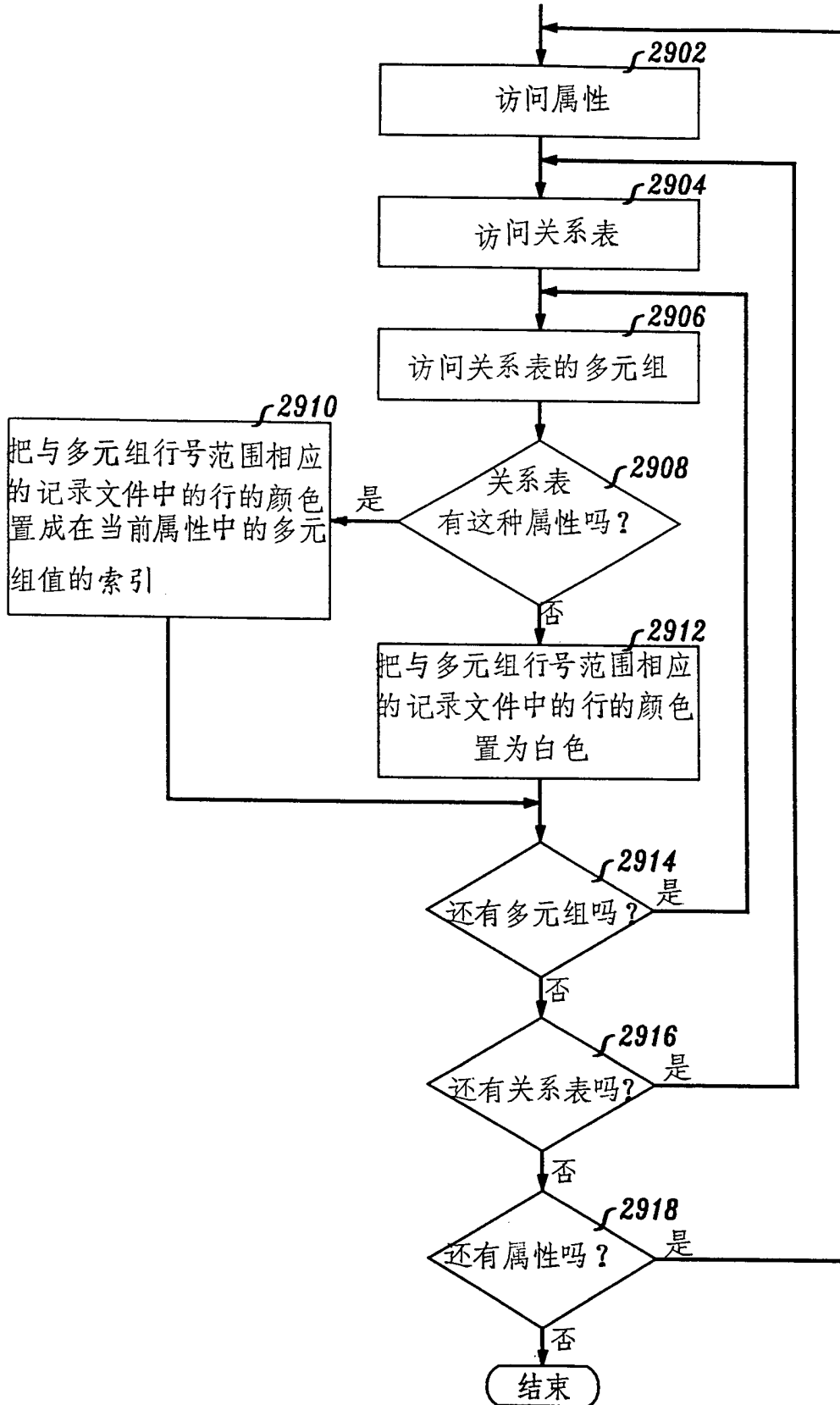


图24

3001

符号的挑选相关

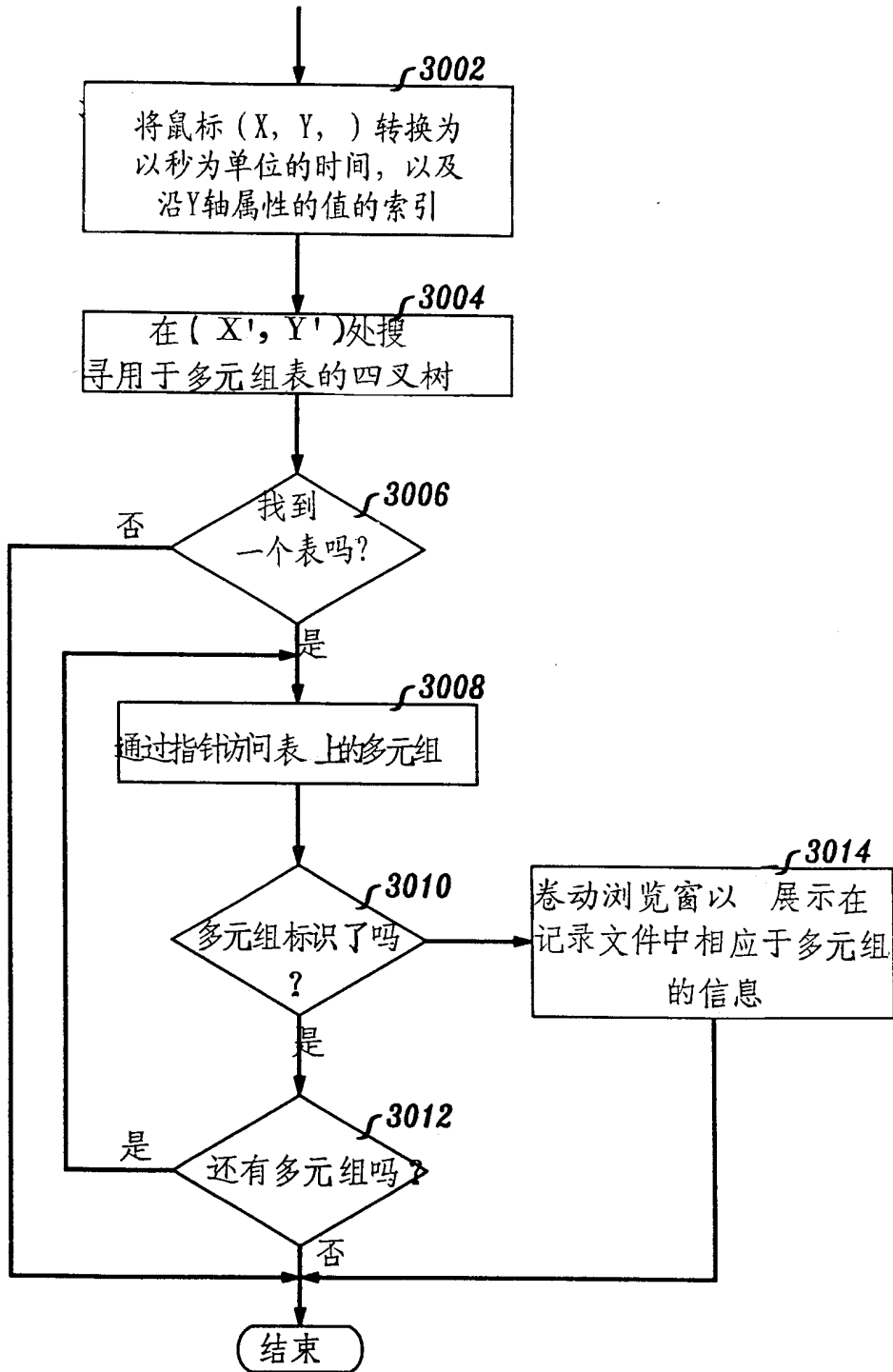


图 25

3101

在屏幕上绘制多元组

