



(12) 发明专利

(10) 授权公告号 CN 1477590 B

(45) 授权公告日 2010.05.26

(21) 申请号 03149300.9

(22) 申请日 2003.06.19

(30) 优先权数据

10/178,443 2002.06.19 US

(73) 专利权人 微软公司

地址 美国华盛顿州

(72) 发明人 张正友 A·古普塔 R·卡特勒

刘自成 贺立为

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

代理人 陈斌

(51) Int. Cl.

G06K 11/06 (2006.01)

G06F 3/16 (2006.01)

G06T 5/00 (2006.01)

(56) 对比文件

KR 20020012699 A, 2002.02.20, 说明书摘

要.

JP 特开平 7-15711 A, 1995.01.17, 全文.

CN 1109695 A, 1995.10.04, 说明书第 16 页

第 2 段, 说明书摘要, 附图 12.

审查员 龚锦玲

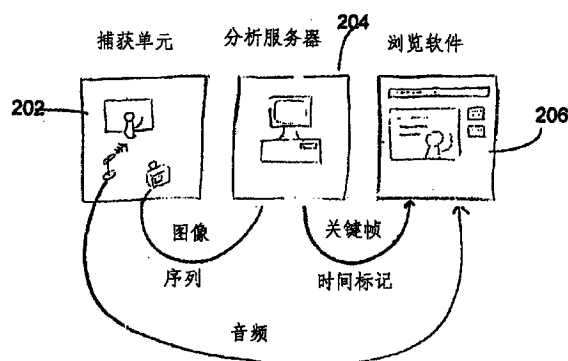
权利要求书 6 页 说明书 19 页 附图 19 页

(54) 发明名称

用于白色书写板和音频捕获的一个系统和方
法

(57) 摘要

使用数字照相机和麦克风来捕获一个会议的白色书写板内容和音频信号的一个系统。系统可能被改进到任何现存的白色书写板。它通过分析被捕获的快照的序列来计算在白色书写板上的笔划的时间标记。它也自动表示在每个擦除之前在白色书写板上所有被书写的内容的一组关键帧。使用白色书写板内容作为一个视觉索引来有效的浏览音频会议。系统不但捕获白色书写板内容,而且帮助用户有效的和安全的观看和管理被捕获的会议内容。



1. 一种用来摘要和索引视听内容的方法,其特征在于,包括以下的处理动作:
用一个照相机来捕获在一个非电子白色书写板上书写的内容的图像序列;
记录与图像序列相关联的音频信号;和
分析图像序列来分离摘要所述书写板内容的关键点的关键帧,其中所述分析处理动作包括下列处理动作:
校正在图像序列中的每个图像中的白色书写板的一个视图;
提取白色书写板背景颜色;
把图像序列的每个图像划分成单元,每个单元表示一个单元图像;
对于随着时间流逝的每个单元,贯穿整个图像序列把单元图像聚集成组;
把每个单元图像分类为笔划、前景物体或白色书写板单元;和使用分类结果提取关键帧图像。
2. 如权利要求 1 所述的方法,其特征在于,还包括把所述音频记录和所述关键帧相关联。
3. 如权利要求 2 所述的方法,其中所述音频信号通过与所记录的音频和图像序列两者均相关的时间标记而与所述图像序列相关联。
4. 如权利要求 3 所述的方法,其中把音频信号和所述图像序列相关联包括处理动作:
在图像被捕获的时间上用一个公共时钟来时间标记所述图像序列;
在音频信号被记录的时间上用一个公共时钟来时间标记所述音频信号;和
使用所述公共时钟的时间标记把图像序列和音频信号相关联。
5. 如权利要求 3 所述的方法,其特征在于,还包括在所述图像序列中的一个关键点上访问所述图像序列和所述相关联的音频信号。
6. 如权利要求 5 所述的方法,其特征在于,所述关键帧用来选择在所述序列中的所述关键点。
7. 如权利要求 1 所述的方法,其特征在于,还包括空间上过滤单元图像来改进分类结果的处理动作。
8. 如权利要求 1 所述的方法,其特征在于,还包括时间上过滤单元图像来改进分类结果的处理动作。
9. 如权利要求 1 所述的方法,其特征在于,还包括对关键帧图像进行颜色均衡来提高图像质量。
10. 如权利要求 1 所述的方法,其特征在于,校正所述白色书写板还包括:
在一个图像中指定所述白色书写板的四个角;
修剪每个图像的任何非白色书写板区域;和
使用双三次插值为每个白色书写板执行一个双线性的扭曲到一个矩形形状以在图像序列的每个被捕获图像中获得一个被修剪和被校正的白色书写板图像。
11. 如权利要求 1 所述的方法,其特征在于,用来提取白色书写板背景颜色的处理动作包括:
确定具有最亮亮度和最小变化的白色书写板单元;
确定在白色书写板颜色中的任何空缺;和
通过搜寻在所述空缺周围的白色书写板单元并且把颜色设置为不是一个空缺的最接

近单元的颜色从而填满每个空缺。

12. 如权利要求 1 所述的方法,其特征在于,用来确定在白色书写板中的任何空缺的处理动作包括处理动作:

指定一个白色书写板单元可能的颜色范围;

检查每个白色书写板单元是否在所述颜色范围之外;并且

指定具有在所述范围之外的一个颜色的任何单元作为一个空缺。

13. 如权利要求 1 所述的方法,其特征在于,用来提取白色书写板背景颜色的处理动作包括:

用直方图表示白色书写板图像亮度;

确定最高的白色书写板亮度;

指定相应于最高亮度的颜色作为初始的白色书写板颜色;

确定任何白色书写板分离物,所述分离物表示误差数据;

把任何分离物标记为空缺;和

通过搜寻在空缺周围的单元并且把颜色设置成不是一个空缺的最接近单元的颜色从而填满每个空缺。

14. 如权利要求 13 所述的方法,其特征在于,还包括在填满每个空间之后过滤白色书写板颜色图像。

15. 如权利要求 14 所述的方法,其特征在于,使用一个最小中值平方技术检测分离物。

16. 如权利要求 15 所述的方法,其特征在于,最小中值平方技术包括下列处理动作:

通过使平方误差的中值最小,在亮度 Y 或 RGB 空间中匹配一个平面;

把不符合这个模型的单元指定作为分离物。

17. 如权利要求 14 所述的方法,其特征在于,过滤所述白色书写板图像包括在红、绿、蓝颜色空间中匹配一个平面。

18. 如权利要求 1 所述的方法,其特征在于,用来把每个图像划分为单元的处理动作包括把每个图像划分为矩形单元。

19. 如权利要求 1 所述的方法,其特征在于,用来把每个图像划分为单元的处理动作包括把每个图像划分为单元,其中单元尺寸大约与书写板上的单个书写字符的尺寸相同。

20. 如权利要求 19 所述的方法,其特征在于,单元尺寸相应于在白色书写板中的一个 1.5 英寸乘 1.5 英寸的区域。

21. 如权利要求 19 所述的方法,其特征在于,单元尺寸由照相机分辨率和白色书写板尺寸所确定。

22. 如权利要求 19 所述的方法,其特征在于,单元尺寸是 1 像素乘 1 像素。

23. 如权利要求 1 所述的方法,其特征在于,用来把单元图像聚集成组的处理动作包括把被认为是随着时间流逝相同的单元图像聚集成组。

24. 如权利要求 23 所述的方法,其特征在于,把被认为是随着时间流逝相同的单元图像聚集成组的处理动作包括通过使用一个标准化交叉相关技术以在一个时间上比较两个单元来确定单元图像是否相同的动作。

25. 如权利要求 24 所述的方法,其特征在于,交叉相关分数范围从对于两个完全不相似图像的 -1 到对于两个完全一样的图像的 1。

26. 如权利要求 25 所述的方法,其特征在于,还包括应用一个 Mahalanobis 距离测试来确定两个单元是否为相同的一个处理动作。

27. 如权利要求 26 所述的方法,其中通过 $d = |\bar{I} - \bar{I}'| / (\sigma + \sigma')$ 给定的 Mahalanobis 距离测试;其中 I 是第一单元图像和 I' 是第二单元图像, $\bar{I}$ 是第一单元图像的中间颜色, $\bar{I}'$ 是第二单元图像的中间颜色, σ 是从 $\bar{I}$ 来的标准偏差和 σ' 是从 $\bar{I}'$ 来的标准偏差并且其中 I 和 I' 是两个单元图像,当且仅当 $d < T_d$ 和 $c > T_c$,它们才被认为是相同的。

28. 如权利要求 27 所述的方法,其特征在于, $T_d = 2$ 和 $T_c = 0.707$ 。

29. 如权利要求 1 所述的方法,其特征在于,单元分类处理动作还包括:

如果 RGB 值近似地相同,把单元图像指定作为白色书写板单元;

如果所述单元用一个或两个主要颜色混和的白色或灰色,则把单元图像指定作为笔划单元;以及

如果单元图像没有白色书写板单元或笔划单元的特征,则把单元图像指定作为前景单元。

30. 如权利要求 1 所述的方法,其特征在于,单元分类处理动作还包括:

确定当前单元图像的颜色分配;

确定相应的白色书写板单元的颜色分配;和

确定当前单元图像和相应的白色书写板单元的颜色分配是相同、还是不相同但具有很强的相似性、或者是完全不同。

31. 如权利要求 29 所述的方法,其中,当且仅当 $|\bar{I} - \bar{I}_w| / (\sigma + \sigma_w) < T_w$ 和 $\sigma / \sigma_w < T_o$ 时,单元图像被分类作为一个白色书写板单元,其中 $\bar{I}_w$ 是白色书写板颜色, σ_w 是从 $\bar{I}_w$ 来的标准偏差, $\bar{I}$ 是当前单元的中间颜色且 σ 是 $\bar{I}$ 的标准偏差;其中 $T_w = 2$ 和 $T_o = 2$;否则,它被分类作为一个前景物体单元。

32. 如权利要求 7 所述的方法,其特征在于,空间上过滤单元图像的处理动作包括下列的处理动作:

识别分离的前景单元;

把分离的前景单元重新分类作为笔划单元;

识别被连接到前景单元的笔划单元;和

把被连接到前景单元的笔划单元重新分类为前景单元。

33. 如权利要求 8 所述的方法,其特征在于,时间上过滤单元图像的处理动作包括下列的处理动作:

估计随着时间流逝的每个单元来确定在它被分类作为一个前景单元之前和之后它是否包含相同的笔划单元;

如果一个单元在它被分类作为一个笔划单元之前和之后被分类作为相同的笔划单元,则把它分类作为一个笔划单元。

34. 如权利要求 1 所述的方法,其特征在于,关键帧提取处理动作包括:

输入分类结果;

在图像序列中为每个帧计数笔划单元;

使用帧的笔划计数来确定笔划计数的峰值和谷值;

如果每个相邻的峰值和谷值之间的差异大于一个规定的阈值,则把谷值之间的数据指定作为章节并且把每个章节中的峰值指定作为表示章节的关键帧。

35. 如权利要求 34 所述的方法,其特征在于,规定的阈值是全部笔划计数的百分之二十。

36. 如权利要求 34 所述的方法,其特征在于,还包括重建所述关键帧图像的处理动作。

37. 如权利要求 36 所述的方法,其特征在于,重建所述关键帧图像的所述处理动作还包括下列处理动作:

输入图像序列的被分类的单元图像和被划分为单元图像的关键帧;

如果一个关键帧单元图像被分类作为一个白色书写板图像或一个笔划图像,它的图像被呈现;

如果一个关键帧前景单元图像在一个笔划的范围之中,这个单元图像用来自序列中的邻近图像的笔划单元图像来被呈现;

如果关键帧单元图像不被分类作为一个白色书写板图像,一个笔划图像或在一个笔划范围内的一个前景单元,它作为一个白色书写板图像被填满呈现。

38. 如权利要求 9 所述的方法,其特征在于,对关键帧图像进行颜色均衡以提高图像质量还包括下列的处理动作:

通过使用平均白色书写板颜色来测量在一个单元中的每个像素的颜色使背景为统一的白色并且增加笔划的颜色饱和度;和减少图像噪声。

39. 一种用来捕获会议的音频和视频内容的一个系统,其特征在于,包括:

一个捕获系统,用来捕获写在一个白色书写板上的一数据图像序列和在一个会议中发生的与声音对应的音频信号;

一个分析服务器,用来分析提取写在白色书写板上的关键数据帧的图像序列并且把音频信号和关键数据帧相关联;和

一个浏览模块,用来观看被分析的会议关键数据帧和相关联的音频,其中分析服务器通过以下来识别关键数据帧:

校正在图像序列中的每个图像中的白色书写板的一个视图;

提取白色书写板背景颜色;

把图像序列的每个图像划分成单元图像;

对于随着时间流逝的每个单元,贯穿整个图像序列聚类相似的单元图像;

把每个单元图像分类为笔划、前景物体或白色书写板单元;和

使用分类结果提取关键帧图像。

40. 如权利要求 39 所述的系统,其特征在于,捕获系统还包括:

一个照相机,被安置来捕获图像序列;

一个麦克风,来记录音频信号;和

一个计算机,用来记录图像序列和音频信号。

41. 如权利要求 40 所述的系统,其特征在于,照相机是静态照相机和摄像机中的至少一个。

42. 如权利要求 40 所述的系统,其特征在于,照相机推向尽可能接近于白色书写板来使分辨率最大。

43. 如权利要求 40 所述的系统,其特征在于,照相机被尽可能与白色书写板平行排列以使景象深度最小。

44. 如权利要求 39 所述的系统,其特征在于,

在记录开始之前一个或多个用户在捕获单元上注册一个用户标识符;

如果至少一个用户在捕获单元上注册,在事件记录和分析之后分析服务器产生一个访问令牌;

访问令牌和被分析的会议数据的计算机存储器位置被提供给注册的用户标识符;和

一个或多个用户访问被分析的事件数据的计算机存储器位置,来回顾被分析的会议数据。

45. 如权利要求 44 所述的系统,其特征在于,用户标识符是一个电子邮件地址。

46. 如权利要求 44 所述的系统,其特征在于,被分析的事件数据的计算机存储器位置是一个因特网网络站点的一个地址。

47. 如权利要求 39 所述的系统,其特征在于,还包括在所述捕获单元中的一个保密特性,它允许一个用户擦除至少以下之一:

序列图像的一部分;和

音频的一部分。

48. 如权利要求 47 所述的系统,其特征在于,所述保密特性通过按动一个图形用户界面按钮或者一个实际按钮来触发。

49. 如权利要求 39 所述的系统,其特征在于,捕获系统是便携式的。

50. 一种用来提取事件内容的方法,包括下列处理动作:

用一个照相机捕获写在一个非电子白色书写板上的内容的图像序列;

记录与图像序列相关联的音频信号;和

通过以下来分析图像序列来分离摘要所述书写板内容的关键点的帧;

校正正在图像序列中的每个图像中的白色书写板的一个视图;

提取白色书写板背景颜色;

把图像序列的每个图像划分成单元;

对于随着时间流逝的每个单元,贯穿整个图像序列把单元图像聚集成组;

把每个单元图像分类为笔划、前景物体或白色书写板单元;和

使用分类结果提取关键帧图像。

51. 如权利要求 50 所述的方法,其特征在于,分析处理动作还包括使用光学字符识别技术来从关键帧图像提供可编辑的文本。

52. 如权利要求 50 所述的方法,其特征在于,分析处理动作还包括使用声音识别技术来把音频转换为文本。

53. 如权利要求 50 所述的方法,其特征在于,捕获处理动作还包括使用手势识别技术。

54. 如权利要求 50 所述的方法,其特征在于,所述捕获处理动作还包括:

安装已知颜色特性的一个颜色补丁在白色书写板的上方;

在所述图像序列中的一个图像中捕获所述颜色补丁;

确定在图像中被颜色补丁展示的颜色;

在所述图像序列中的下一个图像中捕获所述颜色补丁;

确定在所述下一个图像中被所述颜色补丁展示的颜色；

确定被补丁展示的颜色在图像和下一个图像之间是否已经改变了大于一个规定的阈值；和

如果展示的颜色超过了所述规定的阈值，则把照相机的曝光参数调节为一个设置，在照相机已经被设置为所述曝光设置时，所述设置使下一个图像中所述颜色补丁的颜色与所述图像中展示的颜色补丁的颜色相匹配。

55. 如权利要求 54 所述的方法，其特征在于，所述把照相机的曝光参数调节为一个设置的处理动作对于每个被捕获的图像都执行。

56. 一种用来提取一个会议的内容的系统，包括：

一个捕获系统，用于捕获写在一个白色书写板上的一数据序列，所述捕获系统跟踪笔位置以推断写在一个白色书写板上的内容并且记录与写在一个白色书写板上的所述内容相关联的音频信号；

一个分析服务器，用来分析图像序列，所述图像序列提取写在白色书写板上的关键数据帧并把音频信号与关键数据帧相关联。

57. 如权利要求 56 所述的系统，其特征在于，所述分析服务器执行下列动作：

把白色书写板的每个区域划分为单元；

对于随着时间流逝的每个单元，贯穿整个被写的的数据序列中相同的单元聚集成组；

把每个单元分类为笔划或白色书写板单元；和

使用分类结果提取关键帧图像。

58. 如权利要求 57 所述的系统，其特征在于，白色书写板单元被划分成大约一个写入字符尺寸的单元。

用于白色书写板和音频捕获的一个系统和方法

技术领域

[0001] 本发明涉及用来记录会议的一个系统和方法。具体而言,本发明涉及用来捕获一个会议的白色书写板内容和音频的系统和方法。

背景技术

[0002] 会议占用了许多工作人员很大部分的工作时间。更有效的利用在会议中消耗的时间将转化为很大的生产力增长。

[0003] 很多会议现场使用一个白色书写板,广泛的用作集体讨论对话,演讲,项目计划会议,专利发布,等等。在这些会议中,做笔记和拷贝写在书写板上的内容经常会妨碍许多参加者积极的贡献和参与。结果,在一些自动化的方式中,已经采取一些努力来捕获白色书写板的内容。

[0004] 几个技术已经被发展来自动的捕获白色书写板内容,最早的白色书写板复印机是具有一个内置复印机的一个特别的白色书写板。只要单击一个按钮,白色书写板内容就能被扫描和打印。一旦白色书写板内容在纸上,它能被影印,传真,储存在文件柜中,或扫描成数字格式。

[0005] 更多最近的技术试图从开始就在数字格式中捕获白色书写板内容。它们一般分为二类——一些用来捕获白色书写板的图像和一些用来跟踪笔的位置并且从那里推断白色书写板的内容。

[0006] 在第一类中的设备直接捕获白色书写板的图像。经常使用国家电视系统委员会制式 (NTSC) 分辨率的摄影机,因为它们的费用低。由于这些摄影机通常不具有足够的分辨率来清晰的捕获在一个典型会议室大小的白色书写板上所写的内容,几个视频帧必须被订合在一起产生一个单独的白色书写板图像。在这第一类中的另一个设备是数字静态照相机。因为高分辨率数字照相机越来越便宜,用一个数字照相机来对书写板进行快照变成是一个流行的选择。

[0007] 在第二类中的设备跟踪被使用来在白色书写板上高频书写的笔的位置并且从笔坐标的历史记录来推断白色书写板的内容。有时它们包括附加在一个传统白色书写板侧面的一个附加设备并且为干墨水笔和擦除器使用特殊的盒子。当压书写板时,每支笔放射超声波脉冲。在附加设备上的二个接收器使用在音频脉冲到达时间中的差异来对笔坐标作三角测量。由于笔坐标的历史记录被捕获,在任何给定片刻的在白色书写板上的内容可能被随后重建。这种类型的白色书写板记录的用户能像一个电影一样播放白色书写板内容。因为内容以矢量形式被捕获,它能用低带宽和小存储要求来被传送和存档。

[0008] 电子白色书写板也使用笔跟踪技术。它们比使白色书写板成为一个交互式设备的使用先前讨论的附加设备的系统更深了一步。用户用一个被计算机跟踪的特殊铁笔在显示器上书写。计算机在屏幕上呈现铁笔无论在那里接触屏幕的笔划—好像墨水被铁笔沉淀的一样。因为笔划是计算机产生的,它们能被编辑,回画和动画。用户也能对计算机发布手势命令并且在同一屏幕上显示其它计算机应用。

[0009] 然而电子白色书写板由于它们的高成本和小尺寸（一个电子白色书写板在对角线上很少超过 6 英尺）当前还只有有限的安装基础。此外，使用笔跟踪设备的系统具有下列缺点：1) 如果系统没有开机或用户不用特殊的笔来写，内容不能被设备恢复；2) 许多人喜欢使用他们的手指代替特殊擦除器来在白色书写板上改正小错误。这种普通的行为导致额外的笔划出现在被捕获的内容上；3) 人们不得不使用特殊的干墨水笔适配器，对于某些设备它们变得越来越粗和硬来受压。和 4) 笔跟踪的不精确经常导致相邻笔划的记录不准确。

[0010] 除了以上讨论的关于白色书写板捕获方法的工作之外，在多媒体体验的捕获，综合和访问上已经实行了大量的研究，特别在演讲和会议方面。人们已经开发了一些技术和系统，它们使用手写笔记，白色书写板内容，幻灯片，或手工注解来索引被记录的视频和音频，以便于存取。

[0011] 例如，在一个名为教室 2000 项目的项目中，Abowd et al. 使用一个电子白色书写板来按时间标记墨水笔划，所以观众（学生）能使用墨水笔划作为被记录的视频和音频的索引。关键帧（称为页面）基于在由电子白色书写板提供的擦除事件而被计算。教室 2000 项目，不管怎样，都需要一个电子白色书写板。使用一个电子白色书写板，从最终用户的观点来看有许多缺点。首先，绝大多数的办公室和会议室没有安装电子白色书写板。其次，已经显示出人们发现使用一个普通的白色书写板比电子书写板要自然的多。第三，用一个照相机捕获的图像提供太多的前后关系信息例如谁在写和那个主题正在被讨论（通常被手形指示）。除了这些缺点之外，电子白色书写板可能是昂贵的因而不易于可得到和使用。

发明内容

[0012] 本发明涉及一个系统和过程，为了捕获白色书写板内容和相关音频而克服在系统中的上述限制。

[0013] 白色书写板捕获系统与上述系统的不同之处在于直接捕获白色书写板的图像因为它计算笔划的时间标记和通过在被捕获图像上进行分析计算关键帧。关键帧图像包括在白色书写板上的所有重要内容并且适合作为记录的一个摘要。它们能被删节和粘贴到其它文件或作为笔记打印。时间标记和关键帧是被记录的音频的有效的索引。另外，白色书写板捕获系统发明使用一个普通的白色书写板，不是一个电子白色书写板，像一些其它白色书写板捕获系统一样。因而，系统能不加修改的与任何现存白色书写板一起使用。

[0014] 白色书写板系统用一个照相机捕获写在一个非电子白色书写板上的内容的一序列的图像。它同时记录会议的音频信号。一旦记录完毕，图像序列被分析来分离关键帧，它摘要写在白色书写板上内容的关键点。音频记录通过与被记录音频和图像序列关联的时间标记被与在关键帧上的笔划相互关联。这些时间标记通过图像分析被计算。

[0015] 用来获得关键帧的一般分析过程包括在一序列的图像中的每个图像中校正白色书写板的视图。白色书写板背景颜色也被提取和图像序列中的每个图像被分为单元。随着时间流逝相同的单元图像被聚类在一起，这将在随后作更为详细的解释。每个单元图像接着被分类作为一个笔划、一个前景物体或白色书写板单元。接着关键帧图像用分类结果来提取。单元图像在关键帧提取之前被空间上和时间上过滤来改进分类结果。另外，一旦被提取，关键帧图像能被颜色均衡以提高图像质量。

[0016] 更为详细的是,校正白色书写板视图包括修剪每个图像的任何非白色书写板区域。接着白色书写板的四个角在每个图像中被指定。接着为使用双三次插值的每个图像执行一个双线性的扭曲以获得在每个被捕获图像中的一个被修剪和被校正的白色书写板图像。

[0017] 有二种方法可能被用来提取白色书写板的背景颜色。第一个方法包括确定具有最亮亮度和最小变化的白色书写板单元。具有最亮亮度和最小变化的颜色被指定作为白色书写板的背景颜色。一旦白色书写板的背景颜色这样被确定,在白色书写板颜色中的任何空缺被发现并且通过搜寻在每个空缺周围的白色书写板单元来被填满。接着每个空缺的颜色被设置成不是一个空缺的最接近单元的颜色。

[0018] 用来提取白色书写板背景颜色的第二个方法包括用直方图表示白色书写板图像亮度和确定峰值白色书写板亮度。相应于峰值亮度的颜色被指定作为初始白色书写板颜色。接着任何白色书写板颜色分离物(误差的数据)被使用一个最小中值平方技术来确定。这些分离物被用空缺来作标记并且在如上面讨论的确定白色书写板颜色的第一个方法中相同的方式中被填满。在填充每个空缺之后,白色书写板颜色图像可能被过滤。

[0019] 将在输入序列中的每个图像分为单元的过程提高了分析处理速度。代表性地,每个图像被分成单元所以单元尺寸大约是在书写板上的一个单个字符的相同尺寸。对于一个典型的会议尺寸的白色书写板,这相等于 1.5 英寸乘 1.5 英寸,或 25 像素乘 25 像素。然而,另一方面,所有的分析能在按照像素基础的一个像素上被执行。

[0020] 一旦输入图像的序列被校正和白色书写板颜色已经被确定,单元图像被聚类。被认为随着时间流逝相同的单元图像被在一起聚类在组中。一个标准化交叉相关技术和一个 Mahalanobis 距离测试被用来确定二个单元是否相同。

[0021] 单元分类过程确定一个单元图像是否是一个白色书写板单元,一个笔划或一个前景物体。如果红色,绿色,蓝色(RGB)值是近似相同,一个单元图像被指定为一个白色书写板单元。另一方面,如果单元主要是用一个或二个主要颜色混和的白色或灰色,一个单元图像被指定为一个笔划。否则,单元图像被指定为一个前景单元。单元分类过程确定当前单元图像的颜色分配和相应白色书写板单元的颜色分配。接着单元被分类,基于如果当前单元图像和相应白色书写板单元的颜色分配相同,不同但具有一个强烈的相同性,或完全不同。

[0022] 以上分类过程只使用在一个单独单元上的颜色信息。更为准确的结果能通过使用在单元组中空间和时间关系来达到。在空间上过滤中,在每个白色书写板图像上的二个操作被执行。首先,被分离的前景单元被识别和重新分类成笔划。其次,被立即连接到一些前景单元的笔划单元被重新分类成前景单元。关于时间上过滤,基本的观察是当笔划被删除后事实上不可能在完全准确的位置中写相同的笔划。换句话说,如果对于任何给定的单元,二个不同帧的单元图像包含相同的笔划,那么在二帧之间的所有单元图像必须具有相同的笔划除非那儿有一个前景物体阻挡单元。在时间上过滤步骤中,这个单元将被分类作为一个笔划,只要在前景物体阻挡它之前或之后它被曝光到照相机。

[0023] 关键帧接着能被提取。到此为止,分类结果被使用并且在图像序列中为每个图像或帧计数笔划单元。笔划计数的峰值和谷值被确定。如果在笔划计数的每个相邻的峰值和谷值之间的差异大于一个规定的阈值,在谷值之间的数据被指定为章节(每个章节表示一个不同的主题)和在每个章节中的峰值被作为表示章节的关键帧。

[0024] 随后重新构造关键帧图像。这包括输入分类单元图像和被分割为单元图像的关键帧。如果一个关键帧单元图像被分类作为一个白色书写板图像或一个笔划图像,它的图像被分别的呈现作为一个白色书写板图像或一个笔划图像。交替地,如果一个关键帧前景单元图像在一个笔划的范围之中,这个单元图像在序列中从邻近图像中用笔划单元图像来呈现。如果关键帧单元图像不被分类作为一个白色书写板图像,一个笔划图像或在一个笔划范围内的一个前景单元,它作为一个白色书写板图像被呈现。

[0025] 随后使用颜色均衡,通过使用平均白色书写板颜色来测量在一个单元中的每个像素的颜色使背景为统一的白色和增加笔划的颜色饱和度来改进关键帧图像的图像质量。图像噪声也被减少。

[0026] 在分析服务器处理图像序列和产生索引和关键帧图像之后,它将到被处理记录的统一资源定位器 (URL) (一个网络站点或其它因特网服务的“地址”或位置) 发送电子邮件到注册会议参与者。用户可以点击 URL 来登录一个浏览软件。浏览器允许用户浏览关键帧图像和快速访问与一个特殊主题相关联的音频。

[0027] 浏览软件的用户接口 (UI) 具有不同的部件。浏览器 UI 的主要元素包括一个关键帧平面,在那里关键帧略图被显示,并且浏览器的主显示平面显示从照相机来的原始图像的一个合成物和当前关键帧图像。

[0028] 关键帧平面也结合了一个背景透明的滑块,它允许用户在主显示平面上调节被显示的图像从未加工的输入图像到关键帧图像。当前笔划,已经被写在会议回放时间线中的笔划,比将来笔划呈现得更暗和更明显。参与者将来将写的笔划被显示在一个类似于幻影的风格中。这个视觉显示技术使用以下处理来实现。当前的白色书写板内容通过使用当前章节和时间标记信息的关键帧图像来被呈现。然后将来笔划被呈现,使用一个 Gaussian 过滤器使之被转换成灰度色标和模糊不清。这二个图像接着被添加,并且合成图像与从输入序列中被校正的图像一起 alpha 混和。用户能用 GUI 的滑块来控制 alpha 值从 0,只显示被呈现的关键帧白色书写板图像,到 1,正确的显示原始图像。

[0029] 一个 VCR 和标准时间线控制在较低于浏览器 UI 的主显示平面的左下角中被提供。VCR 和标准时间线控制允许用户在图像 / 音序列中顺序向后或向前,慢或快或停止,非常像在一个典型的录像机 (VCR) 上找到的控制。一个时间线条用一个条图形显示音频 / 图像序列的长度,并且提供会议回放的开始时间,结束时间和当前时间的用数字表示的值。在这个条上的一个指针能被选中并且被沿着时间线条向前和向后的拖动来在图像 / 音序列中线性地顺序向前和向后。

[0030] 对记录音频的二个级别的非线性访问在视觉索引的上下文中被提供。非线性访问的第一级别是通过关键帧略图的使用。用户能单击一个关键帧略图从而跳到对于相应关键帧的音频的起始点 (例如,章节的开始)。每个关键帧有与它关联的一个时间范围,用来帮助用户确定与那个特殊关键帧关联的时间范围。对记录音频的非线性访问的第二级别是通过在每个关键帧中笔划的使用。当在主窗口中指针悬在一个笔划单元 (当前笔划单元或将来笔划单元) 上时,指针被变为一个“手”形,显示它是可选的 (例如,用一个鼠标“可单击的”)。用一个鼠标或其它输入设备在单元上双击使应用程序成为音频回放模式。回放开始于当被单击的笔划单元被写入时的会议时间。用户仍能单击其它的笔划单元来跳到会议的其它部分。结合 VCR 和标准时间线控制 1514,这二个级别的视觉索引允许用户在一个非常

有效的方式中浏览一个会议。

[0031] 如先前所述,关键帧图像的略图在关键帧平面中被列出。选择一个略图使相应的关键帧图像到左边的主窗口上并且使应用程序为图像观看模式,其中用户能使用缩放控制按钮来放大和缩小,在图像中读取文本和图表,或切割和粘贴图像的一部分到其它文档。另外,全部的关键帧能被切割和粘贴到其它文档或被打印作为笔记。

[0032] 在白色书写板捕获系统中,会议参与者被要求在会议记录的开始向捕获软件注册。所有的记录会议存在于一个互联网服务器上。如果没有人注册,会议被张贴在一个公开可访问的互联网页面上。如果至少有一个参与者注册,在会议记录和分析之后产生一个访问令牌。令牌是包含一个唯一会议标识符的一个长随机产生的字符串。包含令牌的 URL 被发送电子邮件到注册的参与者。接受者到 URL 启用互联网浏览软件来回顾会议。他们也转寄 URL 到没有参加会议的人。

[0033] 上面描述的基本白色书写板捕获系统能被与许多其它技术和设备结合在一起来呈现另一个实施例。在一个这样的实施例中,传统的光学字符识别(OCR)在关键帧上被执行来提供容易被使用来建立文档或演讲视图的可编辑文本。

[0034] 在另一个实施例中,传统的音频识别软件被使用来转换被捕获数据的音频部分到文本。这允许会议记录和其它文档的简单建立。它也提供一个相对不昂贵的方法来提供会议信息到听力损伤者。

[0035] 白色书写板捕获系统通过,例如,使用具有一个麦克风的一台笔记本电脑和装配在三角架上的一个照相机也能被制成是便携式的。这种设置只需要一个附加的初始校准来确定相对于白色书写板的照相机的位置。这个校准能通过手工确定在图像中平面的四个角手工的,或通过使用例如边缘检测的方法自动的被执行。

[0036] 白色书写板捕获系统的分析软件通过使用笔跟踪的白色书写板系统也能被使用来确定关键帧从而推断白色书写板内容。使用具有这样一个系统的白色书写板捕获系统分析软件简化了分析过程。白色书写板背景颜色的确定是不必要的,也不必要白色书写板区域的校正,不需要空间或时间上过滤,并且白色书写板单元的分类变得较为简单因为单元图像将或者是笔划或者是白色书写板,因为没有前景物体将干扰写在白色书写板上的文本。

[0037] 另外,为了达到较高的帧频,一个高分辨率视频摄像机例如一个 HDTV 摄像机能代替一个静态照相机被使用。

[0038] 还在另一个实施例中,白色书写板捕获系统结合手势识别以使用手势命令。例如,在白色书写板上的某处一个命令方框能被书写。当用户移动或指向方框时,系统使用手势识别来时间标记在做手势的特殊时间上的图像。

[0039] 白色书写板捕获系统减轻了会议参与者烦琐的作笔记的工作,因而他们能集中注意力在会议中提出和吸收观点。通过提供摘要白色书写板内容的关键帧图像和结构化的视觉索引到音频,系统帮助参与者在稍后时间上回顾会议。此外,没有参加会议的人常常能在一小部分时间中理解会议要旨。

附图说明

[0040] 本发明特殊的特征,方面,和优点结合下列描述,所附的权利要求书,和附图将变

得能较好的理解,其中:

[0041] 图 1 是描述为了实现发明组成一个典型系统的一个通用计算设备的一个图表。

[0042] 图 2 是描述白色书写板捕获系统的三个主要成分 - 捕获单元,分析服务器和浏览软件的一个图表。这个图表使用一个原型的白色书写板捕获系统来被捕获。

[0043] 图 3 是根据本发明的一个白色书写板捕获系统的一个示意图。

[0044] 图 4 是从一个输入图像序列来的显示被选择帧的一系列的图像。

[0045] 图 5 是根据发明描述系统和方法的图像分析过程的一个流程图。

[0046] 图 6A 是计算白色书写板颜色的一个第一技术。

[0047] 图 6B 是计算白色书写板颜色的一个第二技术。

[0048] 图 7 是显示白色书写板颜色提取结果的一系列的图像。左边的图像是计算白色书写板颜色的第一策略的结果,中间的图像是计算白色书写板颜色的第二策略的结果,和右边的图像显示实际空缺的白色书写板图像。

[0049] 图 8 是根据本发明描述系统和方法的单元分类过程的一个流程图。

[0050] 图 9 是分类结果的一系列采样。在修剪和校正之后上面的图像相应于在图 5 中的图像。

[0051] 图 10 是对于在图 4 的序列中笔划数目对比时间的一个图。

[0052] 图 11 是用来从一序列的输入图像中选择关键帧的一般过程的一个流程图。

[0053] 图 12 是根据本发明描述在系统和方法中识别章节和关键帧的过程的一个流程图。

[0054] 图 13 是根据本发明描述在系统和方法中重建关键帧图像的过程的一个流程图。

[0055] 图 14 是根据本发明描述在系统和方法中颜色均衡关键帧图像的过程的一个流程图。

[0056] 图 15 描述白色书写板捕获系统的浏览器界面的一个图像。每个关键帧图像表示在记录中的一个关键时刻的白色书写板内容。

[0057] 图 16 是根据本发明描述在系统和方法中显示当前和将来笔划的过程的一个流程图。

[0058] 图 17 是根据本发明描述在系统和方法中使用的安全处理的一个流程图。

[0059] 图 18A 提供在发明的一个工作实施例的三个安装地点上拍摄的白色书写板内容的采样图像。

[0060] 图 18B, 18C 和 18D 是一系列的图,用来描述输入(图 18B)和输出本发明的一个工作实施例的关键帧图像(图 18C 和 18D)。

具体实施方式

[0061] 在本发明首选实施例的下列描述中,组成本文中一部分的附图被参考,并且其中被显示作为说明特殊实施例,在其中,发明可能被实行。其它实施例可能被使用和不偏离本发明范围的情况下做出结构的变化是能够被理解的。

[0062] 1.0 典型操作环境

[0063] 图 1 举例说明发明可能在其上被实现的一个合适的计算系统环境 100 的一个例子。计算系统环境 100 只是一个合适的计算环境的一个例子并且不被用来建议如作为使用

范围或发明功能的任何限制。计算环境 100 将不被解释为具有任何相依性或在典型的操作环境 100 中涉及举例说明的部件的任何一个或结合的要求。

[0064] 发明与许多其它通用的或专用的计算系统环境或配置一起操作。可能适合与发明一起使用的著名的计算系统,环境和 / 或配置的例子包括,但不限于,个人计算机,服务器计算机,手持型或膝上型设备,多处理器系统,基于微处理器的系统,机顶盒,可编程的消费者电子产品,网络 PC,小型机,大型计算机,包括任何上述系统或设备,和相似的分布式计算环境。

[0065] 发明可能在计算机可执行指令的一般环境中被描述,例如正在被一个计算机执行的程序模块。一般的,程序模块包括执行特殊任务的或实现特殊抽象数据类型的例行程序,程序,对象,部件,数据结构,等。发明也可能在分布式计算环境中被实施,其中通过一个通讯网络连接的远程处理设备执行任务在一个分布式计算环境中,程序模块可能位于包含存储器存储设备的当地的或远程的计算机存储介质中。

[0066] 参考图 1,用来实现发明的一个典型系统包括以一个以计算机 110 形式表示的一个通用计算设备。计算机 110 的部件可能包括,但不局限于,一个处理单元 120,一个系统存储器 130,和连接不同的系统部件包括系统存储器到处理单元 120 的一个系统总线 121。系统总线 121 可能是几种类型的总线结构中的任何一种,包括一个存储器总线或存储器控制器,一个外围总线和使用多种总线结构中任意一种的一个局域总线。作为例子,和没有限制的,这种结构包括工业标准结构 (ISA) 总线,微通道结构 (MCA) 总线,扩展工业标准结构 (EISA) 总线,视频电子标准协会 (VESA) 局域总线,和周边组件扩展接口 (PCI) 总线也被称为 Mezzanine 总线。

[0067] 计算机 110 典型的包括许多计算机可读介质。计算机可读介质可能是能被计算机 110 访问的任何可利用的介质并且包括易变的和不易变的介质,可移动的和不可移动的介质。作为例子,和没有限制的,计算机可读介质可能包括计算机存储介质和通讯介质。计算机存储介质包括在任何方法或技术中实现的易变的和不易变的,可移动的和不可移动的介质,用来存储信息例如计算机可读指令,数据结构,程序模块或其它数据。计算机存储介质包括,但不局限于, RAM, ROM, EEPROM, 闪存,或其它存储器技术, CD-ROM, 数字化多用途光盘 (DVD) 或其它光学磁盘存储,磁性磁带盒,磁性磁带,磁性磁盘存储或其它磁性存储设备,或能被使用来存储所需信息和能被计算机 110 访问的任何其它介质。通讯介质典型的包含计算机可读指令,数据结构,程序模块或在一个被调制数据信号中的其它数据例如一个载波或其它传送机制并且包括任何信息传递介质。术语“被调制的数据信号”意指一个信号,它具有一个或多个它的特征集或在如信号中编码信息这样一种方式中被改变。作为例子,和没有限制的,通讯介质包括有线介质例如一个有线网络或直接有线连接,和无线介质例如声音的, RF, 红外线和其它无线介质。上面任何的结合也将被包括在计算机可读介质的范围内。

[0068] 系统存储器 130 包括以易变的和 / 或不易变的存储器例如只读存储器 (ROM) 131 和随机存取存储器 (RAM) 132 的形式存在的计算机可读介质。一个基本的输入 / 输出系统 133 (BIOS), 包含帮助在计算机 110 中的组件之间转移信息的基本例行程序例如在启动时, 被典型的存储在 ROM 131 中。RAM 132 典型的包含数据和 / 或能被立即访问和 / 或当前正被操作在处理单元 120 上的程序模块。作为例子,和没有限制的,图 1 举例说明了操作系统

134, 应用程序 135, 其它程序模块 136, 和程序数据 137。

[0069] 计算机 110 也可能包括其它可移动的和不可移动的, 易变的和不易变的计算机存储介质。只作为例子, 图 1 举例说明一个硬盘驱动器 141 从不可移动的, 不易变的磁性介质中读取或写入, 一个磁盘驱动器 151 从一个可移动的, 不易变的磁盘 152 中读取或写入, 和一个光盘驱动器 155 从一个可移动的, 不易变的光盘 156 中读取或写入例如一个 CD ROM 或其它光学介质。使用在典型的操作环境中的其它可移动的 / 不可移动的, 易变的 / 不易变的计算机存储介质包括, 但不限于, 磁带盒, 闪存卡, 数字化多用途光盘, 数字化录像磁带, 固 (体状) 态 RAM, 固 (体状) 态 ROM, 和相似的。硬盘驱动器 141 典型的通过一个不可移动的存储器接口例如接口 140 被连接到系统总线 121 上, 和磁盘驱动器 151 和光盘驱动器 155 通过一个可移动的存储器接口, 例如接口 150 被连接到系统总线 121 上。

[0070] 驱动器和上面讨论的在图 1 中举例说明的与它们关联的计算机存储介质, 为计算机 110 提供计算机可读指令, 数据结构, 程序模块和其它数据的存储。在图 1 中, 例如, 硬盘驱动器 141 被举例说明为存储操作系统 144, 应用程序 145, 其它程序模块 146, 和程序数据 147。值得注意的是这些部件可能相同于或不同于操作系统 134, 应用程序 135, 其它程序模块 136, 和程序数据 137。操作系统 144, 应用程序 145, 其它程序模块 146, 和程序数据 147 在这里被给定不同的数字来至少阐明它们是不同的拷贝。一个用户可能通过输入设备输入命令和信息到计算机 110, 输入设备例如一个键盘 162 和指示设备 161, 一般指一个鼠标, 轨迹球或触垫。其它输入设备 (未显示) 可能包括一个麦克风, 操纵杆, 游戏垫, 圆盘式卫星电视天线, 扫描仪, 或相似的。这些和其它的输入设备经常通过被连接到系统总线 121 上的一个用户输入接口 160 被连接到处理单元 120, 但可能通过其它接口和总线结构被连接, 例如一个并行口, 游戏口或一个通用串行总线 (USB)。一个显示器 191 或其它类型的显示设备也通过一个接口例如一个视频接口 190 被连接到系统总线 121。除显示器之外, 计算机也可能包括其它外围输出设备例如扬声器 197 和打印机 196, 这可能通过一个输出外围接口 195 被连接。本发明特殊的重要之处在于, 能够捕获一序列图像 193 的一个照相机 192 (例如一个数字 / 电子静态或视频摄像机, 或胶卷 / 照相扫描仪) 也能被包括作为到个人计算机 110 的一个输入设备。此外, 虽然只有一个照相机被描述, 多个照相机能被包括作为到个人计算机 110 的一个输入设备。从一个或多个照相机来的图像 193 通过一个适当的照相机接口 194 被输入到计算机 110。这个接口 194 被连接到系统总线 121, 从而允许图像被发送到或存储在 RAM 132 中, 或与计算机 110 关联的一个其它数据存储设备中。然而, 值得注意的是图像数据在不需要使用照相机 192 的情况下也能从任何上述的计算机可读介质中被输入到计算机 110。

[0071] 计算机 110 可能操作在一个使用逻辑连接到一个或多个远程计算机的网络环境中, 例如一个远程计算机 180。远程计算机 180 可能是一个个人计算机, 一个服务器, 一个路由器, 一个网络 PC, 一个对等设备或其它通用网络节点, 并且典型的包括关于计算机 110 的上述的许多或所有的组件, 尽管只有一个存储器储存设备 181 在图 1 中已经被举例说明。在图 1 中描述的逻辑连接包括一个局域网 (LAN) 171 和一个广域网 (WAN) 173, 但也许还包括其它网络。这种网络环境在办公室, 企业范围计算机网络, 企业内部互联网, 因特网中是普遍的。

[0072] 当被使用在一个 LAN 网络环境中时, 计算机 110 通过一个网络接口或适配器 170

被连接到 LAN 171。当被使用在一个 WAN 网络环境中时,计算机 110 典型的包括一个调制解调器 172 或用来在 WAN 173 上建立通讯的其它办法,例如因特网。调制解调器 172,可能是内置或外置的,通过用户输入接口 160 可能被连接到系统总线 121,或其它合适的机制。在一个网络环境中,被描述的关于计算机 110 的或在其中一部分的程序模块可能被存储在远程存储器存储设备中。作为例子,和没有限制的,图 1 举例说明如存在于存储器设备 181 上的远程应用程序 185。值得欣赏的是所示的网络连接是典型的和在计算机之间建立一个通讯链接的其它方法可能被使用。

[0073] 典型的操作环境现在已经被讨论,这个描述部分的剩余部分将被专用于具体表达发明的程序模块的一个描述上。

[0074] 2.0 白色书写板捕获系统和方法

[0075] 2.1 系统结构

[0076] 概念上,白色书写板捕获系统包括三个主要成分:一个捕获单元 202,一个分析/处理服务器 204,和浏览软件 206,如在图 2 中所示。

[0077] 1. 捕获单元:捕获单元被使用来捕获白色书写板内容的图像并且记录与白色书写板内容产生联系在一起的音频。捕获单元被安装在会议发生的一个房间中。如在图 3 中所示,它包括一个数字照相机 302,一个麦克风 304,和一台个人计算机 (PC) 306。捕获单元拍摄白色书写板 308 的图像并且通过麦克风 304 记录存储到 PC 306 的音频。被拍摄的图像和音频都被做了时间标记。图像和音频采样在一个普通时钟上被获得,通常是系统时钟。普通时钟的时序与图像和音频采样相关联并且被存储作为它们的时间标记。

[0078] 2. 分析服务器:分析服务器 204 位于一个中心位置并且分析和存储被记录的图像数据。在一个实施例中,在用户在捕获单元中停止记录之后,一个分析程序被自动的启用。在处理被记录数据之后,包含 URL,其中会议记录可得到的电子邮件被发送到注册的参与者。如果没有人注册,会议记录可能被张贴到一个公开可访问的互联网站点。

[0079] 3. 浏览软件:浏览软件 206 允许用户观看和回放记录的和被分析的会议数据。浏览软件 206 更适宜于作为一个环球网插件程序被提供使希望观看会议记录的用户能够安装。一旦安装完,用户能单击上述的 URL 启用软件来访问在分析服务器上的数据。

[0080] 2.2 图像获得

[0081] 输入到白色书写板捕获系统的是一系列的静态数字图像。图 4 显示了这种图像的一个典型集合。图像序列被分析来确定何时和何地用户在书写板上书写和提取一序列的摘要贯穿一个会议的白色书写板内容的关键帧图像。

[0082] 允许照相机通过一个计算机控制的任何相对高分辨率的照相机为了图像获得能被使用。照相机更适宜于安放在一个会议室的侧面或者背面。照相机被推向尽可能的接近白色书写板来得到有效分辨率的最大值。在安装之后照相机是静止的并且作出假设白色书写板不移动,所以白色书写板图像在整个捕获序列中是静止的。

[0083] 如果一个被使用的照相机只有自动对焦模式,如果在白色书写板前的一个物体触发了照相机的自动对焦机制的关注,白色书写板变成了焦点没对准。这个问题能被减轻,通过尽可能的平行排列照相机的图像平面到白色书写板来将景象深度减到最少和/或将光圈减到最小来增加现场深度。实际上,在白色书写板捕获系统的一个工作实施例中只有 1-2% 的帧被观察到焦点没对准。

[0084] 照相机尽可能快的拍摄图片并且传递图像到 PC,较适宜于通过一个 USB 连接。一个 JPEG 图像在白色书写板捕获系统的一个工作实施例中大约每 5 秒被获得。曝光和白色均衡参数典型的保持不变。假设光线设置在一个会议中不改变,白色书写板背景的颜色将在一个序列中保持恒定。

[0085] 发现略微曝光下的图像给出较好的颜色饱和度,将使稍后讨论的笔划提取过程更为准确。在记录之后的一个颜色均衡步骤可能被执行使浅灰色的白色书写板图像更为吸引人。

[0086] 2.3 图像序列分析

[0087] 因为在书写板上书写的一个人是在数字照相机和白色书写板之间的视线中,他/她经常遮蔽白色书写板的一些部分并且在其它部分上投射阴影。有必要的在笔划,前景物体(例如,在书写板上书写的人),和白色书写板之间区别。一旦分类结果已知,关键帧图像和一个索引可能被浏览软件使用。

[0088] 不在一个每个像素级别上分析图像(尽管这可能被完成了),白色书写板区域被划分为矩形单元以降低计算成本。单元尺寸被选择为大约与在书写板上的一个单一字符的期望尺寸相同(在一个工作实施例中,大约是 1.5 英寸乘 1.5 英寸,或 25 像素乘 25 像素)。因为单元栅格划分在输入序列中的每个帧为单元图像,输入可能被认为是单元图像的一个三维矩阵(例如,x,y,时间)。在输入图像已经被校正之后,典型的执行每个图像被分割为单元。

[0089] 如在图 5 中和如下所示的,分析输入图像序列的程序的一般处理动作如下:

[0090] 1. 校正在序列中每个图像的白色书写板区域(处理动作 502)。

[0091] 2. 提取白色书写板背景颜色(处理动作 504)。

[0092] 3. 划分在序列中的每个图像到相应的单元图像之后贯穿整个序列对于相同单元把单元图像聚集成组(处理动作 506)。如果二个单元图像随着时间流逝被认为是相同的,它们被聚类在相同的组中。

[0093] 4. 分类每个单元图像作为一个笔划,一个前景物体,或白色书写板(处理动作 508)。

[0094] 5. 空间上和时间上过滤单元图像以改进分类结果(处理动作 510)。

[0095] 6. 使用分类结果提取关键帧图像(处理动作 512)。

[0096] 7. 颜色均衡关键帧图像(处理动作 514)。

[0097] 在下面的段落中,如在图 4 中所示的运行例子被使用来举例说明输入图像序列分析过程。

[0098] 2.3.1 校正白色书写板图像

[0099] 在传递图像序列到笔划提取过程之前,非白色书写板部分被修剪和图像被校正。因为在工作实施例中使用的照相机的镜头有相当低的光线失真,只不过有必要识别白色书写板的四个角(否则在校正图像之前通过传统的方法对于光线失真的纠正是必要的)。这通过在一个以前的校准步骤中单击在一个被捕获图像中的在白色书写板的四个角的位置上来被手工完成,尽管这可能被自动完成(例如,通过使用边缘检测)。用四个角,为在使用双三次插值的序列中的每个图像执行一个简单传统的双线性扭曲以获得在每个被捕获图像中的一个被修剪和被校正的白色书写板图像。

[0100] 2.3.2 计算白色书写板颜色

[0101] 为了单元的分类,有必要知道对于每个单元的白色书写板的颜色是什么(就是说,没有任何东西写在它上的白色书写板本身的颜色)。白色书写板颜色也被使用为了在产生关键帧中白色均衡,所以它可能被准确的估计以确保关键帧图像的高质量。

[0102] 为了计算白色书写板颜色二个策略已经被使用。第一个策略,在图 6A 中概括论述,是基于白色书写板单元具有随着时间流逝的最亮亮度和具有最小变化的(例如,在每个单元中几乎一致)假设。这是合理的因为笔划的颜色(红色,绿色,蓝色或黑色)将降低亮度。如在处理动作 602 中所示的,具有最亮亮度和最小变化的白色书写板单元被计算。无论如何,这可能产生在最终白色书写板颜色图像中的空缺。例如,如果一个单元或者包含一个笔划或者在整个序列中被一个前景物体遮蔽,为这个单元计算的白色书写板颜色将是不正确的(这个单元看来与白色书写板的其它不同,从而看上去像一个空缺)。到此为止,如在处理动作 604 中所示的,通过使用一个称为最小中值平方的技术(相似于如在下一个段落中描述的分离物检测方法)在白色书写板颜色图像中的任何空缺被发现。接着空缺被填满。(处理动作 606)。为了填满一个空缺,它的邻近区域被搜索并且白色书写板颜色被设置成不是一个空缺的最接近单元的那个。这种策略通常运作的非常好,但是当一个人身穿白色圆领汗衫和/或手持一张白纸时,它会失败。图 7 的左图显示了从图 4 中的输入序列中计算的白色书写板颜色图像的结果,其中在一些帧中一个人手持一张白纸。可见被计算的白色书写板颜色被白纸所破坏。

[0103] 确定白色书写板颜色的第二个策略在图 6B 中被显示并且更为复杂。假设随着时间流逝在每个单元中的像素的一个重要部分属于白色书写板。通过为每个单元搭建一个亮度直方图,具有一个高亮度值的相应于峰值的颜色很可能是对于这个单元的白色书写板颜色。因此,第一步骤为每个单元建立一个直方图并且计算峰值亮度(处理动作 610 到 614)来在这种方式中计算一个初始的白色书写板颜色。这种技术工作即使在序列中一个单元包含一个笔划,但是当一个人身穿白色圆领汗衫和/或手持一张白纸或当一个单元总是被人或其它物体隐藏的情况下,它会失败。在这种情况下,被计算的白色书写板颜色图像包括分离物。下一个动作是检测任何分离物(处理动作 616)。分离物检测基于一个称为最小中值平方的加强的技术。假设跨越白色书写板的颜色平滑的变化,通过求最小平方误差的中值,一个平面被适合在亮度 Y 或 RGB 空间中。颜色不跟随这个模型的单元被认为是分离物并且因此被抛弃,例如,它们被以空缺来作记号(处理动作 618)。感兴趣的读者可以查阅附录来详细了解本技术。接着,如在处理动作 620 中所示,通过使用如在第一白色书写板颜色计算策略中相同的过程(处理动作 620)来填满空缺。最后,为了更进一步改进结果,白色书写板颜色图像可能在 RGB 空间中通过局部适合一个平面来被过滤(处理动作 622)。感兴趣的读者可以再一次查阅附录来详细了解。在相同的例子上用这种新技术获得的结果在图 7 的中图中被显示。如在左图所示的在用第一策略获得的结果上清楚的改进是可见的。作为对比,实际空缺的白色书写板在右图中也被显示。

[0104] 2.3.3 随着时间流逝把单元图像聚集成组

[0105] 在会议中,每个单元的内容通常随着时间流逝而变化。对于每个单元,任何人都愿意聚类在时间序列中的所有单元图像到组中,其中每个组包含被认为随着时间流逝相同的单元图像。一个改进的标准交叉相关算法被使用来确定二个单元图像是否相同或不同。在

随后,标准交叉相关技术被描述成使用图像的一个颜色成分,但它应用于所有的 RGB 成分。

[0106] 假定有二个单元图像 I 和 I'。让 I 和 I' 成为它们的中间颜色和 σ 和 σ' 是它们的标准偏差。标准交叉相关分数通过 $c = \frac{1}{N\sigma\sigma'} \sum_i (I_i - \bar{I})(I'_i - \bar{I}')$ 被给定,其中总和是遍历每个像素 i, N 是像素的总数。分数范围从 -1 (对于二个完全不相似图像) 到 1 (对于二个完全一样的图像)。因为分数是在中间颜色被减去之后计算,它仍可能给定一个高值即使二个图像具有非常不相同的中间颜色。所以一个附加的测试被使用在基于 Mahalanobis 距离的

中间色差上,它是通过 $d = \left| \bar{I} - \bar{I}' \right| / (\sigma + \sigma')$ 被给定。概括的说,二个单元图像 I 和 I' 如果和仅仅如果 $d < T_d$ 和 $c > T_c$, 才被认为是相同的因而它们应该被放到相同的组中。在白色书写板捕获系统的一个工作实现中, $T_d = 2$ 和 $T_c = 0.707$ 被成功使用。

[0107] 2.3.4 分类单元

[0108] 单元分类过程的动作确定一个单元图像是否是一个白色书写板,一个笔划,或一个前景物体。下列探索法被使用:1) 一个白色书写板单元在颜色中是统一的并且是灰色的或白色的(例如,RGB 值几乎是相同的);2) 一个笔划单元主要是用一个或二个主要颜色混和的白色或灰色;3) 一个前景物体没有上面的特性。因此分类确定当前单元图像的颜色分配和白色书写板颜色分配是否相同,或不相同但具有强的重叠,或完全不同。Mahalanobis 距离再一次被使用,如下所述。

[0109] 注意到白色书写板颜色已经如先前描述的一样被计算。RGB 的一个颜色成分再一次被作为一个例子使用。让 $\bar{I}_w$ 成为白色书写板颜色和 σ_w 成为标准偏差(它是一个小的值,因为一个白色书写板单元几乎是统一的)。让 $\bar{I}$ 和 σ 成为当前单元图像的中间和标准偏差。如果和仅仅如果 $\left| \bar{I} - \bar{I}_w \right| / (\sigma + \sigma_w) < T_w$ 和 $\sigma / \sigma_w \geq T_o$, 单元图像被分类作为一个白色

书写板单元;如果 $\left| \bar{I} - \bar{I}_w \right| / (\sigma + \sigma_w) < T_w$ 和仅仅如果和 $\sigma / \sigma_w \geq T_o$, 单元图像被分类作为一个笔划单元;否则,它被分类作为一个前景物体单元。在白色书写板捕获系统的一个工作实施例中, $T_w = 2$ 和 $T_o = 2$ 被成功使用。

[0110] 2.3.5 过滤单元分类

[0111] 上面分类过程只使用在一个单个单元中的颜色信息。通过使用在单元组中空间和时间关系能够获得更为准确的结果。

[0112] 2.3.5.1 空间过滤 关于空间上过滤,基本的观察是前景单元不应该空间上看起来是分离的因为一个人通常阻挡了白色书写板的一段连续的区域。在空间上过滤中,在每个单个的白色书写板图像上二个操作被执行,如在图 8 中所示的。首先,被分离的前景单元被识别和重新分类作为笔划(处理动作 802)。第二,立即被连接到一些前景单元的笔划单元被重新分类作为前景单元(处理动作 804)。第二操作的一个主要目的是在前景物体的边缘上处理单元。如果这样一个单元包含笔划,第二操作将不正确的分类这个单元作为一个前景物体。幸运的是,无论如何,下面的时间上过滤改正这个潜在的误差。

[0113] 2.3.5.2 时间过滤 关于时间上过滤,基本的观察是当笔划被删除后事实上不可

能在完全相同的位置中写相同的笔划。换句话说,如果对于任何给定的单元,二个不同帧的单元图像包含相同的笔划,那么在二帧之间的所有单元图像必须具有相同的笔划除非那儿有一个前景物体阻挡单元。这个观察对于分割前景物体是非常有用的。考虑到在前面部分中的例子其中在前景物体的边缘上的一个笔划单元不正确的被分类作为一个前景单元。在时间上过滤步骤上,这个单元将被分类作为一个笔划只要在前景物体阻挡它之前和之后它被曝光到照相机。

[0114] 图 9 显示对于在图 4 中的采样图像的分类结果,其中笔划是在绿色中,前景是在黑色中,和白色书写板是在白色中。

[0115] 2.3.6 关键帧图像提取

[0116] 关键帧图像包括在白色书写板上的所有重要内容并且适合作为记录的一个摘要。用户将期望关键帧图像具有下列特性:1) 它们应该捕获在书写板上的所有重要内容;2) 关键帧的数目应该保持在一个最小值;3) 它们应该只包含笔划和白色书写板,但在前面没有人;4) 为了容易剪贴和打印,它们应该具有统一的白色背景和饱和的笔的颜色。

[0117] 关键帧提取过程使用从先前描述的处理动作中的单元图像分类结果。过程首先决定在序列中的那个帧应该被选择作为关键帧;接着它重建关键帧图像。这在下面被详细描述。

[0118] 2.3.6.1 关键帧选择 在选择关键帧中没有唯一的解决办法——正像没有单一的方法来摘要一个会议。在大多数一般意义上而言,参考图 11,已经被分类为笔划,前景或白色书写板的输入图像单元被使用(处理动作 1102)。会议被首先分为几个章节(主题)(处理动作 1104)。一个书写板内容的一个重要部分的擦除通常表示主题的一个变化所以它被使用作为章节的一个分割物。接着对于那个章节建立表示白色书写板内容的一个关键帧图像(处理动作 1106)。正好在一个重要擦除开始之前的帧被选择作为关键帧来确保内容被保存在这些帧中。在图 12 中显示的详细过程依下列各项运作:

[0119] 1. 序列中对于每个帧的笔划单元的数目被计数(处理动作 1202)。一个笔划单元图像可能横越多帧——对于那些帧的每一个,它都被包括在计数中。图 10 显示在一个实例会议中(图 4)笔划计数对比于帧数的图形表示。在图中的一个上升表示更多的笔划被写在书写板上,在那里在图中的一个下降表示一些笔划被擦除。图表有不少噪声。有二个原因:1) 用户正在书写板上不断的作出小的调整;2) 分类结果包含小的误差。

[0120] 2. 使用对于不同的帧的笔划计数。笔划计数的峰值和谷值被确定(处理动作 1204)。如果在每个下降上产生一个关键帧,十二个关键帧将形成。为了保持关键帧的数目为最小值,数据被过滤只保留重要的擦除事件。过程忽视了在数据中的波动除非在相邻的峰值和谷值之间的差异大于一个规定的阈值(处理动作 1206)。最大笔划计数的百分之二十被成功的使用在系统的一个工作实施例中。

[0121] 3. 接着在数据中的谷值被使用来将会议划分成章节,在一个章节中包含峰值的帧被选择作为表示章节的关键帧(处理动作 1208)。

[0122] 2.3.6.2 图像重建 一旦帧被选择,有必要及时地重建对应于在此时的这些点上白色书写板所相像的图像。然而,人不能简单的使用从输入序列中来的未加工的图像,因为它们可能包含前景物体。通过收集在帧中的单元图像,图像被重建。参考图 13,输入划分成为单元图像的帧和划分成为单元图像的关键帧(处理动作 1302)。依赖于单元分类法,有三

种情况：

[0123] 1. 如果一个关键帧单元图像是白色书写板或笔划，它自己的图像被使用（处理动作 1304, 1306）。

[0124] 2. 如果关键帧前景单元图像在一个笔划的范围中（例如，人遮蔽了在书写板上的笔划。这在分析阶段中通过时间上过滤被确定），这个单元图像被从邻近帧中来的笔划单元图像所代替（处理动作 1308, 1310）。

[0125] 3. 否则，如在处理动作 1312 和 1314 中所示的，一个前景物体必须正覆盖在这个单元中的白色书写板背景，并且如前面讨论的在被计算的白色书写板颜色中被填满。

[0126] 2.3.7 关键帧颜色均衡

[0127] 重建过程从白色书写板图像中移除了人物，但是图像仍看上去像从输入序列中来的未加工的图像：浅灰色的和噪声的。它们能被颜色均衡以产生一个更好的图像。过程包括二个步骤：

[0128] 1. 使背景为统一的白色，并且增加笔划的颜色饱和度。对于每个单元，如前面讨论的被计算的白色书写板颜色， $\bar{I}_w$ 被使用来计算在单元中每个像素的颜色。

$$I_{out} = \min(255, \frac{I_{in}}{\bar{I}_w} \cdot 255) \quad (\text{处理动作 1402}).$$

[0129] 2. 减少图像噪声。根据一个 S 形曲线在关键帧中的每个像素的每个颜色区域的值被重新映射（处理动作 1404）。小于 255/2 的亮度被按比例缩小到 0 而大于 255/2 的亮度被按比例提高到 255。

[0130] 章节的开始和结尾时间和它们的关键帧图像的文件名被存储在连同笔划的时间标记一起的索引中。一个笔划的时间标记是这个笔划出现的第一个帧。这个信息已经在节 2.3.3 中被计算。

[0131] 2.4 浏览器操作和用户界面

[0132] 2.4.1 概述

[0133] 在分析服务器处理图像序列和产生索引和关键帧图像之后，它将到被处理记录的 URL 发送电子邮件到注册会议参与者。用户可以点击 URL 来启用一个浏览软件。浏览软件的目的在于允许用户浏览关键帧图像和快速访问与一个特殊主题相关联的音频。

[0134] 浏览软件的用户界面 (UI) 在图 15 中被显示。UI 的主要区域包括一个关键帧平面 1504，其中关键帧略图 1502（关键帧图像的图形表示）被显示，浏览器的主显示平面用来显示从照相机来的未加工的图像 1512 的一个合成和当前的关键帧图像 1502。关键帧平面 1504 也结合了一个背景透明的滑块 1516，它允许用户在主显示平面 1506 中调节被显示的图像从未加工的输入图像到关键帧图像。当前笔划 1510（已经被写在会议回放时间线中的笔划），比将来笔划 1508 呈现得更暗和更明显，将来笔划还没有被写在主显示平面的会议回放时间线中。参与者将写的将来笔划 1508 被显示在一个类似于幻影的风格中。这个视觉显示技术在稍后将作更为详细的描述。

[0135] 一个 VCR 和标准时间线控制 1514 在浏览器 UI 的较低于主显示平面 1506 的左下角中被提供。VCR 和标准时间线控制 1514 允许用户在图像 / 音频序列中顺序地向后或向前、慢或快、或者停止，非常像在一个典型的 VCR 上发现的控制。时间线条 1518 按图形方式把音频 / 图像序列的长度显示为一个条，并且提供会议回放的开始时间，结束时间和当前

时间的用数字表示的值。在这个条 1518 上的一个指针 1520 能被选中,并且被向前和向后拖动以线性地在图像 / 音频序列中顺序向前和向后。

[0136] 值得注意的是即使上述的一些 UI 组件的位置被给定,这不是指被限定。这些 UI 组件可能或者单独的或者与其它组件结合地在显示的任何位置中被呈现。

[0137] 2.4.2 对会议数据的非线性访问

[0138] 对记录音频的二个级别的非线性访问在视觉索引的上下文中被提供。

[0139] 非线性访问的第一级别是通过关键帧略图 1502 的使用。每个关键帧略图在显示上具有与它关联的一个时间范围。用户能单击一个关键帧略图从而跳到对于相应关键帧的音频的起始点(例如,章节的开始)。

[0140] 对记录音频访问的第二级别是通过在每个关键帧中笔划的使用。当在主窗口 1506 中指针悬在一个笔划单元(当前笔划 1510 或将来笔划 1508)上时,指针被变为一个“手”以表示它是可选的(例如,用一个鼠标“可单击的”)。用一个鼠标或其它输入设备在单元上双击使应用程序成为音频回放模式。回放开始于当被单击的笔划单元被写入时的会议时间。当相同图案的单元图像在序列中出现时,被单击笔划被写的时间是最早的时间。主窗口在那个时间开始显示图像。用户仍能单击其它的笔划单元从而跳到会议的另一部分。

[0141] 结合 VCR 和标准时间线控制 1514,这二个级别的视觉索引允许用户在一个非常有效的方式中浏览一个会议。

[0142] 2.4.3 图像观看

[0143] 如在图 15 中所示,关键帧图像的略图(例如,1502)被列在关键帧平面 1504 中。用一个鼠标指针或其它输入设备选择的略图 1502 的一个使相应的关键帧图像到在左边的主窗口 1506 并且使应用程序成为图像观看模式,其中用户能使用缩放控制按钮 1522 来放大和缩小,在图像中读取文本和图表,或切割和粘贴图像的一部分到其它文档。另外,全部的关键帧能被切割和粘贴到其它文档或被打印作为笔记。

[0144] 2.4.4 白色书写板内容可视化

[0145] 给定关键帧图像和时间标记信息,在任何给定时间上相应于白色书写板内容的一个图像可能被重建。如果使用时间线控制 1514 根据音频回放时间每个帧的图像被呈现,主窗口像一个电影一样回放白色书写板内容。使用这种方法,用户具有会议的听觉的和视觉的上下文内容。但他们不能及时单击使它们向前的任何笔划(将来笔划 1508)因为这些笔划还没有在主窗口中被呈现。

[0146] 在白色书写板捕获系统的最初的实现中,将来笔划在一个淡色的缺少颜色或密度的模式中被显示。然而,在一个短的试验时期之后,浏览器的用户经常混淆将来笔划和没有被干净擦除的笔划。关于界面的另一个抱怨之处在于虽然用户喜欢没有人在前的白色书写板图像,但他们有时也想知道谁在书写笔划。

[0147] 在反复设计之后,在图 16 中显示的下列可视化过程,专注于对所有上述的关注点做出决定。这个过程的处理动作如下:

[0148] 1. 使用当前章节和时间标记信息的关键帧图像呈现当前的白色书写板内容(处理动作 1602)。

[0149] 2. 呈现将来笔划,转换结果到灰度色标,并且使用一个 Gaussian 过滤器把它们弄得模糊不清(处理动作 1604)。

[0150] 3. 累加从步骤 1 和步骤 2 中来的图像（处理动作 1606）。

[0151] 4. Alpha 混和从步骤 3 来的图像和从输入序列中来的被校正的图像（处理动作 1608）。被校正图像是从输入序列中来的相应图像（如在图 4 中所示的）但有被修剪的非白色书写板区域，跟随着一个重新映射到一个矩形形状。用户能用一个 GUI 滑块（图 15 的 1516）从 0（只显示被呈现的关键帧白色书写板图像）到 1（正确的显示原始被校正的图像）来控制 alpha 值。被呈现的关键帧白色书写板图像是前景物体被移除的，并且由被它遮蔽的笔划代替的关键帧图像。

[0152] 值得相信的是这是可视化的一个非常有帮助的方法，因为 1) 当前的和将来的笔划在呈现的白色书写板图像上被显示，允许用户向后跳到过去和向前跳到将来，和 2) 混和的被校正的输入图像和关键帧，增加了前景物体因而给出更多的上下文。参看图 15，它是用 $\alpha = 0.8$ 的这样一个可视化的例子。

[0153] 2.5 安全性

[0154] 会议参与者经常忧虑记录一个会议，因为机密的信息可能被非计划内的人观看。为了他们，保持被记录数据的安全性是一个关注点。为了专注于这个关注点，发展了一个简单的基于令牌访问的安全模型。这个过程的处理动作被显示在图 17 中。

[0155] 在白色书写板捕获系统中，会议参与者被要求在会议记录的开始向捕获软件注册（处理动作 1702）。他们能或者在计算机屏幕上的一个对话框中填写他们的电子邮件别名或者为了加速过程，插入他们的公司识别卡到一个智能卡读卡机来注册。

[0156] 所有记录的会议存在于一个互联网服务器上。如果没有人注册，会议被张贴在一个公开可访问的互联网页面上（处理动作 1706）。如果至少有一个参与者注册，在会议记录和分析之后产生一个访问令牌（处理动作 1708）。令牌是包含一个唯一会议标识符的一个长随机产生的字符串。包含令牌的 URL 被发送电子邮件到注册的参与者（处理动作 1710）。接受者到 URL 启用互联网浏览软件来回顾会议（处理动作 1712）。他们也能转寄 URL 到没有参加会议的人。

[0157] 这种简单的用模糊安全性的模式看上去工作的很好。其它安全性方法仍然能够被使用。

[0158] 除上面讨论的白色书写板捕获系统的安全性特征之外，当记录会议时，一个保密模式也是可以用到的。会议参与者说的或写的一些东西如果不希望被记录，将存在一个特征以擦除先前的 15 秒（尽管另一个指定的时间时期可能被使用）的图像和音频数据。这种擦除通过按或者一个实在的或者 GUI 按钮来开始。

[0159] 2.6 另一个实施例

[0160] 上面描述的基本白色书写板捕获系统可以与许多其它技术和设备结合在一起呈现另一个实施例。下面讨论的不同实施例可能被单独的或结合在一起使用。

[0161] 在一个这样的实施例中，在关键帧上实行传统的光学字符识别（OCR），来提供可编辑的文本，很容易被用来建立文档或演讲视图。

[0162] 在另一个实施例中，传统的音频识别软件被使用来转换被捕获数据的音频部分到文本。这允许很容易地建立会议记录和其它文档。它也提供一个相对不太昂贵的方法以提供会议信息到听力损伤者。

[0163] 白色书写板捕获系统通过，例如，使用具有一个麦克风的一台笔记本电脑和装配

在三角架上的一个照相机也能被制成是便携式的。这种设置只需要一个附加的初始校准来确定相对于白色书写板的照相机的位置。这个校准能通过手工确定在图像中平面的四个角被手工的,或通过使用例如边缘检测这样的传统方法被自动的执行。

[0164] 白色书写板捕获系统的分析软件通过使用笔跟踪的白色书写板系统也能被使用来确定关键帧从而推断白色书写板内容。因为在这些系统中笔坐标的历史记录被典型的捕获在矢量形式中,在任何给定时刻上的在白色书写板上的内容可能在稍后被重建。使用具有这样一个系统的白色书写板捕获系统分析软件简化了分析过程。白色书写板背景颜色的确定是不必要的,也不必要白色书写板区域的校正,不需要空间或时间上过滤,并且白色书写板单元的分类变得较为简单因为单元图像将或者是笔划或者是白色书写板,因为没有前景物体将干扰写在白色书写板上的文本。单元“图像”现在被从在白色书写板区域上笔的位置推断的内容中得到。这个发明的实施例主要是把单元“图像”聚集成组,如在图 5 中被讨论的处理动作 506,分类每个单元作为一个笔划或白色书写板单元相似于处理动作 508 除了没有前景单元之外,并且使用分类结果来提取关键帧图像(处理动作 512)。结果能用低带宽和小存储要求来传送和存档。另外,OCR 也能在这个实施例中被使用来将被捕获的关键帧转化为文本。

[0165] 另外,在白色书写板捕获系统的一个工作实施例中,系统的帧频被商业可用的静态照相机的帧频所限制。为了达到较高的帧频,一个高分辨率视频摄像机例如一个 HDTV 摄像机能被使用。

[0166] 还在另一个实施例中,白色书写板捕获系统结合手势识别来使用手势命令。例如,在白色书写板上的某处一个命令方框能被写入。当用户移动或指向方框时,系统使用手势识别来时间标记在做手势的特殊时间上的图像。

[0167] 在基本应用中,分析过程假定白色书写板背景的颜色在一个输入序列中保持恒定。然而,一个已知的颜色补丁能被安装在白色书写板的顶部上,在那里没人能从摄像机来遮蔽它。接着基于这个容易检测补丁的已知颜色特性,软件能在每个帧的基础上为不同的照明条件调节照相机曝光参数。这依照下列来完成。如果曝光参数保持恒定,在房间中用不同的照明条件补丁的颜色在被捕获图像中可能是不同的。已知在先前帧中补丁颜色是什么,照相机可能调节它的曝光参数。补丁的颜色可能保持在一个指定的范围内和白色书写板区域也将是这样的。

[0168] 3.0 系统性能和使用

[0169] 3.1 背景

[0170] 白色书写板捕获系统的设计目的在于它将 1) 与任何现存的白色书写板一起工作;2) 自动的和可靠的捕获白色书写板内容;和 3) 使用白色书写板内容作为一个视觉索引来有效的使用系统浏览一个记录的会议。

[0171] 被比作使用一个传感机械装置或一个电子白色书写板的白色书写板捕获系统,白色书写板捕获系统也有一系列的独特的技术挑战。首先,白色书写板背景颜色不是典型的被预先校准的(例如,通过拍摄一个空白的白色书写板的一个照片)因为每个房间有几个会议进行中会改变的照亮设置。其次,经常地,人们在数字照相机和白色书写板之间移动,这些前景物体遮蔽了白色书写板的一些部分并且在它上面投射了阴影。在一个序列中,可能没有一个帧是完全不遮蔽的。这些问题不得不被处理以计算时间标记和提取关键帧。

[0172] 3.2 系统部件

[0173] 在白色书写板捕获系统的设计中,原型系统被建立和反复改进。三个会议室被配备了一个白色书写板捕获系统。关于这三个房间的信息被列在下面的表 1 中。采样图像(80 乘 80 像素,大约在书写板上的 96 点字体)在图 18A 中被显示(图像从左到右分别相应于房间 1,房间 2 和房间 3)。

[0174] 表 1 关于三个安装地点的信息

[0175]

	房间 1	房间 2	房间 3
书写板面积(英尺)	4×3	8×5	12×5
关键帧图像面积(像素)	1200×900	2400×1500	2400×1000
分辨率(dpi)	25	25	16.7

[0176] 在那些房间中的白色书写板尺寸不同,并且所产生的关键帧图像的质量也一样不同。如从采样图像中能被可见的(图 18A),在一个 12'×5' 书写板上的笔迹(很右边)比在其它二个书写板上的笔迹更为模糊失真因为对于一个四百万像素的输入图像分辨率是最高的了。不过,它们仍相当清晰的。使用一个 12'×5' 白色书写板(图 18B)从一个会议中来的几个被选择的帧和相应的关键帧图像(图 18C 和 18D) 也被显示。

[0177] 因为系统和任何现存的白色书写板一起工作,不需要特殊的笔和擦除器,一个直接捕获装置,一个静态照相机,被选择来捕获白色书写板内容。在白色书写板捕获系统的典型的工作实施例中,一个使用四百万像素的Canon® PowerShot G2 数字静态照相机被使用。这个照相机提供 2272 像素乘 1704 像素的图像——相等于对于一个 6'×4' 书写板的 31.6dpi。选择这个照相机的一个重要的理由是由于其软件开发工具包的实用性,使客户的软件解决方案被写得能从 PC 来控制照相机。软件能在一个每次拍照的基础上可视地指定所有的照相机参数。因为系统直接拍摄白色书写板的照片,笔划的登记不会不准确。只要在擦除之前用户打开系统,内容就能被保存。

[0178] 分析服务器在一个 Pentium III 800Hz 双 CPU 的 PC 上运行。分析过程对于每个小时的会议时间大约花费 20 分钟。使用 MP3 编码对于 16 比特 11KHz 单音频的存储要求大约是 15Mb 每小时。使用运动 JPEG 压缩的输入图像序列需要大约每小时 34Mb。

[0179] 安装在三个会议室的系统频繁的被不同的小组使用。在 6 个月过程内,108 个会议总共 48 小时被记录 - 平均每个会议 27 分钟和每个星期 4.5 个会议。每个会议的关键帧平均数目是 2.7。关键帧图像被存储在 JPEG 格式中。平均图像大小是 51.8Kb。大小范围从 17Kb 到 150Kb。因为在统一的白色背景上的 JPEG 压缩工作非常好,图像大小更涉及用户在书写板上书写的多少而不是图像大小。

[0180] 系统的所有用户相信对于广泛使用一个白色书写板的会议系统是非常有用的。关键帧图像和视觉索引的性能不但允许参与者在随后的时间上回顾一个会议,也允许没有参加会议的用户从实际会议时间的一个片断中理解会议的要点。

[0181] 一些用户发现了新的方法来使用最初没有打算使用的系统。举通常不需要在白色

书写板上书写的情况的会议为例。人们仍然打开白色书写板捕获系统。当轮到某人讲话时,管理者在书写板上写下他 / 她的名字,所以通过单击在关键帧图像中的名字,讲话片断就能够随后在记录的音频中被很容易的发现。另一个例子是在一个集体讨论会议中,当有人想到一个好的想法,他在书写板的一方写一个星号并且大声的讲述它。因而音频通过在星上单击就能被随后重新找到。

[0182] 前面描述的发明已经展现了,以达到图解和描述的目的。并不在意于详尽的或限制发明在被披露的精确的形式中。按照以上的理论许多更改和变化是可能的。意图在于发明的范围没有被这个详细的描述所限制,而是更适合于被所附的权利要求书所限制。

[0183] 附录:基于平面的白色书写板颜色估计

[0184] 只有颜色图像的一个成分被考虑,但以下描述的技术应用在所有成分 (R, G, B, 或 Y)。每个单元 i 被它的图像坐标 (x_i, y_i) 定义。它的颜色被 z_i ($z = R, G, B$, 或 Y) 指定。颜色被计算如在节 2.3.2 中描述的,并且因此是干扰的甚至是不正确的。从用会议室的经验上,白色书写板的颜色经常的变化。它通常在上面的部分中特别明亮而在趋向下面的部分时变暗,或者在一个上面的角上特别明亮而在趋向相对下面的角时变暗。这是因为灯光是倚着天花板安装的。因此,对于一个局部区域 (例如, 7×7 单元),一个平面能准确的适合颜色;对于整个图像,一个平面适合依然是合理的,并且提供一个很有效的指示一个单元颜色是否是一个分离物。

[0185] 一个平面能被表示为 $ax+by+c-z=0$ 。一组只在 z_i 中具有噪音的 3D 点 $\{(x_i, y_i, z_i) | i = 1, \dots, n\}$ 被给定。平面参数 $p = [a, b, c]^T$ 能通过求以下目标函数的最小参数值来被估计: $F = \sum_i f_i^2$, 其中 $f_i = ax_i + by_i + c - z_i$ 。最小平方解由 $p = (A^T A)^{-1} A^T z$ 给定,其中 $A = \begin{bmatrix} x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots \\ x_n & y_n & 1 \end{bmatrix}$ 和 $z = [z_1, \dots, z_n]^T$ 。一旦平面参数被确定,单元 i 的颜色由 $\hat{z}_i = ax_i + by_i + c$ 所代替。

[0186] 最小平方技术对于误差数据 (分离物) 不是很鲁棒性。如先前提及的,被最初计算的白色书写板颜色不包含分离物。为了检测和抛弃分离物,对于整个白色书写板图像的一个适合一个平面的鲁棒性技术被使用。最小中值平方 [11], 一个非常鲁棒性的能够容错几乎一半数据是分离物的技术被使用。这个主意是通过求最小中值来估计参数,而不是平方误差的和,例如, $\min_p \text{median}_i f_i^2$ 。 m 随机二次采样被绘制 (3 是确定一个平面的最小数字)。每个二次采样给出平面的一个估计。数字 m 应该足够大,所以 m 二次采样的至少一个是好的,可能性接近于 1, 就是说 99%。如果假设一半的数据可能是分离物,那么 $m = 35$, 因此,随机采样能非常有效的被完成。对于每个二次采样,平面参数和平方误差的中值 $fi2$ 被计算。给出平方误差的最小中值的平面参数被保留,用 M 表示。接着被称为鲁棒性的标准偏差 $\sigma = 1.4826 \sqrt{M}$ (当没有分离物出现时,系数被使用来达到相同的效率) 被计算。一个点 i 当它的误差 $|f_i| > 2.5 \sigma$ 时,它被认为是一个分离物并且被抛弃。最后,使用先前描述的最小平方技术使一个平面匹配于好的点。一个分离物单元 i 的颜色由 $\hat{z}_i = ax_i + by_i + c$ 所代替。

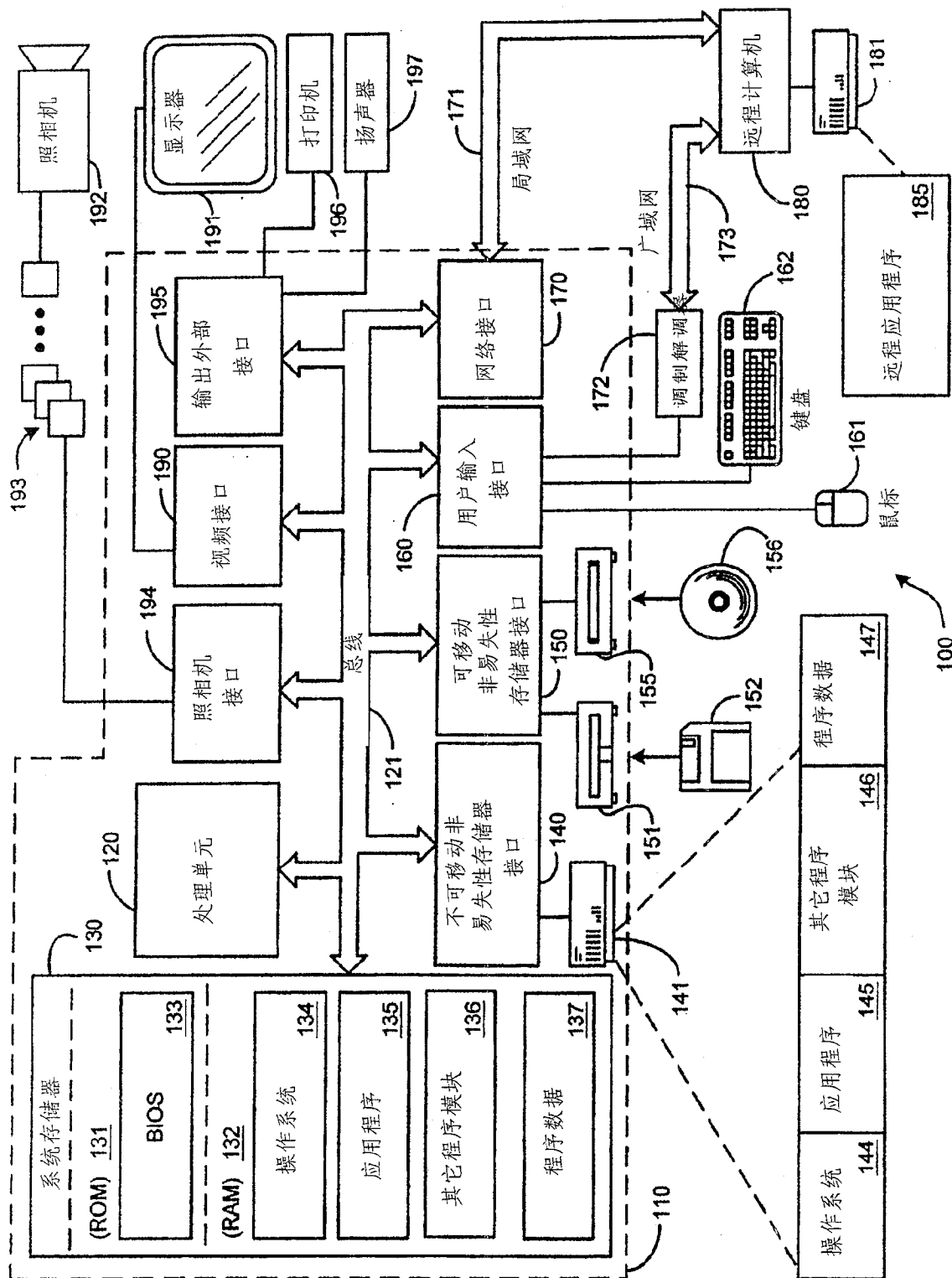


图 1

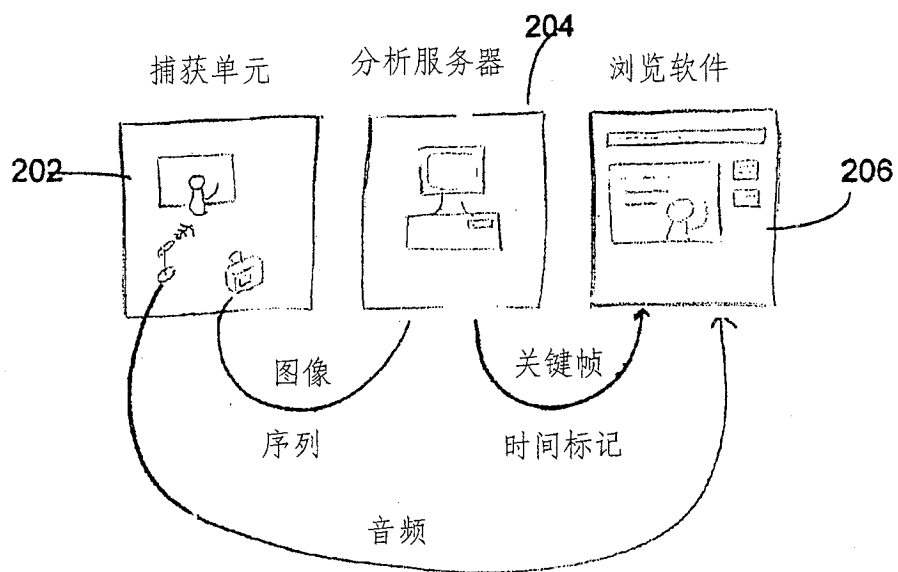


图 2

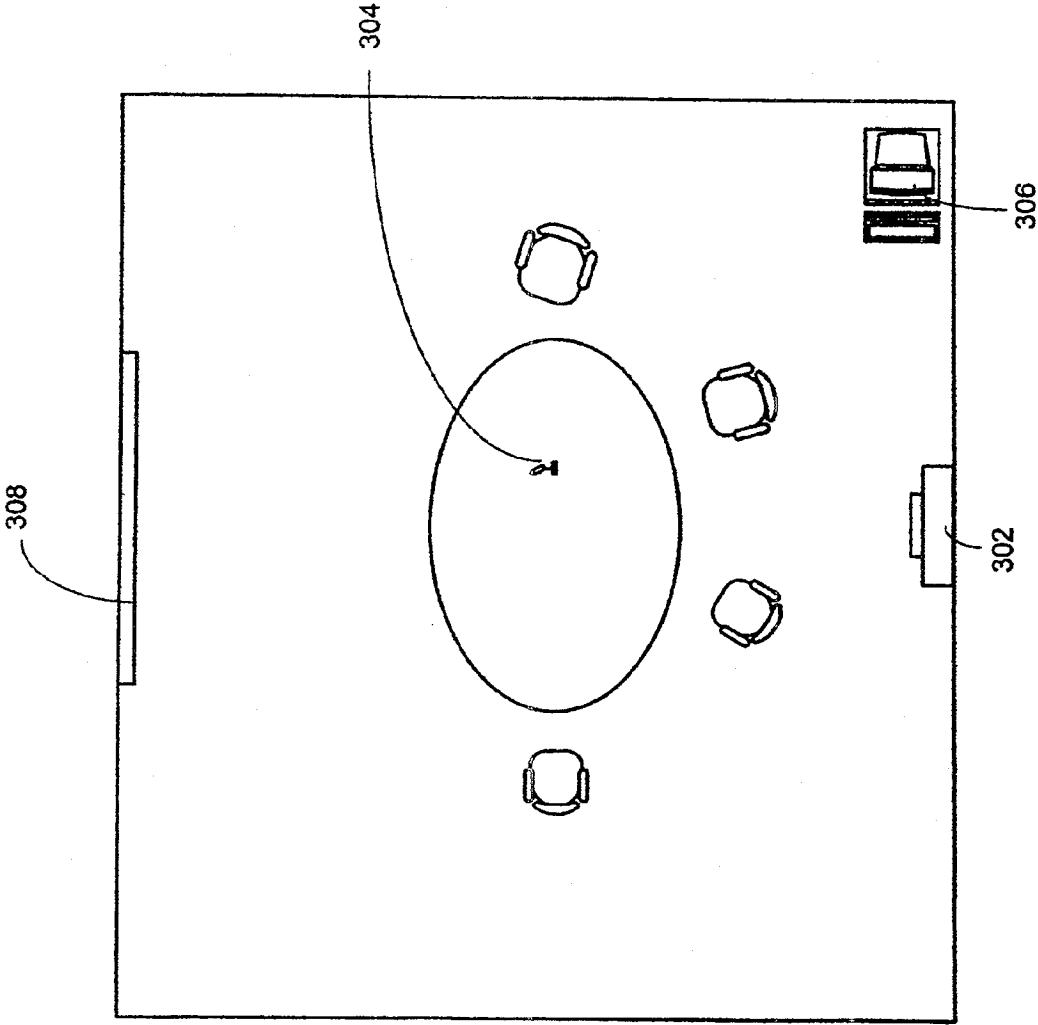


图 3

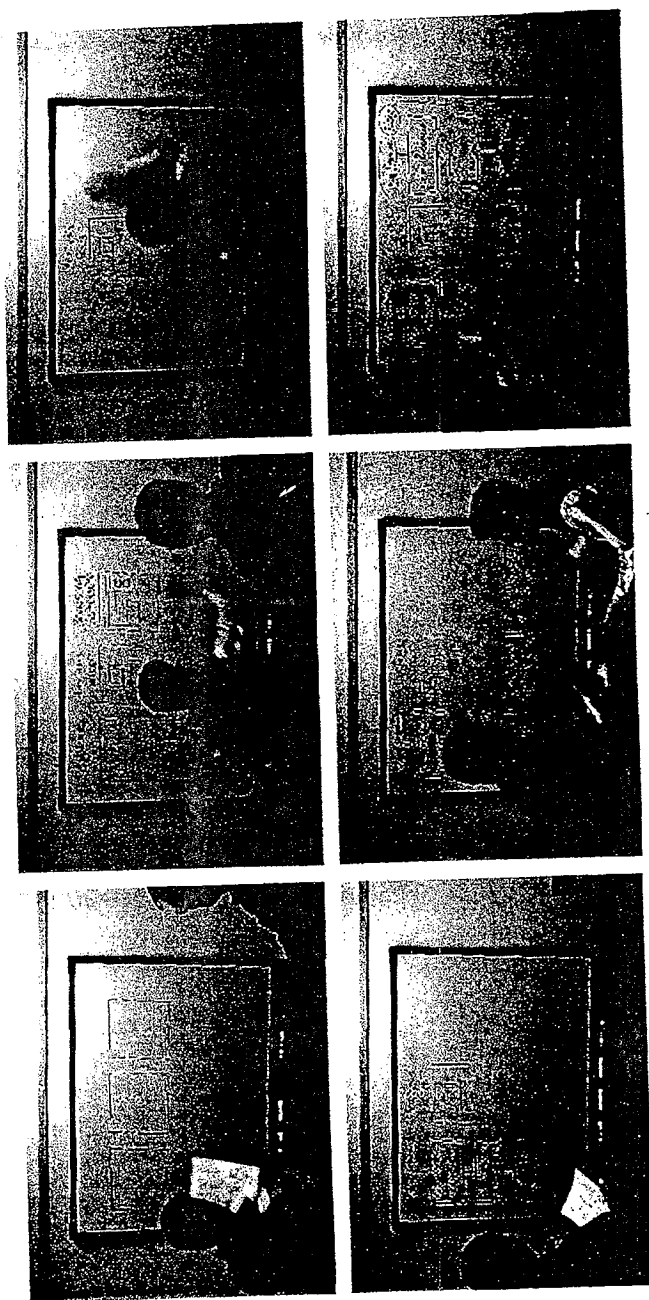


图 4

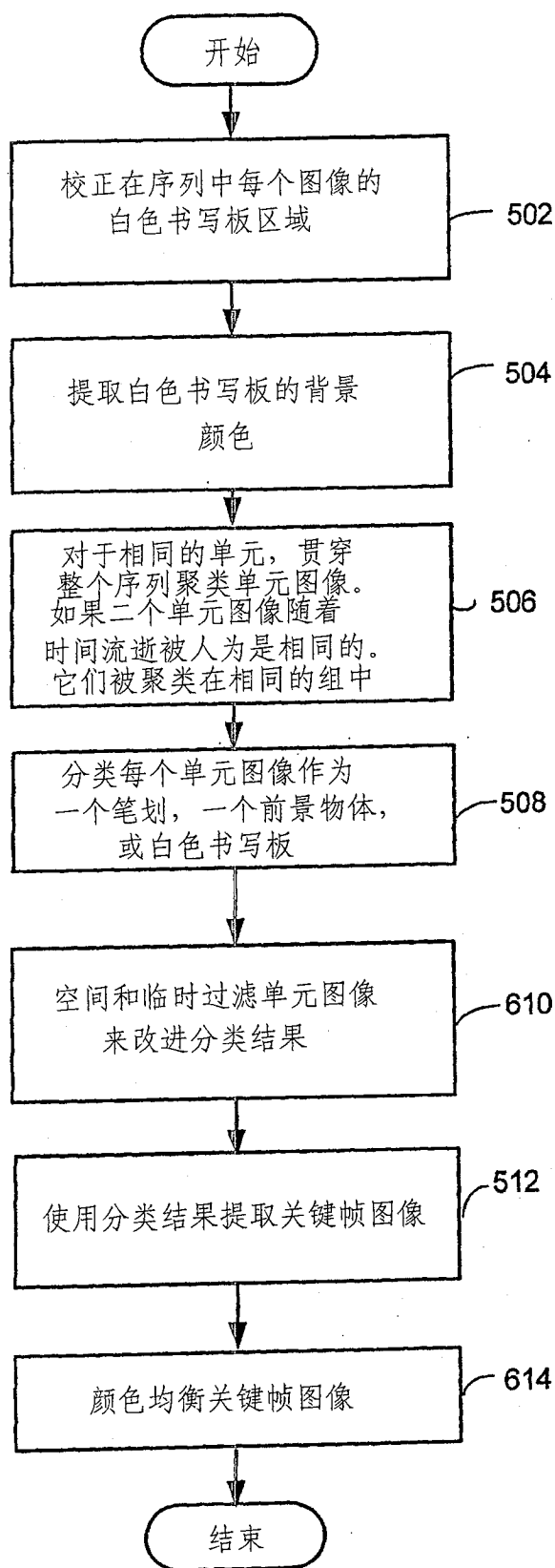


图 5

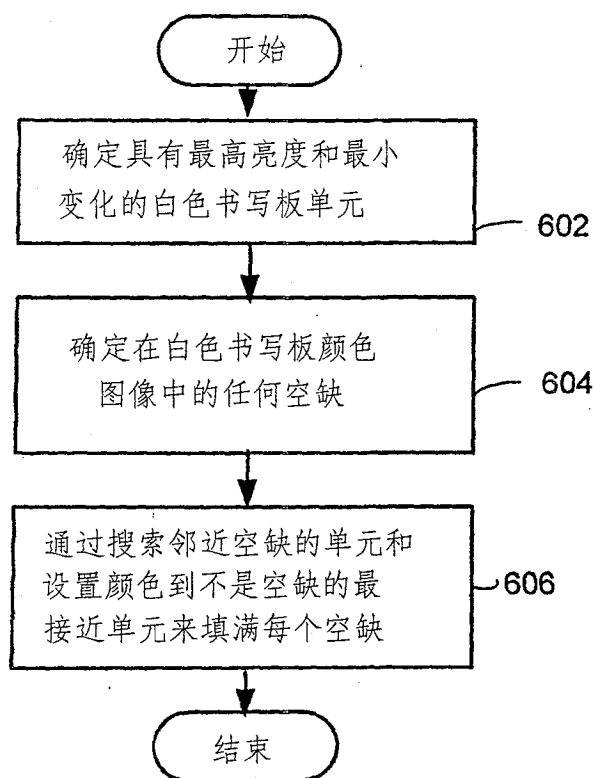


图 6A

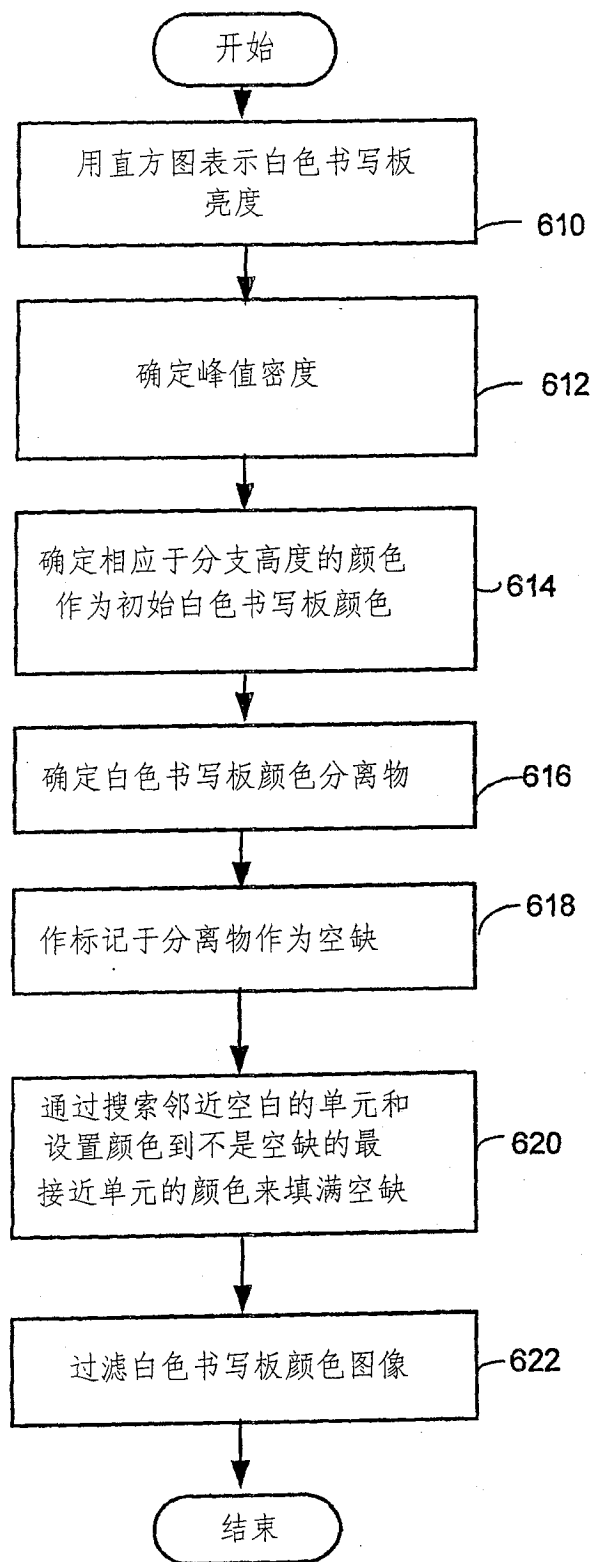


图 6B

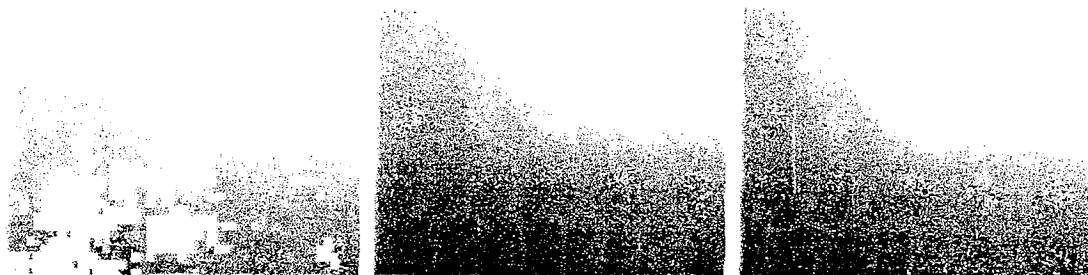


图 7

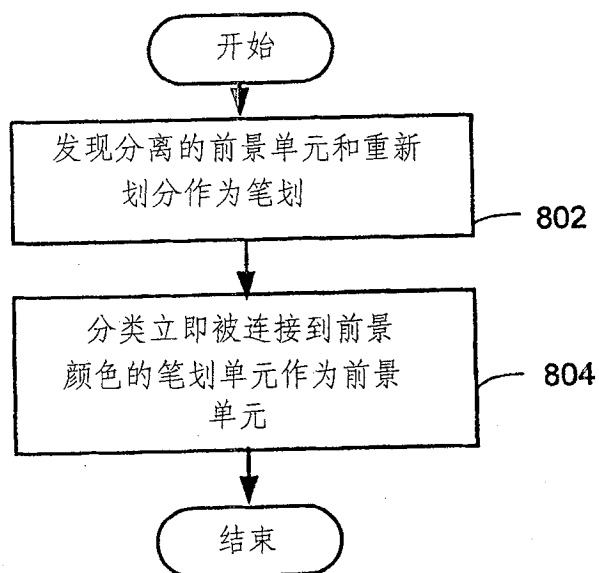


图 8

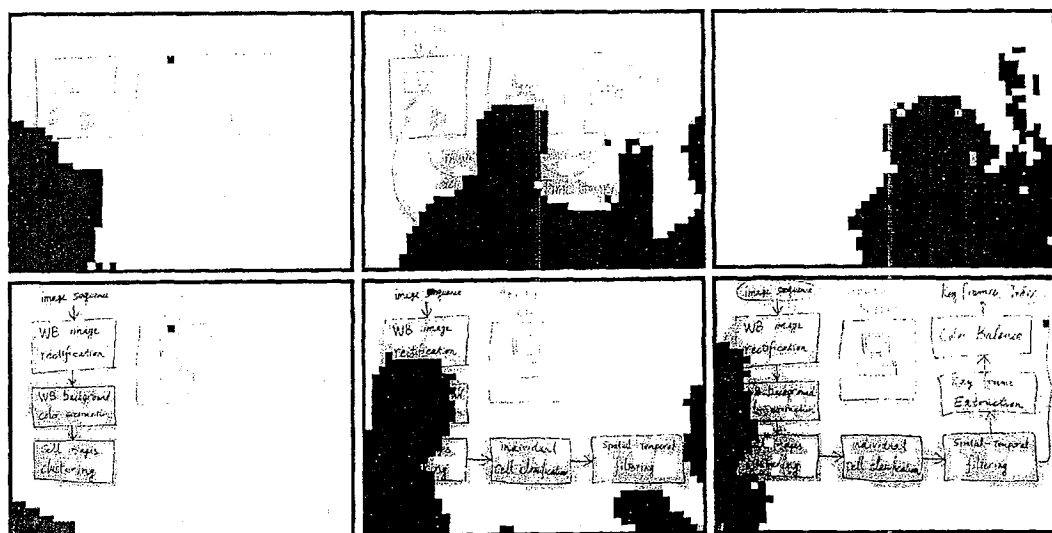


图 9

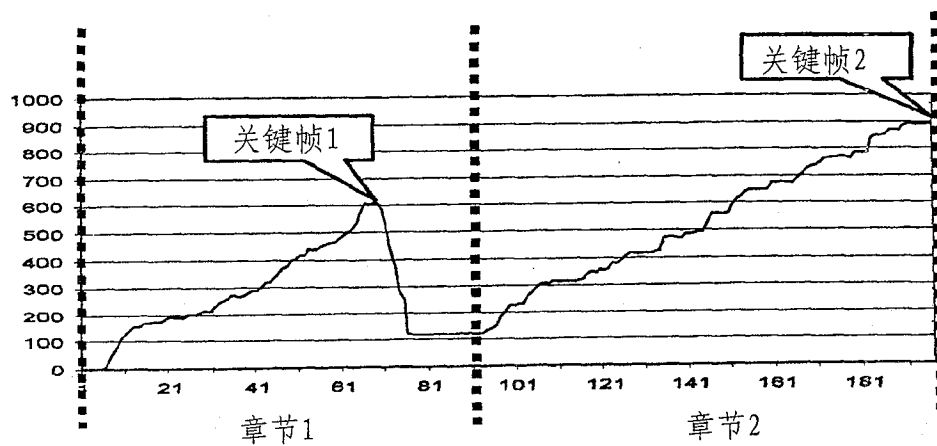


图 10

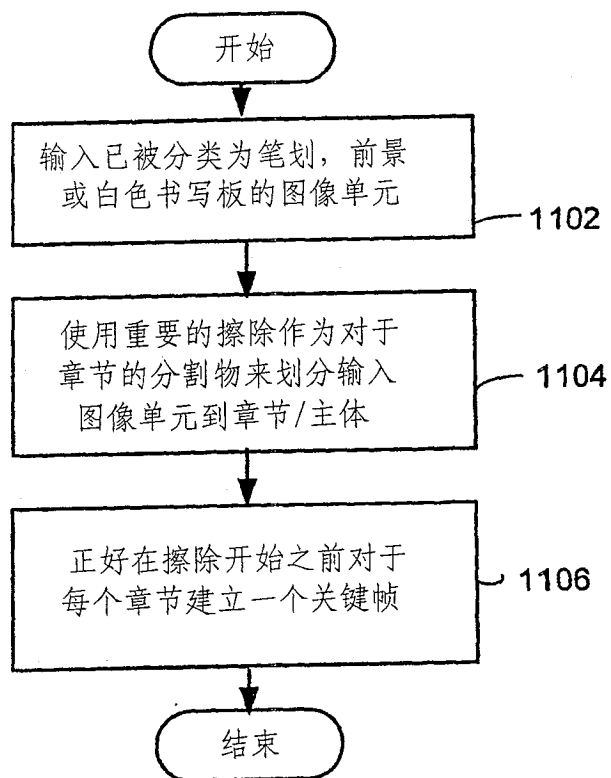


图 11

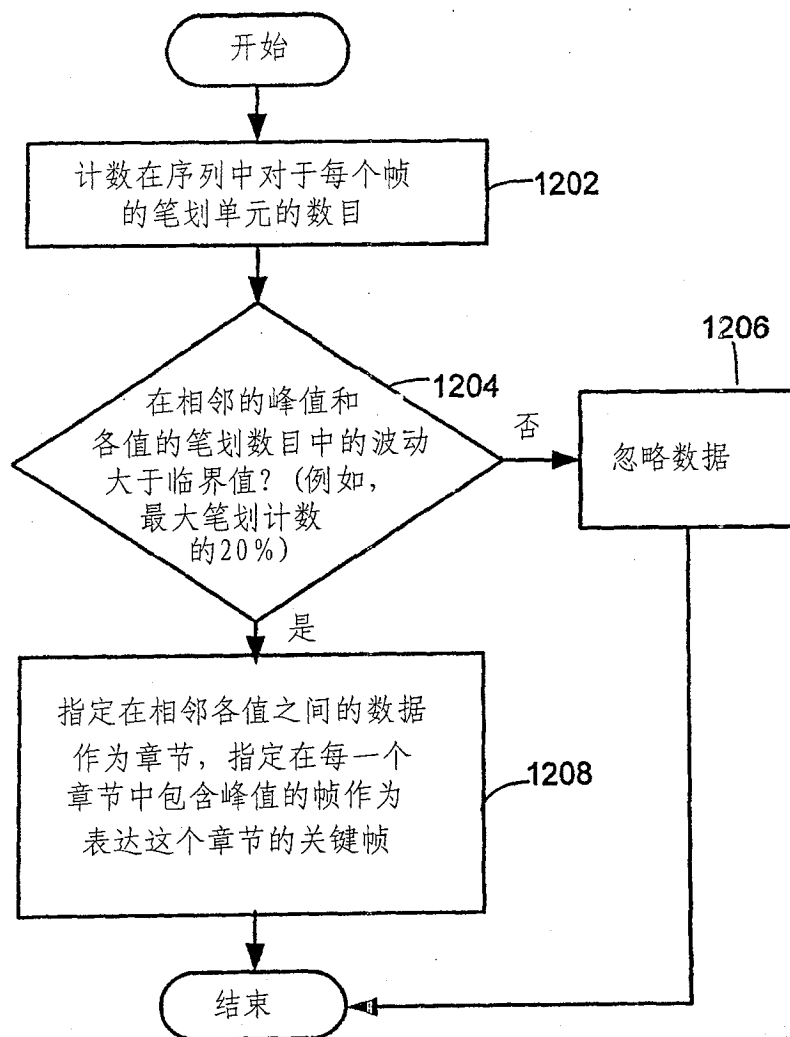


图 12

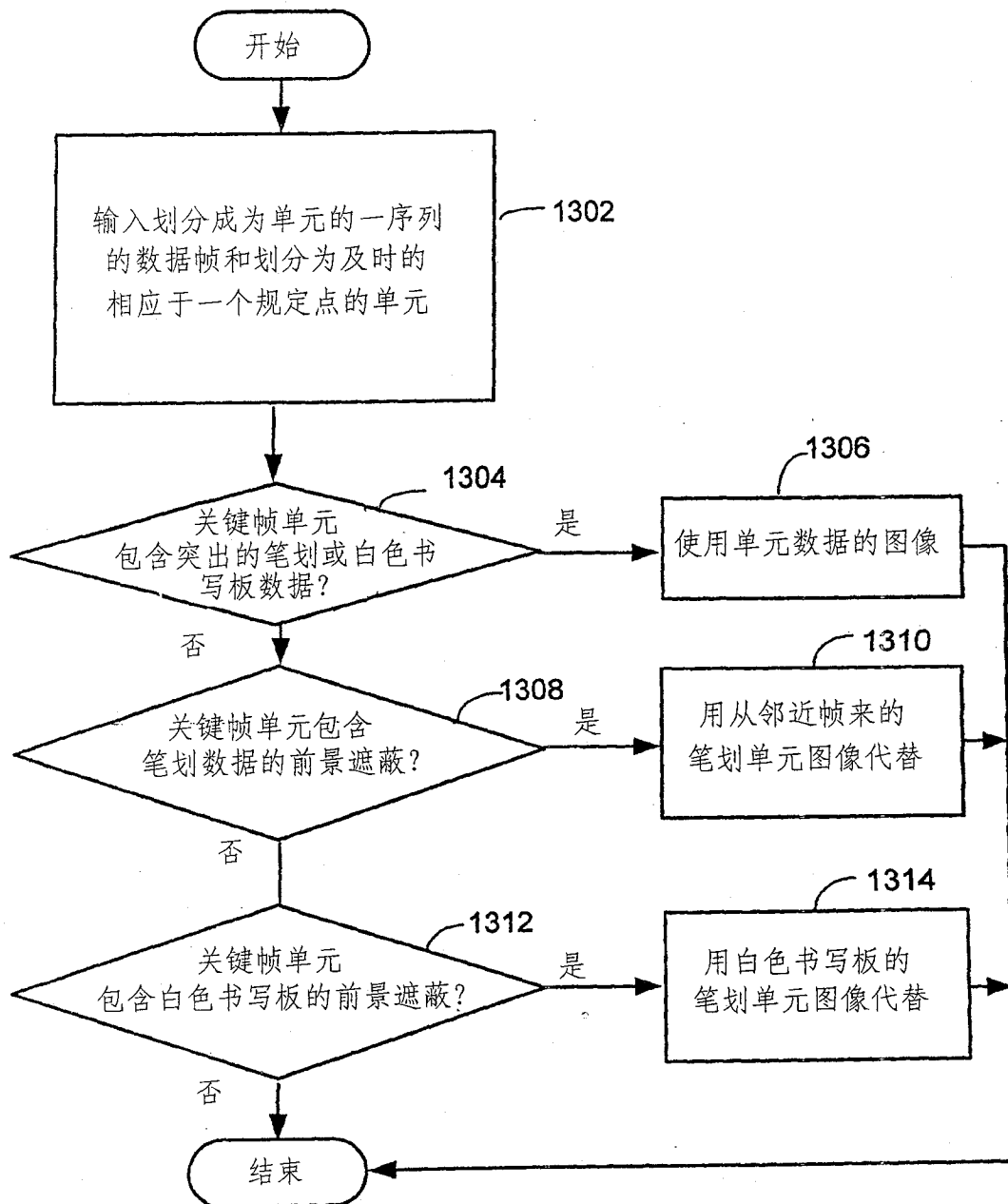


图 13

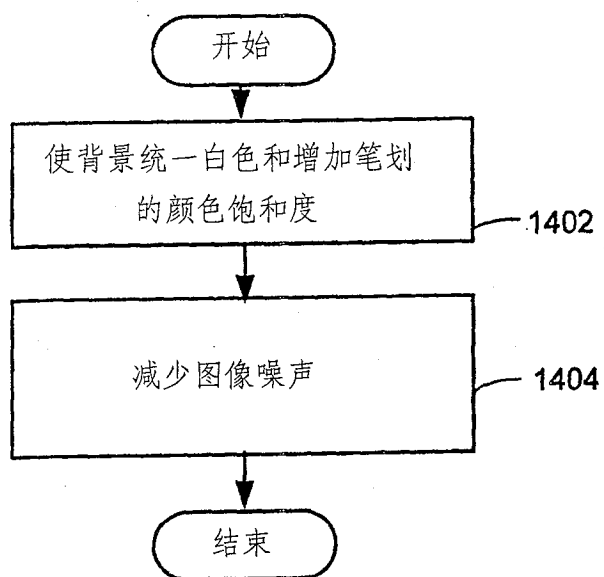


图 14

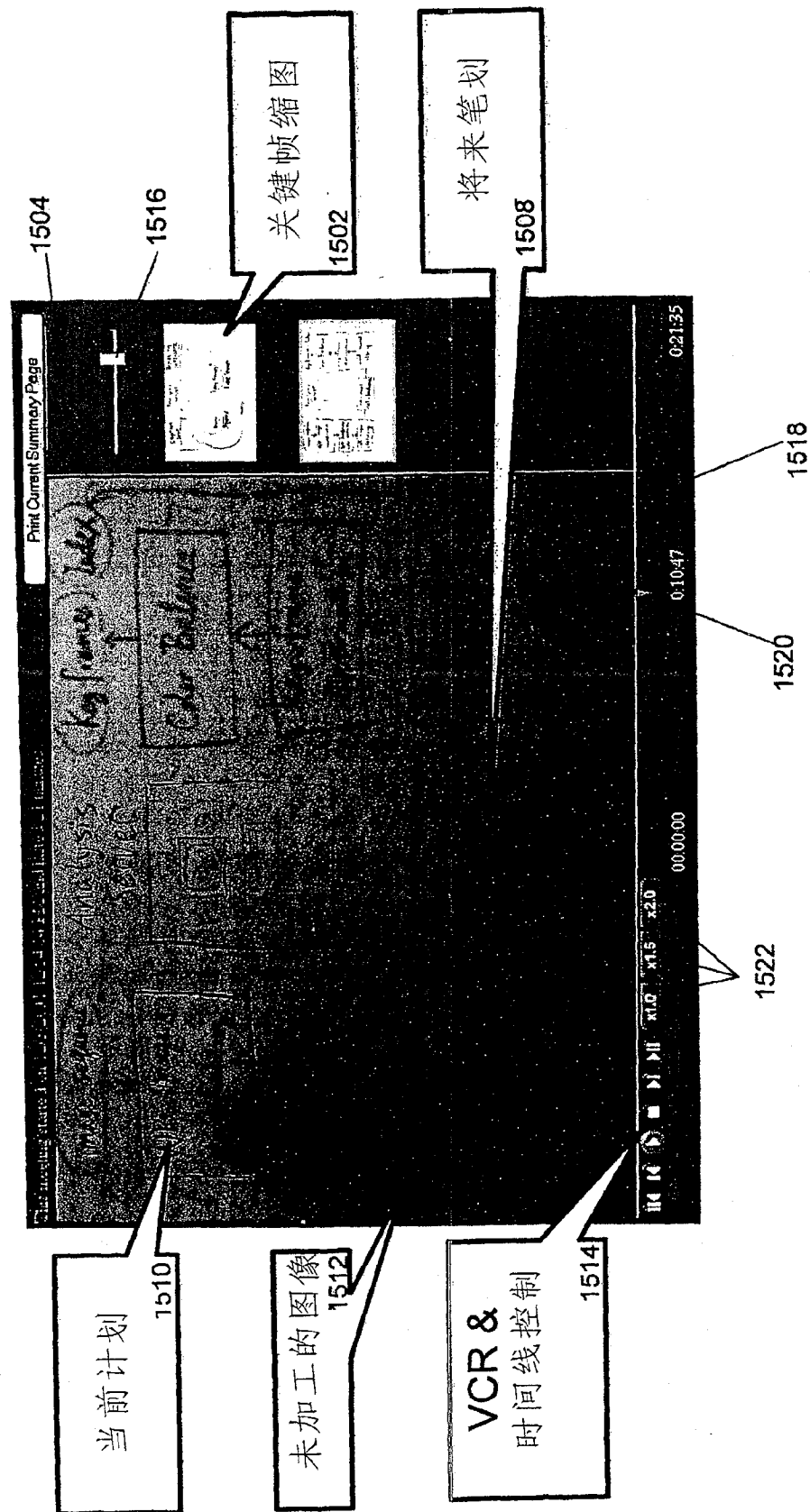


图 15

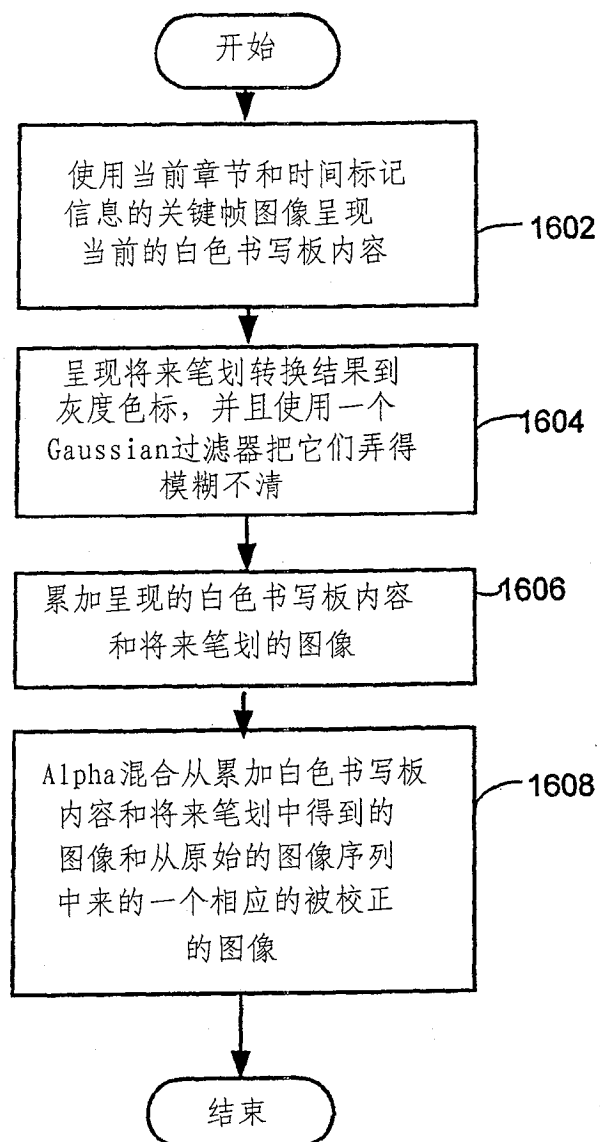


图 16

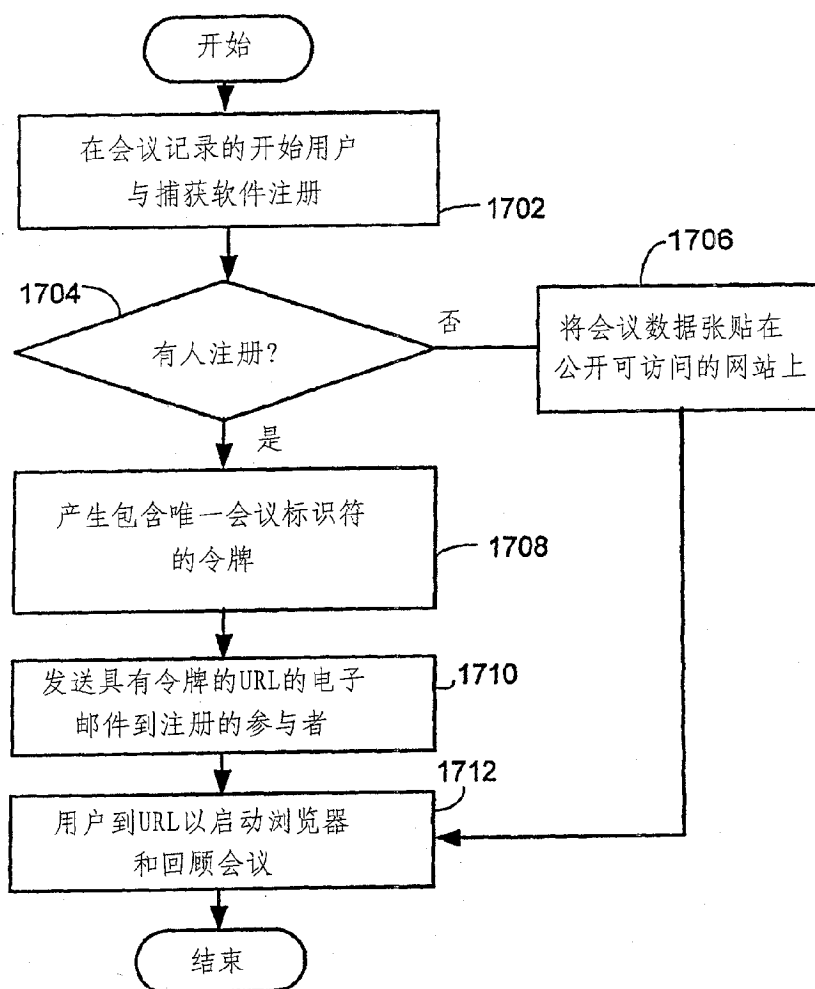


图 17

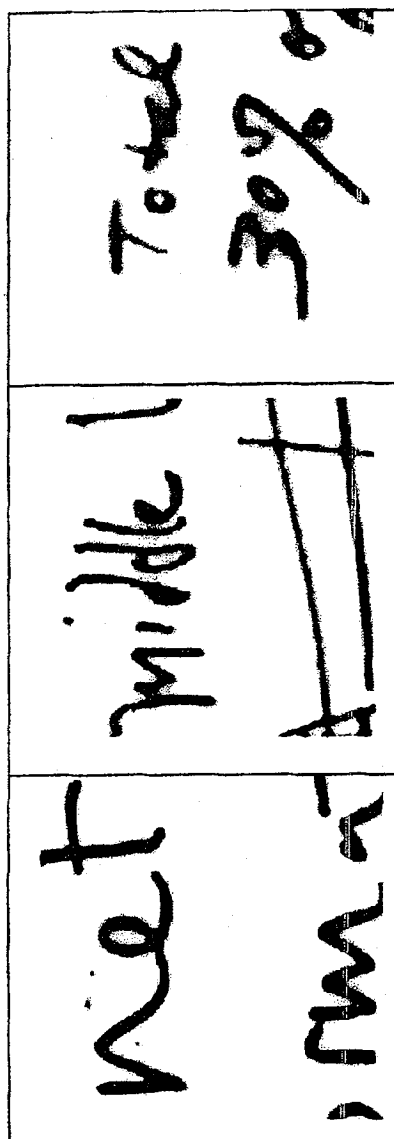


图 18A

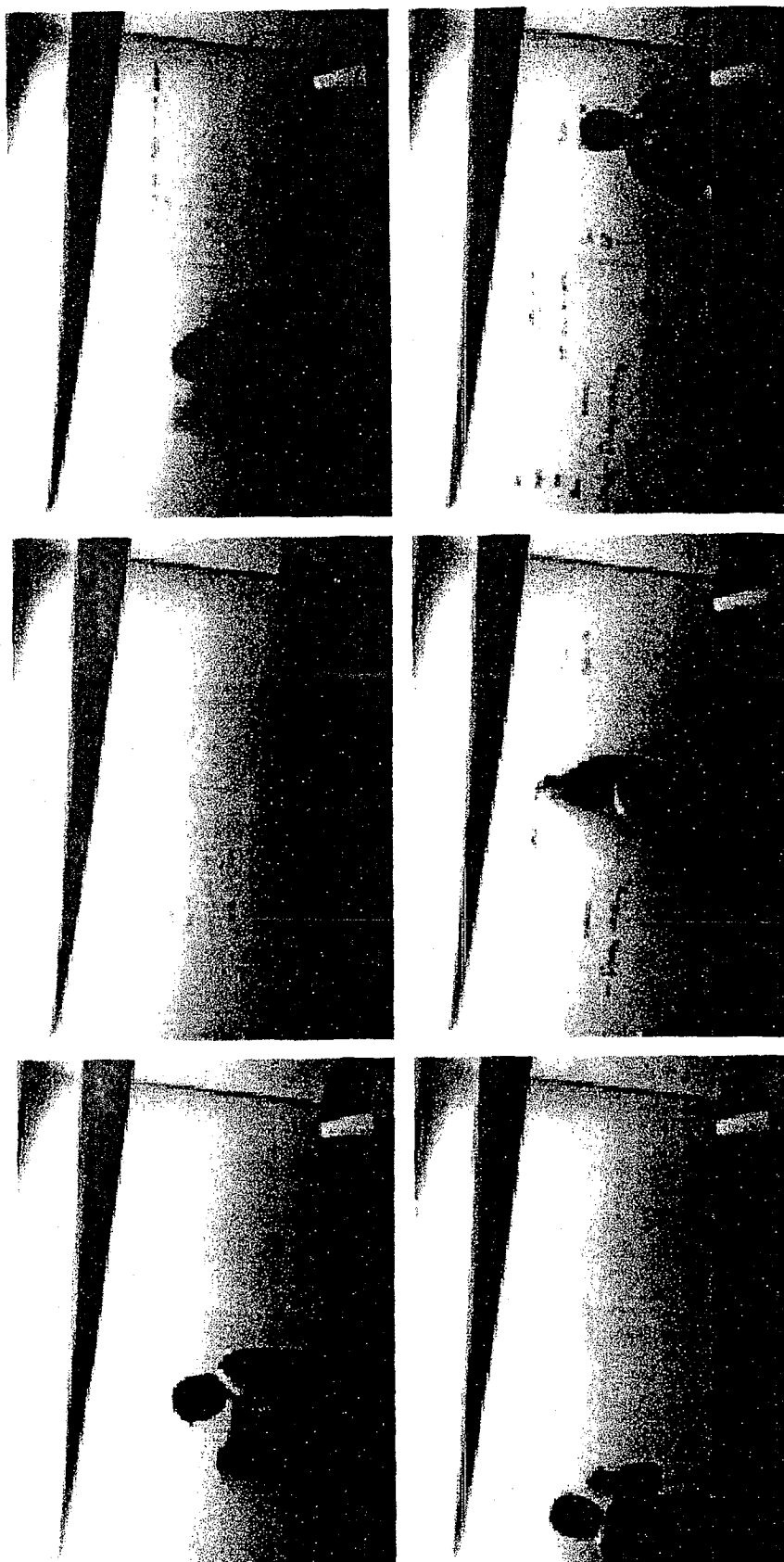


图 18B

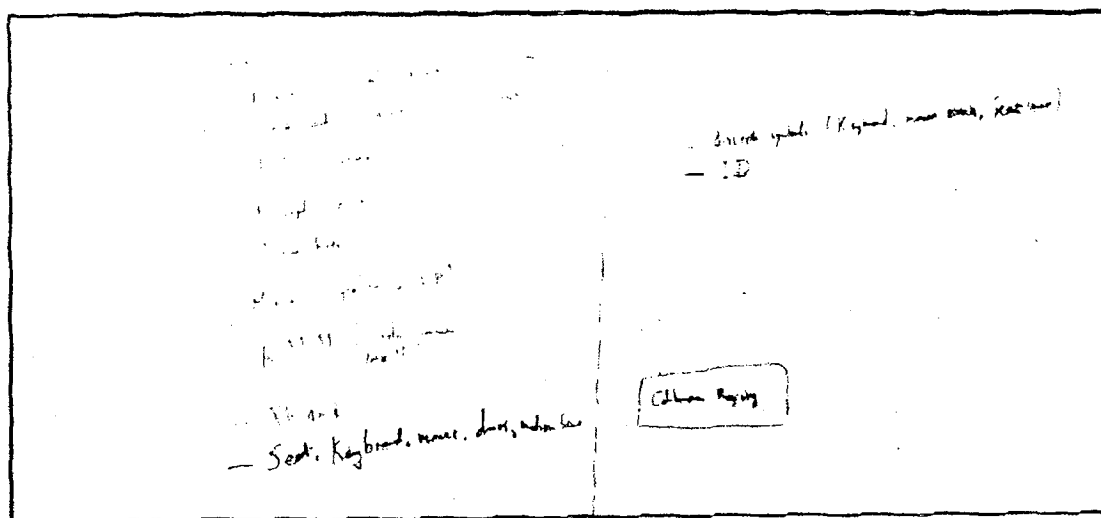


图 18C

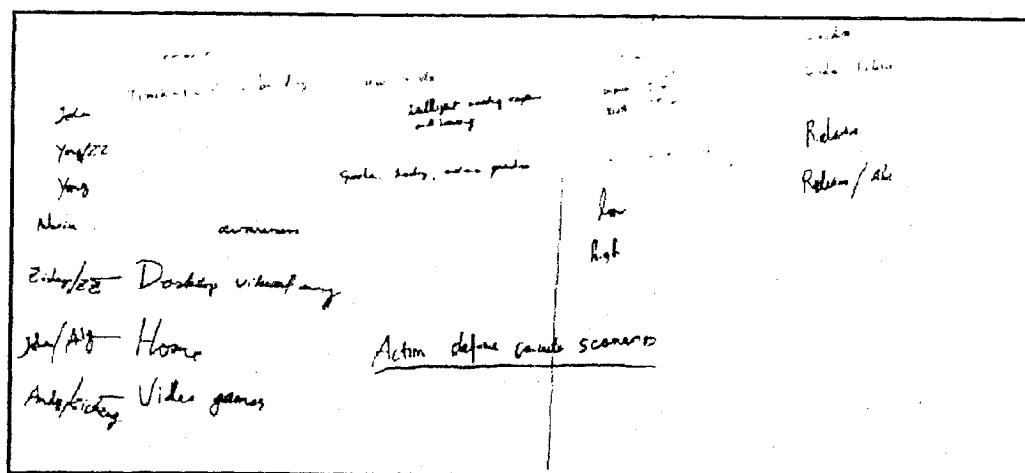


图 18D