



[12] 发明专利申请公开说明书

[21] 申请号 200480002864.1

[43] 公开日 2006年3月1日

[11] 公开号 CN 1742300A

[22] 申请日 2004.1.27

[21] 申请号 200480002864.1

[30] 优先权

[32] 2003. 1. 30 [33] US [31] 10/356,166

[32] 2003. 11. 10 [33] US [31] 10/705,186

[86] 国际申请 PCT/US2004/002287 2004. 1. 27

[87] 国际公布 WO2004/070536 英 2004. 8. 19

[85] 进入国家阶段日期 2005. 7. 26

[71] 申请人 电影教学系统股份有限公司

地址 新加坡新加坡市

[72] 发明人 M·J·G·格莱斯纳

M·S·奈顿 T·C·莫耶尔

P·J·德劳伦蒂斯

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 杨 凯 梁 永

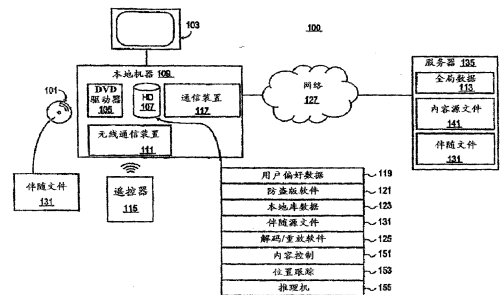
权利要求书 14 页 说明书 22 页 附图 9 页

[54] 发明名称

通过单个媒体上的嵌入内容学习语言的系统

[57] 摘要

结合扩充语言学习内容(500)和娱乐媒体(707)使用预先存在的娱乐媒体、如 DVD 上的故事片或音乐或 CD(101)的学习系统。用于观看或收听扩充内容和娱乐媒体的播放器(250)。播放器可包括诸如父母控制(151)、位置跟踪(153)和推理机(155)之类的特征。



1. 一种方法，包括：
获取包含声音记录的原始数字音频内容；
5 提供包含所述声音记录中存在的单词的文本的附加数字内容；
以及
提供单词的文本与单词在其中发声的原始内容的段之间的链接。
2. 如权利要求 1 所述的方法，其特征在于，所述附加数字内容
10 在所述原始内容的重放过程中向用户显示。
3. 如权利要求 1 所述的方法，其特征在于，所述附加数字内容
还包括有关所述单词的信息。
4. 如权利要求 1 所述的方法，其特征在于，所述附加内容和所
述原始数字音频内容在数据库中被链接。
- 15 5. 如权利要求 1 所述的方法，其特征在于，所述附加内容与所
述原始数字音频内容的重放在时间上同步地向用户显示。
6. 如权利要求 1 所述的方法，其特征在于还包括：
播放与单词的文本关联的所述原始数字音频内容，其中所述单
词的文本的长度和起始点随用户输入而定。
- 20 7. 如权利要求 1 所述的方法，其特征在于还包括：
播放来自所述单词的文本的多个顺序相邻单词，其中重放速度
响应用户输入而调整。
8. 如权利要求 7 所述的方法，其特征在于还包括：
相对于所述重放速度来调整声音重放的音调，从而提高所讲单
25 词的可理解性。
9. 如权利要求 7 所述的方法，其特征在于还包括：
在重放中相对于所述重放速度来调整所讲单词之间的时间间
隔，从而改进对所讲单词的识别。

10. 如权利要求 9 所述的方法，其特征在于：

所述时间间隔之间的各个所讲单词保留其原始的自然音调和语音速率。

11. 如权利要求 1 所述的方法，其特征在于还包括：

5 分析用户输入、所述用户输入的上下文、所述数字音频内容的数据库、所述附加数字内容的数据库以及用户信息的数据库其中至少一项；

指定起始点和结束点、重放的时间顺序、附加数字内容以及重放的修改类型其中至少一项；以及

10 播放符合所述指定的段。

12. 如权利要求 1 所述的方法，其特征在于，所述附加数字内容包括所述音频数字内容中的单词的索引，所述方法还包括：

响应用户输入而调整所述音频数字内容的重放速度；

15 调整所述数字音频内容中单词的音调和时间间隔其中至少一项，以便改进可理解性和识别其中至少一项；以及

通过引用所述索引保持所述音频数字内容中的单词与所述音频数字内容中的特定点的相关性。

13. 如权利要求 1 所述的方法，其特征在于，所述附加数字内容包括所述数字音频内容中可听的单词的索引，所述方法还包括：

20 为所述索引中的多个单词提供可听发音的库；以及

响应用户输入而播放所述发音。

14. 如权利要求 1 所述的方法，其特征在于还包括：

25 分析用户输入、所述用户输入的上下文、所述音频数字内容的数据库、所述附加数字内容的数据库以及用户信息的数据库其中至少一项，以便标识与所述原始内容的段相关的受关注信息；以及

在播放所述段之前呈现所述受关注信息。

15. 如权利要求 1 所述的方法，其特征在于还包括：

分析用户输入、所述用户输入的上下文、所述音频数字内容的

数据库、所述附加数字内容的数据库以及用户信息的数据库其中至少一项，以便标识与所述原始内容的段相关的受关注信息；以及

提示用户进行附加输入，所述附加输入导致对所述重放的进一步修改。

5 16. 如权利要求 1 所述的方法，其特征在于还包括：
提供到通过分布式网络可访问的其它内容的链接。

17. 如权利要求 11 所述的方法，其特征在于，所述修改类型包括播放可听的附加内容。

18. 如权利要求 1 所述的方法，其特征在于还包括：
10 根据对用户授予的权限来控制对内容和功能其中至少一项的访问。

19. 如权利要求 18 所述的方法，其特征在于，根据所接收的付款来授予权限。

20. 一种方法，包括：
15 定义音频和视频数字内容其中至少一项中的段；
 向所述段分配至少一种属性；
 经由相同类型的媒体传递所述段和属性分配信息；
 提供接受关于所述属性的用户指定的界面；以及
 提供访问权以修改符合所述指定的所述媒体的呈现。

20 21. 如权利要求 20 所述的方法，其特征在于还包括：
 根据所述段的属性为多个段编索引。

22. 如权利要求 21 所述的方法，其特征在于还包括创建将所述段和属性相关的数据库。

23. 如权利要求 20 所述的方法，其特征在于还包括将附加内容
25 链接到所述段。

24. 如权利要求 20 所述的方法，其特征在于，所述属性与暴力内容、性内容、裸露及语言内容其中至少一项相关。

25 25. 如权利要求 21 所述的方法，其特征在于还包括：

提供审阅特征以便允许根据所述指定来呈现内容。

26. 如权利要求 20 所述的方法，其特征还在于还包括：

提供包括在所述音频和视频数字内容的声迹中所讲单词的索引的附加内容；

5 响应用户输入而调整所述音频和视频数字内容其中至少一项的重放速度；

 调整单词的音调和时间间隔其中至少一项，以便改进可理解性和识别其中至少一项；以及

10 通过引用所述索引保持所讲单词与所述音频和视频数字内容其中至少一项中的特定点的相关性。

27. 如权利要求 20 所述的方法，其特征还在于还包括：

提供包括在所述音频或视频内容中所讲单词的索引的附加内容；

15 为所述索引中的多个单词提供可听发音的库；以及
 响应用户输入而播放所述发音。

28. 如权利要求 20 所述的方法，其特征还在于还包括：

分析用户输入、所述用户输入的上下文、所述音频和视频数字内容的数据库、附加内容的数据库以及用户信息的数据库其中至少一项，以便标识与所述音频或视频数字内容的段相关的受关注信息；
20 以及

 在播放所述段之前呈现所述受关注信息。

29. 如权利要求 20 所述的方法，其特征还在于还包括：

分析用户输入、所述用户输入的上下文、所述视听数字内容的至少一个的数据库、附加内容的数据库以及用户信息的数据库其中
25 至少一项，以便标识与所述音频和视频数字内容中至少一个的段相关的受关注信息；以及

 提示用户进行附加输入，所述附加输入导致对所述重放的进一步修改。

30. 如权利要求 20 所述的方法，其特征在于还包括：
提供到通过分布式网络可访问的其它内容的链接。
31. 如权利要求 20 所述的方法，其特征在于还包括：
根据对用户授予的权限来控制对内容和功能其中至少一项的访问。
- 5
32. 如权利要求 31 所述的方法，其特征在于，根据所接收的付款来授予权限。
33. 一种方法，包括：
获取原始内容，其中包括最初主要为了非语言学习目的而制作的
10 的视频和音频内容其中至少一项；
经由相同数字媒体传递所述原始内容与附加内容；
其中所述附加内容包括所述原始内容中存在的单词的文本数据库；以及
其中所述附加内容还包括有关所述单词的信息。
- 15
34. 如权利要求 33 所述的方法，其特征在于还包括向用户呈现所述原始内容和所述附加内容其中至少一项，以帮助语言学习。
35. 如权利要求 33 所述的方法，其特征在于，所述数字媒体是 DVD、分布式网络、因特网、电缆传输以及无线电传输其中之一。
36. 如权利要求 33 所述的方法，其特征在于，所述附加内容与
20 所述原始内容的重放在时间上同步地向用户显示。
37. 如权利要求 33 所述的方法，其特征在于还包括：
播放与多个顺序相邻单词关联的所述原始内容，其中单词序列
的长度和起始点随用户输入而定。
38. 如权利要求 33 所述的方法，其特征在于还包括：
25 播放多个顺序相邻单词，其中重放速度响应用户输入而调整。
39. 如权利要求 38 所述的方法，其特征在于还包括：
相对于所述重放速度来调整声音重放的音调，从而提高所述原始内容中存在的单词的可理解性。

40. 如权利要求 38 所述的方法，其特征在于还包括：

在所述重放过程中相对于所述重放速度来调整所述原始内容中存在的单词之间的时间间隔，从而改进对所述原始内容中存在的单词的识别。

5 41. 如权利要求 40 所述的方法，其特征在于，所述时间间隔之间的所述原始内容中存在的各个单词保留其原始的自然音调和语音速率。

42. 如权利要求 33 所述的方法，其特征在于还包括：

10 分析用户输入、所述用户输入的上下文、所述原始内容的数据库、所述附加内容的数据库以及用户信息的数据库其中至少一项；

指定起始点和结束点、重放的时间顺序、附加内容以及所述重放的修改类型其中至少一项；以及

播放符合所述指定的段。

15 43. 如权利要求 33 所述的方法，其特征在于，所述附加内容包括所述原始内容的声迹中讲述的单词的索引，所述方法还包括：

响应用户输入而调整所述原始内容的重放速度；

调整单词的音调和时间间隔其中至少一项，以便改进可理解性和识别其中至少一项；以及

20 通过引用所述索引保持所讲单词与所述原始内容中的特定点的相关性。

44. 如权利要求 33 所述的方法，其特征在于，所述附加内容包括所述原始内容中讲述的单词的索引，所述方法还包括：

为所述索引中的多个单词提供可听发音的库；以及

响应用户输入而播放所述发音。

25 45. 如权利要求 33 所述的方法，其特征在于还包括：

分析用户输入、所述用户输入的上下文、所述原始内容的数据库、所述附加内容的数据库以及用户信息的数据库其中至少一项，以便标识与所述原始内容的段相关的受关注信息；以及

在播放所述段之前呈现所述受关注信息。

46. 如权利要求 33 所述的方法，其特征在于还包括：

分析用户输入、所述用户输入的上下文、所述原始内容的数据库、所述附加内容的数据库以及用户信息的数据库其中至少一项，
5 以便标识与所述原始内容的段相关的受关注信息；以及

提示用户进行附加输入，所述附加输入导致对所述重放的进一步修改。

47. 如权利要求 33 所述的方法，其特征在于还包括：

提供到通过分布式网络可访问的其它内容的链接。

10 48. 如权利要求 42 所述的方法，其特征在于，所述修改类型包括播放可听的附加内容。

49. 如权利要求 33 所述的方法，其特征在于还包括：

根据对用户授予的权限来控制对内容和功能其中至少一项的访问。

15 50. 如权利要求 49 所述的方法，其特征在于，根据所接收的付款来授予权限。

51. 一种方法，包括：

呈现原始内容，其中包括最初主要为了非语言学习目的而制作的视频或音频内容其中至少一项；

20 向用户提供辅助，以帮助语言学习；

观察用户的活动；

推导用户的语言知识程度；以及

自动调整对用户的辅助形式。

52. 如权利要求 51 所述的方法，其特征在于还包括：

25 经由相同数字媒体传递所述原始内容与附加内容；

其中所述附加内容包括所述原始内容中存在的单词的文本数据库；以及

其中所述附加内容还包括有关所述单词的信息。

53. 如权利要求 51 所述的方法，其特征在于还包括：
将来自分开的数字媒体的附加内容与所述原始内容结合；
其中所述附加内容包括所述原始内容中存在的单词的文本数据库；以及
- 5 其中所述附加内容还包括有关所述单词的信息。
54. 如权利要求 51 所述的方法，其特征在于还包括：
播放与多个顺序相邻单词关联的所述原始内容，其中所述单词序列的长度和起始点随用户输入而定。
55. 如权利要求 51 所述的方法，其特征在于还包括：
10 播放多个顺序相邻单词，其中重放速度响应用户输入而调整。
56. 如权利要求 55 所述的方法，其特征在于还包括：
相对于所述重放速度来调整声音重放的音调，从而提高可听单词的可理解性。
57. 如权利要求 55 所述的方法，其特征在于还包括：
15 在所述重放中相对于所述重放速度来调整可听单词之间的时间间隔，从而改进对所述可听单词的识别。
58. 如权利要求 57 所述的方法，其特征在于，所述时间间隔之间的各个可听单词保留其原始的自然音调和语音速率。
59. 如权利要求 51 所述的方法，其特征在于还包括：
20 在重放过程中，基于所述知识程度在某个点自动暂停所述内容并持续一段时间。
60. 如权利要求 59 所述的方法，其特征在于还包括：
在暂停过程中根据所述知识程度自动提供附加内容。
61. 如权利要求 51 所述的方法，其特征在于还包括：
25 提示用户指明他们是否希望更多或更少辅助。
62. 如权利要求 51 所述的方法，其特征在于还包括：
提供包括在所述原始内容的声迹中所讲单词的索引的附加内容；

响应用户输入而调整所述原始内容的重放速度;

调整单词的音调和时间间隔其中至少一项, 以便改进可理解性和识别其中至少一项; 以及

通过引用所述索引保持所讲单词与所述内容中的特定点的相关性。

5

63. 如权利要求 51 所述的方法, 其特征在于还包括:

提供包括在所述原始内容中所讲单词的索引的附加内容;

为所述索引中的多个单词提供可听发音的库; 以及

响应用户输入而播放所述发音。

10

64. 如权利要求 51 所述的方法, 其特征在于还包括:

分析用户输入、所述用户输入的上下文、所述原始内容的数据库、附加内容的数据库以及用户信息的数据库其中至少一项, 以便标识与所述原始内容的段相关的受关注信息;

在播放所述段之前呈现所述受关注信息。

15

65. 如权利要求 51 所述的方法, 其特征在于还包括:

分析用户输入、所述用户输入的上下文、所述原始内容的数据库、附加内容的数据库以及用户信息的数据库其中至少一项, 以便标识与所述原始内容的段相关的受关注信息; 以及

提示用户进行附加输入, 所述附加输入导致对所述重放的进一步修改。

20

66. 如权利要求 51 所述的方法, 其特征在于还包括:

提供到通过分布式网络可访问的其它内容的链接。

67. 如权利要求 51 所述的方法, 其特征在于还包括:

根据对用户授予的权限来控制对内容和功能其中至少一项的访问。

25

68. 如权利要求 67 所述的方法, 其特征在于, 根据所接收的付款来授予权限。

69. 一种方法, 包括:

获取原始内容，其中包括视频和音频可被动播放内容其中至少一项；

经由相同类型的数字媒体传递所述原始内容与附加内容，所述附加内容包括所述原始内容中存在的多个单词的文本数据库；

5 在所述数据库中包含单词与它们在所述原始内容中出现的点之间的链接；以及

提供访问权以根据所述数据库中的单词来修改所述原始内容的重放。

70. 如权利要求 69 所述的方法，其特征在于还包括：

10 播放与多个顺序相邻单词关联的所述原始内容，其中所述单词序列的长度和起始点随用户输入而定。

71. 如权利要求 69 所述的方法，其特征在于还包括：

播放多个顺序相邻单词，其中重放速度响应用户输入而调整。

72. 如权利要求 71 所述的方法，其特征在于还包括：

15 相对于所述重放速度来调整声音重放的音调，从而提高所讲单词的可理解性。

73. 如权利要求 71 所述的方法，其特征在于还包括：

在所述重放中相对于所述重放速度来调整所讲单词之间的时间间隔，从而改进对所讲单词的识别。

20 74. 如权利要求 73 所述的方法，其特征在于：

所述时间间隔之间的各个所讲单词保留其原始的自然音调和语音速率。

75. 如权利要求 69 所述的方法，其特征在于还包括：

25 分析用户输入、所述用户输入的上下文、所述原始内容的数据库、所述附加内容的数据库以及用户信息的数据库其中至少一项；

指定起始点和结束点、重放的时间顺序、附加内容以及所述重放的修改类型其中至少一项；以及

播放符合所述指定的段。

76. 如权利要求 69 所述的方法, 其特征在于, 所述附加内容包括所述视频或音频内容的声迹中讲述的单词的索引, 所述方法还包括:

响应用户输入而调整所述内容的重放速度;

5 调整单词的音调和时间间隔其中至少一项, 以便改进可理解性和识别其中至少一项; 以及

通过引用所述索引保持所讲单词与所述内容中的特定点的相关性。

10 77. 如权利要求 69 所述的方法, 其特征在于, 所述附加内容包括所述原始内容中讲述的单词的索引, 所述方法还包括:

为所述索引中的多个单词提供可听发音的库; 以及

响应用户输入而播放所述发音。

78. 如权利要求 69 所述的方法, 其特征在于还包括:
提供到关于所述原始内容段中存在的单词的信息的链接。

15 79. 如权利要求 69 所述的方法, 其特征在于还包括:

分析用户输入、所述用户输入的上下文、所述原始内容的数据库、所述附加内容的数据库以及用户信息的数据库其中至少一项, 以便标识与所述原始内容的段相关的受关注信息;

在播放所述段之前呈现所述受关注信息。

20 80. 如权利要求 69 所述的方法, 其特征在于还包括:

分析用户输入、所述用户输入的上下文、所述原始内容的数据库、所述附加内容的数据库以及用户信息的数据库其中至少一项, 以便标识与所述原始内容的段相关的受关注信息; 以及

25 提示用户进行附加输入, 所述附加输入导致对所述重放的进一步修改。

81. 如权利要求 69 所述的方法, 其特征在于还包括:
提供到通过分布式网络可访问的其它内容的链接。

82. 如权利要求 75 所述的方法, 其特征在于, 所述修改类型包

括播放可听的附加内容。

83. 如权利要求 69 所述的方法，其特征在于还包括：

根据对用户授予的权限来控制对内容和功能其中至少一项的访问。

5 84. 如权利要求 83 所述的方法，其特征在于，根据所接收的付款来授予权限。

85. 一种方法，包括：

在非易失性存储器中存储可被动播放的视频内容的重放中最近播放的点；

10 允许终止重放会话；以及

在相同内容的后续重放时，返回到所述重放中的相同点。

86. 一种其中存储了指令集的机器可读媒体，所述指令集被执行时，使机器完成包括以下各项的一组操作：

获取包含声音记录的原始数字音频内容；

15 提供包含所述声音记录中存在的单词的文本的附加数字内容；
以及

提供单词的文本与单词在其中发声的原始内容的段之间的链接。

20 87. 一种其中存储了指令集的机器可读媒体，所述指令集被执行时，使机器完成包括以下各项的一组操作：

定义音频和视频数字内容其中至少一项中的段；

向所述段分配至少一种属性；

经由相同类型的媒体传递所述段和属性分配信息；

提供接受关于所述属性的用户指定的界面；以及

25 提供访问权以修改符合所述指定的所述媒体的呈现。

88. 一种其中存储了指令集的机器可读媒体，所述指令集被执行时，使机器完成包括以下各项的一组操作：

获取原始内容，其中包括最初主要为了非语言学习目的而制作

的视频和音频内容其中至少一项;

经由相同数字媒体传递所述原始内容与附加内容;

其中所述附加内容包括所述原始内容中存在的单词的文本数据库; 以及

5 其中所述附加内容还包括有关所述单词的信息。

89. 一种其中存储了指令集的机器可读媒体, 所述指令集被执行时, 使机器完成包括以下各项的一组操作:

呈现原始内容, 其中包括最初主要为了非语言学习目的而制作的视频或音频内容其中至少一项;

10 向用户提供辅助, 以帮助语言学习;

观察用户的活动;

推导用户的语言知识程度; 以及

自动调整对用户的辅助形式。

90. 一种其中存储了指令集的机器可读媒体, 所述指令集被执行时, 使机器完成包括以下各项的一组操作:

15 获取原始内容, 其中包括视频和音频可被动播放内容其中至少一项;

经由相同类型的数字媒体传递所述原始内容与附加内容, 所述附加内容包括所述原始内容中存在的多个单词的文本数据库;

20 在所述数据库中包含单词与它们在所述原始内容中出现的点之间的链接; 以及

提供访问权以根据所述数据库中的单词来修改所述原始内容的重放。

91. 一种其中存储了指令集的机器可读媒体, 所述指令集被执行时, 使机器完成包括以下各项的一组操作:

25 在非易失性存储器中存储可被动播放的视频内容的重放中最近播放的点;

允许终止重放会话; 以及

在相同内容的后续重放时，返回到所述重放中的相同点。

92. 一种其中存储了指令集的机器可读媒体，所述指令集被执行时，使机器完成包括以下各项的一组操作：

5 获取原始内容，其中包括视频和音频可被动播放内容其中至少一项；

经由相同类型的数字媒体传递所述原始内容与附加内容，所述附加内容包括在所述原始内容中存在的多个单词的文本数据库；

在非易失性存储器中存储所述原始内容的重放中最近播放的点；

10 允许终止重放会话；以及

在相同内容的后续重放时返回到所述重放中的已定义点，其中所述已定义点根据对所述内容的分析来确定。

15 93. 如权利要求 92 所述的机器可读媒体，其特征在于，所述已定义点在重放的最后点之前，并且通过确定句子、对话交流、场景、主题或内容的其它逻辑段其中至少一项的开始位置来确定。

94. 如权利要求 92 所述的机器可读媒体，其特征在于，所述已定义点在重放的最后点之前，通过考虑自上一次重放会话以来经过的时间来确定。

通过单个媒体上的嵌入内容学习语言的系统

5 相关申请的交叉引用

本申请是 Michael J.G.Gleissner 等人于 2003 年 1 月 30 日提交的、序号为 10/356166、标题为“基于视频的语言学习系统”的共同未决申请的部分继续申请。

10 发明领域

本发明涉及媒体管理和语言学习工具。具体来说，本发明涉及一组媒体管理工具，它们采用与娱乐内容关联的音频、视频和文本来提供用于访问与音频和/或视频内容关联的文本和信息的增强服务，以及控制对内容的访问。

15

背景

诸如 CD、DVD、盒式录音磁带、录像带及类似媒体之类的音频和/或视频内容提供诸如音乐、电影、电视节目、电台节目和类似内容的内容。大部分媒体的重放限制为呈现媒体上的录制素材。例如，听音乐 CD 的用户可能使用小型盘播放器或类似装置来听录制的音频。用户的选项通常限制为选择曲目、快退、快进和暂停。

20

大部分媒体素材是为娱乐而制作的。这些素材不是设计成有助于学习素材中所使用的语言。这种娱乐素材不是初级和中级学习者易于理解的，因为这些素材节奏过快，并且充满了惯用语、俚语和不规范

25

的句子结构。这些娱乐素材还可能包含不适合部分观众、如儿童的某些素材。父母必须直接监督或限制观看或听这类素材。

附图简介

在附图的各图中，作为实例而不是限制来说明本发明的实施例，附图中，相同的参考标号表示相似的元件。应当注意，本公开中，“一个”实施例的不同说法不一定表示同一个实施例，这类说法表示至少一个。

- 5 图 1 是音频和/或视频重放系统的示意图。
图 2A 是重放界面的图示。
图 2B 是音频播放器的图示。
图 3 是音频和/或视频重放速度调整系统的流程图。
图 4 是音频和/或视频重放扩充系统的流程图。
10 图 5 是伴随源格式的示意图。
图 6 是内容控制系统的流程图。
图 7 是内容控制界面的图示。
图 8 是推理机的流程图。
图 9 是存储器暂停功能的流程图。

15

详细说明

- 在一个实施例中，一组音频和/或视频重放增强特征包括便携式媒体上存储的或者通过网络或广播可访问的原始内容的附加内容。增强特征可包括语言学习、内容控制、使附加内容适应用户需求的推理机以及重放位置保存功能。这些增强特征可与诸如音乐、电影、电视节目、有声书、小应用程序、解说词和类似内容之类的娱乐内容配合使用。娱乐内容可以是可被动播放的。本文所使用的术语“可被动播放媒体”或“可被动播放内容”表示在典型重放过程中不需要用户与内容交互的内容。例如，音乐 CD 可以是可被动播放的，因为它在重放过程中不需要用户交互，除非用户想要跳过某个声迹或停止重放。这些特征可利用附加内容，其中包括伴随文件中所存储的数据。伴随文件可存储在相同媒体、分开的媒体中，或者作为娱乐内容利用相同媒体或不同媒体来分发。
- 20
25

在一个实施例中，增强特征可与交互音频和/或视频语言学习系统配合使用，该系统包括播放器软件应用程序，允许用户播放 CD、DVD 或者包含娱乐素材(例如音乐或故事片)的类似音频和/或视频媒体，其中具有帮助语言学习的扩充特征和附加内容。本文所用的“或者”用于表示具有它的非排他含义，如果“或者”用作排他含义，则使用“或者...或者...”结构。扩充特征和附加内容可包括要学习的语言的转录，诸如词典、语法信息、语音发音信息及类似的语言相关信息的语言学习工具。播放器应用系统采用包含附加内容以及对扩充特征的支持的伴随文件，它可与关联娱乐素材分开存储或者与其结合。伴随文件包含创建可适合语言学习的娱乐素材的扩充特征所需的信息。

图 1 说明使用户能够利用本地机器 109 和显示装置 103 观看或听媒体 101 中存储的音频和/或视频内容的系统 100。本地机器 109 可以是台式计算机或膝上型计算机、因特网设备、控制台系统(例如 Microsoft® Corporation 制造的 Xbox®)、DVD 播放机、专用装置或类似装置。结合了增强特征的音频和/或视频播放器可从连接到本地机器 109(例如在 DVD、CD、硬盘驱动器或类似媒体上)的随机存取或顺序存储装置 105 或者经由远程服务器 135 访问和播放音频和/或视频内容，以及将其中的音频和/或视频内容与伴随文件 131 关联，该文件提供附加内容来扩充音频和/或视频内容。

在一个实施例中，伴随文件 131 可独立于音频和/或视频内容或者与其结合，并且可源自分开的媒体、同一媒体或类似配置。这个系统可用于使用现成的 CD、DVD 和类似媒体来帮助语言学习。在各种实施例中，存储音频、视频和类似内容的随机存取存储媒体可以是 CD、DVD、磁盘、光存储媒体、本地硬盘文件、外围设备、固态存储媒体、网络连接的存储资源或因特网连接的存储资源其中之一。在另一个实施例中，音频和/或视频内容可经由广播、流式传输或类似方法由用户用于重放。伴随文件 131 可驻留在分开的存储媒体中、与

娱乐内容相同的媒体 101 中，或者可采用娱乐媒体来分发，例如通过诸如 FTP、流式媒体、广播媒体之类的网络连接或类似的分发方法进行。音频和/或视频内容、附加内容和伴随文件也可以临时保留在相同或不同的媒体类型中以便于重放。例如，音频内容可以是现成的 CD 101，以及附加内容可位于该 CD 上，或者附加内容可位于分开的 CD 上。来自 CD 101 的音频内容以及附加内容可以存储或缓存在本地机器 109 中，以便于增强特征的重放速度或响应性。在另一个实施例中，内容可包含视频和/或音频、如 DVD 或类似媒体。

在一个实施例中，伴随文件 131 可在媒体制作时或者在媒体销售之前放置在与音频和/或视频内容相同的媒体上。例如，电影制片厂或分销商可制作和销售包含影片以及那部影片的适当伴随文件 131 的 DVD。在一个实施例中，这个伴随文件 131 或附加内容可以“被解锁”并且不对用户通过播放器访问设置阻碍。在另一个实施例中，伴随文件 131 或附加内容可“被锁定”或者在限制情况下可访问。访问伴随文件 131 或附加内容可能要求密码或其它安全机制。访问伴随文件 131 或附加内容可能要求通过网络到服务器或类似关守的连接。在一个实施例中，获取密码以访问全部或部分附加内容可能要求对制片厂或分销商的附加付费。

在一个实施例中，显示装置 103 可以是基于阴极射线管的装置、液晶显示器、等离子体屏幕、数字投影系统或者能够与本地机器 109 交互的类似装置。本地机器 109 可包括可移动媒体读取装置 105，以便访问媒体 101 的音频和/或视频内容。读取装置 105 可以是 CD、DVD、VCD、DiVX 或类似驱动器。在一个实施例中，本地机器 109 包括存储系统 107，用于存储播放器软件、解码/视频软件、伴随源数据文件 131、本地语言库软件 123、防盗版软件 121、用户偏好和跟踪软件 119 以及与播放器软件配合使用的其它资源文件。本地驱动器 107 也可存储数据和应用程序，其中包括内容控制 151、位置跟踪 153 和推理机 155。本地驱动器 107 也可以是存储装置，例如 ROM、RAM

或类似装置。或者媒体 101 或者存储系统 107 可以是 CD、DVD、磁盘、硬盘、外围设备、固态存储媒体、网络连接的存储媒体或因特网连接的装置。在一个实施例中，本地机器 109 包括与遥控器 115 通信的无线通信装置 111。遥控器 115 可产生播放器软件的输入，以便访问语言信息以及调整视频内容的重放。通信装置 117 可将本地机器 109 连接到网络 127 和服务器 135。

在一个实施例中，防盗版软件 121 包括一个系统，其中音频和/或视频内容被唯一标识以确保用户具有那个内容的合法副本。在一个实施例中，伴随文件 131 或其中某个部分被加密，或者在检验用户具有访问此文件的适当许可(例如，音频和/或视频内容的合法副本，向语言学习服务登记及类似标准)之前是不可访问的。在一个实施例中，防盗版软件 121 管理音频和/或视频内容和伴随文件 131 的本地副本，以便确保单个本地副本在授权时被使用，以及当授权丢失或者授权媒体从系统 100 中移去时被删除。在一个实施例中，盗版软件 121 确定音频和/或视频内容的授权副本是否通过在媒体 101 上对它进行访问而成为可用的。在一个实施例中，防盗版软件可强迫使用网络连接，以便允许对附加内容访问以及验证内容的使用。如果媒体 101 不可用，则对本地副本的访问可受到限制或被排除。

在一个实施例中，服务器 135 可为播放器软件提供对全局语言库软件和数据库 113、基于万维网的可下载内容、广播和流式内容以及类似资源的访问。在一个实施例中，播放器软件能够浏览基于万维网的内容，支持聊天室以及服务器 135 所提供的其它资源。

图 2A 是用于播放声迹、MP3 和类似格式的播放器软件的示范说明。类似的播放器界面可用于其它音频和/或视频数据、如影片和类似内容。在一个实施例中，音频和/或视频内容从例如本地驱动器 105 中的 CD 或 DVD 之类的媒体 101 获取，以及伴随文件 131 从例如本地硬盘 107 之类的分开的媒体获取。在另一个实施例中，伴随文件 131 位于媒体 101 中。在又一个实施例中，音频和/或视频内容以及伴

随文件 131 可通过网络经由文件传输协议、流式传输或类似技术来获取。因此，例如，在一个实施例中，原始音频内容、如 MP3 文件可通过因特网来获取，以及附加内容文件(伴随文件)也可通过因特网来获取。音频和/或视频内容可通过网络从与伴随文件 131 相同的源或不同的源来访问。播放器软件在重放过程中将伴随文件 131 与音频和/或视频内容关联，从而扩充音频和/或视频内容的重放。播放器软件界面可包括窗口或观看区域 201，用于显示例如声迹的歌词或单词等附加内容。单词在被讲述时可被突出显示。单词的突出显示被认为包括强调单词文本的一部分或文本周围的观看区域的任何可视机制。这可包括例如改变当前单词或背景的颜色、当单词被讲述时加下划线、当单词被讲述时加上阴影、对所讲述单词加粗或者类似的技术。突出显示可伴随到当前单词的指针 211。在另一个实施例中，使用指针 211 而没有突出显示。下面详细论述从伴随文件 131 导出的其它附加内容、如序言和后文素材。

在一个实施例中，伴随文件 131 通常包括可在重放过程中用来扩充音频和/或视频内容的附加内容。附加内容可没有限制地包括在与所讲述的帧或时间点关联的音频和/或视频内容中所讲述单词的索引、跟踪音频和/或视频内容的剧本的一种或多种语言的文本、用于音频和/或视频内容中的带有或没有发音辅助的任何或全部单词的定义、用于音频和/或视频内容中的带有或没有定义的惯用语、单词和/或惯用语的用法示例、现有说明字幕的翻译及类似内容其中的任一项或全部。所显示文本可包括说明字幕、对话气球以及类似的可视显示。发音辅助可包括例如传统词典中所见的基于文本的发音音标(例如语音拼写习惯的使用)，或者先前录制或由计算机程序生成的“正确”发音单词的音频。

在一个实施例中，如果音频和/或视频内容的文本形式存在，则可直接处理它以准备伴随文件 131。在另一个实施例中，伴随文件的剧本可由自动过程产生。系统可利用光学字符识别实用程序来使用与

视频内容关联的说明字幕获取粗略剧本,或者对于声迹采用语音识别实用程序。翻译实用程序则可用于将剧本翻译成预期语言。然后,人类编者可审阅输出并纠正错误。在另一个实施例中,伴随文件 131 的剧本可由审阅原始内容的编者人工预备。

5 在一个实施例中,人类编者可采用音节检测软件应用程序来审阅内容,并将单词的文本与在其中讲述它们的音频和/或视频内容段中的点相关。本文所用的术语“段”表示两个定义点之间的内容的一部分。在另一个实施例中,系统可通过估计在某一段中所讲述的单词的近似数量并在段的时间长度上分配剧本中的单词,来尝试预备将与音频和/或视频内容对齐的剧本。在一个实施例中,以这种方式预先对齐的文本的单词可经过审阅,从而更精确地将文本的单词与音频和/或视频内容对齐。在一个实施例中,搜索单词含义、惯用语和类似数据
10 的数据库,以便分类和检查所产生的剧本。

 在一个实施例中,播放器软件提供图形用户界面(GUI),从而允许用户更深入地掌握附加内容。例如,用户能够单击字幕中的某个单
15 词,并从伴随文件 131 的词典中获得该单词的定义。示范实施例包括窗口 203,用于显示与音频和/或视频内容相关的附加内容和转录。还可配备导航装置,使得例如单击词典中的某个单词将使用户转到音频和/或视频内容中单词被使用的位置。在一个实施例中,播放器软件
20 可自动识别可用的媒体,并访问或检索诸如艺术家名称、发行商、章节或声迹信息及类似数据之类的相关数据。播放器可允许用户选择要与播放器结合使用的附加内容的方法或位置。

 在一个实施例中,GUI 还可为用户提供重复所观看或收听的内容的任意部分的能力。例如,可提供软按钮,从而引起前一行、先前的
25 歌词、对话交流、场景或音频和/或视频内容的相似段的重复。音频和/或视频内容以及附加内容的随机存取性质允许用户指定关于要观看或收听音频和/或视频内容及关联附加内容的什么部分到任意粒度等级。这样,用户可选择观看或收听场景、对话交流或者只是音频和

/或视频内容中的一行。以任意粒度重复的能力增强了学习体验。GUI 还可为用户提供控制音频和/或视频的速度和/或音调以便于理解口语的能力。速度可以通过在保持所讲述的实际单词的正常音调和速度的同时在单词之间插入间隔来调整。

5 在一个实施例中，播放器支持全屏幕和窗口模式。在全屏幕模式中，播放器根据音频和/或视频内容的尺寸限制、如纵横比以及显示装置的限制来显示音频和/或视频内容。在一个实施例中，GUI 包括一组图标或导航选项 213。在一个实施例中，图标或导航选项 213 允许用户通过使用诸如鼠标、键盘、遥控器或类似装置的外围输入装置
10 来访问附加语言内容。在一个实施例中，重放选项可根据用户的要求被启用或禁用。

 在一个实施例中，图标和导航选项将音频和/或视频内容链接到词典、目录和指南以及类似的语言参考和导航工具。这些链接可使播放器显示专用屏幕，向用户显示相关内容。在一个实施例中，图标或
15 导航选项链接到解释屏幕，它以多种语言列出一段音频和/或视频内容中的惯用语。通过图标和导航选项 213 可访问的专用屏幕还可显示与单词定义、俚语、语法、发音、语源和语音训练有关的信息以及访问菜单、人物信息菜单及类似特征。在另一个实施例中，备选导航技术用来访问特殊内容，例如热键、超链接或类似技术及其组合。在一个
20 实施例中，当访问专用屏幕时，音频和/或视频内容被最小化或者减小尺寸，以便在显示器中创建空间来观看或收听附加内容，同时在适当时候仍然允许观看或收听音频和/或视频重放。当用户完成审阅专用屏幕的素材时，音频和/或视频内容用作返回到全屏幕模式的图标或选项。在另一个实施例中，当显示专用内容时没有显示音频和/
25 或视频内容。

 在一个实施例中，单词和/或惯用语的词典可在通过图标、导航选项或者直接突出显示或选取所显示文本可访问的专用屏幕上显示。词典数据可以是音频和/或视频内容特定的。例如，它可包括特

定的音频和/或视频内容中使用的单词或惯用语的定义而不是该单词或惯用语的全部定义。词典数据可包含不同于音频和/或视频内容的语言的某种语言的定义及相关单词或惯用语。词典数据可包括通用的或者只有特定音频和/或视频内容才有的其它受关注数据。受关注数据可包括单词和/或惯用语到另一种语言的翻译、单词的用法示例、惯用语与单词之间的关联、惯用语的定义、惯用语到另一种语言的翻译、惯用语的用法示例、音频和/或视频内容中讲述单词的人物、在其中讲述单词或惯用语的场景的标识符、与在其中讲述单词或惯用语的场景相关的主题或者类似信息。这种数据可保持在数据库、平面文件或伴随源文件段中,其中具有关联的链接以允许用户从数据库中的内容直接跳转到音频和/或视频内容的相干部分。

播放器可具有与被播放音频和/或视频内容的类型有关的附加特征。在示范实施例中,播放器可标识带有字幕 205 的音频和/或视频作品的标题或部分(例如声迹或场景)。播放器可列出音频和/或视频内容的其它部分 209,用于为每个选择提供标题或标签。播放器还可产生可视表示或相随的图形显示 207 来伴随音频内容。

图 2B 是音频内容的示范便携式播放器的图示。在一个实施例中,便携式播放器装置 250 可在内部存储器或便携式存储装置中存储音频内容和伴随文件。便携式装置 250 可以是系统 100 的缩小形式。在一个实施例中,便携式播放器 250 可具有系统 100 的每个组件。在另一个实施例中,便携式播放器 250 可能具有组件的简化集,其中包括播放选项 253 和显示 257。显示 257 可标识被播放的内容 251 以及与内容关联的文本。便携式播放器可支持当前可听文本的突出显示 255。在一个实施例中,便携式播放器可以是 MP3 播放器、CD 播放器、手持装置、个人日常/数字助理(PDA)、蜂窝电话、平板 PC 或类似的装置。在又一个实施例中,类似的便携式视频内容观看器、如便携式 DVD 播放器还可支持具有特征的全集或简化集的播放机。

图 3 是流程图,说明调整音频和/或视频内容的重放的过程。用

户可利用与本地机器 109 直接或者无线连接的外围设备来调整音频和/或包括与视频内容关联的音频部分的视频内容的重放。外围设备可以是鼠标、键盘、轨迹球、操纵杆、游戏控制器、遥控器 115 或类似装置。播放器软件接收来自外围设备 115 的输入(框 315)。在一个
5 实施例中，播放器软件确定这个输入与音频和/或视频内容的重放相关，其中包括确定预期重放速度和重放的起始点(框 317)。播放器软件将音频和/或视频内容排队到预期起始位置，并开始重放音频和/或视频内容。播放器软件根据来自外围设备的输入来调整音频和/或视频内容的重放速率。

10 在一个实施例中，播放器软件还调整音频和/或视频内容的音频部分中所讲述的单词的音调(框 319)。在一个实施例中，播放器软件以调整后的速度来调整被重放的单词的定时和间隔，以便增强与每个单词关联的声音的离散集，从而帮助用户理解单词(框 321)。时间间隔被调整而没有影响说话者的语音的音调。在一个实施例中，播放器
15 软件以调整后的速度在内容与伴随源数据文件之间将数据相关，其中包括以调整后的速度显示字幕、以调整后的速度突出显示字幕中的单词以及对扩充重放的类似速度相关调整(框 323)。在一个实施例中，用户可根据各个单词、句子、段或者划分视频内容的声迹的类似方式来选择重放的类型。

20 在一个实施例中，外围设备 115 向播放器软件提供输入，它确定要提供的调整重放的类型。在接收到来自外围输入设备 115 的第一输入(例如按钮的单击)时，播放器软件以正常速度重复一段音频和/或视频内容。如果在预定周期内接收到两个输入，则播放器软件可采用时间间隔和音调调整技术以较慢的速率再播放音频和/或视频内容段。
25 如果在预定周期内接收到三个输入，则播放器软件可采用来自清晰发音单词库的音频重放该音频和/或视频内容段。如果在预定时间周期内接收到四个输入信号，则播放器可显示与相干音频和/或视频内容段中的句子相关的下拉屏幕。下拉屏幕可包括语音、语法以及与句子

相关的类似信息，并且可与减速音频或来自库的音频结合显示。在又一个实施例中，包括播放器装置的输入机制的图标、导航选项的使用可用来发起这些调整重放特征。在一个实施例中，在一段音频和/或视频内容的重放过程中，在预定初始时间周期中所接收的输入信号可
5 发起音频和/或视频内容的前一段的重放。

在一个实施例中，播放器软件包括语音训练子程序，以帮助用户纠正发音。语音训练程序提供与调整重放特征结合工作的接口，以便以降低的速度重放音频和/或视频内容的音频部分的各段，以便于用户理解音频部分。在一个实施例中，语音训练程序允许带有音频外围
10 输入装置(例如话筒或类似装置)的用户重复所选音频段。在一个实施例中，语音训练程序向用户提供建议、评定等级或类似的反馈，以便帮助用户校正其语音，从而匹配来自音频部分的语音。在一个实施例中，用户可访问已经预先录制的一组变化的发音，听某个人物对一行的发音或者收听剧本的相干部分的计算机语音朗读。在一个实施例中，
15 显示某个单词或一组单词的正确语音发音。如果用户记录某个发音，则等效于用户所记录的内容的语音将被显示，用于比较和反馈。语音训练程序显示正确发音的图形表示，使得用户可将他的录制发音与正确发音进行比较。例如，这个图形表示可以是在正确发音附近或与其重叠显示的用户的录制音频的波形。在另一个实施例中，图形表示
20 是录制音频的计算机生成的语音转录，允许用户查看他的发音与所录制单词的正确语音拼写进行比较的情况。录制的用户音频和正确的发音也可作为条形图、彩色码映射、动画生理仿真或类似表示来显示。

在一个实施例中，播放器软件包括备选重放选项，它允许音频和
25 /或视频内容的剧本与诸如演员的语音或计算机生成的语音之类的另一种语音一起播放。这个特征可与调整重放特征和语音训练特征结合使用。这在音频部分不清晰或者未使用正确发音时帮助用户。

在一个实施例中，播放器软件显示在音频和/或视频内容以及音

频和/或视频内容段的开始和结束处附加的介绍屏幕、序言屏幕和后文屏幕。介绍屏幕可能是允许用户选择在重放中预期的选项的菜单。在一个实施例中，用户可选择在重放过程中要跟踪或使用的一组偏好。在一个实施例中，用户可选择在重放过程中突出显示剧本中的单词的所选集合的‘热字标记’。这些单词被突出显示，以及‘提示’单词也可显示，它帮助解释或阐明突出显示单词的含义。在一个实施例中，用户感到有困难的单词被标记为‘热字’，并被索引或编目录以供用户参考。用户可启用书签，它允许用户在重放过程中标记将被返回或索引以供以后观看或收听的场景。在一个实施例中，介绍屏幕允许对语言、用户水平、特定用户标识和类似参数的选择，以便使语言学习内容适应用户的需要。在一个实施例中，用户水平分为初级、中级、高级和精通。在另一个实施例中，用户的这些水平基于数值标度、如 1-5，具有不断增加的难度等级和预计流畅性。每个较高等级比较低等级显示更多高级内容或更少辅助内容。在一个实施例中，介绍屏幕可包括其它产品的广告或者音频和/或视频内容。

在一个实施例中，序言屏幕可附加到音频和/或视频内容段的开始处(例如歌曲或电影场景)。在一个实施例中，与某个段关联的单词和惯用语可在序言屏幕中显示。所显示的单词和信息将符合指定的用户水平。在一个实施例中，序言屏幕在音频和/或视频段之前引入包括以下项目的素材：段中的单词，单词解释，单词发音，与音频和/或视频内容或语言相关的问题，与用户的先前体验相关的信息以及类似素材。序言中的链接允许用户在特定帧开始重放。例如，序言可具有序言与场景中出现的单词之间的链接，以便允许用户直接跳转到单词在其中被使用的音频和/或视频内容中的帧。在一个实施例中，用户可设置偏好，它阻止显示部分或全部序言屏幕，或者只在接收到另一个输入时才显示它们。在一个实施例中，屏幕截图或其它图像或动画用于序言屏幕中说明单词或概念或者标识关联场景。在一个实施例中，用于序言屏幕的一组预先渲染图像被包装为播放器软件的一部

分。在一个实施例中，在用户“选择参加”之前不显示序言屏幕，以免中断音频和/或视频内容的自然流动。

5 在一个实施例中，序言屏幕包括将为学习过程突出显示的特定单词、短语或语法结构。来自与场景相关的伴随文件 131 的相干素材由播放器软件来编译。播放器软件分析与场景中的各数据项关联的用户水平数据，并构造与用户水平对应的或者满足用户指定偏好或标准的相干数据类型的列表。在一个实施例中，与场景相关的附加素材可添加到列表中、例如“热字”，而与它所表明的用户水平无关。从列表中删除由播放器软件所存储的跟踪数据表明用户完全理解或者已经
10 由先前的序言屏幕检测过的素材。随机或伪随机函数则用于从序言屏幕中要使用的汇编列表中选择单词、短语、语法结构等。在另一个实施例中，序言屏幕上显示的单词或信息由编者选择或者从所收集的关于用户的数据来推理。

15 在一个实施例中，后文屏幕为交互测试或小程序，它测试用户对于与音频和/或视频内容相关的语言和内容的理解。在一个实施例中，问题被定时，以及正确和不正确的答案使不同的屏幕或者音频和/或视频内容被显示。在一个实施例中，如果出现超时，则显示正确答案。

20 在一个实施例中，后文素材位于场景或音频和/或视频内容的结尾。在一个实施例中，内容和问题根据观看或收听音频和/或视频内容的过程中所跟踪的用户输入自动产生。例如，基于再播放次数、用户感到困难的音频和/或视频内容段在后文中按照难度顺序再播放。在一个实施例中，根据相似的语言内容、人物、主题、演员或类似标准，可使用来自另一个音频和/或视频内容的内容或者与来自所观看
25 或收听的音频和/或视频内容的内容进行对照。在一个实施例中，后文屏幕显示包括与序言屏幕相似的链接的语言和词汇信息。后文屏幕可由用户以与序言屏幕相同的方式来去活或部分激活。在一个实施例中，屏幕截图或其它图像或动画用于后文屏幕中以说明单词或概念或

者标识关联场景。在一个实施例中，用于后文屏幕中的一组预先渲染图像被包装为播放器软件的一部分。播放器软件访问伴随文件 131 以确定何时插入序言和后文屏幕以及关联内容。在一个实施例中，所有后文屏幕‘选择参加’，但是，一旦音频和/或视频内容已经结束，
5 例如在影片结束时，在此情况下，后文将被提供，除非用户通过提供输入‘选择不参加’。

在一个实施例中，如上所述，播放器软件跟踪用户偏好和动作，以便根据用户需求更好地调整扩充重放信息。用户偏好信息包括用户精通水平、暂停和调整重放使用、训练表现、书签和类似信息。在一个实施例中，播放器软件根据用户输入编译可定制的单词数据库作为词汇表。
10

在一个实施例中，用户偏好可从播放器软件导出到其它装置和机器，以便与其它机器上的其它程序和播放器软件配合使用。在一个实施例中，服务器存储用户偏好，以及允许用户登录到服务器 135 以便获取和配置本地播放器软件来加入偏好。
15

图 4 是将伴随文件 131 与音频和/或视频内容相关的播放器软件过程的流程图。播放器软件识别用户希望观看或收听的音频和/或视频内容(框 413)。在一个实施例中，播放器软件访问音频和/或视频内容以查找标识数据序列，并且利用本地或远程数据库或者通过搜索本地可访问的伴随文件 131 将那个序列与伴随文件 131 相关。一旦音频和/或视频内容已经被识别，播放器软件确定适当的伴随源文件的副本是否本地可用。
20

在一个实施例中，伴随文件 131 可存储在可移动媒体存储产品、如 CD、DVD 或类似存储媒体中。在一个实施例中，如果伴随文件 131 不是本地可用的，则播放器软件通过网络 127 访问服务器 135，以便下载适当的伴随源文件。在一个实施例中，音频和/或视频内容的伴随文件 131 还可设置在相同媒体中，与音频和/或视频内容结合传送，或者从相同的远程存储位置传送。在又一个实施例中，伴随文
25

件 131 可存储在本地驱动器 105 或者存储装置 107 中。播放器可根据它与音频和/或视频内容的并存来标识适当的伴随文件 131(框 415)。在一个实施例中，播放器软件则开始对音频和/或视频内容的访问及重放(框 419)。本文所用的术语“媒体”用来表示传送内容的产品、管道及方法，例如 CD、DVD、网络流、广播和类似的传送方法。对位于相同媒体中的两个项目的说法表明，两个项目位于相同的产品或流(例如媒体的单个示例)中，对位于相同的媒体类型中的项目的说法表明这两个项目可能位于一个或多个产品中，例如一对 CD 或一对 DVD 或者网络流(或者可能位于单个媒体中)。

10 在一个实施例中，播放器软件逐帧或者逐个时间点地将音频和/或视频内容和伴随文件 131 相关(框 421)。在一个实施例中，伴随文件 131 以顺序方式包含基于与音频和/或视频内容中的各帧或时间点关联的一组索引、关于音频和/或视频内容的信息。根据准备用于显示的音频和/或视频内容的帧或时间点，播放器软件访问伴随文件 131 中的相关数据以产生扩充重放。相关数据可包括剧本、词汇、惯用表达以及涉及音频和/或视频内容的对话的其它语言相关素材。

20 在一个实施例中，伴随文件 131 可以是平面文件、数据库文件或类似的格式化文件。在一个实施例中，伴随文件 131 的数据以 XML 或类似的计算机解释语言来编码。在另一个实施例中，伴随文件 131 将以具有各单词、行、场景示例和适当的类的对象的示例所表示的相似段的面向对象的范例来实现。

25 在一个实施例中，播放器利用伴随文件 131 的数据来扩充音频和/或视频内容的重放(框 423)。扩充可包括显示文本、语音发音、链接到附加菜单的图标以及与音频和/或视频内容相关的特征，例如与音频和/或视频内容相关的指南、菜单以及类似信息。在一个实施例中，通过播放器软件和伴随文件 131 可得到的其它资源包括：剧本中句子结构的语法分析和解释，语法相关课程，惯用表达的解释，人物及内容相关索引以及类似资源。在一个实施例中，播放器将访问初始行或

场景部分,并使用其中的信息来查找单词索引中的起始位置及相应的起始帧。重放将依次连续通过每个部分,除非通过请求访问特定信息或者跳转到音频和/或视频内容中不同位置的用户输入来转向。

图 5 是示范伴随文件格式的示意图。在这个实施例中,伴随文件
5 131 配置成与音频和/或视频内容、如影片、有声书籍、电视节目及类似情况配合使用。在一个实施例中,伴随文件 131 分为剧本相关数据和元数据。在一个实施例中,剧本相关数据主要是依次存储或索引的数据,包括与其中包含单词、行和对话交流的剧本相关的数据以及场景相关数据。元数据主要是根据用户请求来访问的辅助或参考相关数据,
10 例如词典数据、发音数据和内容相关索引。

在一个实施例中,剧本数据以平面顺序二进制格式 500 存储。平面格式 500 包括按照所定义分级结构分组的与剧本相关的多个部分。每个部分中的数据以按照剧本的顺序的顺序方式来组织。在一个实施例中,格式中的字段具有固定长度。在一个实施例中,这些部分
15 包括单词部分、行部分、对话交流部分、场景部分和其它类似部分。单词部分包括标识单词部分序列中的单词位置的单词示例索引、单词文本、将单词链接到定义数据的单词定义标识或指针、将单词链接到相关发音数据的发音标识字段或指针、以及标识来自音频和/或视频内容、单词与其关联的起始和结束帧的起始和结束帧字段。在一个实
20 施例中,行部分包括标识行部分序列中各行的位置的行索引、表明单词部分中与行关联的第一单词的起始单词索引、表明与行关联的最后一个单词的结束单词索引、表明或指向与剧本行的语言解释相关的数据的行解释索引、指向行或者将行与音频和/或视频内容中的人物链接的人物标识字段、起始和结束帧指示符以及类似信息或者指向与行
25 相关的信息的指针。在一个实施例中,对话交流部分包括标识对话交流部分的索引中的位置的交流索引、与对话交流关联的起始帧和结束帧以及类似指针和信息。在一个实施例中,场景部分包括标识场景部分中的场景的位置的索引、序言标识字段或指针、后文标识字段或指

针、起始和结束帧以及类似指示符和与场景相关的信息。

5 在一个实施例中，元数据部分包括行解释部分、单词词典部分、单词发音部分以及与涉及音频和/或视频内容及其中语言的辅助和参考类型信息相关的类似部分。在一个实施例中，解释部分将包括表明行解释部分中的行解释的位置的索引、表明对应行的行索引、与为给定行提供的各种语法和语义解释数据相关的一组解释数据字段以及
10 与对应行解释的数据相关的类似字段。在一个实施例中，单词发音部分包括表明单词发音部分中的示例位置的索引、指向音频数据的指针、音频数据字段的长度、音频数据类型字段以及类似的发音相关数据和指针。

15 在一个实施例中，指针用于字段中来表明大于二进制文件中的字段大小的数据。这允许所使用数据的大小方面的灵活性，同时保持二进制文件中的字段的标准格式和长度。在一个实施例中，伴随文件 131 具有用于编辑和文件创建的备选格式，例如 XML 和其它标记语言、数据库(例如关系数据库)或面向对象的格式。在一个实施例中，
20 伴随文件 131 以不同格式存储在服务器 135 中。在一个实施例中，伴随文件 131 作为关系数据库文件来存储，以便于文件在被创建或编辑时的动态修改。数据库被平展为平面文件格式，以便于重放过程中由播放器软件进行访问。

25 在另一个实施例中，伴随文件 131 的格式可被修改或重新定义以用于其它内容类型，例如相册、歌曲、音乐视频、教育素材、记录片、访谈及类似的内容。例如，相册的伴随文件 131 可根据声迹中的时间点而不是场景和行来组织。打算在便携式装置中使用的伴随文件 131 可具有基于便携式播放器装置的功能的字段的简化集。例如，与发音或者剧本的详细分析相关的字段可被省略或忽略。

图 6 是内容控制系统的操作的流程图。在一个实施例中，内容控制系统可允许用户选择音频和/或视频内容中要过滤或改变的内容的类型。例如，父母可能想过滤他们的小孩将观看或收听的影片或歌曲

的亵渎语言。这个控制内容系统可用于语言学习系统的上下文中，或者可用于在传统观看或收听娱乐及类似媒体期间控制内容。

内容控制系统根据伴随文件 131 来起作用，伴随文件 131 包含对于与音频和/或视频内容关联的转录的单词和短语分类的信息。仅与
5 内容控制系统配合使用的伴随文件 131 可具有专用格式，它包括单词和短语的索引脚本及分类，但可省略与其它增强特征相关的其它数据和字段。可对于随机或顺序存取优化伴随文件 131。在另一个实施例中，伴随文件 131 中的附加内容的索引可能不是基于剧本而是可能基于帧、时间参考或者索引音频和/或视频内容的类似方法。在一个实
10 施例中，这种索引有助于非语言内容控制、例如裸露。

内容控制系统取决于伴随文件 131，伴随文件 131 包含音频和/或视频内容所用的剧本中的段、单词和短语的每个的分类的标识(框 601)。剧本的每个段、单词、短语或类似部分可根据它是否涉及用户可能希望过滤的性内容、暴力内容、亵渎内容、邪恶内容或类似内容
15 来分类(框 603)。伴随文件 131 与分类数据和剧本可在相同媒体、分开的媒体上或者通过相同或分开的分发方法提供给具有播放器程序的用户的本地机器(框 605)。伴随文件 131 可包含与媒体的单词、帧或段关联的属性。例如，分配给单词的某个属性可能是表明厌恶程度的数值等级。

20 用户可采用播放器所提供的界面来确定要过滤的内容集合(框 607)。图 7 是内容控制系统的示范界面屏幕。界面屏幕包括选择用户希望观看、收听或变更的分类集合的一组导航选项或图标 705。在示例界面中，内容分为语言、暴力、性、裸露和道德分类。所示的语言
25 屏幕的界面屏幕包括与所选分类关联的单词或短语的列表。在示例界面屏幕中，在本例中称作亵渎语言的语言中的所有单词和短语均被显示。用户可选择在重放中被显示或者例如要被忽略的单词或短语。在一个实施例中，选择触发布尔值，它在激活过滤器时标记是否要重放、改变或类似地审查音频和/或视频内容的单词、短语、场景或类

似部分。在另一个实施例中，更大粒度选择可允许用户应用可能影响音频和/或视频内容的过滤的一系列选项。可能选项的一部分实例包括使某段静音、跳过段、跳过相关段和类似的可能审查技术。

在示例界面屏幕中，在一个实施例中，选择可通过滑动指示器 5 703 来实现。当滑块朝“冷”移动时，厌恶性的门限变得更低。这样，在最低处，所有厌恶单词将被忽略。如果假设零与十之间的亵渎标度，其中十为最亵渎，则当滑块位于中间时，具有大于五的亵渎属性的单词将被选择用于变更。类似的属性等级可结合其它分类来指定。在一个实施例中，单词旁边的单选按钮随着滑块的移动而改变，因此 10 用户可看到在选择时滑块的移动的效果。在一个实施例中，属性可以是与特定分类的某个单词或短语(场景、帧或段)关联的值，它标识单词或短语可能被过滤的条件。属性通常包含在伴随文件 131 中，但在部分实施例中可以是用户定义的。

在示例屏幕界面中，范围从‘热’到‘冷’的滑杆指示器 703 可 15 用来设置单词组或分类的过滤器等级。与属性值以及一组单词或短语的滑杆指示器 703 的位置有关的信息可由播放器软件结合诸如当前用户身份、时刻、内容类型(例如音乐或视频)以及可能影响哪个过滤等级适当的类似数据之类的其它信息使用。

界面屏幕可具有附加特征以便于选择修改的内容。在一个实施例中， 20 界面屏幕可包括观看屏幕 707，以便观看或收听单词或短语在其中出现的一段音频和/或视频内容。如果内容只是音频，则可视表示可伴随此音频。例如，用户可从分类‘语言’的单词列表中选择单词‘流产’。这个单词在其中出现的影片或音乐段则可排队，用于在观看屏幕 707 中进行审阅。界面屏幕还可包括导航选项和图标 709，以 25 便继续进行播放或者访问附加信息或选项。

在一个实施例中，在重放过程中，播放器连续检查被播放的当前段，以便确定过滤器是否应当应用到将要播放的单词或短语(框 609)。在一个实施例中，播放器可跳过包含要过滤内容的音频和/或视

频内容的场景或段。在另一个实施例中，内容可变模糊、静音、发出
哗哗声或者以阻碍对过滤内容的观看或收听的类似方式来审查。在一个
实施例中，播放器软件允许用户从用于过滤的这些选项中选择要过滤
5 的单词或短语的不同类别或示例。用户偏好可保存供以后使用。偏
好可结合到单个内容或者在内容的各分类上一般化。用户可完全禁用
内容控制。在一个实施例中，禁用控制的能力被限制到主用户，并且
可能具有密码保护或类似保护。

图 8 是用于增强用户为了语言学习而观看或收听音频和/或视频
内容的学习体验的推理机的流程图。在一个实施例中，播放器
10 可跟踪与音频和/或视频内容的重放相关的用户输入。播放器通过以
缺省重放模式或者按照播放器的当前设定向用户呈现音频和/或视频
内容开始(框 801)。播放器还根据用户能力的缺省水平或者用户的语
言能力的当前估计水平来提供对附加内容的访问(框 803)。

在一个实施例中，在音频和/或视频内容及附加内容的重放过程
15 中，播放器跟踪用户的响应和输入的类型(框 805)。所跟踪的输入和
响应的类型可包括用户回溯音频和/或视频内容的特定单词、短语或
段的播放的次数、用户观看或收听某个段的速度、对用户提供的问题
的响应、使用对于提示或问题的帮助信息响应所耗用的时间、诸如红
外相机读取、控制器使用、用户移动、静止等的生理反馈以及类似的
20 信息和数据。推理机分析所收集数据，以便确定用户的主观语言的知
识水平(框 807)。

在一个实施例中，对用户语言方面的能力的这种确定则用来选
择或调整向用户呈现音频和/或视频内容的设定。推理机可利用可变
加权和类似计算来评估用户能力。推理机可实现为专家系统、神经网
25 络或类似系统。在一个实施例中，推理机可被设计或训练成供不同语
言及文化背景的用户使用。

在一个实施例中，播放器可变更播放单词或短语的速度，可改变
序言或后文段中的问题的类型或数量，可改变剧本的显示，变更背景

音乐的等级, 提供附加内容, 提供动画人物, 通过各种音调变化提供剧本的文本的发声, 提供词典定义, 以及可将重放调整到适合用户的学习需求的类似动作。在一个实施例中, 在音频和/或视频内容的重放过程中, 可提供画外音来帮助用户理解内容。画外音可以是剧本的文本的发声、内容的解释(例如场景、对话交流、概念、短语、单词或类似内容的解释)或者在伴随文件 131 中提供的类似素材。对重放的其它调整可包括调整音频的各种方面的音量(例如背景音乐、对话和类似声迹)、静音、速度调整、暂停和类似动作。被确定具有高能力水平的用户一般将接收到较少辅助或者更复杂的帮助, 以及具有较低能力水平的用户一般将接收到更多辅助以及更简单的辅助类型。

用户可忽略推理机的设定, 而选择获取更高或更低能力水平的辅助。在一个实施例中, 系统存储推理机跟踪和状态数据供将来使用。数据和状态可用于将来使用特定内容, 或者用作具有新内容的通用模板。所存储数据可包括加权因子、神经连接数据、历史记录及类似数据。

图 9 是用于跟踪播放器的重放位置的系统的流程图。被跟踪的重放会话位置信息可用来保存‘书签’, 让用户从他或她在音频和/或视频内容中较早前停止的位置继续进行。这个系统以会话开始来开始(框 901)。本文所用的“会话”可以是用户开始音频和/或视频内容的重放直到重放被停止的一段时间。重放可由用户的直接选择或者通过某种系统故障或类似事件、如电力失去来停止。重放监测系统按照一定的间隔存储重放位置(框 903)。在一个实施例中, 间隔可小于三十秒间隔。在一个实施例中, 间隔小于一秒。在部分实施例中, 系统的状态在每个间隔被存储。状态存储可通过存储自上一个间隔以来的状态的 Δ 来实现。只要会话过程中的重放继续进行, 重放监测系统可按照一定间隔连接存储重放位置(框 905)。在一个实施例中, 如果重放被中断或终止, 则在重放重新开始时, 重放将在先前停止的点自动恢复(框 907)。用户可利用外围设备或类似输入装置来选择不参加。用

户可通过偏好设定来改变自动重新开始。在另一个实施例中，如果重放被中断或终止，则在重放重新开始或者在新会话开始时，播放器可提供在上一次保存的位置开始重放。在又一个实施例中，重放的重新开始可在音频和/或视频内容中比上次播放点稍前的点开始。重放也可在5 当前段的起始处、在前一个句子或对话交流结束之后或者在类似的起始点开始。在一个实施例中，自上一次重放会话以来经过的时间量可被分解为播放应当重新开始的位置的确定。例如，如果重放因例如两分钟电话呼叫而被中断，则以最近句子的起始处开始可能是足够的。但是，如果经过数天，则可能希望返回到例如当前对话交流的开头。

10 在一个实施例中，播放器利用特殊存储器或存储装置来跟踪重放位置。在另一个实施例中，与播放器分开的装置可管理重放位置的存储。存储器可以是诸如 EPROM、闪存、电池支持的 RAM 或类似存储装置等非易失性存储器、固定光盘媒体、磁媒体、物理媒体或类似15 存储装置。重放的位置可通过相对音频和/或视频内容的起始的重放时间点、通过使用索引、段标识信息或类似的位置标识信息来确定。在一个实施例中，系统可存储多个重放位置。不同的音频和/或视频内容的重放位置可同时存储。在一个实施例中，系统的附加状态信息可被跟踪和存储，其中包括附加素材重放位置、推理机、变更记录、20 当前设定和偏好以及类似数据。

在一个实施例中，播放器应用、服务器应用和其它要素以软件(例如微码、汇编语言或高级语言)来实现。这些软件实现可存储在机器可读媒体中。“机器可读”媒体可包括能够存储或传递信息的任何媒体。机器可读媒体的实例包括 ROM、软盘、CD-ROM、DVD、闪存、25 硬盘驱动器、光盘或类似媒体。

在以上说明中，参照特定实施例对本发明作了说明。但是显然可对其进行各种修改及变更，而未背离所附权利要求提出的本发明的广义精神和范围。因此，说明和附图应看作是说明性而不是限制性的。

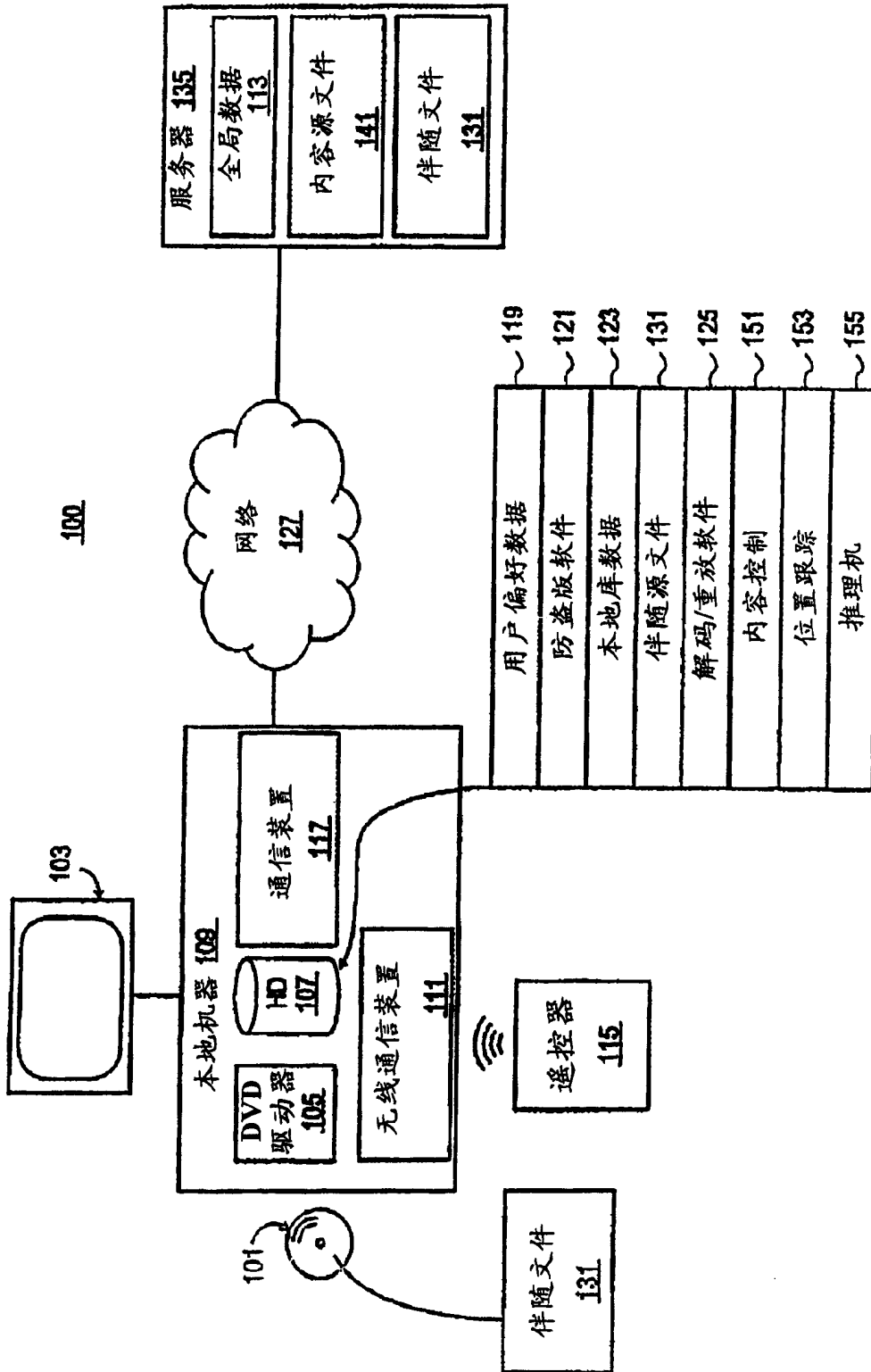


图 1

图 2A

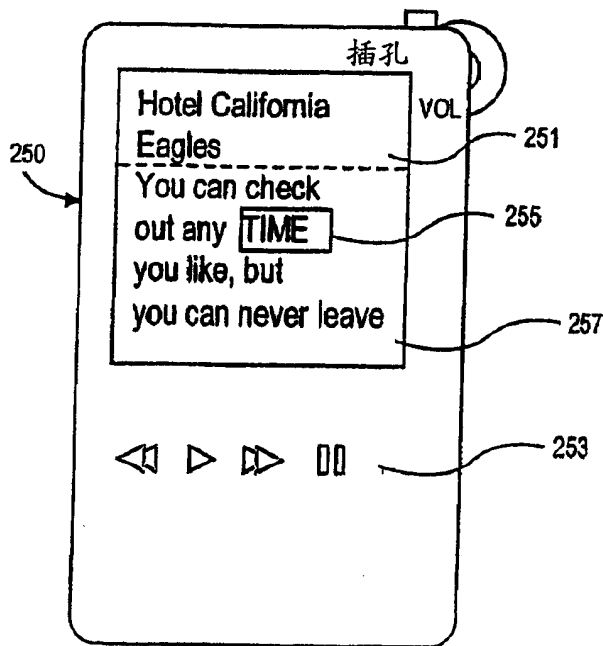
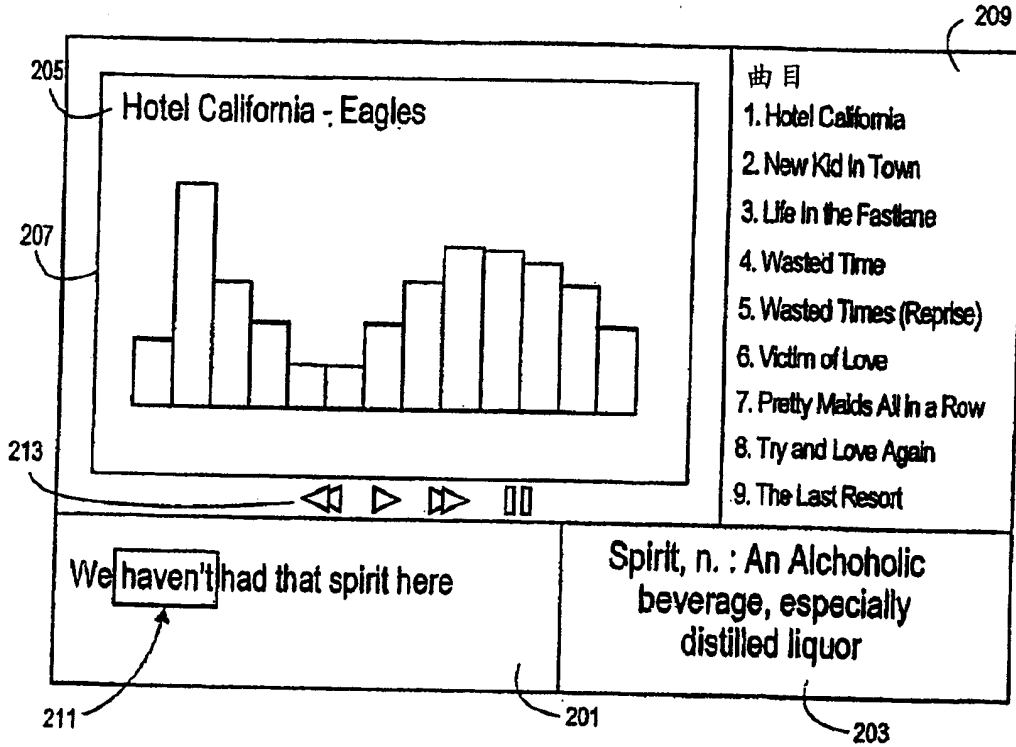


图 2B

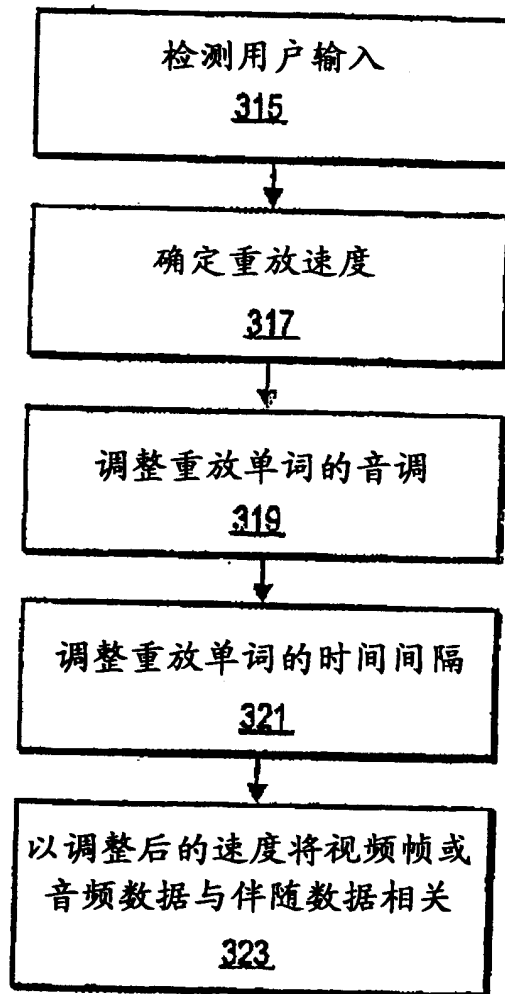


图 3

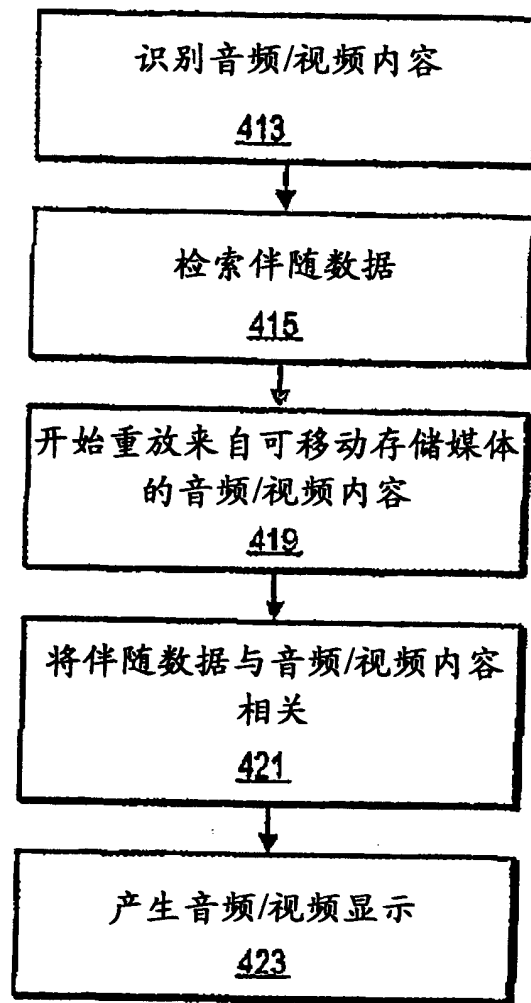


图 4

500

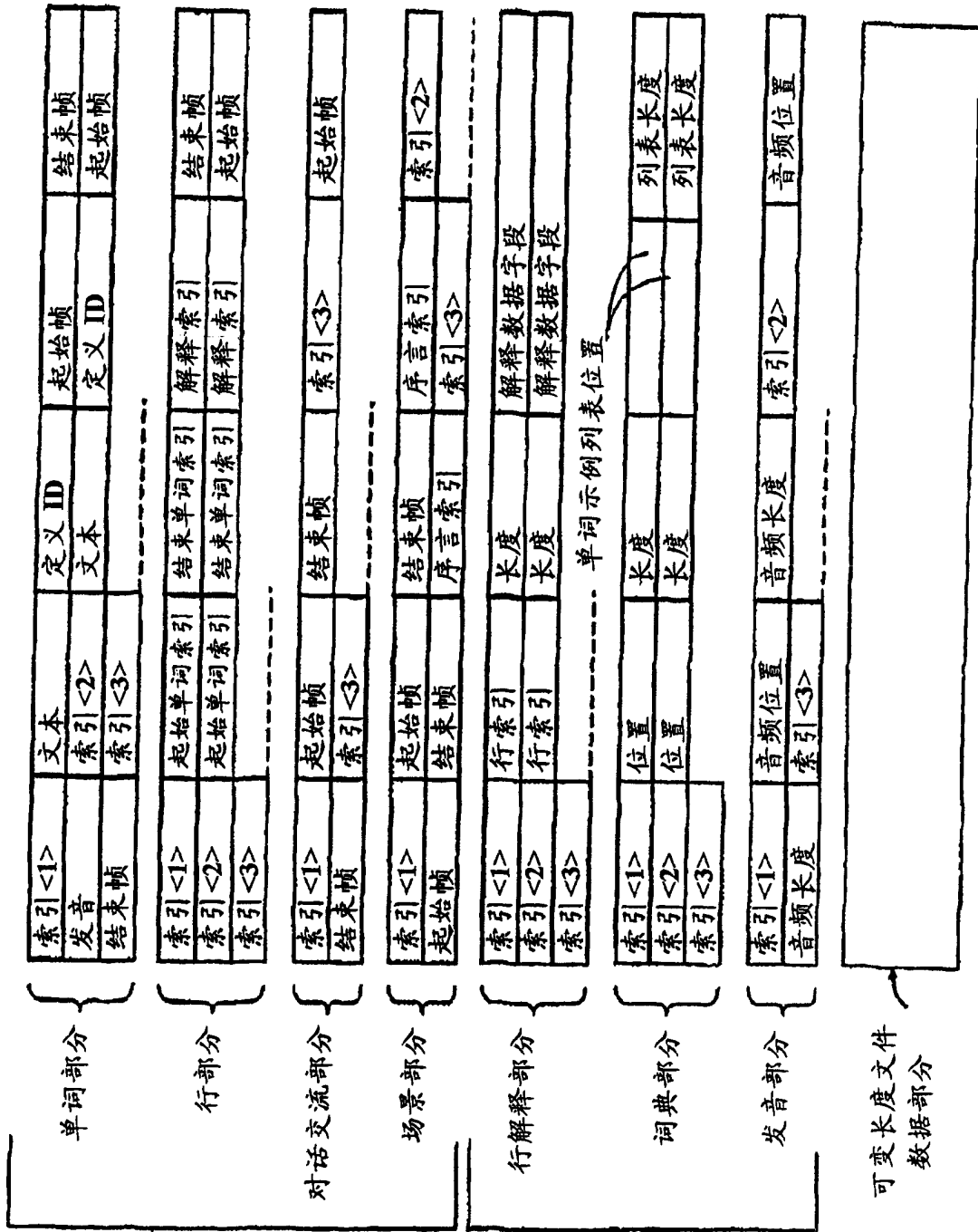


图 5

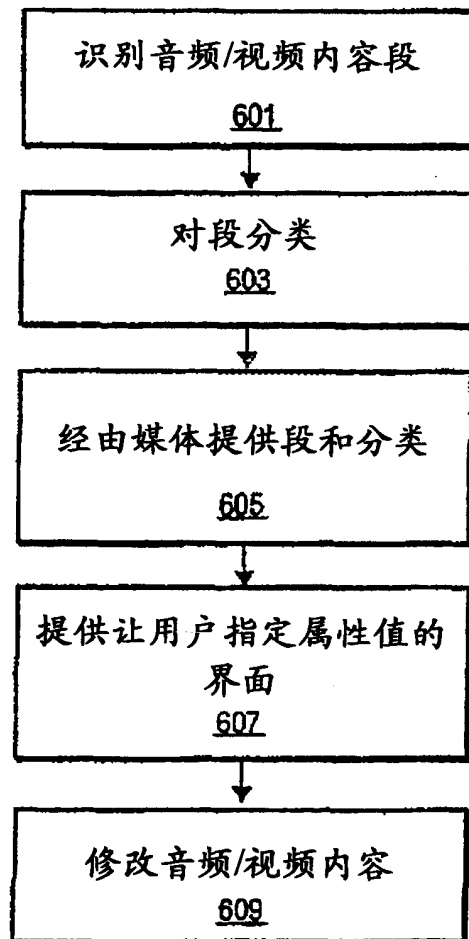


图 6

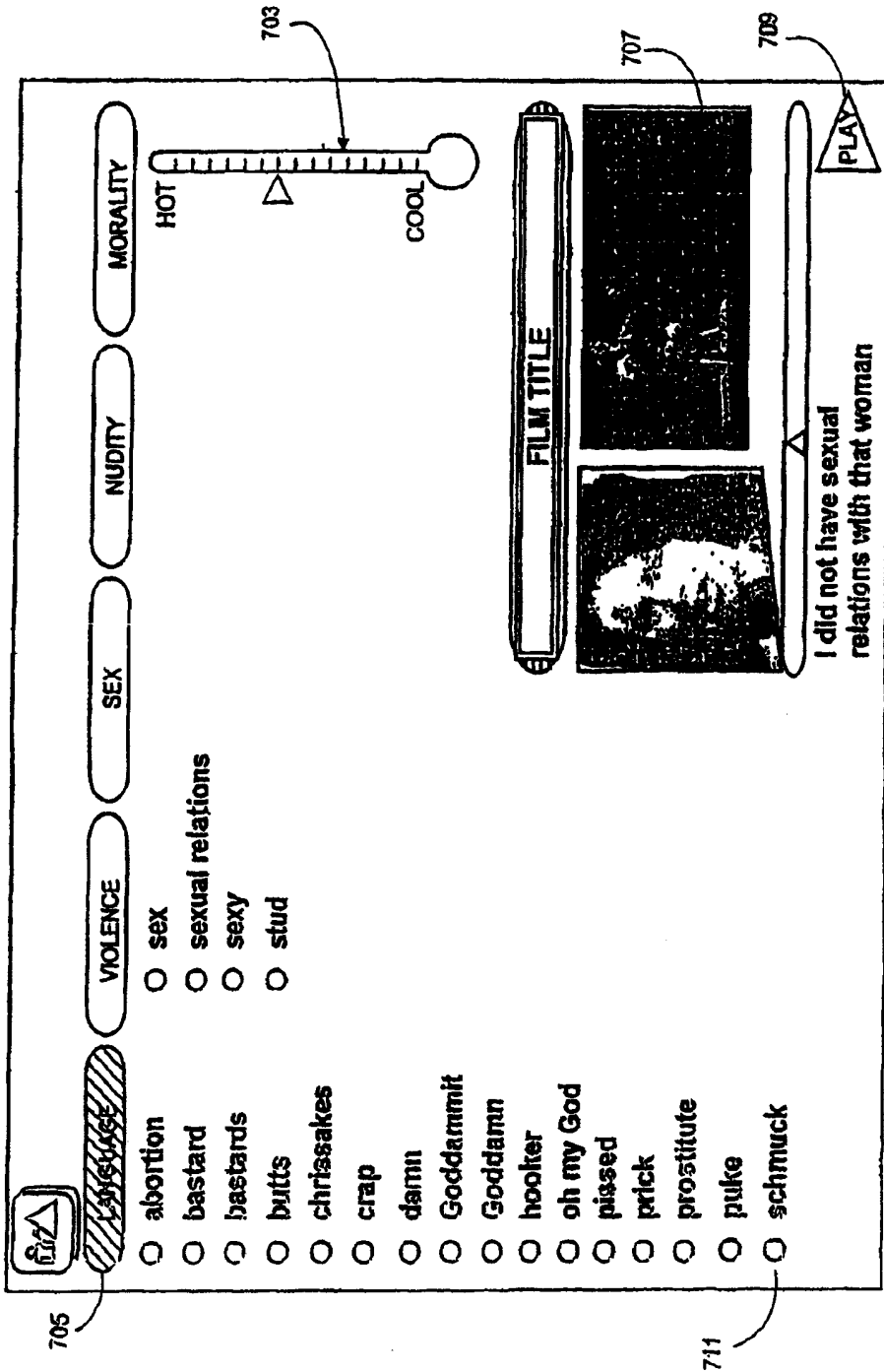


图 7

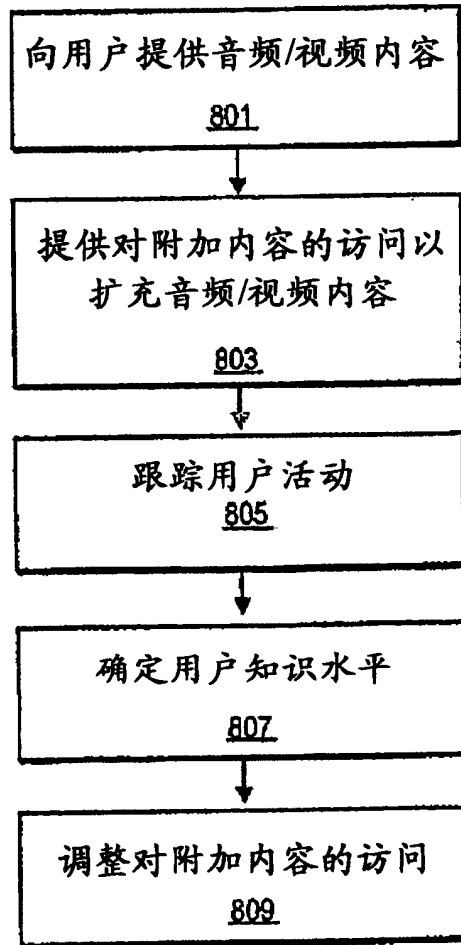


图 8

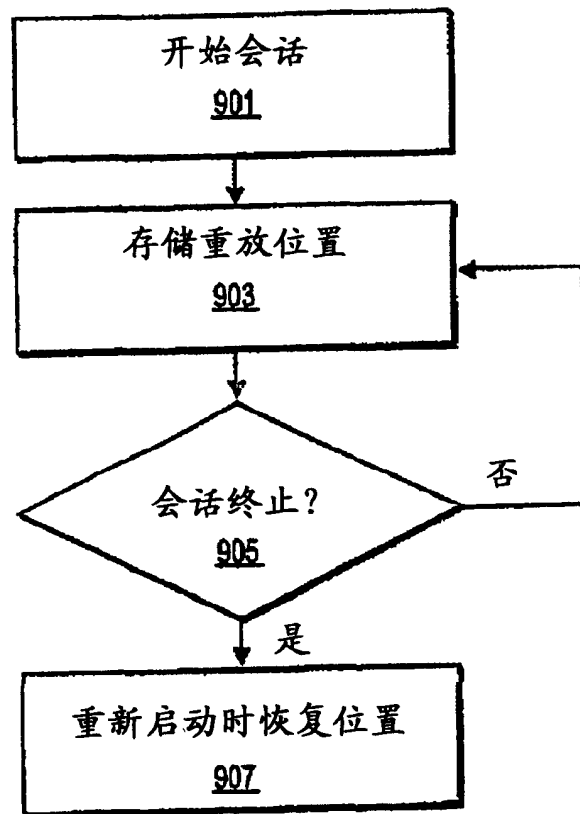


图 9