[19]中华人民共和国国家知识产权局

[51]Int.Cl[7]
G09B 5/00

[12]发明专利申请公开说明书

[21]申请号 98806601.7

[43]公开日 2000年8月23日

[11]公开号 CN 1264484A

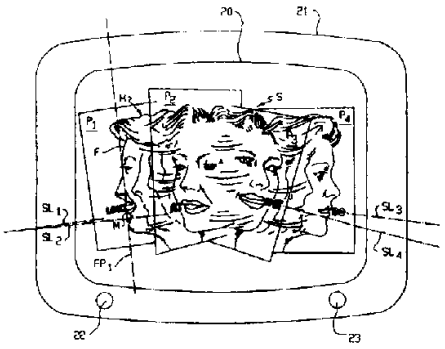
[22]申请日 1998.4.25 [21]申请号 98806601.7
[30]优先权
[32]1997.4.25 [33]US [31]08/845,863
[86]国际申请 PCT/US98/08447 1998.4.25
[87]国际公布 WO98/49666 英 1998.11.5
[85]进入国家阶段日期 1999.12.27
[71]申请人 里德斯皮克公司
地址 美国纽约州
[72]发明人 W·E·柯克塞
[74]专利代理机构 中国专利代理(香港)有限公司
代理人 王 勇

权利要求书9页 说明书18页 附图页数38页

[54]发明名称 用于制作具有一系列可视字符与相配合的语音的一个音像制品的方法与系统及其这种音像制品

[57]摘要

本发明有关音像制品及其生产的方法,该音像制品具有放置在制品图象上的文 字,以使随着每一单词或其它发音被听到,一个与所听内容关联的文字通过加 重该文字或其它方式与所看见的内容相配合,使得发音或者文字的未来播放将 唤起视听者头脑中的文字或者发音。每一单词将以可读的透视的形式适当地出 现在发声者(S)的嘴部(M)的旁边。制品能够线性地显示或在视听者的计 算机控制之下连同附加教育的材料一同显示。可以使用加亮的词组的选择,或 不使用加亮的词组。



ISSN1008-4274

权　利　要　求　书

1.一种教学音像制品，包括具有自然的交流情形的图象景物以播放给视听者，该音像制品包括：

(a)至少一个图象景物，包括在一个屏幕上显示的并且其伴音能够被听到的多个画面；

(b)由至少一个发音者发出的一系列单词发音，每一发声者的头部在这种景物播放的连续画面中出现；

(c)一系列可视单词符号，一个或者多个单词符号在靠近该发声者头部的一个画面部分中显示一个时间期间，以一个或者多个这种单词符号在有差异的该邻近画面部分中显示一个有差异的时间期间，其中每一有差异的单词符号：

(i)对应于每一这种单词发音并且与每一这种单词发音关联，

(ii)能够教导视听者被发音的单词的直观外形；

(iii)在该靠近画面部分中显现这种差异，以使视听者具有能够做出这种关联的机会，并且

(iv)在该有差异单词的发音被听到的同时在图象景物的这种画面中显现该差异，以使得同时把每一这种发音和每一这种相应的有差异的单词符号相关联，并且关联由该视听者所视听的时间长度，

从而该视听者把每个这种发音与每个这种有差异符号的单词关联，以便记住说出的单词的直观形状和该单词符号的发声。

2.权利要求1的音像制品，其中随着所选择的单词被发声之时，该一个或者多个单词符号显现是从该嘴部按顺序地引出的，所选择的引出单词之间有差异。

3.权利要求1的音像制品，其中每一单词符号在首先出现在嘴部之后就使之有差异。

4.权利要求1的音像制品，其中的差异是由对单词进行加重实现的，其单词在差异时间期间结束时被解除加重，并且一个其次显示的单词被加重一个第二加重时间期间。

5.权利要求1的音像制品，其中单词符号从发声者的嘴引出并且同时从该嘴部移走而保持接近头部，并且随后移到远距该头部的画面的一个部分。

6.权利要求1的音像制品，其中该发声者具有耳朵，并且其中该

单词符号出现在实际上垂直于穿过该发声者的耳朵的一条直线的一个对话平面中。

7. 权利要求1的音像制品，其中该播放包含非教程景物，以这种单词符号叠加在这种景物上。

8. 权利要求1的音像制品，其中一个或多个单词出现在一个对话平面中，并且这种对话平面相距平板型屏幕平面至少六十(60)度。

9. 权利要求1的音像制品，其中至少一个单词符号是一个标记。

10. 权利要求1的音像制品，其中至少一个单词符号是一个图符。

11. 权利要求1的音像制品，其中该图象景物以一线性方式播放而没有中断。

12. 权利要求1的音像制品，其中通过使用可由视听者操作的计算机，该图象景物以一个中断的非线性方式播放。

13. 权利要求1的音像制品，其中该音像制品的一部分被扩展到可看到一个或者多个单词符号的时间长度。

14. 权利要求8的音像制品，其中一对话平面通过一个发声者的头部定位，其中至少一个单词符号出现在这种对话平面或接近这种对话平面。

15. 权利要求1的音像制品，其中一个发声者的头部包含面孔和嘴唇，而其中与单词发音相关的面部的、嘴唇的和头部的运动被包含在这种画面中。

16. 权利要求1的音像制品，其中针对一个选择的时间出现有差异的单词符号，并且在该时间期间，该单词符号随着画面变化由在该图象景物中的一个目标遮掩。

17. 权利要求1的音像制品，其中该单词符号被整形和呈一定角度，以便显现为三维的单词符号。

18. 权利要求1的音像制品，其中发音系列是具有含意的句子形式中的一系列单词。

19. 用于对观众播放的一个音像制品，包括：

a) 多个场面；

b) 至少一个图象景物，包括具有在一个屏幕上显示的并且它们的伴音能够被听到的多个画面；

c) 由至少一个发声者发音的一个或者多个单词，每一发声者的头

部出现在景物播放的这种连续的画面中，其单词发音由该视听者在一个发音时间期间听到；

d)在一个或多个单词符号期间的发音时间期间，一个或者多个单词符号显现在发声者头部的近处；

i)与所说的一个或者多个单词发音对应并且与所说的一个或者多个单词发音关联；

ii)相对于发声者的嘴部定位；

iii)传达该单词发音的含意；

从而，观众可以把说出的一个单词或多个单词与看到的一个单词或多个单词相关联，并且从而该看到的一个或多个单词把该含意提供给该说出的一个或多个单词的观众。

20.权利要求19的音像制品，其中说出的一个或多个单词是具有含意的单词组，并且其中该一个或多个单词被显示具有相同的含意。

21.权利要求19的音像制品,其中该发声者具有耳朵，并且其中该单词符号出现在实际上垂直于穿过该发声者的耳朵的一条直线的一个对话平面中。

22.权利要求19的音像制品，其中该播放包含非教程景物，以这种单词符号叠加在这种景物上。

23.权利要求19的音像制品，其中有一个对话平面，并且这种对话平面相对平板型屏幕平面至少六十(60)度。

24.权利要求19的音像制品，其中至少一个单词符号是一个标记。

25.权利要求19的音像制品，其中至少一个单词符号是一个图符。

26.权利要求19的音像制品，其中该图象景物以一线性方式播放而没有中断。

27.权利要求19的音像制品，其中通过使用可由视听者操作的计算机，该图象景物以一个中断的非线性方式播放。

28.权利要求19的音像制品，其中该音像制品的一部分被扩展到可看到的一个或者多个单词符号的时间长度。

29.权利要求19的音像制品，其中一对话平面通过一个发声者的头部定位，其中至少一个单词符号出现在这种对话平面中或接近这种

对话平面。

30. 权利要求19的音像制品，其中一个发声者的头部包含面孔和嘴唇，而其中与单词发音相关的面部的、嘴唇的和头部的运动被包含在这种画面中。

31. 权利要求19的音像制品，其中该单词符号显示一个所选择的时间，并且在该期间中一个或者多个该单词符号随着画面变化由在该图象景物中的一个目标遮掩。

32. 权利要求18的音像制品，其中该单词符号被整形和呈一定角度，以便显现为三维的单词符号。

33. 使得包括若干成员的一组人员懂得在一地区中说的语言的一组书面单词的一种方法，包括步骤:

(a)使得产生多个音像制品，每一制品包括多个图象片段，还包括伴随每一图象片段的一系列发音;

(b)使得在这种片段上出现一个或者多个对应于这种发音的书面单词，其出现的方式是通过使之有所差异或利用相应发音的别的方式把每一发音和每一书面单词相关联;

(c)在所说的制品中至少提供一个片段用于在书面单词组中的每一单词; 和

(d)在一定程度上把这种制品播发并以别的方式引入到该地区一个时间期间，从而使得该人员组懂得这种语言的书面单词。

34. 一种教学学生的方法，包括步骤:

(a)产生包括自然交流情形的一个或者多个音像制品，每一制品包括在多个同时地具有对应文字的发音情形过程中的播放， 每个视听制品包括一个学生的注意力被自然吸引的区域，该区域包括: (i) 一个第一部分，发音看上去从其发出或学生的视线由发音的含意引导被吸引到该部分，和(ii) 一个第二部分，显示同时具有每一发音的一个或者多个书面单词，使得在学生的记忆中出现在发音与一个有差异的相应书面单词之间的关联;

(b)使学生可利用这种制品; 和

(c)允许学生选择进行观看和收听的时间安排，直到多个发音的每一个都已经被听到，连同它们相应的文字的观看足够的次数，以便在对应于某些文字的某些发音的学习中帮助学生。

35. 权利要求34的方法，其中的第一和第二部分重叠。

36. 用于播放给视听者的包括具有自然的交流情形的图象景物的一种教学音像制品，包括下列部分:

(a)由一个人或其它发声者在场景的播放中产生的一系列发音;

(b)与这种系列发音相关的一系列文字，通过初始外形或加重所特定化的文字，简要地位于图象景物中，这种文字对应听到的发音，以使每一发音和文字在视听者的头脑中相关联。

37. 权利要求36的方法，具有在一个对话平面中放置一个或者多个单词符号的另外步骤。

38. 权利要求36的方法，其中发音系列是具有含意的句子形式中的一系列单词。

39. 权利要求36的方法，具有进一步的步骤: 使该发声者具有耳朵，并且使该单词符号看上去在实际上垂直于穿过该发声者的耳朵的一条直线的一个对话平面中。

40. 权利要求36的方法，具有附加的步骤: 以一个或者多个单词符号叠加在非教学景物上而在该图象景物中展现非教学景物。

41. 权利要求36的方法，其中该视听制品在一个平板型屏幕平面上观看并且具有把这种对话平面相对于这种平板型屏幕平面至少六十(60)度定位的步骤。

42. 权利要求36的方法，其中至少一个单词符号是一个标记。

43. 权利要求36的方法，其中至少一个单词符号是一个图符。

44. 权利要求36的方法，具有以一线性方式播放该图象景物而没有中断的步骤。

45. 权利要求36的方法，具有通过使用可由视听者操作的计算机以中断的非线性方式播放该图象景物的步骤。

46. 权利要求36的方法，具有扩展该音像制品的一部分，以便延长被看到的第一区域单词符号时间长度的步骤。

47. 权利要求36的方法，其中一对话平面通过一个发声者的头部定位，其中至少一个单词符号出现在这种对话平面中或接近这种对话平面。

48. 权利要求36的方法，其中一个发声者的头部包含面孔和嘴

唇，而其中与单词发音相关的面部的、嘴唇的和头部的运动被包含在这种画面中。

49. 权利要求36的方法，其中一个或多个单词符号出现持续一个所选择的时间长度，在该期间该一个或多个单词符号随着画面变化由在该图象景物中的一个目标遮掩。

50. 权利要求36的方法，其中该单词符号被整形和呈一定角度，以便显现为三维的单词符号。

51. 权利要求36的方法，其中发音系列是具有含意的句子形式中的一系列单词。

52. 在图象、画面或一个视听播放的片断上定位一系列文字的方法，其图象、画面或其它片段在上面具有人或没生命的发声者产生发音，该方法包括步骤：

(a)选择包括一种视频文本装置的设备的一个操作器控制单元；

(b)使得这种视频文本装置显示多个单词，每一单词有各种尺寸、形状和用于重现的方向；

(c)通过单元的操作器定位用于观看的一系列图象的每一个；

(d)操作该视频文本装置，以便选择该显示的单词；

(e)把所选择的单词输送到一个所选择的图象、画面或片段上用于在那上面永久定位；并且

(f) 使得所选择的单词被加重，以使显现在一个片段上的单词与该单词的发音在位置和时间上都关联。

53. 权利要求52的方法，其中单词接近发声者的头部放置。

54. 一种基于计算机的数据处理系统，用于包括伴随每个图象片段的发音的多个图象片段的改进的字幕，该数据处理系统包括：

(a)用于处理数据的计算机处理器装置；

(b)用于存储数据的存储装置；

(c) 用于在表示包括伴随每一图象片断的发音的多个图象片段的数据、表示书面的一个或多个单词的数据上进行叠加的装置，对应于这种发音加重该一个或多个单词，其方式是使得每一发音和每一书面单词被关联。

55. 权利要求54的系统，其中用于叠加的装置中包括：

(a)用于输入第一模拟视频信号的装置；

(b)用于把该模拟视频信号转换到视频数字数据的装置;

(c)用于每次把该视频数字数据显示为一个画面的装置;

(d)用于从用户输入单词数据和象限数据的装置;

(e)用于根据象限数据和其它预定的准则把单词数据结合在该视频数字数据中的装置;

(f)用于把视频数字数据转换到一个第二模拟视频信号的装置;

(g)用于输出该第二模拟视频信号的装置。

56.权利要求54的系统，其中用于从用户输入单词数据和象限数据的装置包括:

(a)用于允许该用户选择一个象限的装置;

(b)用于允许该用户绘出一条源线的装置;

(c)用于计算针对该源线的一个角度的装置;

(d)用于允许该用户输入单词数据的装置；和

(e)用于自动地沿着该源线进行定位的装置。

57.一个基于计算机的用于允许用户交互操作一个改进的字幕的音像制品的系统，包括:

(a)用于显示该音像制品的装置;

(b)用于停止该音像制品的显示的装置;

(c)用于允许该用户指定一个单词的装置;

(d)用于处理关于由该用户指定的单词的进一步数据的装置。

58.权利要求57的系统，其中允许该用户指定一个单词的装置允许该用户指定当音像制品的显示被停止时显示的一个单词。

59.权利要求57的系统 其中用于处理关于该单词的进一步数据的装置包括:

(a)用于存储表示针对多个单词的正确的发音的标准数据的装置;

(b)用于从该用户输入表示该用户的单词发音的话音数据的装置;

(c)用于把该话音数据与针对该单词的标准数据相比较的装置;和

(d)响应用于比较该话音数据的装置，用于向该用户指示该用户的单词发音是否是正确的装置。

60. 一种由视听者在收听其所不理解的音像制品期间在该音像制品中提供听到的一个或者多个单词的含意的方法，包括步骤：

1)在这种音像制品中展现多个场面，在每个场面中存在具有一个头部和嘴部的发声者；

2)在收听期间的全部或部分中使得一个或者多个单词符号出现在所说的发声者的嘴部的近处，收听期间的一个或多个单词表达该听到的一个或者多个词汇的含意。

61. 权利要求60的方法，具有在一个对话平面中放置一个或者多个单词符号的附加步骤。

62. 权利要求60的方法，其中发音系列是具有含意的句子形式中的一系列单词。

63. 权利要求60的方法，具有进一步的步骤：使该发声者具有耳朵，并且该单词符号看上去在实际上垂直于穿过该发声者的耳朵的一条直线的一个对话平面中。

64. 权利要求60的方法，具有附加的步骤：以一个或者多个单词符号叠加在非教学景物上而在该图象景物中展现非教学景物。

65. 权利要求60的方法，其中该视听制品在一个平板型屏幕平面上观看并且具有把这种对话平面相对于这种平板型屏幕平面至少六十(60)度定位的步骤。

66. 权利要求60的方法，其中至少一个单词符号是一个标记。

67. 权利要求60的方法，其中至少一个单词符号是一个图符。

68. 权利要求60的方法，具有以一线性方式展现该图象景物而没有中断的步骤。

69. 权利要求60的方法，具有通过使用可由视听者操作的计算机以中断的非线性方式展现该图象景物的步骤。

70. 权利要求60的方法，具有扩展该音像制品的一部分，以便延长被看到第一区域的单词符号时间长度的步骤。

71. 权利要求60的方法，其中一对话平面通过一个发声者的头部定位，其中至少一个单词符号出现在这种对话平面中或接近这种对话平面。

72. 权利要求60的方法，其中一个发声者的头部包含面孔和嘴唇，而其中与单词发音相关的面部的、嘴唇的和头部的运动被包含在

这种画面中。

73. 权利要求60的方法，其中显现差异的一个或多个单词符号出现一个所选择的时间长度，在该期间，该单词符号随着画面变化由在该图象景物中的一个目标遮掩。

74. 权利要求60的方法，其中该单词符号被整形和呈一定角度，以便显现为三维的单词符号。

75. 权利要求 60 的方法，其中发音系列是具有含意的句子形式中的一系列单词。

说　明　书

用于制作具有一系列可视字符与相配合的语音的一个音像制品的方法与系统及其这种音像制品

本申请是1994年9月22日提交的序列号为08/310,458、标题是"具有配合语音的一系列可视字符的音像制品"的未决美国申请的部分继续申请；该申请是1994年2月24日提交的序列号为08/209,168的、现在被放弃的美国申请的部分继续申请，而后者又是1993年9月23日提交的序列号为08/126,341的、现在被放弃的美国申请的部分继续申请。

发明背景

早先的视听播放就已经包含配置说明字幕(美国专利3,199,115及5,097,349或气泡形图注(美国专利1,240,774及2,524,276)，为的是帮助对播放的口述内容作语言的文字解释。

虽然先前的说明字幕时常是与一种不同语言中的单词的说出同时出现，但是这种出现是偶然、少见的，并且没有一个实现把一系列发声与一系列文字做特定关联的控制模式。而且，说明字幕单词的定位远离图象表演动作。

已有技术的、每一个显示一个单词的闪现卡片试图通过反复强调及不自然的练习来教阅读。虽然最终具有一些效果，但是这种卡片的使用需较长的学习周期，并且需要母亲或教员之类的文字家教的面授。而且这种卡片不提供本发明通过在叙述性的视听媒质内提供涉及的对象而具有的关联能力，而这种视听媒质在读写能力指教功能外，还对于学生有吸引力。

美国专利5,241,671揭示了在一个计算机屏幕以某些字加下划线而某些字不加下划线的方式显示任何文章的文本。当用户从该文本选择一个单词时，其解说在一个窗口上出现，并且出现该词汇的发音。另外还可以显示一个声音的图符。

封闭字幕制品提供独立的区域或相邻的框，其中显示分组的单词。封闭字幕系统在屏幕的底部或其它远距说话人或演员的位置上显示成组的单词。封闭字幕中的单词以不同的背景出现在视频画面场景的两侧、下面或上面，该背景通常是白色的。为了帮助聋哑人而具有

音象的手势语言符号也显示在单独的邻近方框中。这些方框显示技术可能对观众构成一种妨碍。

已经播放的教学用视听节目包括有一位面对摄象机镜头进行讲解的教师，进行讲解的同时在讲解者的前面显示出对应的书面文字。视听者厌倦这种形式，尤其是年轻的视听者会因此而失去对所讨论的主题内容的兴趣。

发明概要

简单地说，本发明包括一种视听制品及其制作方法，其中利用通常所见制品的正常设置，即伴随言语或其它口语发音的图象帧或段的一般常见序列，另外叠加上与视听者顺序听到的发音序列相关或相应的文字序列。我们称其为“改进的”（euthetic）（恰当定位的）字幕。上下文中，如果语音词汇和书面词汇是同一个单词，则二者是相对应的。一种语言的语音词汇与另一种语言中具有相同含义的书面词汇是该上下文中的相关的词汇。

根据本发明的某些实施例，其特点是每一文字书写出现在说话人的头部附近、上面或相关的地方，这样视／听者可以同时观察到书面词汇、语音词汇以及相伴的面部、唇部和头部运动表情，通过靠近和正对嘴部的方式而产主一种其单词是从嘴中发出的印象。根据另外一些实施例，每一书面文字都出现在使用手语的人的手的附近、上面或相关的地方，根据本发明的另外一些实施例，盲文被“显示”在与说话人的语音词汇相关的某一单独装置上。

本发明与不需指导的通常为娱乐、信息或其它非文化教育意图的目的而制作的视听制品一起使用。它可以作为新产品的组件或者先前制作的视听制品的改进与这些资料一起使用。本发明还可以用于新制作的文化教育意图的为应用本发明而设计的资料。这种新制作的文化教育意图的实施本发明的资料对学生来说要比现有的用于文化教育目的的视听资料少一些厌烦和胁迫感。

本发明的特征是：视听制品可以线性播放，也可以通过编程以及使用多媒体计算机平台集成制作成一种可以由视／听者提供附加指令而进行交互式操作的制品。

本发明方法的另一特征是：在社会工作中所具有的一定的实用性，因其中有大量的工作需要使用某一语言中的基本词汇如本文中所

述在视听制品上编配文字，并向某一群体反复广播或展示该制品，教会其中的一部分人在阅读时能够认识这些词汇。

最后，本发明包含对于接近文字声源或视在声源的文字组的选择使用，其文字可以是或可以不是被有选择地加亮，或以别的方式使文字之间有差异，以便提供听到的具有相同含意的一个或多个单词的关联性，或使得看到的和听到的一个或多个单词被关联。

附图简介

图1是先有技术；

图2是某一讲演者的正面视图序列，其中书面文字出现在讲演者嘴边的不同平面上；

图3是某一讲演者的正面视图序列，其中书面文字出现在讲演者嘴边的同一平面上；

图4是一流程图，其中示意本发明中所使用的步骤和设备部件；

图5是另一流程图，其中表示本发明的一个交互式制品的制作过程，该制品中包括同时的声象的发音/文字；

图6是一流程图，其中示意使用在本发明中的另一些步骤和设备部件；

图7是一流程图，其中示意扩展视听制品中的音频部分以助于协调声音和文字的方法；

图8是一嘴部附近有一单词的讲演者的前正面视图；

图9是图8的一个部分示意性平面视图，其中画出了对话平面；

图10是一电视机屏幕的透视图，上面有处在不改进的置的一个讲演者；

图11是另一讲演者的另一透视图；

图12a-b是依据本发明实施改进的字幕的一种方法的流程图；

图13a-b是依据本发明实施改进的字幕的另一系统和方法的流程图；

图14是依据本发明实施改进的字幕的另一系统和方法的一个流程图；

图15a-b是根据本发明的一个方面的波形延展图；

图16是依据本发明实施改进的字幕的另一系统和方法的一个流程图；

图17是图16中所述系统和方法关于计算机工作站的更详细一些的流程图；

图18a-b是关于图17中所述计算机工作站的更详细一些的流程图；

图19a-b是应用改进的字幕的表示；

图20a-b是依据本发明应用改进的字幕所得到的四象限方位的表示；

图21a-b是描述依据本发明改进的字幕直观应用的流程图；

图22是依据本发明的一个多媒体平台的示意图；

图23是依据本发明的一种交互能力的流程图；

图24是图23中所描述的交互式单词发音的流程图；

图25是依据本发明的一个阻视区（blockout zone）的示意性表示；

图26是本发明的一个使用手势语言的实施例的示意性表示。

图27是具有接近嘴部的单词"WOW"的一个人的透视画；

图28是类似于图27的具有两个接近嘴部的单词的图示；

图29是类似于图27的具有三个单词的一个图示；

图30是一个类似于图27的具有一个单词显示得比其它两个高的三单词的图示；

图31是一个类似于图27的显示具有一个较高单词相对于其它两个所显示单词不同定位的三单词的图示；

图32是在具有单词的四个(4)不同对话平面中的一个人头的透视画；

图33是类似于图32的图，其中一组单词出现在每一对话平面中，所选择的文字被加重；

图34是类似于图32的图，其中该单词出现在平行于平板型屏幕平面的几个平面中，所选择的文字被加重；

图35是类似于图33的图，其中该单词出现在平行于平板型屏幕平面的一个平面中，所选择的文字被加重；

图36是类似于图32的图示，没有单词加重，其中通过最初的出现使单词有差异；

图37是类似于图36的图示，单词出现在平坦的平面中；

图38是类似于图33的没有单词被加重的一个图示； 和

图39是类似于图35的没有单词被加重的一个图示。

较佳实施例的描述

图1表示已有技术的一个屏幕1，其上有一个典型的视听画面2（阴影区），画面2中有一已有技术的封闭字幕框3，里面写有"in the house"几个单词；画面中还有已有技术的一个手语框4以及位于画面下部的其中写有"at eleven"几个单词的延长单词运行区5。其中区域5中单词沿箭头A的方向移动。包括与画面2相关的对话在内的声音多半从声源区6中发出。

依据本发明的单词或其它符号通常安置在声源区6中视听的画面部分；不过，单词也可以叠加在画面2的其它部分上，其中视听者的注意力由其对该视听的兴趣来引导，如哪里有动作、该动作的位置与声源是否一致等。

在一优选实施例中，本发明把单词放置在视听中的讲演者的参考画面（即不与观看屏幕的平面平行的平面）中。因为观察者的参考画面是所视的屏幕的平面；单词从该平面中移动到演员所处的环境的过程就更容易和直观地观察到，对观察者来说，看起来就象是三维观测对象一样。

参照图2-3，视听制品中的讲演者S具有头部H和嘴M，书面文字"look"在其被说出的同时出现在平面P_1里。平面P_1近似地和连接说话人两只耳朵的直线（图中未画出）垂直。每个单词最好出现在其被说出或被发声的那一短暂的时间段里；但是，该单词还可以刚好出现在其被说出的之前或之后[只要它的出现与前面或随后说出的单词或声音不发生干扰即可]。判断标准是当每个单词被说出时，有一相关的相应书面单词或文字提供给视听者。本发明在为演播提供一有意义的语音单词（或其它声音）序列的同时，还提供一相应的书面单词、数字或其它文字的序列，以此在视听者的头脑中一次一个地完成语音词汇或声音与文字的关联。

演播中把许多书写单词或文字一次一个地与相应的词汇语音对应，其目的在于为视听者提供一个在自然的情景中相应的声音和情景关联的机会，以便能够记住这种声音和情景之间的关联，这样，在将来再次演播该声音或情景中的任一个时，视听者都应能记起对应的另

外一个。当把这种目的用于文化、外语学习和教育时，本发明还有另外一个优点，即增加理解过程的清晰性，其中视听者可以从视觉、听觉的一方面或两方面的刺激来接收和理解该单词，这依赖于该视听者的能力、注意力或相对于显示该视听制品的单元的位置。第二个优点是外语声道的翻译，通过在嘴边或附近的位置插入书面译文来加强理解。第三个优点是可以通过同时显示两串字符，提供双语言的书面演播，其中一串用说话者的语言，另一串用另外一种不同的语言，二者都与相应的发音同时出现。当语音序列播放的速度太快以至于无法在视觉上为其编配相应的文字时，视／听介质中受其影响的那部分可以对声音进行数字化扩展，而在视觉上通过数字或模拟手段进行扩展，以获得可理解的关联。

二字词、三字词或四字词（二个、三个或者甚至是四个单词的序列）可以同时显示，其目的是在单个单词显示的速度过快时，使聋子或者非说话人能够理解视听制品声道上的语言。在两种况情中的意图都是在提供字幕时应使用观察者已经理解的语言。在这种情况下，应当使用尽可能少的单词，但是对有关关注的词组必需保证有足够的阅读理解时间的窗口。对于其目的是局限于对叙述性或娱乐性节目的理解而不是与声音的关联的封闭字幕或外文电影字幕来说，上述方法是它们的一种替方案。

根据本发明的另外一个实施例，一个或多个单词定位在与使用手语的人的手的位置相关的地方，以便在这些单词和某个手语元素之间能有一关联。通过这种方法，可以为视听者提供一个在自然的情景中关联相应的单词的机会，其目的是要记住该单词和相应手语语言元素之间存在的关联。

根据本发明的另外一个实施例，视听制品中的单词设置的方式可以使得只对于使用特殊阅读设备的观察者来说才是可见的。这有些类似于三维立体电影，只有当观众戴上一种特殊类型的眼镜时才能看见三维立体显示。实际上使用一种特殊类型的眼镜正是实现该实施例的一种较佳方法。

本发明中的单词作为制品的必备部分叠加在制品的画面情景上一起显示。画面场景包括人物、家具、天空、背景市景等部分。单词可以叠加在一个或多个画面部分上，这样可以防止对画面元素的某一

部分的观察或防止对画面元素的某一部分的单独观察，其中的书面文字是透明或半透明的，或者由线框字母组成。

在说多个单词的同时，由于是一次向视听者显示多个书面单词，所以使得这种显示难于实现，不是不可能，就是很难将声音与其相应的书面单词正确地关联。因此，重要的是，在利用每一声音与其相应的书面单词时，采用的方式应尽可能使得视听者能够更容易地关联这两个元素。为避免分神和混淆，除了单词的附加象形图、手语表示和外语译文外，每个语音词汇只应与其唯一的一个书面相关文字相伴。只要这些显示不在说出前面或后面的单词期间产生，该单词或这些单词可以在其被说出之前、期间或之后显示。

只要每个单词和符号都要与其相应的发音关联，就可以在发音期间出现多个单词或符号。例如，在说“thank, you”时，单词“thank you”和“merci”可以同时出现。

在视听播放期间，通常存在其间单词发声说话的时间期，和在这种发声说话的时间期之间没有发声的时间段。在本发明的一次一词过程中，书面文字只在下述时间段里出现：（1）说出前面单词后的没有说话声的时间段；（2）说当前单词时；（3）接在说出当前单词后，说出下一单词前的没有话声的时间段。

通过在画面场景中显示基于字母的单词，在不识字的学生的头脑中建立起这些需要记住的文字的标志图符，这需要调动大脑记忆整个视觉图象的那部分机制，正如胶卷接收光线生成相片一样。同步伴随语音或图象对象的字母文字的具有创造性的分段显示能够在学生的头脑中建立起书面单词（虽然“正常”书写，但作为标志图符来理解）与同时显示的一个或多个对象之间的可恢复的关联。经过多次重复后，在显示该基于字母的单词的随后显示（标志图符）时，将会在学生的头脑中回忆起相应的对象，即发音单词，这显然定义了一个阅读过程，而对其能力的训练则是本发明的目的之一。

配对语音和书面单词的同一过程还可以反过来训练认识某一给定语言的学生学会说该语言。在这种情况下，对象就是书面单词或标志图符，而学习目标就是发音单词。

尤其是对于婴儿来说，学习强度的关键在于要在一种“自然”的环境中提供知识的关联，该环境与小孩学说话的环境相似。在正常的

环境中的小孩学说话一般不包括正式的演讲课程，与本发明音视提供的环境是同一类型的环境。在本发明较佳的直观实施例中，向视听者提供某一故事的环境或者其它演示环境，其主要目的并不在于传授文化。当一个人在学说话时，容易受到直观的图象和动作的影响，学习者所在环境中的直观的图象和动作由代理人（例如父母）分别演示或暗示，并且充当将与并行出现的语音相关联的对象。本发明的环境中直观的图象和动作在学习者（例如儿童）的环境中由代理人（如父母）分别演示或暗示，并且充当将与并行出现的发音相关联的对象。由于是在学习者的环境中的代理人，所以这种环境中包括有意义的语音序列的编排，而不是制造随机发音的规则。这种自然的语言学习环境出现在通常的电影中，其中描绘了自然的交流情景，在电影播放期间通过自然的重现向观众重复展现同一个单词。自然环境和电影仿真都提供了动作和对象之间的关联以及与其相应的解说语音，本发明在自然环境的视听仿真中将这种关联机会扩展到了书面单词上。

本方法在训练阅读时，向学生演示的是整个单词而不是单词的一个个音节或字母。观察和记忆一个完整的单词类似于观察和记忆一个符号，如一张猫的画面或一个汉语方块字等，其中的单词据信是由人的大脑以相同的方式进行处理。作为一个整体来观察每个单词（或情景阅读）提供了基于在大脑中开发关联的一种训练方法，这种关联是被直观地记住或铭记并通过关联而非人脑的分析（基于字母、音节或语音的阅读所要求的）来回忆的。

相应于同一语音词汇显示有两种书写形式，即一个为文字形式，另一个为图示形式，这两种形式可以融合为一体或由其中之一变成另外一个，以示二者是相关的或甚至是同一个。例如，当一个人在电视中说出单词“cat”时，书面单词“cat”可以演变成一张猫的画片。

由本发明所建立的关联不管是在已有视听的上下文还是将要创建的视听的上下文中，都以一种常规的、自然的画面情景出现。例如，这种关联可以出现在电影场景中，其中一个侦察员正在与一个嫌疑犯进行交谈；也可以出现在电视中，其中一个演员正在演唱；还可以出现在电视新闻广播中，其中一个气象预报员正一边解说一边在地图上指指点点。在所有上述这些事例中，其目的都可不必包括文化目的。

本发明还可以用于训练唇语阅读，其中当发音和文字显示出现时

还可同时观察到作为制品的视频部分的一部分的嘴唇的运动。

将单词放在位于所显示画面的区域中的上下文的发源（contextual source）之地或其附近的一个好处是，在观察者听到某个单词时能够更容易地看到该单词，而与此同时还能保持对制品中所发生的动作的注意力。虽然人眼可以看到位于观察屏幕的外围边缘处甚至是画面区域之外的单词，但在阅读这些单词时，同时保持对视听制品中的动作元素的注意还是比较困难的。而这一点正是本发明比封闭字幕要优越的原因之一。再者，封闭字幕一次要同时显示多个单词，这妨碍了单个单词对单个声音的关联。另外，本发明按与说话人的空间关系显示单词，这减轻了胁迫感并减小了书面单词所占用的屏幕区域。

当两个人进行交谈时，不管他们是否是面对面，二人之间的单一平面可以作为用来显示书面文字的平面。当其中的一个人不在摄象机镜头里时，该技术同样适用，其中观众能意识到镜头外的那个讲话人的相关位置。

每个书面单词的字母的颜色、形状和其它特性在设计时应注意到不要引人注目。例如，如果单词所在的背景画面部分的颜色是深蓝色的，单词的字母的颜色则应是浅蓝色的或具有其它亮度的蓝颜色。再者，书面单词还可以以透明或半透明的形式显示，这样可以部分继续看到背景被遮挡的那部分画面。另外，一个单词还可以以颜色、字体或以其它形式编码到其源中。

再次考虑图2-3，当说话人（S）的头（H）转动时，平面P_1，近似地和通过说话人两只耳朵的直线垂直，移动到三个（3）新的位置P_2－P_4。在说单词“AT”时透视地出现在平面P_2中；在说单词“SPOT”时透视地出现在平面P_3中；最后在说单词“GO”时出现在平面P_4中，每个单词在说出时位于说话者头部或其附近甚至上面，而且最好位于说话人的嘴部（M）或其附近，请注意，当说话人（S）的头部（H）转动时，同时也发生了倾斜来抬高下巴（参见平面P_4）。书写方向最好反映头部从一边到另一边以及上和下的方向。

在图3中，所有说出的单词都出现在平面PP上，而平面PP则与播放视听作品的屏幕在同一平面上或与其平行。

图4中描述了制作视听制品的装置，其中包括：一个操作台；一

个用来产生所需文字（如单词“look”）的视频文本发生器；用于提供上面尚无文字的制品的视频制品输入装置；一台数字光学操纵器，用于提供结合文本和视听制品的装置，从而在适当的平面方向上提供本发明的发音／文字关联。这种操纵制作出一种具有创造性的视听制品，其中在制品中从头到尾都有这种关联，并且在观看和听的时候可以不必中断由本发明的直观实施例所制作的演播。

在说出单词的时候，成组的字母贴在、印在、叠加在或者以其它方式位于画面中最容易观察到的部分。当说话人的头部可见时，单词通常应当位于嘴部或其附近，以暗示单词是从该嘴巴中发出的。该序列持续于整个制品（如果需要，制品中的一段）中所有或相当数量的语音。字母可以是任意大小、字体或颜色。在一较佳实施例中，大小、字体、颜色或其它图形属性在选择时要能反映背景颜色以及每一语音的情绪和意向内容。对于背景来说，每一书面单词缺省地都应是透明的、半透明的、线框状的或者取某一亮度的背景颜色，并要足以与背景颜色区分开来，以便在保持可见性的情况下避免在单词消失之后留下视网膜晕影（retinal halo）或鬼影图象。对于情绪、意向或含义，举例来说，恼怒的词汇应当具有红颜色和尖锐边界的字样，而催眠曲的歌词则应具有软色调和柔软、卷曲的字样。情绪上中性的单词应当使用缺省的颜色来显示。图形属性的目的就在于通过音量、语气、音调或其它声音属性的变化为视听者提供一种与语音的细微差别平行的动态图形，从而达到增强将来可以由大脑回忆起的关联的目的。

自然的交流情景流行于视听制品中。这些情景包括上面所提到的一个侦察员质问一个嫌疑犯的情景。在包括自然的交流情景的场景中安放单词提供了一种用于建立声音和所期望的文字之间的关联的工具，而同时视听者仍能保持对制品中自然的交流情景的注意力。

在图5中，本发明的线性实施例可通过编制一种计算机程序用于创建一个交互式的实施例，该程序允许视／听者停止视听播放，于正在播放视听的屏幕上弹出一个菜单，通过选择提供单词的定义、语法和句子的上下文用法或其它信息。交互式的制品由视／听者的操作来播放，该视／听者使用一种可编程的教学装置，该装置用于使用上述程序显示上述制品，停止上述制品而查看所选的文字并获得有关该文

字的附加信息。

在图6中，视听制品由计算机图形设计者在其工作站上创建，工作站的屏幕上向设计者显示的是制品的视频信号（模拟或数字形式）。在对帧（每秒30幅的画面序列）进行处理时，设计者创建一计算机图形或文本（即一个单词）并将其叠加到一帧或多帧的视频信号上，帧数的多少依赖于说出相应单词所需要的时间长短。说一个单词所需要的时间长短，在日常英语（或其它语言）会话中，对于很多单词说出一个单词的时间都不相同，从六十分之一秒到二分之一秒不等，借助于动画技术和使用画盒软件可以将其它字符加到该单词中字母的字体以及在选定的平面上该单词的方向上。

图7示意延长声像制品中说出某个单词的时间的方法，其目的是为了给相关的书面文字提供更长的显示时间。这种对一个单词能够听到的时间延长或延展过程首先在硬盘上将单词的声音数字化为一个波形，然后再对该波形重新进行配置。这种技术不会改变音调和语气。

图8中的头部（H'）面向观察者的右侧，如图中虚的源线（SL）所示。直线（SL）位于说话人的参考对话平面（P_5）（未画出）上。垂直的观察者参考平面（A）对于观察者来说所看到的只是一条直线。该平面保持固定不变。直线（SL）象一根烤肉的叉子串起单词“WOW”。

单词（WOW）的始点与说话人的头部（H'）之间的距离最好在（d_2）之内，而距离（d_2）则是说话人具有鼻子（N）（见图8）的脸部（F）的宽度（d_1）的两倍（如图8所示）。将单词（WOW）置于范围$2d_1$之内为说话人的头部处于比较近的位置的场景提供了一个比较好的结果。当头部处于远景中距离很远时，单词可以比头大，但仍然需要靠近头部（H）或移至观察者感兴趣的对象处，在这种情况下，距离（d_2）可以是距离（d_1）的3或4倍。

图9是图8的一个示意平面视图，其中给出了对话平面（P_5）、平面A（180° 观察者参考平面）和平面B（90° 平面）。具有源线（SL）的对话平面（P_5）中包括出现于该方向的单词“WOW”。出现在其它对话平面（P_6）和（P_7）（离观察者平面（A），即180° 的观察者参考平面25° ）上的单词也能阅读，但是因为放在靠近观察者参考平面（A）（观察者的主视平面）的对话平面上的单词读起来比较困难，

所以本发明的实际应用中很少使用这种放置方法（在该“阻视区”内）。

图10给出了电视机21的屏幕20，其上带有控制按钮22和23。说话人的头／脸的位置跟图2中一样是以多视图的形式显示的。屏幕20左边的视图显示了头（H）、脸（F）和上面带有源线（SL_1）的对话平面（P_1）。脸平面（FP_1）与对话平面（P_1）垂直，源线（SL_1）与脸平面（FP_1）垂直。当说话人站着或坐着时，脸平面通常位于与水平面垂直的平面上。源线（SL_1）直线地等分单词“look”。其它源线（SL_2）、（SL_3）和（SL_4）分别位于它们各自的对话平面（P_2）、（P_3）和（P_4）中，每条线都直线地等分其各自的单词。

最后，图11给出的一个有头（H_3）、脸（F3）和嘴（M）的说话人（S_2）。脸平面（FP_2）与对话平面（图中未画出）垂直。位于对话平面（图中未画出）上的源线（SL_3）等分单词“fast”。因为当说话人（S_2）向后靠或将其背部转向观察者时头部（H_3）可能会沿任何方向移动，源线（SL_3）上的单词在该方向上由头部（H_3）说出时，在各种情况下都应放在对话平面上，除非对话平面的方向不利于观看者对单词的阅读时。例如，假设说话人（S_2）处于站立位置并且背对观察者，单词“fast”如果此时放在对话平面上的话，则将处于观察者无法阅读的一个角度。为避免这种情况，在保证可识别的情况下，单词“fast”应尽可能地放在离对话平面较近的平面上。这种情况下，单词“fast”将位于如此选定的平面中的某一透视方向上，给人一种该单词正远离头部（H_3）的印象。

如果时间允许，单词“fast”可以先在使嘴部（M）的某一部分模糊的位置上出现，然后迅速地沿对话平面的直线（SL_5）移动。也可以是，例如，如果单词将在屏幕上总共停留0.024秒，则该单词可在嘴部（M）出现0.008秒，然后沿直线（SL_3）移动0.008秒，最后在消失以前再在该直线上停留0.008秒。

将单词放在源线（SL）和透视平面上的目的在于造成单词是从嘴部中说出来的感觉，从而使得单词能以视听场景的三维空间中的物体的形式出现。作为一个物体，该书面单词同样要受到其它物体所受到的物理定律的约束。

因此，在使用本发明的视听制品中，如果有人走到说话人的前

面，则说话人的声音有可能被遮住，将暂时看不到书面单词的一部分或全部。本发明这方面的目的是为了让单词看起来象是真的物体，这是一种年轻人普遍接受的概念，对于那些要求单词对用户友好而不是抽象的人来说，尤其如此。

单词的位置看起来应具有透视感，其中的字母大小应递增或递减（参见图8，其中左边的字母“W”比字母“0”要小，而字母“0”又比它右边的字母“W”要小）。具有透视感的单词看起来要有一定的方向，包括在该方向上的移动的状态。具有透视感靠近说话人嘴部的单词要看起来象是从该嘴中说出的。在不影响作为交流的一部分的说话人的面部表情的情况下，单词要尽可能地位于靠近嘴部的地方。

在一制品当中并不是所有的语音词汇都需要显示与其相应的书面单词，因为制品中所选定的某一使用阶段中在文字方面可能提供特殊的难度或由于其它原因不需要文字。

本发明的最佳用法是将单词、其它字母数字符号或其它文字按照一有意义的顺序放在磁带、胶片、计算机磁盘、光盘或其它介质上，按以上所述方法提供了它们与相应的磁带、胶片、光盘或计算机磁盘的语音部分的关联。这种顺序可能持续贯穿胶片或磁带的首尾。本发明的视听制品最好具有一定的娱乐性，或者存在具有上下文关系的有意义的主题和内容。视／听者在其进行娱乐或其它内容时，不需要作出特别的努力即可完成学习过程。

本发明在制品的画面区中建立语音词汇的视觉感，好象每个单词都可见，并且具有维度、颜色、字体、运动和其它属性。单词的维度就是单词在显示屏幕平面或与之有一定角度的平面上的方向。在该方向上的单词如画面中的其它元素一样是三维的。

文字可以包括字母、单词、图象或其它符号。

根据本发明的另外一个实施例，文字是以盲文符号来显示的，并且最好在一台单独的设备上可以由人们（例如有视觉障碍的人）在听声频节目时使用。类似于其它实施例，可以在盲文和语音表示之间建立其一一对应关系。这样便可以为用户提供一个在自然的情景中关联相应的文字的机会，其目的是要记住该文字与相应语音之间存在着关联。

实例

在此举一个娱乐性视频游戏的例子。玩此游戏的人在其中可以利用一定数量的图示（文字图表）对象。各图表可以用光标拖动。每个对象都可以用鼠标单击，该对象随即渐变（“变形”）成其相应的书面单词（标志图符），同时出现一画外音或一正说话的头象说出该单词。

本游戏的目的是要创建一排图示对象，并由它们形成一个有意义的序列。一旦玩游戏的人安排好了这排对象，只要双击鼠标即可激发各对象从左到右、一次一个地变形成其相应的书面单词（标志图符），同时伴有相应的语音对象。然后该有意义的序列由说话人从左到右大声地重复，其间每个单词突然由一对象“伴音”。

在玩该游戏过程中，如果由拖动所创建的对象没能形成一个有意义的序列，双击鼠标时将没有输出，并且不能得分。

颜色和字体图形的差别可以随对话的有意义的表述的自然进展而变化。这样，本发明的整体“有机外观”将创建出一种新型的标准化的“情感图形词汇表”。例如，下述颜色和图形可以用于以下情感：

情感	颜色	图形
高兴	白色或粉红色	闪烁／闪耀
悲伤	蓝色或黑色	哥特体
恼怒	红色	黑体
性	紫色	波浪形

字体和颜色的差别也可以用来关联物理现实，如在自然界中的发现。

物理性质	颜色	图形
冷	灰色／冰蓝色	冰条
热	橙黄色／红色	火苗
湿	乳白色	水滴

这种关联基于一般感觉或/和已有的关于人的情感和学习记忆的颜色、纹理等的图形效果之间的关联的研究。另外，现有的计算机图形软件的某些功能，包括视觉现象，如“闪光”和“辐射”等，可以进行分层以增强其它关联的效果。

根据本发明，叙述性上下文中的改进的字幕可以通过多种途径来实现。图12a-b示意了手工应用改进的字幕的步骤。

图13a-b描述了一个电视导播系统以及应用改进的字幕的方法。

图14描述了一个在不影响音调或语气、不造成明显失真的情况下放慢语音速度的系统和方法。图15a-b分别是单词“future”的正常和扩展波形，扩展时使用的是图14中所描述的系统和方法。图15b中波形的音调与图15a中的相同，因为在扩展时波形的振幅保持不变。

当使用小数放大器而不是全整数放大器对波形进行扩展时，使用本发明的另外一个实施例比较好。例如，当需要将波形的长度延长一半（增加50％）而不是一倍（增加100％）时，已知的方法只是随机地选择波形的某些部分进行扩展，而根据本发明的特点，对于波形各部分的随机选择仅局限于元音部分。这可以由该领域中普通技术人员公知的装置来完成。

图16描述一个应用改进的字幕的数字系统和方法，其中使用已有的字符动画软件来放置单词。

图17-21涉及本发明的另外一个实施例，是一个关于直观改进的字幕的系统和方法。图17是图16中所描述的系统和方法，其中根据本发明的这一实施例应用了改进的字幕软件。图18是一流程图，是对图16和17中所描述的系统和方法中所使用的计算机工作站的进一步的详细描述。

图19a-d展示根据本发明应用改进的字幕的细节。图20a-b描述某一改进的字幕单词所处的虚拟的三维空间的四个象限。图21是一软件模块的流程图，该软件模块用来在虚拟三维空间的任一象限中为改进的字幕单词选择一个较好的方向。这种软件最好在计算机工作站系统上运行。这里可以使用本领域普通技术人员已知的各种输入设备，而比较好的一种是由用户使用鼠标来指定象限和画出一条源线，然后使用键盘进行输入。运行上述软件模块的计算机系统应自动地将单词放置在源线上，其中源线最好能穿过小写字母主体的中心处（排印工人

所熟知的“x高度”），如字母“e”中的水平横线。系统和软件安放好单词之后，相应的源线即被删除。

图22是一示意图描述一根据本发明制作交互式改进的字幕制品的多媒体平台，该多媒体平台包括一交互式的多媒体计算机工作站。图23是一实现交互能力的软件的流程图。

图24是图23中所描述的交互式单词发音的流程图。根据本发明的这一特征，当用户在某一特定的单词上停下改进的字幕视听制品的播放时，可以就此上一堂发音课。用户面前最好能有一跟某计算机相连的麦克风，该计算机上装有音波分析软件，能将用户对该单词发音的波形与存储在计算机文件里的正确发音的标准波形进行比较。然后计算机向用户提供反馈，或者对正确的发音进行肯定（如“很好”或“棒极了”），或者提示用户把单词再说一遍。

图23中所描述的其它选项最好作成用户可以选择的交互式应用程序的菜单。例如，用户可以选择一书写应用程序，该程序允许用户对所显示的单词进行摹仿，摹仿的手段可以是通过键盘键入该单词，也可以使用电子书写板书写并将输出传给手写体识别软件。该交互式系统最好能提供一反馈通知用户其所键入／书写的单词是否正确。

图25示意如何在图9中所描述的前向“阻视区”中放置单词。作为一种方式，单词可以选择放在该区域中（图25中通过直线AA的平面上），使得单词看起来象是从讲演者的嘴中发出的。

一个或多个单词也可以出现在与使用手语的人的手的位置相关的地方，以便在这些单词和某个手语元素之间能有一关联。图26中描述了本发明的这一实施例的一个例子，其中电视屏幕上有一讲演者和一嵌入的矩形区域，该区域中有一人手正作同步的手语翻译，当讲演者说出单词“Future”时，手语翻译人将对该单词作出相应的比划，同时在嵌入的矩形区域中将显示出与该手语元素相关联的单词“Future”。

转到图27-31，其中示出一系列从说话人的嘴部(M)来的单词。在图27中，单词"WOW"是沿着话音线(SL)在嘴部(M)的近处。话音线(SL)在一个方向上从嘴部(M)伸出，使得单词看上去象刚刚从该嘴部发出的物体。在图28中，第二个单词"HE'S"已经出现在已经移动从而提供用于单词"HE'S"的空间的单词"WOW"的后面。通过加重单词"WOW"而使

得单词"WOW"已经不同于单词"HE'S"，在图29中第三个单词"FAST"已经出现并且通过粗体字母的加重而有差异。在图30中，单词"WOW"现在被加重，向上移动，但仍在头的附近。相对于其它显示的单词，图31示出单词"WOW"远距离定位，并且在头的右上方。图27- 31的所有的单词都在头的近处并且图27- 30的单词是在嘴的近处。

本发明的进一步的实施例在图32中示出，其中按照播放的顺序每次相继出现一个新词。随着单词积聚靠近嘴部与在其它单词上面的某些单词聚集，所有的单词都被显示在对话平面P_1- P_4中的说话人的嘴部邻近，这些平面随着说话人的头的移动而具有不同方向。单词的聚集定位这些单词以便占据某些区域，其具有的高和宽的尺度不同于一串单词的一个区域的尺度，该一串单词的区域尺度将占据宽度实际上大于高度的延长矩形区域。包括使用两条或多条线或单词行的集束的目的是把所有的词汇靠近该说出单词的源点或视在源点定位。

回到图32，显示在平面P_1中的单词首先出现，跟随的是在平面P_2-P_4中的显示。在平面P_2中，单词"LOOK AT"就在说话者嘴唇上面成一排。在随后的平面P_2中，在说话者转动她的头之后，单词"LOOK AT"类似地出现成一排，以"LOOK"靠近嘴唇而"AT"远离嘴唇。单词"SPOT"以靠下的一排粗体字母出现在平面P_3中。平面P_2-P_4是对话平面，其中每一单词放置并有别于其它单词，并且在该有差别词汇由视听者聆听期间，图形目标由粗印刷体字表示一个时间期间。

通过使用把一个单词与其它单词相差别并且同时由视听者听到所差别的单词，而在该视听者关注的该差别单词和听到的单词之间建立关联。差别可以通过能使得视听者把单词与其它单词或显示目标区区别开的任何方式实现。在本实施例中，所有的显示的单词都是在说话者嘴部近处和头的附近。用于形成单词差别的技术实例是使得单词或单词的任何部分放大到一个特定的尺寸，或具有特定的描影、颜色、形状或发光，或使得该单词显得在移动或真正地移动，或以别的方式创建一种视觉效果，这效果引起对出现在其它词汇中的某些词汇的注意从而形成这种单词与其它单词的差异。

在图33的另一个实施例中，针对所选择的时间期间，单词的整个组或整个集束被示出在每一对话平面P_1-P_4中，并且在这种时间期间的一个部分期间，每一被以粗体字母加重而有差别的单词被说出。该

差别和说出或发声同时出现。在本实施例中，所有的显示的单词都是在说话者嘴部近处。图34和35示出类似于图32和33的实施例，只是单词出现在平行于该观看屏幕的平面(PP)中。

在图36的实施例中，引入单词差别技术。说单词的同时，该单词首先引入或出现在屏幕上。当单词"LOOK"首先出现在屏幕上时，视听者因此听到单词"LOOK"以便创建期望的关联。在单词"LOOK"仍然在屏幕上的同时，单词"AT"出现并且听到单词"AT"。在第三序列中，视听者看到"LOOK AT"接近说话者的嘴，并且随着视听者听到单词"SPOT"，单词"SPOT"首先被看到。图37示出图36的技术，只是单词出现在平坦的平面中。

图27- 37的实施例的一个或者多个单词最好出现在说话者的嘴近处以及出现在接近头的画面部分。而且对于该视听者，该单词最好以关于术语、成语、或词组的次序出现。在视听者看该接近头部的词汇期间，单词以一种有序的方式被发声和被听到。

本发明的最后实施例使用在屏幕上的单词的组显示技术(见图38-39)。本实施例具有两个特定的目的。 一个目的是提供给视听者对于该视听者来说是不了解的一种正在说的语言的大意。第二目的是提供由于视听者听觉或其它损伤或由于听不见这种声音而造成其不理解所说单词的大意。所集束的单词可以放置于对话平面或平行于第一屏幕平面的平面中。该实施例中没有单词需要加重。

说 明 书 附 图

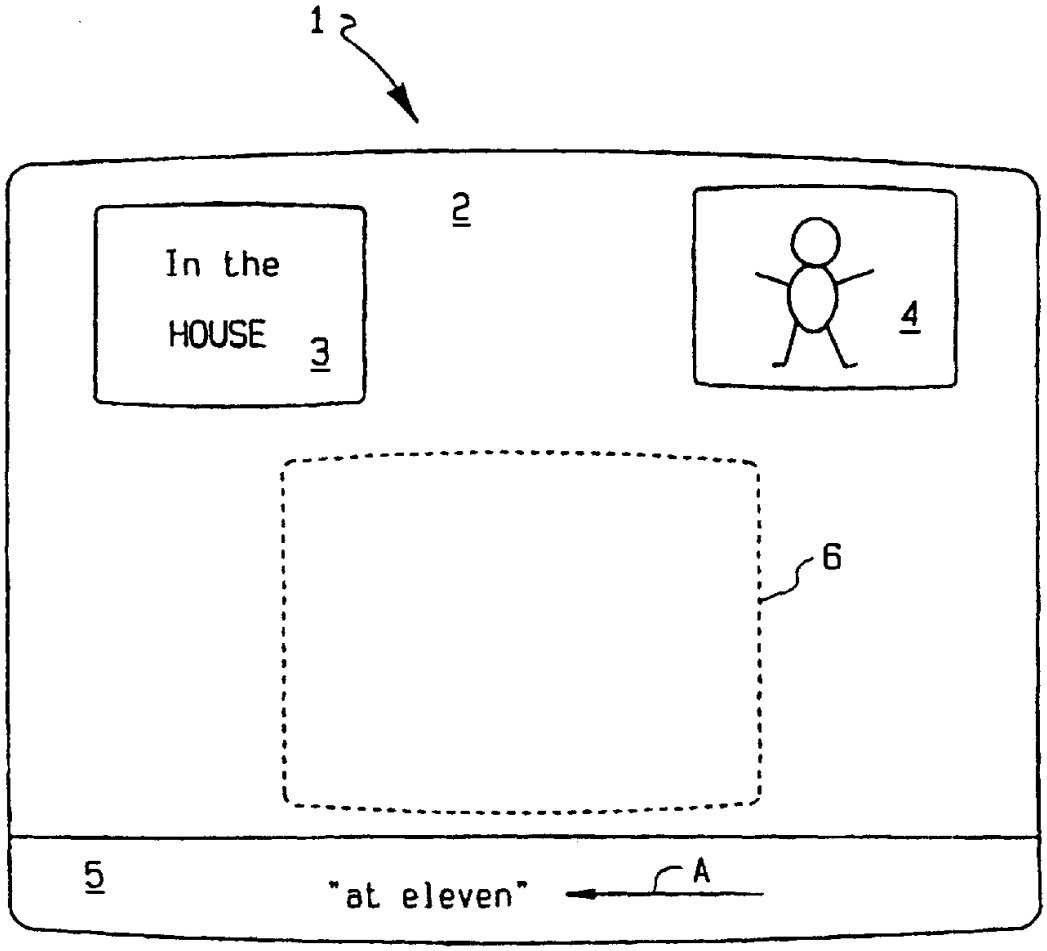


图 1

现有技术

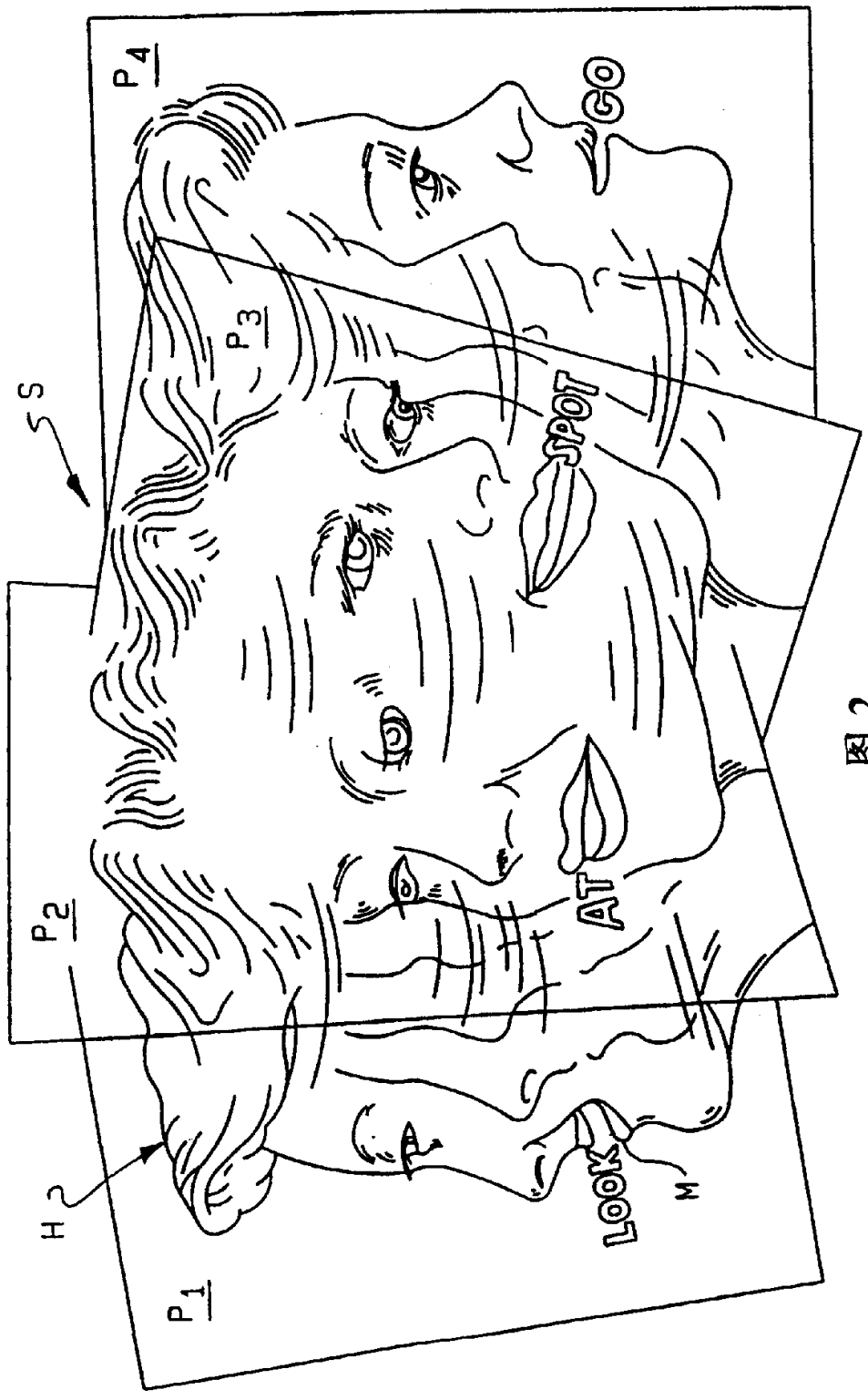


图 2

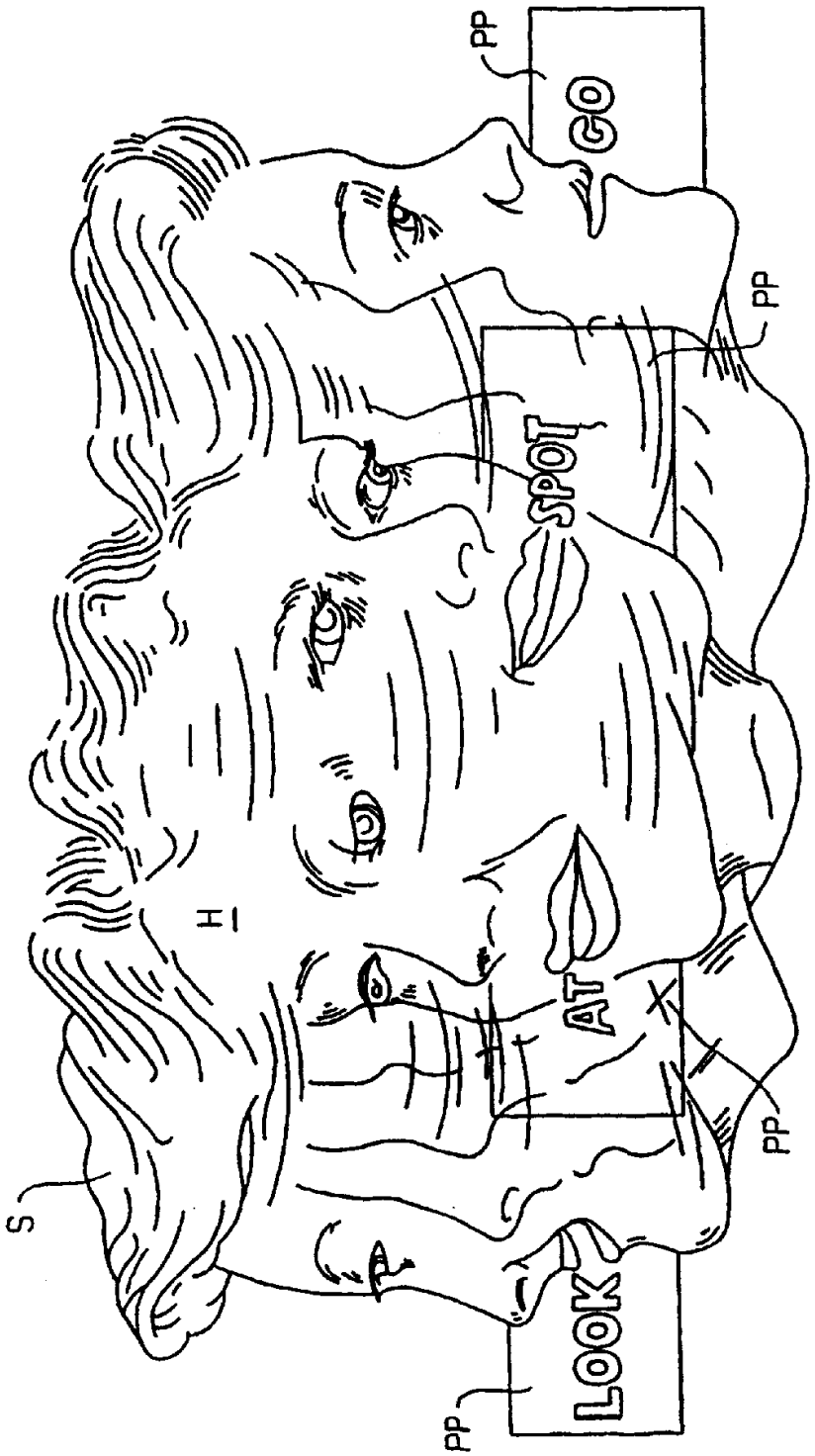


图 3

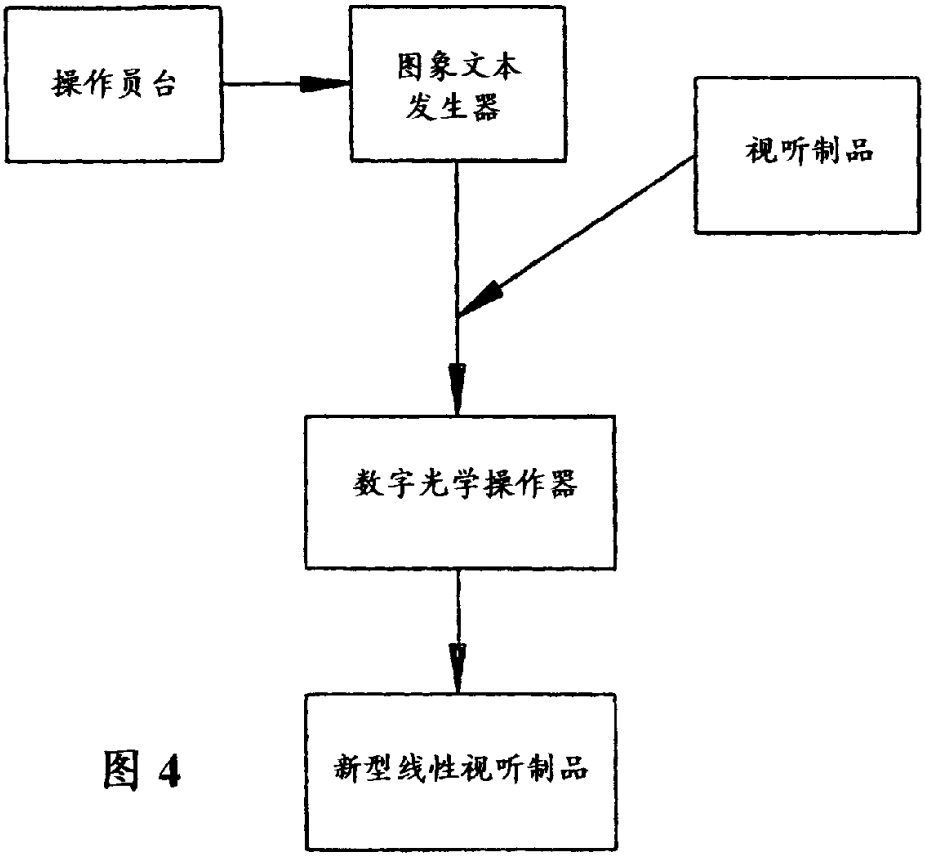

操作员台
图象文本
发生器
视听制品
数字光学操作器
新型线性视听制品

图 4

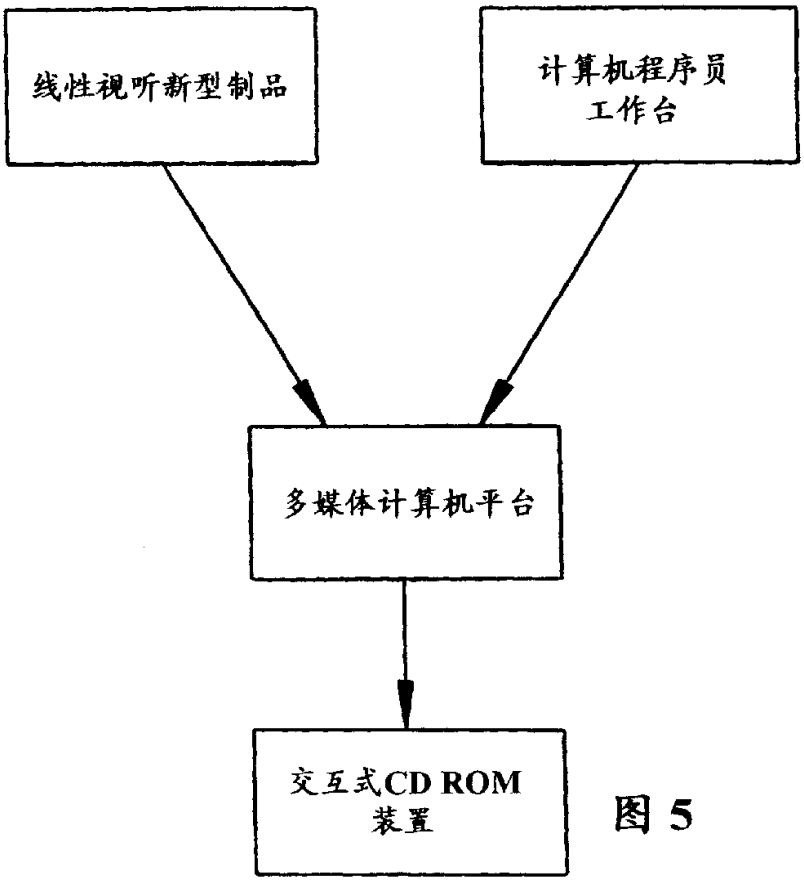

线性视听新型制品
计算机程序员
工作台
多媒体计算机平台
交互式CD ROM
装置

图 5

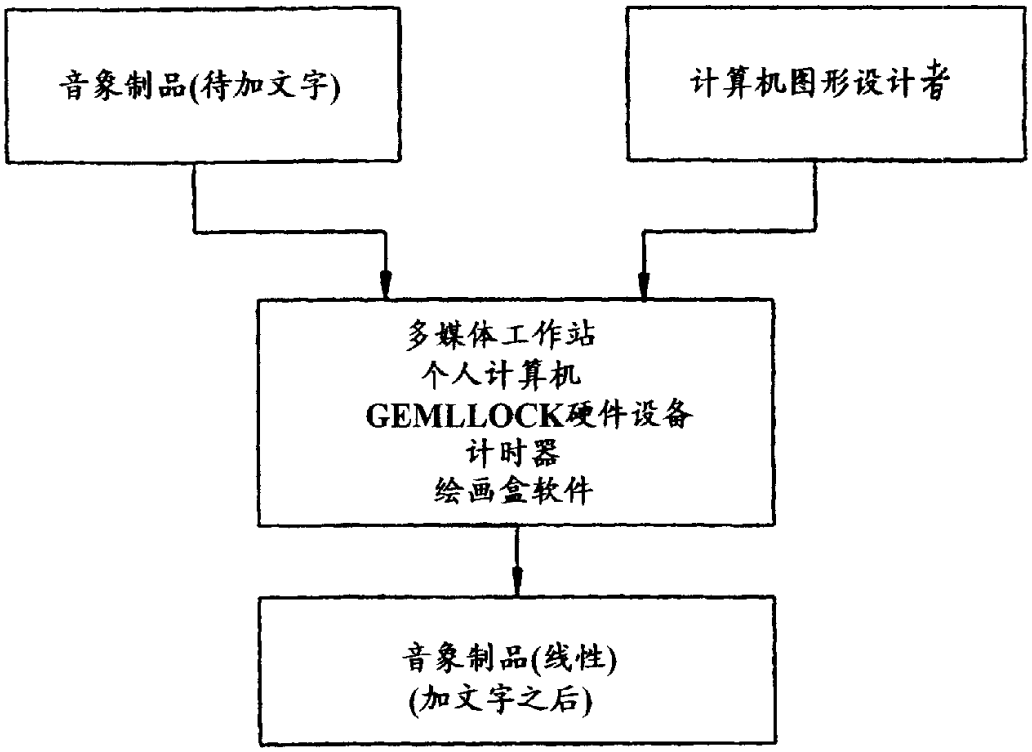

音象制品(待加文字)
计算机图形设计者
多媒体工作站
个人计算机
GEMLLOCK硬件设备
计时器
绘画盒软件
音象制品(线性)
(加文字之后)

图 6

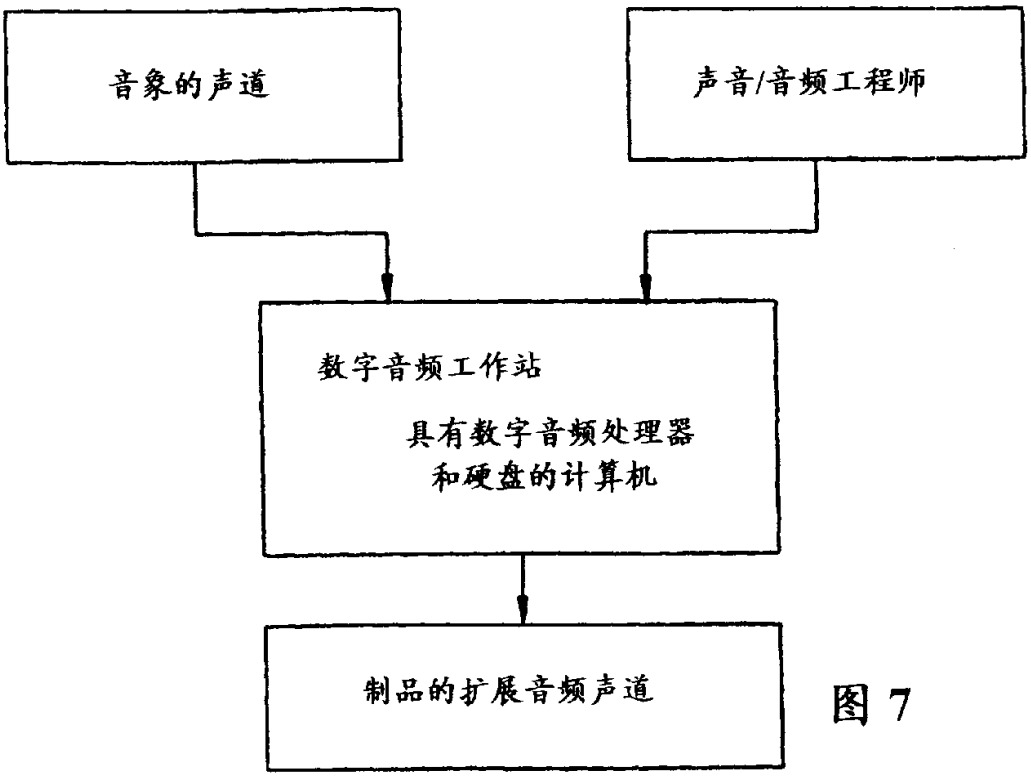

音象的声道
声音/音频工程师
数字音频工作站
具有数字音频处理器
和硬盘的计算机
制品的扩展音频声道

图 7

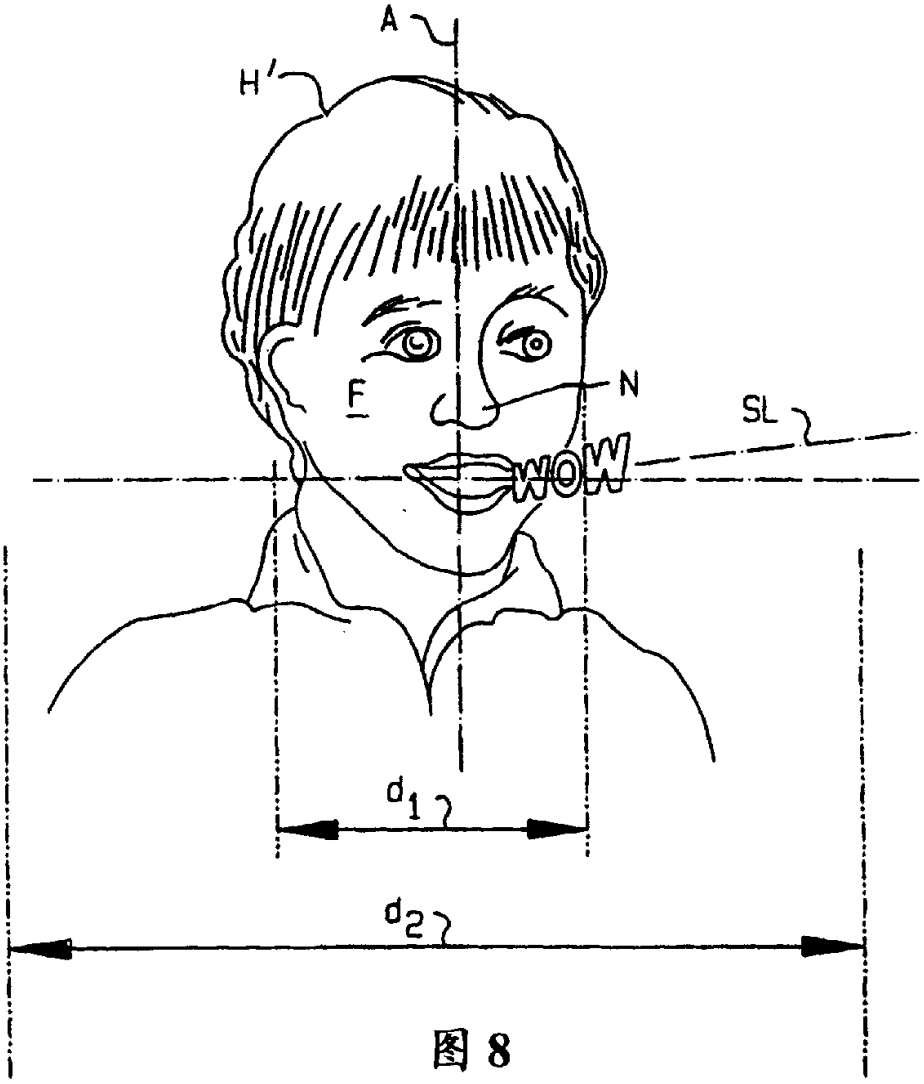


图 8

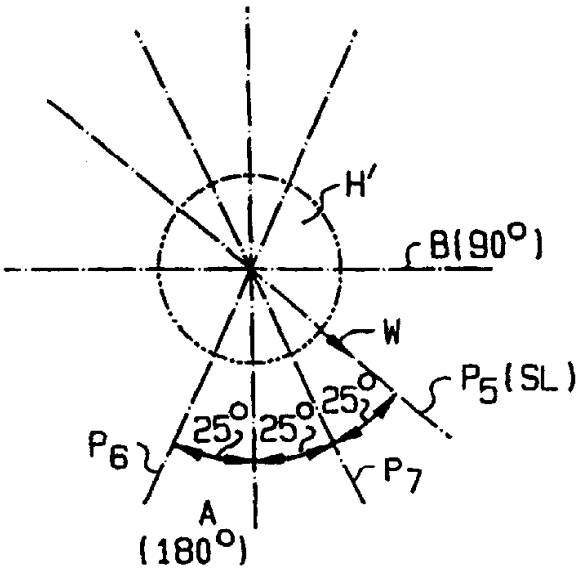


图 9

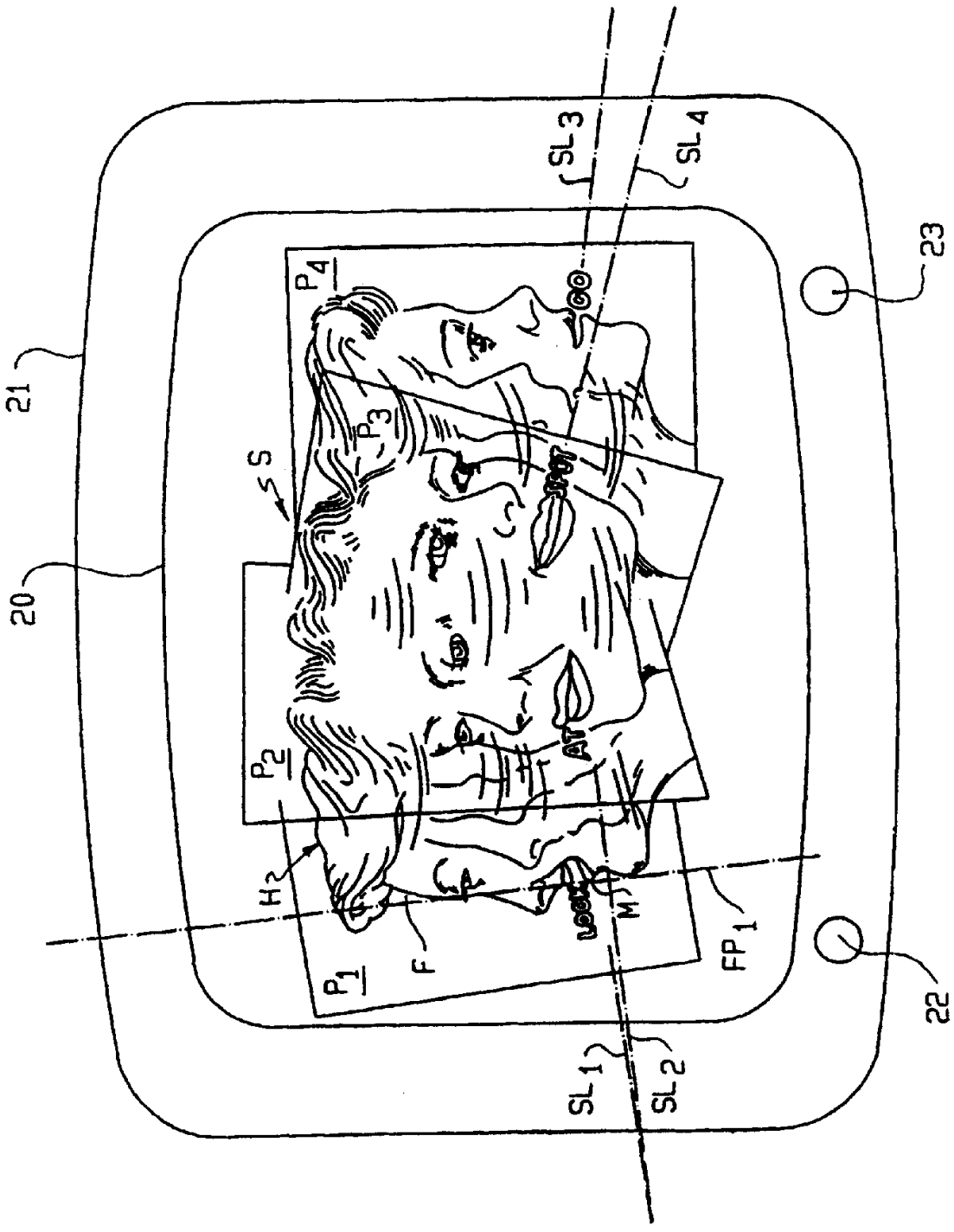


图 10

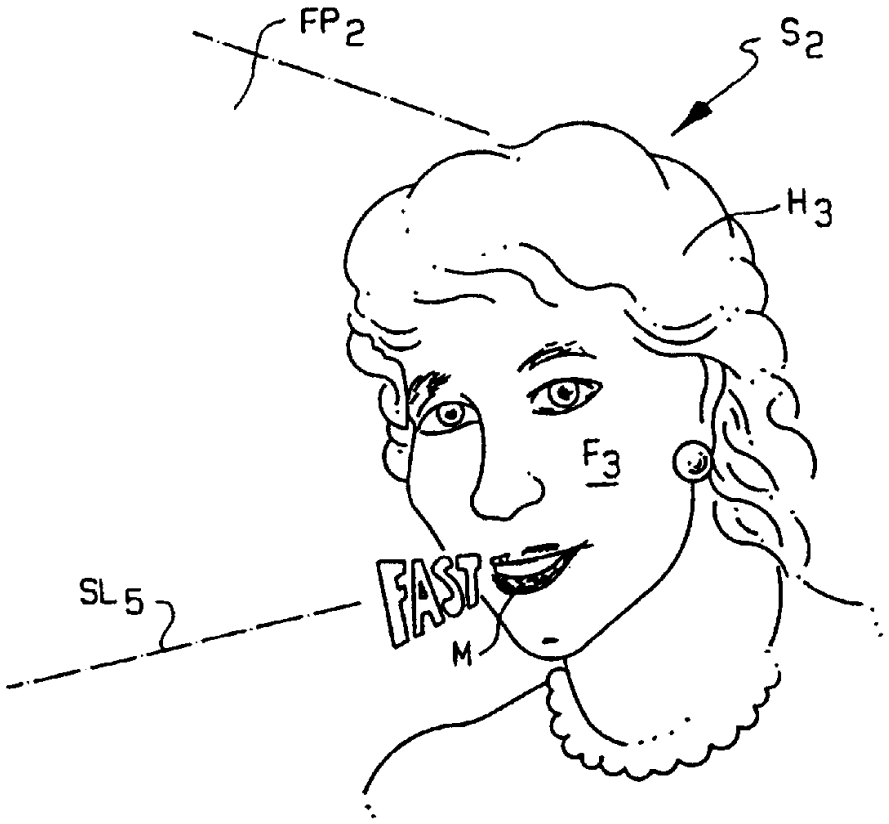


图 11

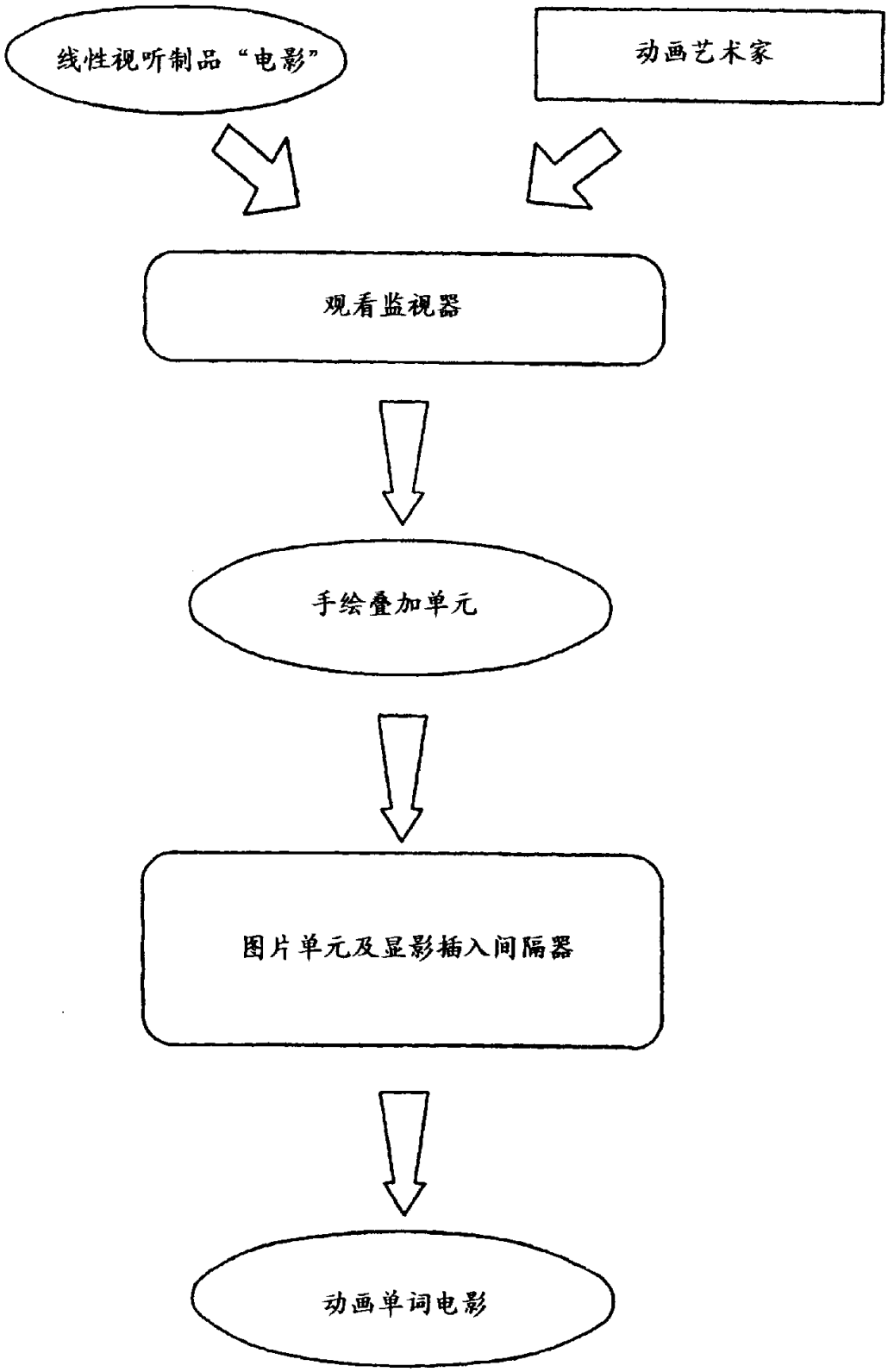


图 12a

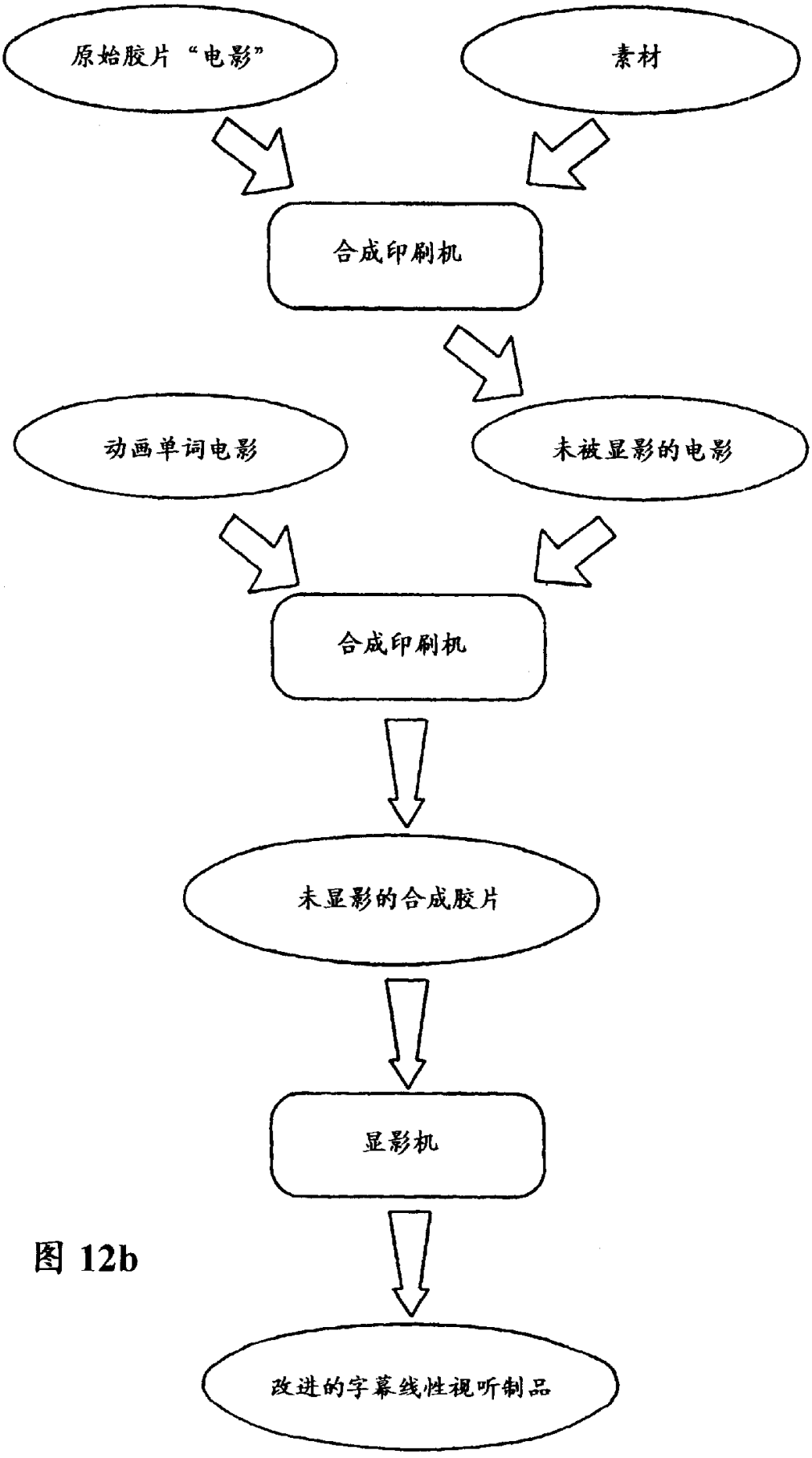

原始胶片“电影”
素材
合成印刷机
动画单词电影
未被显影的电影
合成印刷机
未显影的合成胶片
显影机
改进的字幕线性视听制品

图 12b

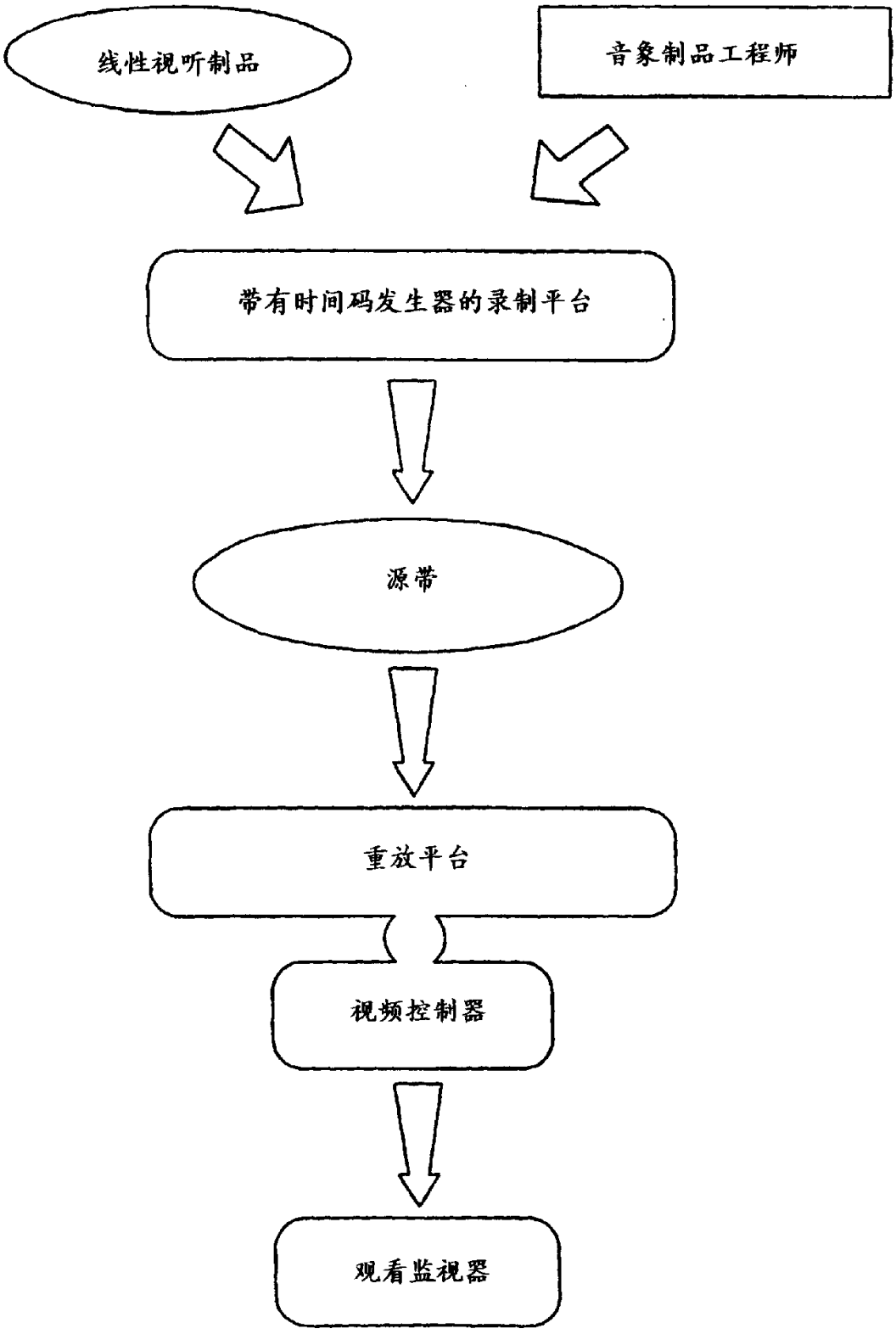


图 13a

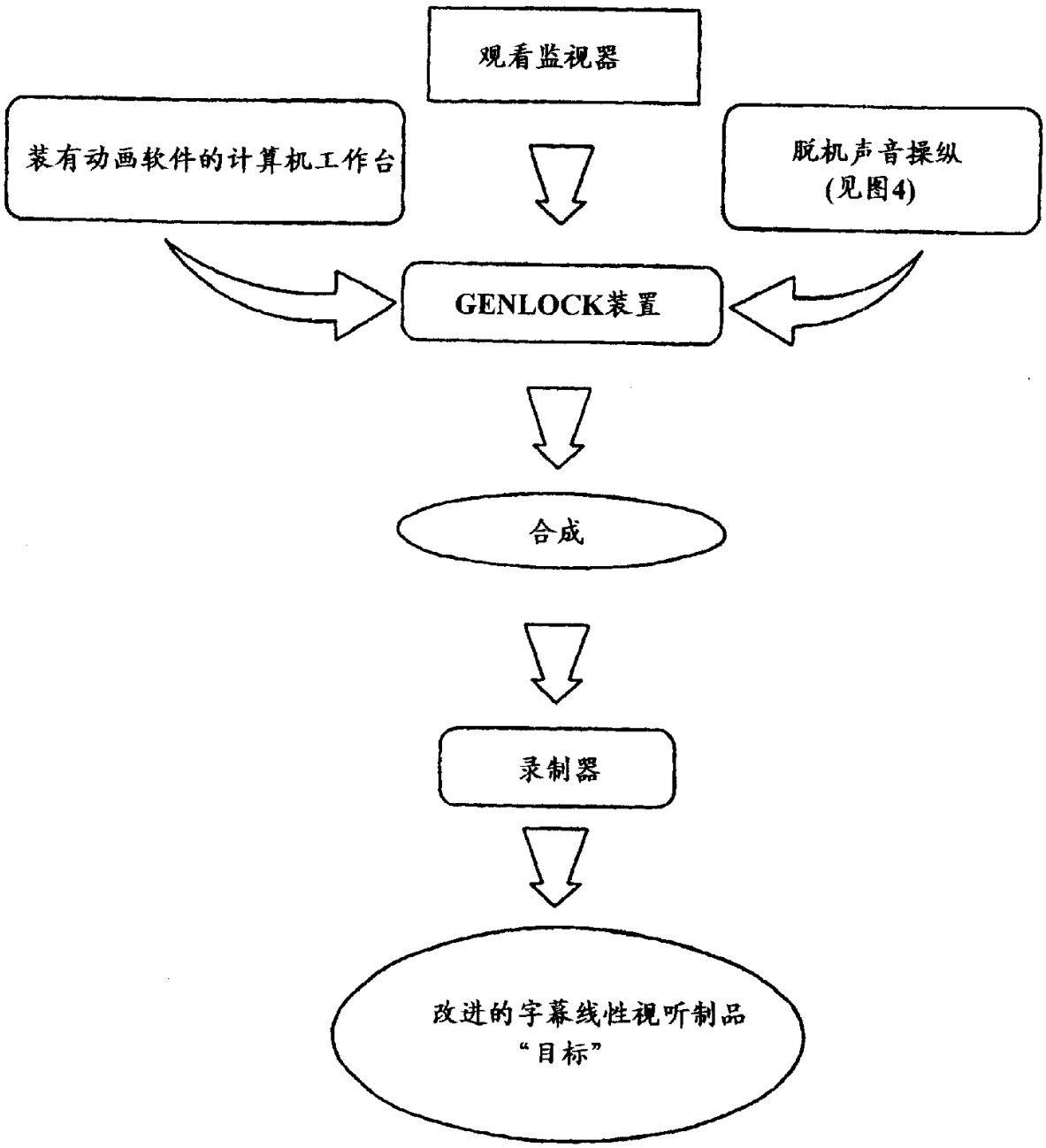

观看监视器
装有动画软件的计算机工作台
脱机声音操纵
(见图4)
GENLOCK装置
合成
录制器
改进的字幕线性视听制品
“目标”

图 13b

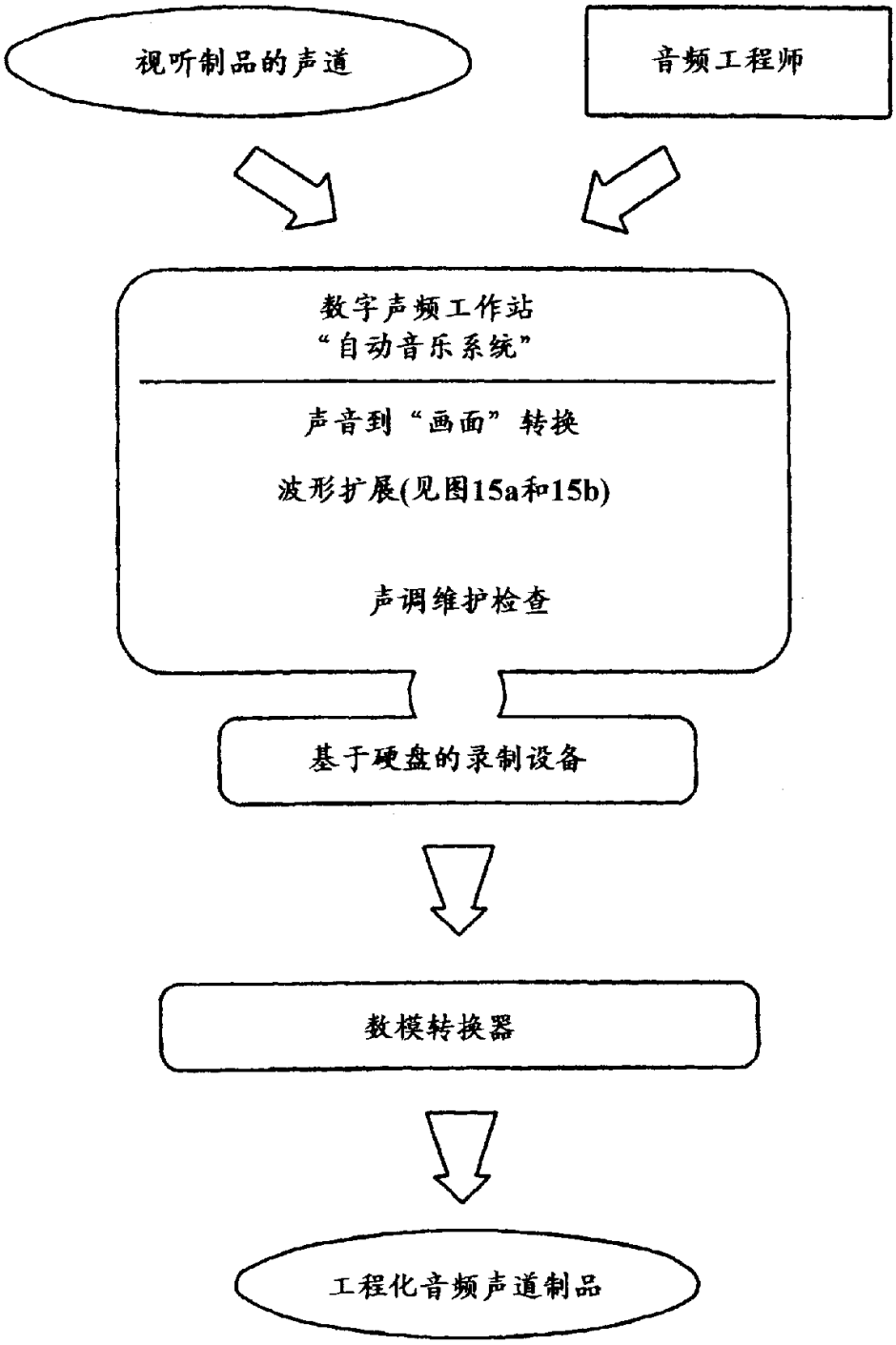

视听制品的声道
音频工程师
数字声频工作站
“自动音乐系统”
声音到“画面”转换
波形扩展(见图15a和15b)
声调维护检查
基于硬盘的录制设备
数模转换器
工程化音频声道制品

图 14

原始波形



图 15a

增强波形



图 15b

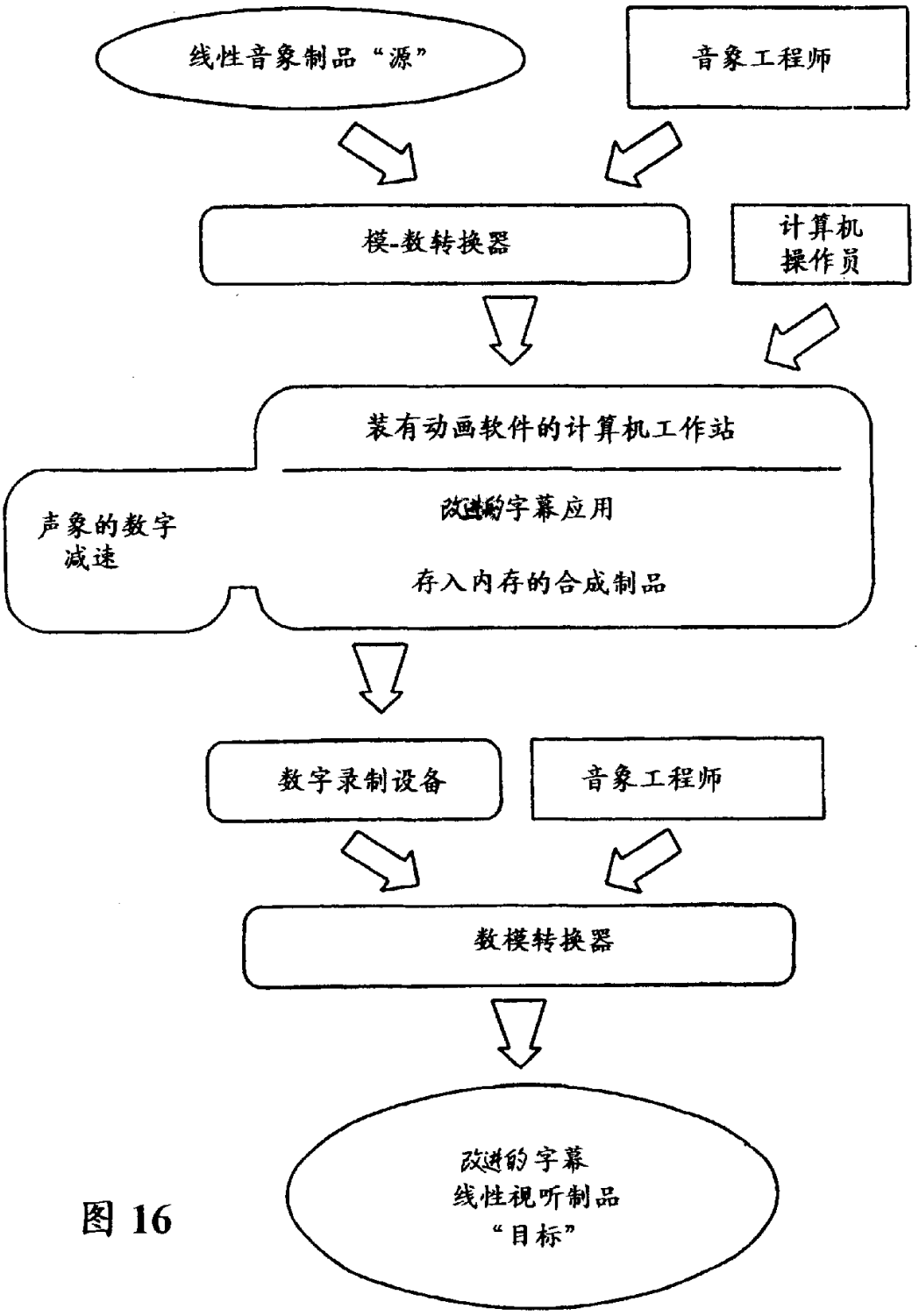

线性音象制品“源”
音象工程师
模-数转换器
计算机
操作员
装有动画软件的计算机工作站
声象的数字
减速
改进的字幕应用
存入内存的合成制品
数字录制设备
音象工程师
数模转换器
改进的字幕
线性视听制品
“目标”

图 16

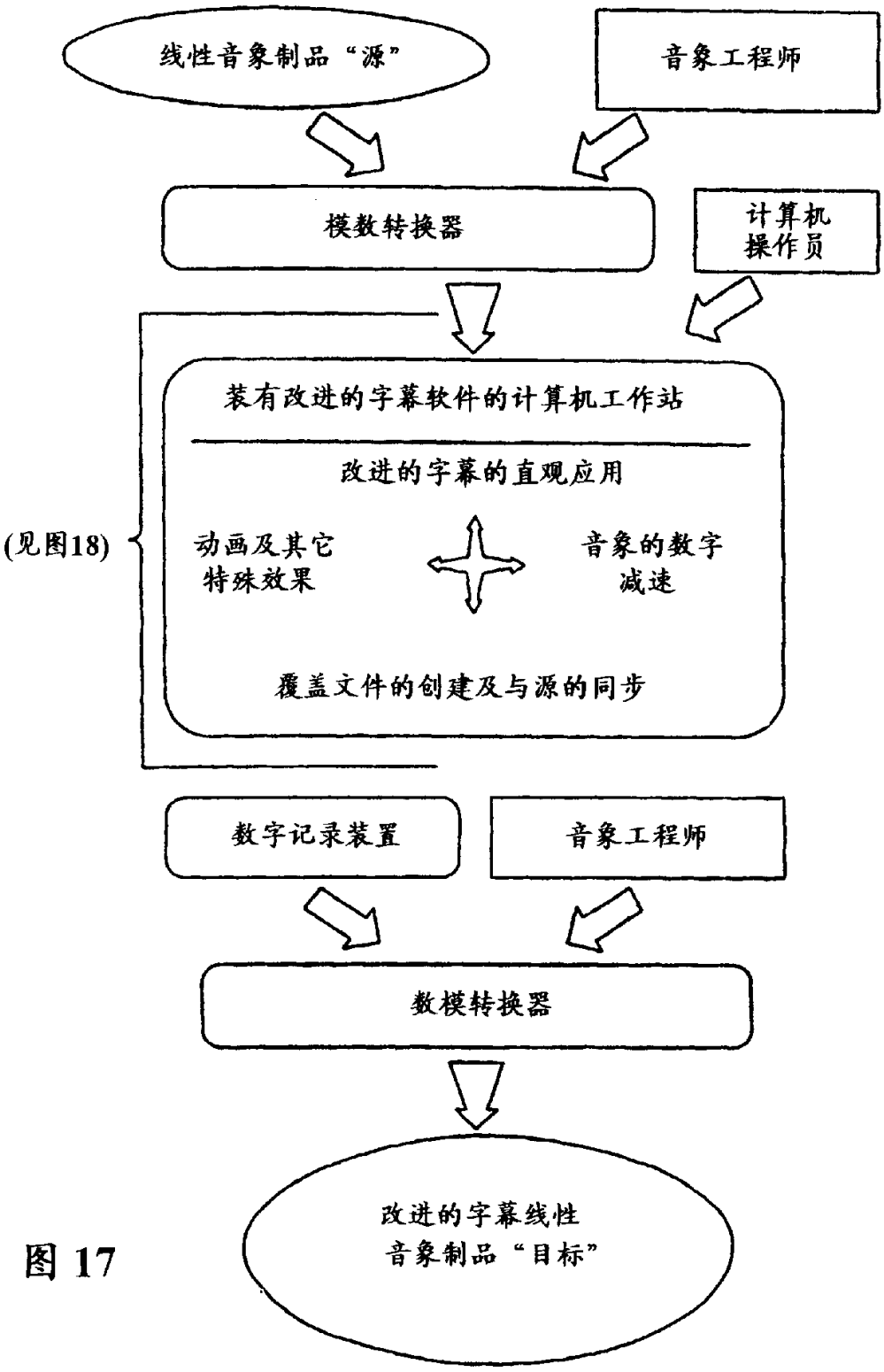

线性音象制品“源”
音象工程师
模数转换器
计算机
操作员
装有改进的字幕软件的计算机工作站
改进的字幕的直观应用
(见图18)
动画及其它
特殊效果
音象的数字
减速
覆盖文件的创建及与源的同步
数字记录装置
音象工程师
数模转换器
改进的字幕线性
音象制品“目标”

图 17

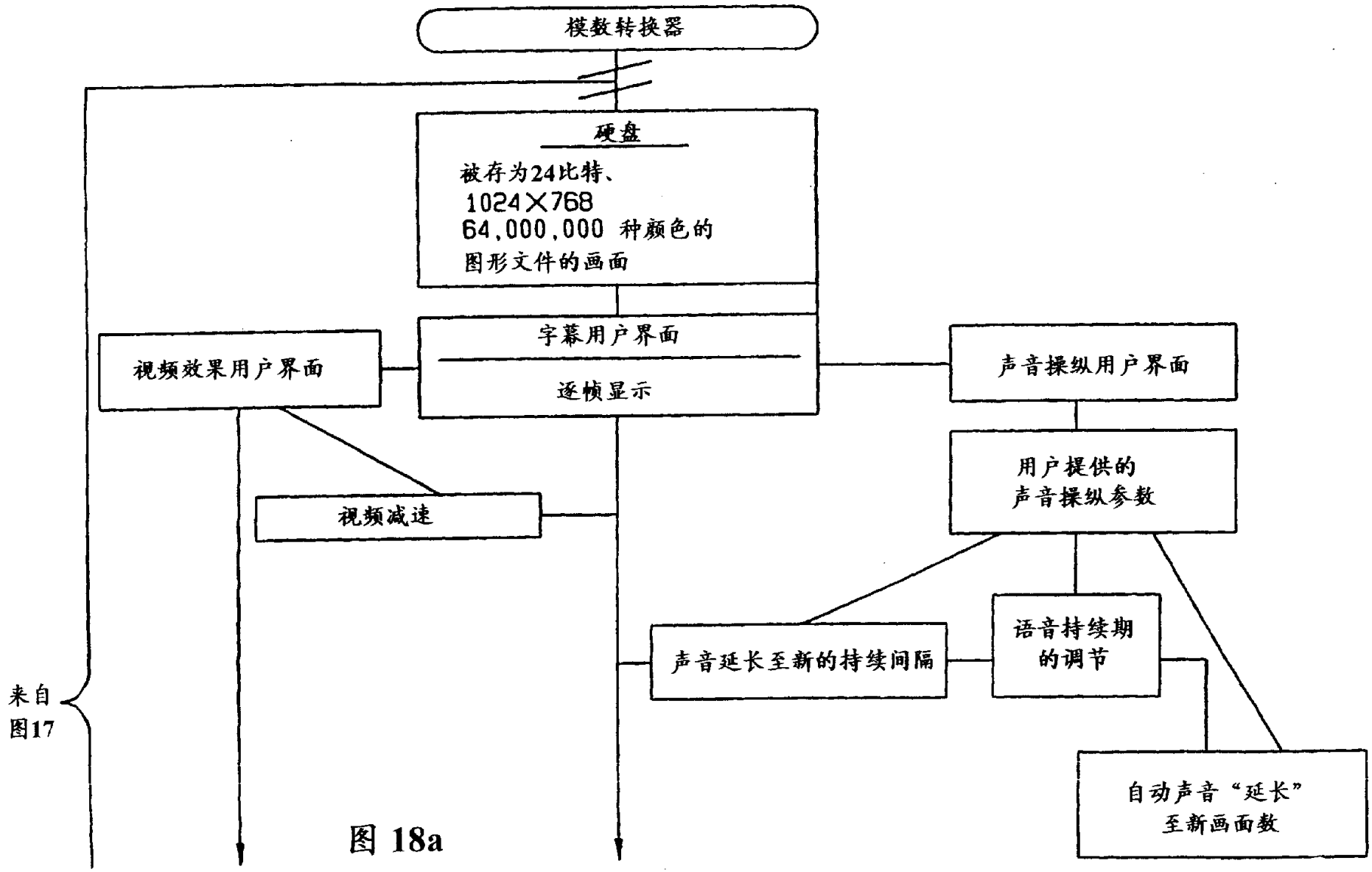

模数转换器
硬盘
被存为24比特、
1024×768
64,000,000 种颜色的
图形文件的画面
字幕用户界面
逐帧显示
视频效果用户界面
声音操纵用户界面
用户提供的
声音操纵参数
视频减速
声音延长至新的持续间隔
语音持续期
的调节
自动声音“延长”
至新画面数
来自
图17

图 18a

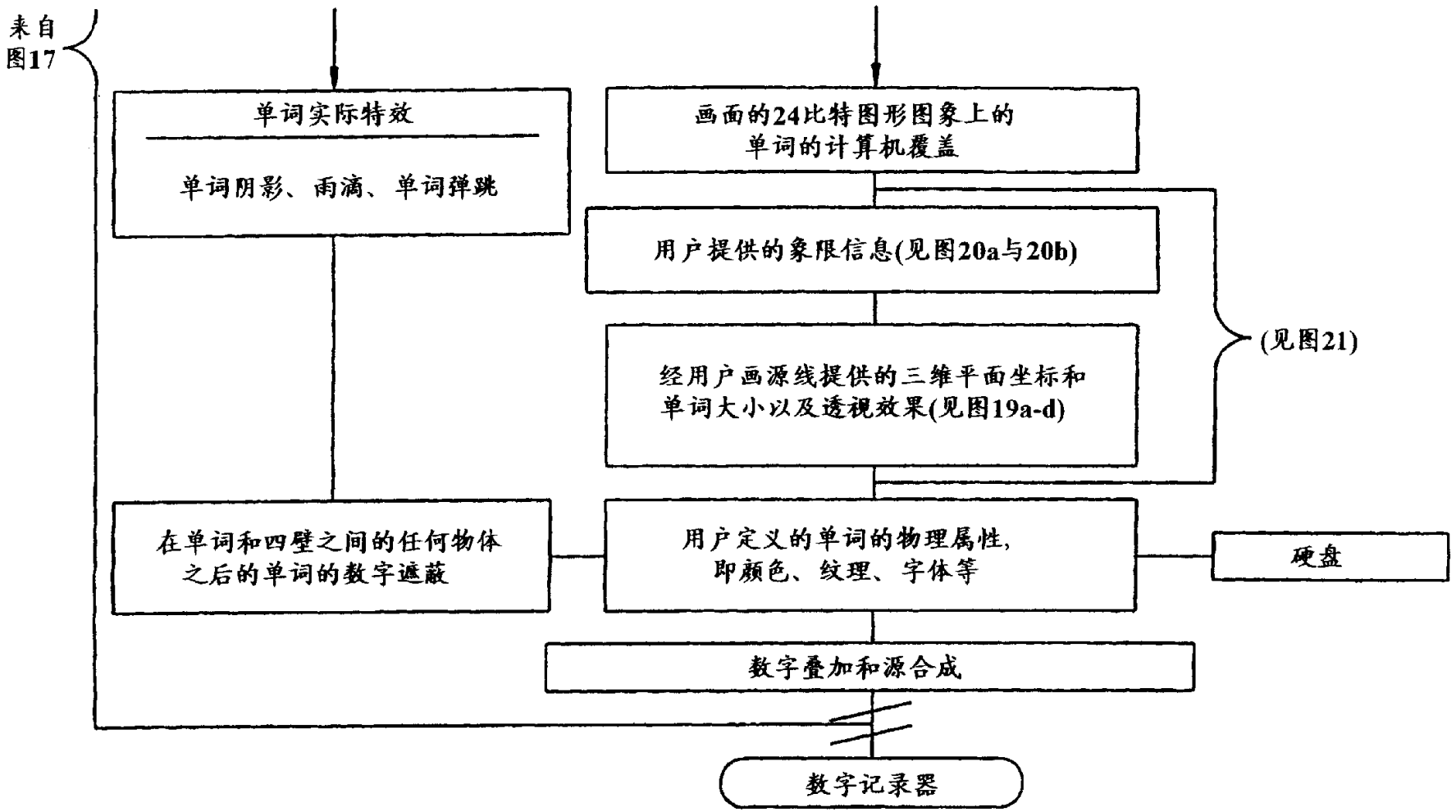

来自
图17
单词实际特效
单词阴影、雨滴、单词弹跳
画面的24比特图形图象上的
单词的计算机覆盖
用户提供的象限信息(见图20a与20b)
经用户画源线提供的三维平面坐标和
单词大小以及透视效果(见图19a-d)
(见图21)
在单词和四壁之间的任何物体
之后的单词的数字遮蔽
用户定义的单词的物理属性,
即颜色、纹理、字体等
硬盘
数字叠加和源合成
数字记录器

图 18b

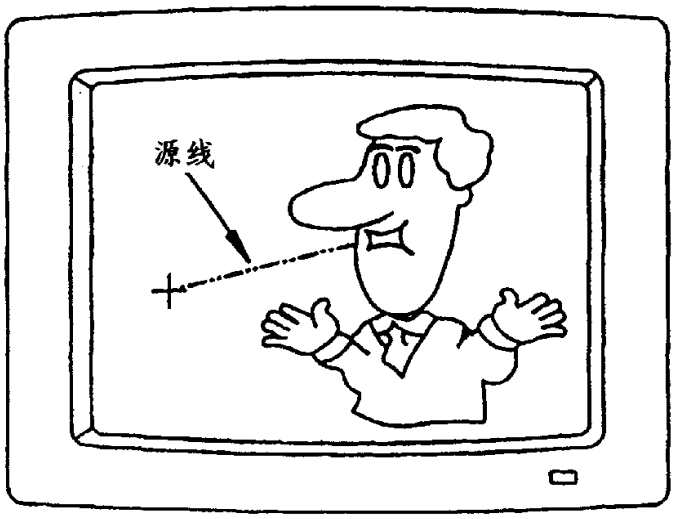


图 19a

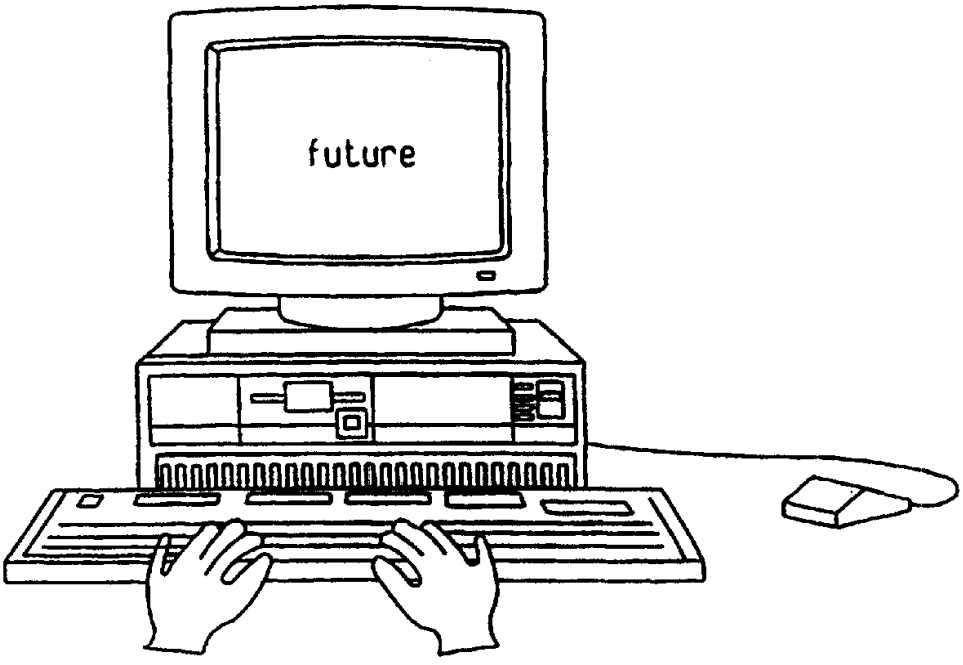


图 19b

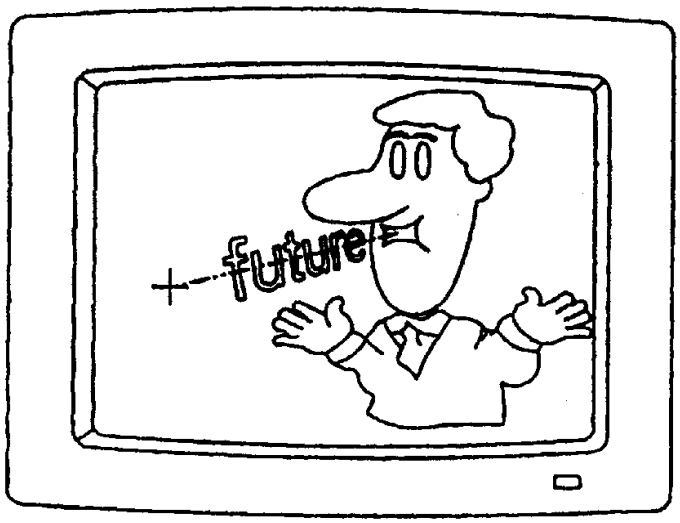


图 19c

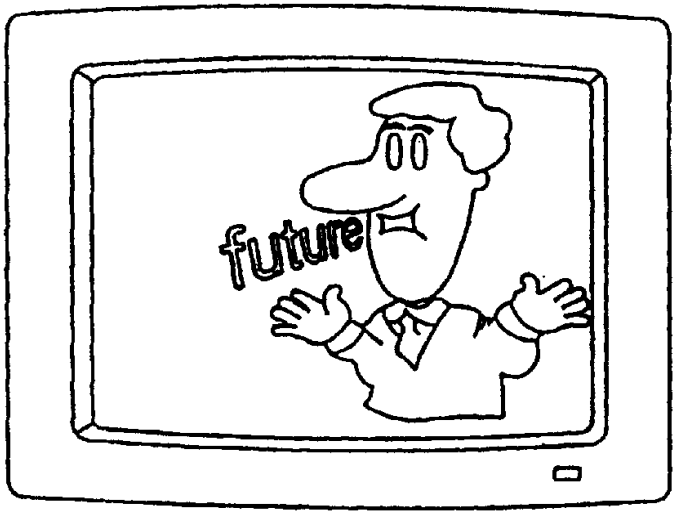


图 19d

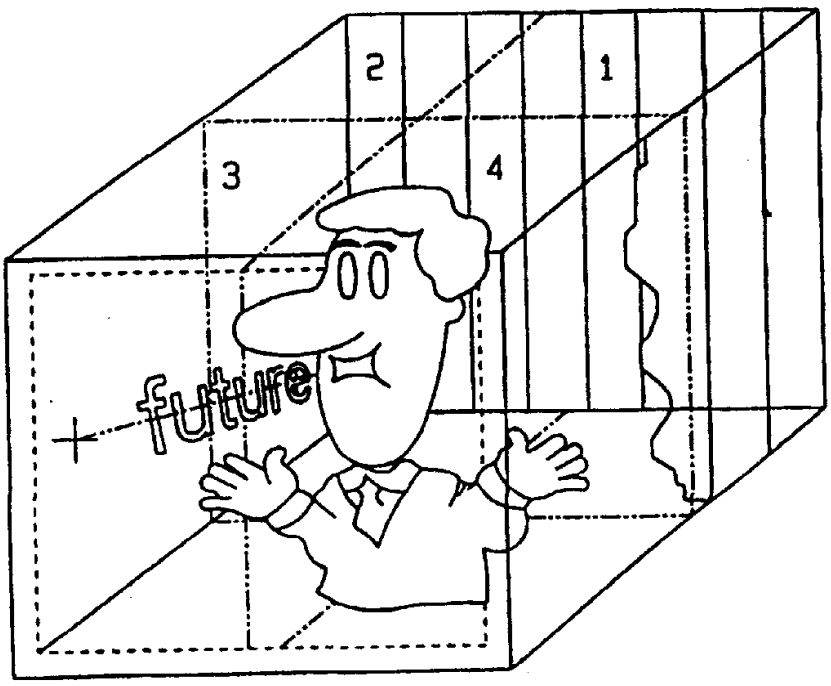


图 20a

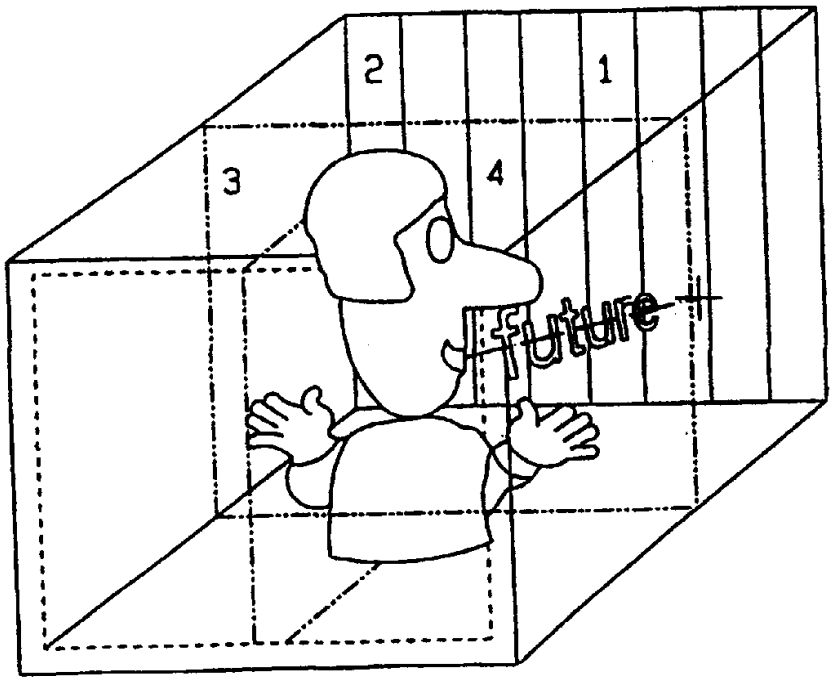


图 20b

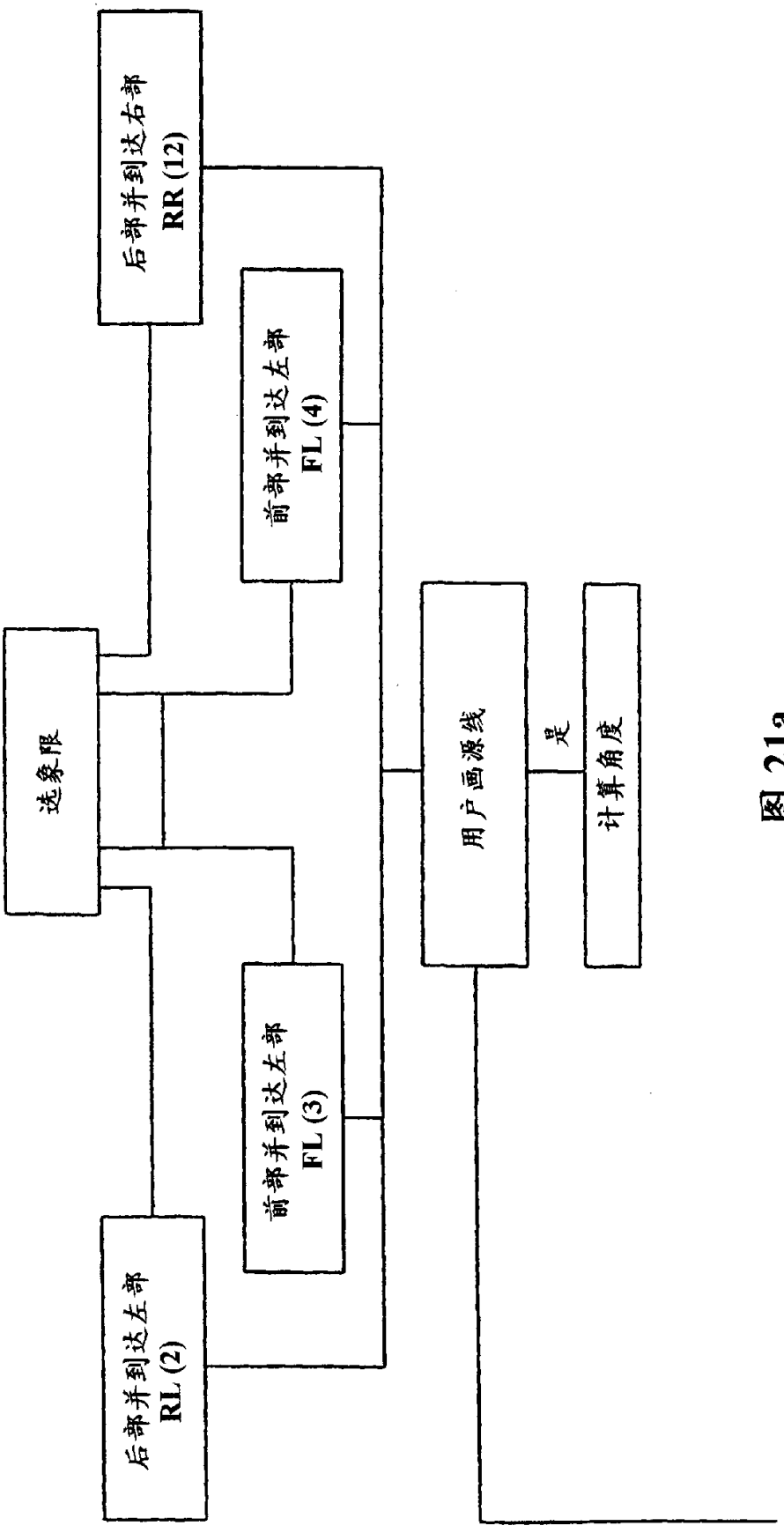

后部并到达右部
RR (12)
前部并到达左部
FL (4)
选象限
用户画源线
是
计算角度
前部并到达左部
FL (3)
后部并到达左部
RL (2)

图 21a

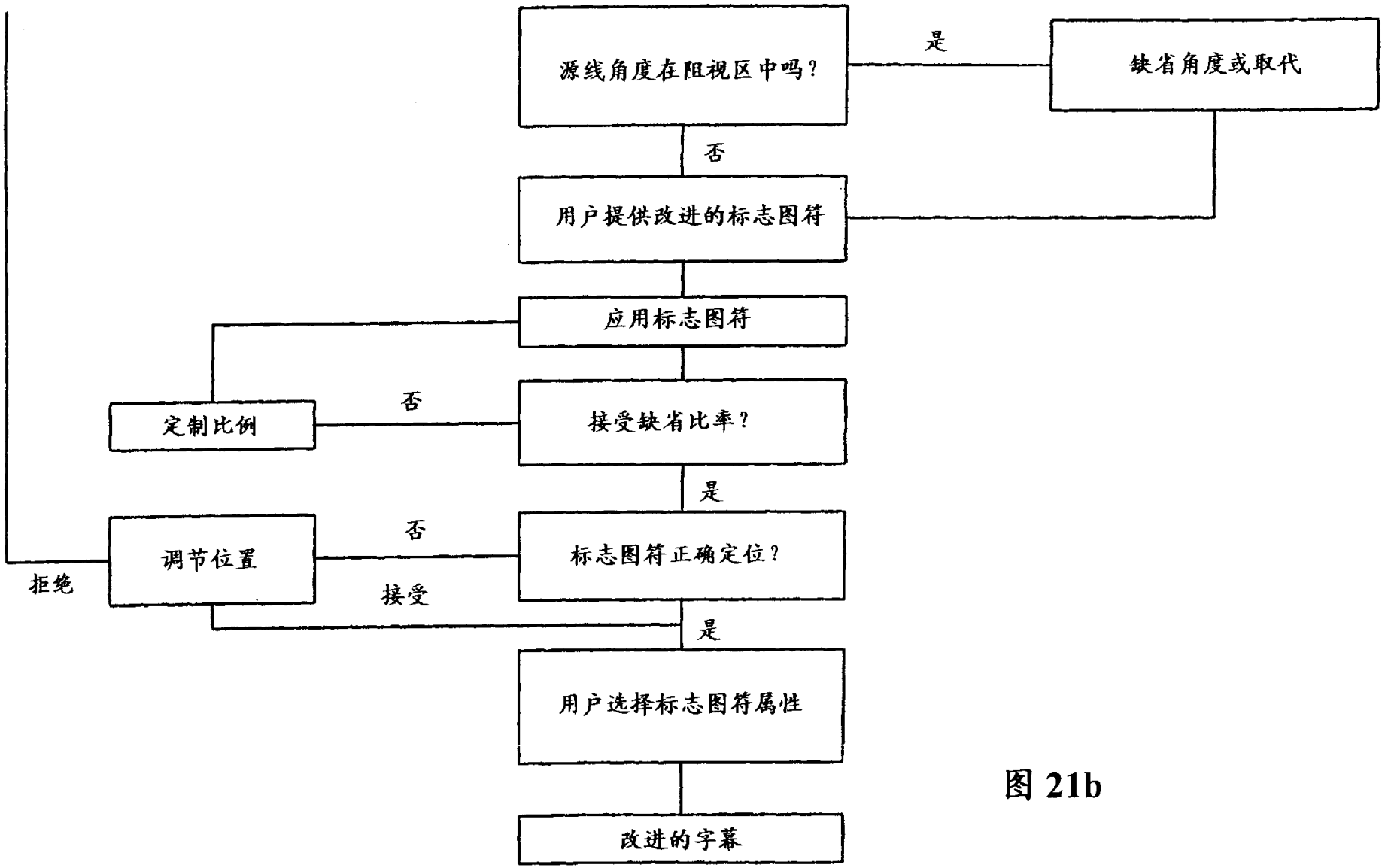

源线角度在阻视区中吗?
是
缺省角度或取代
否
用户提供改进的标志图符
应用标志图符
否
定制比例
接受缺省比率?
是
否
调节位置
标志图符正确定位?
拒绝
接受
是
用户选择标志图符属性
改进的字幕

图 21b

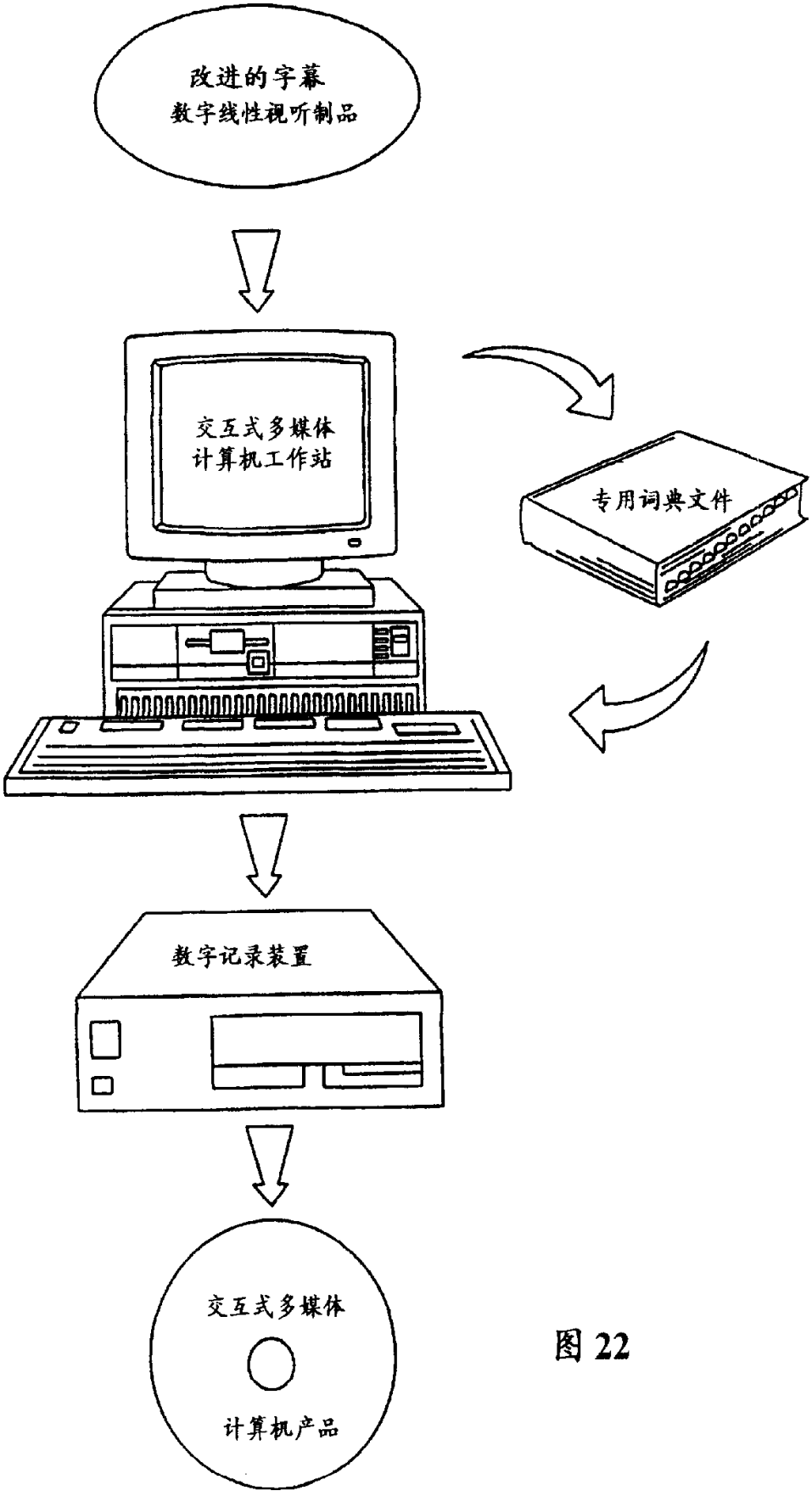


图 22

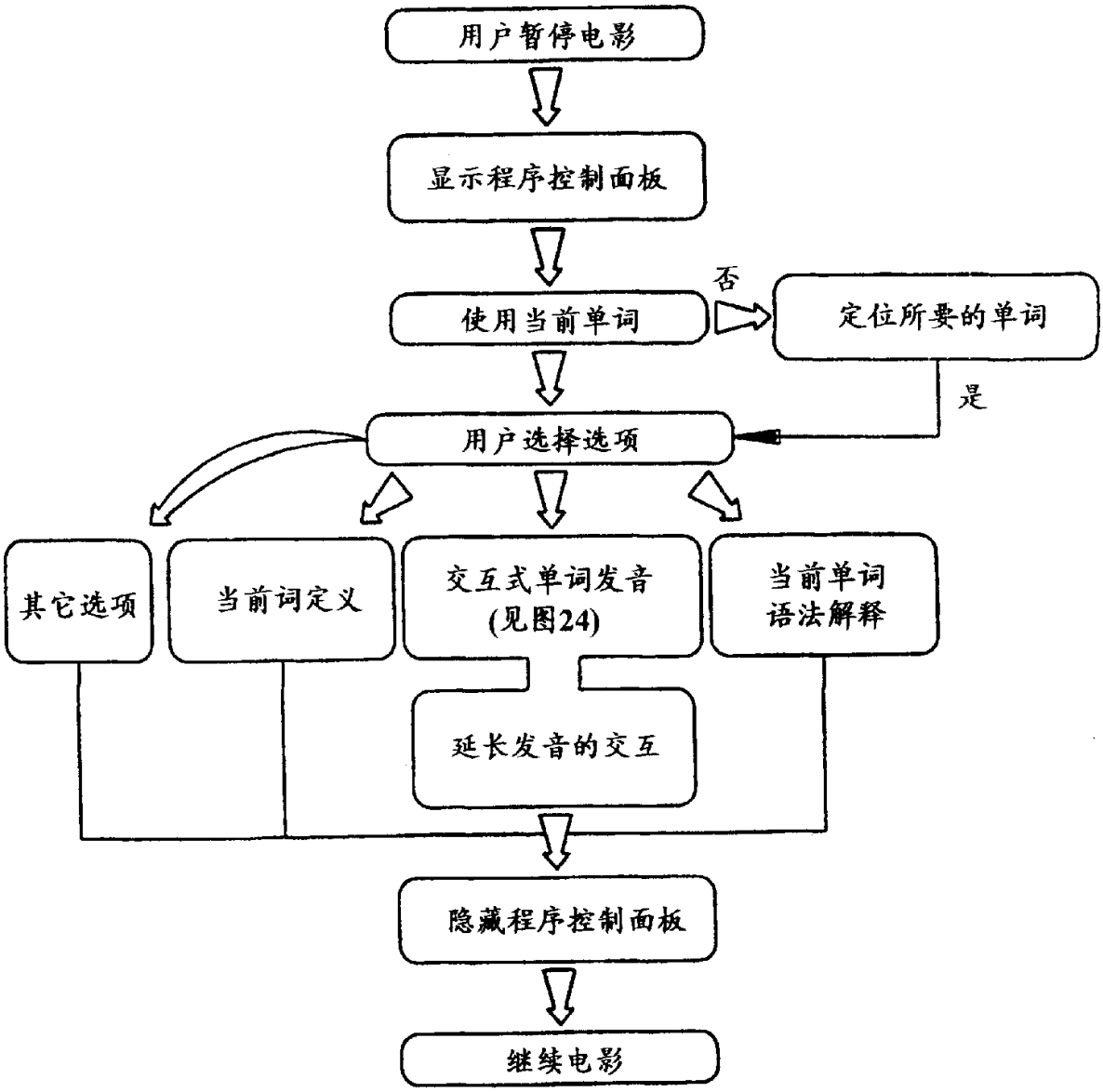

用户暂停电影
显示程序控制面板
使用当前单词
否
定位所要的单词
是
用户选择选项
其它选项
当前词定义
交互式单词发音
(见图24)
当前单词
语法解释
延长发音的交互
隐藏程序控制面板
继续电影

图 23

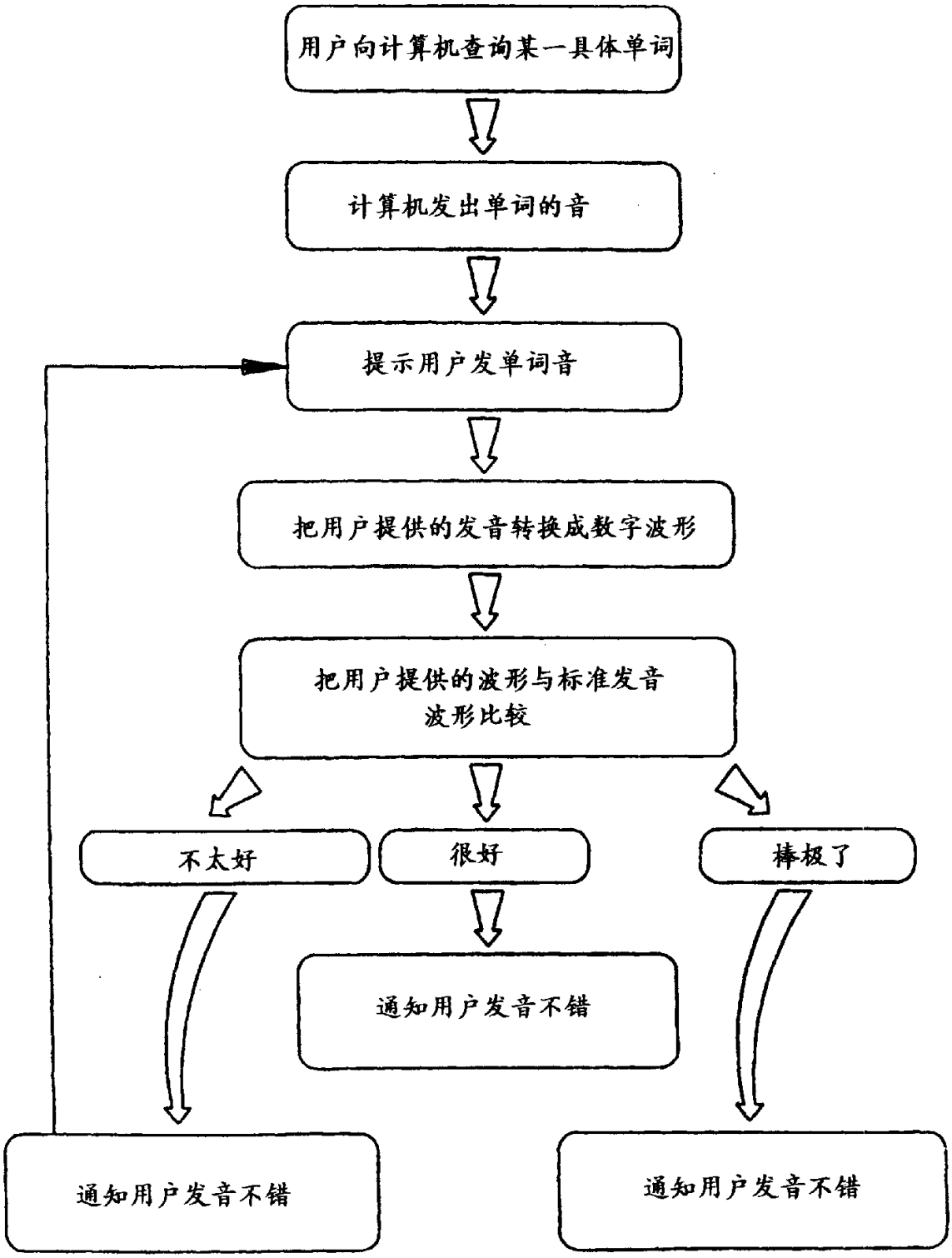

用户向计算机查询某一具体单词
计算机发出单词的音
提示用户发单词音
把用户提供的发音转换成数字波形
把用户提供的波形与标准发音
波形比较
不太好
很好
棒极了
通知用户发音不错
通知用户发音不错
通知用户发音不错

图 24

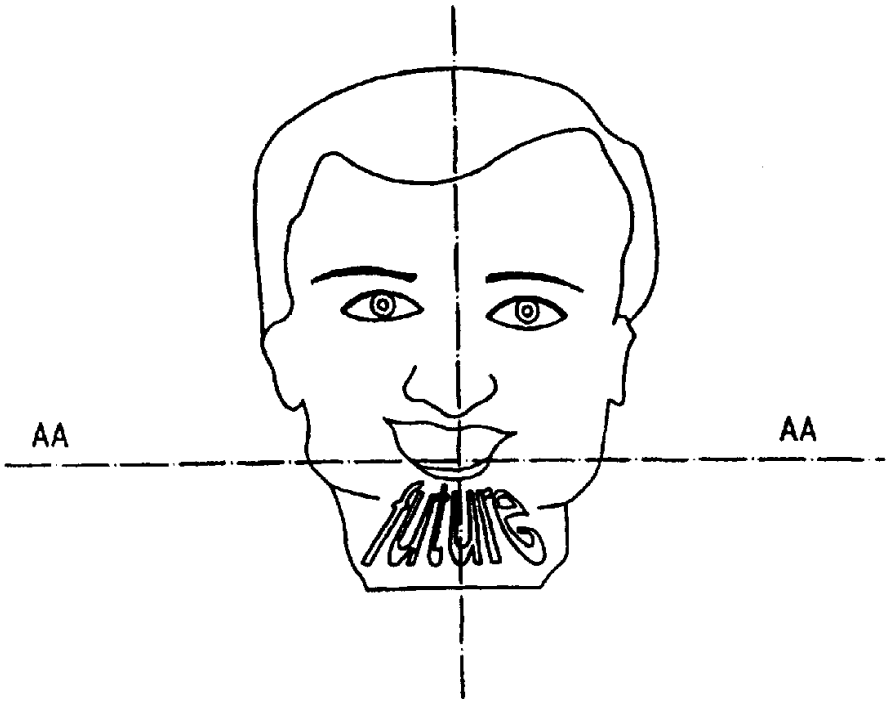


图 25

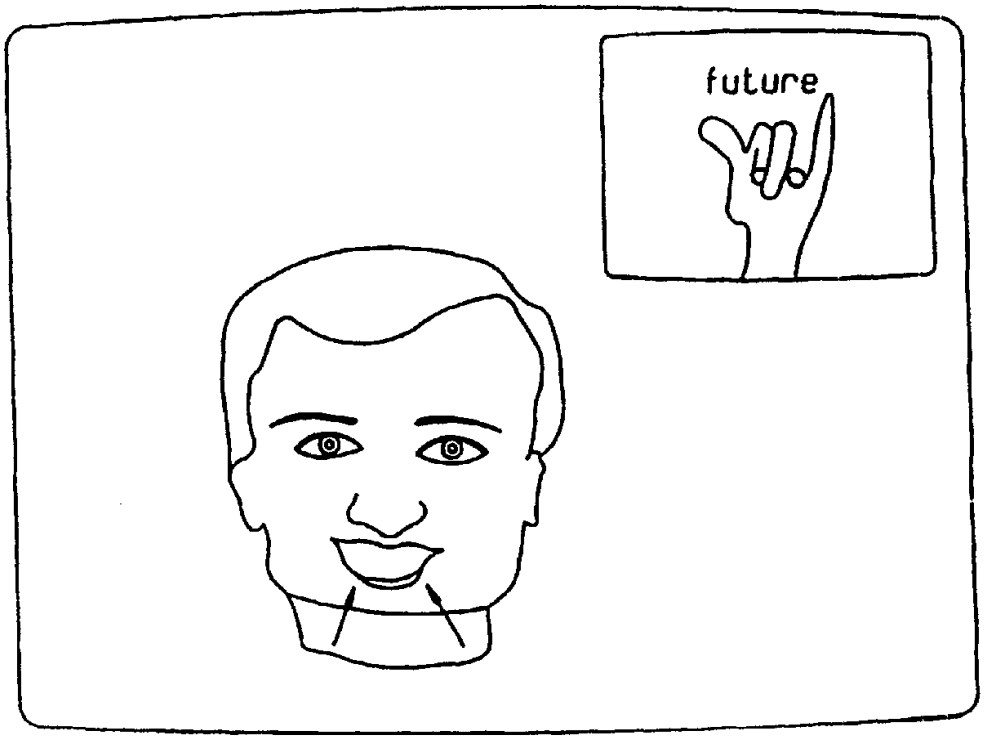


图 26



图 27

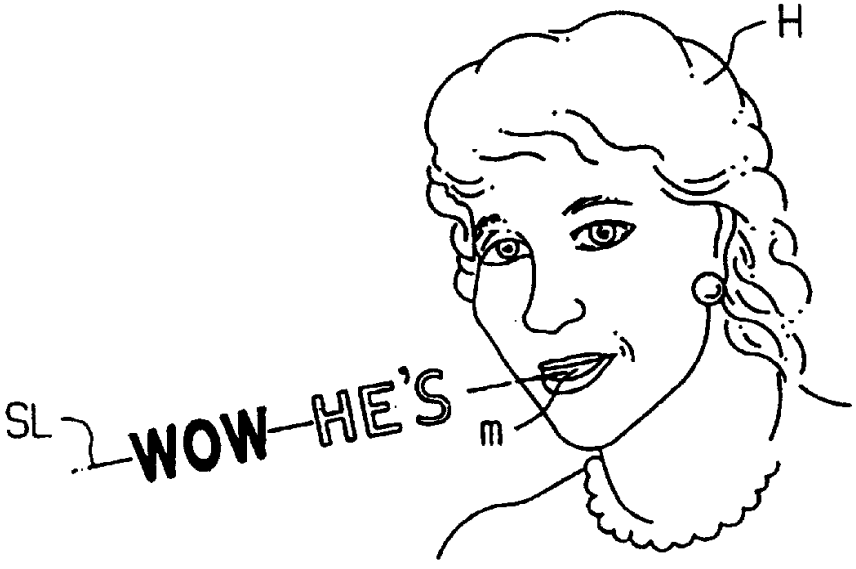


图 28



图 29



图 30



图 31

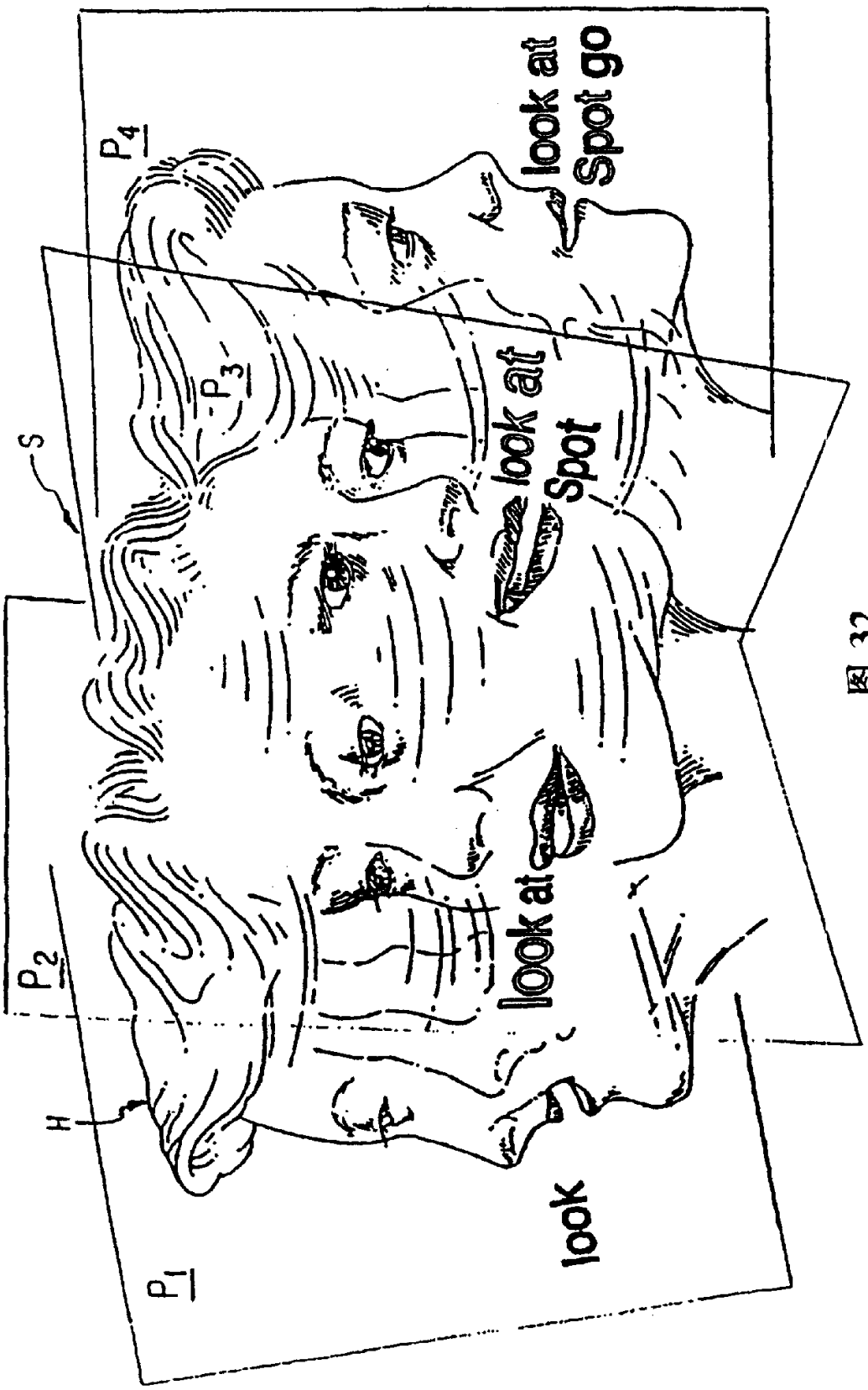


图 32

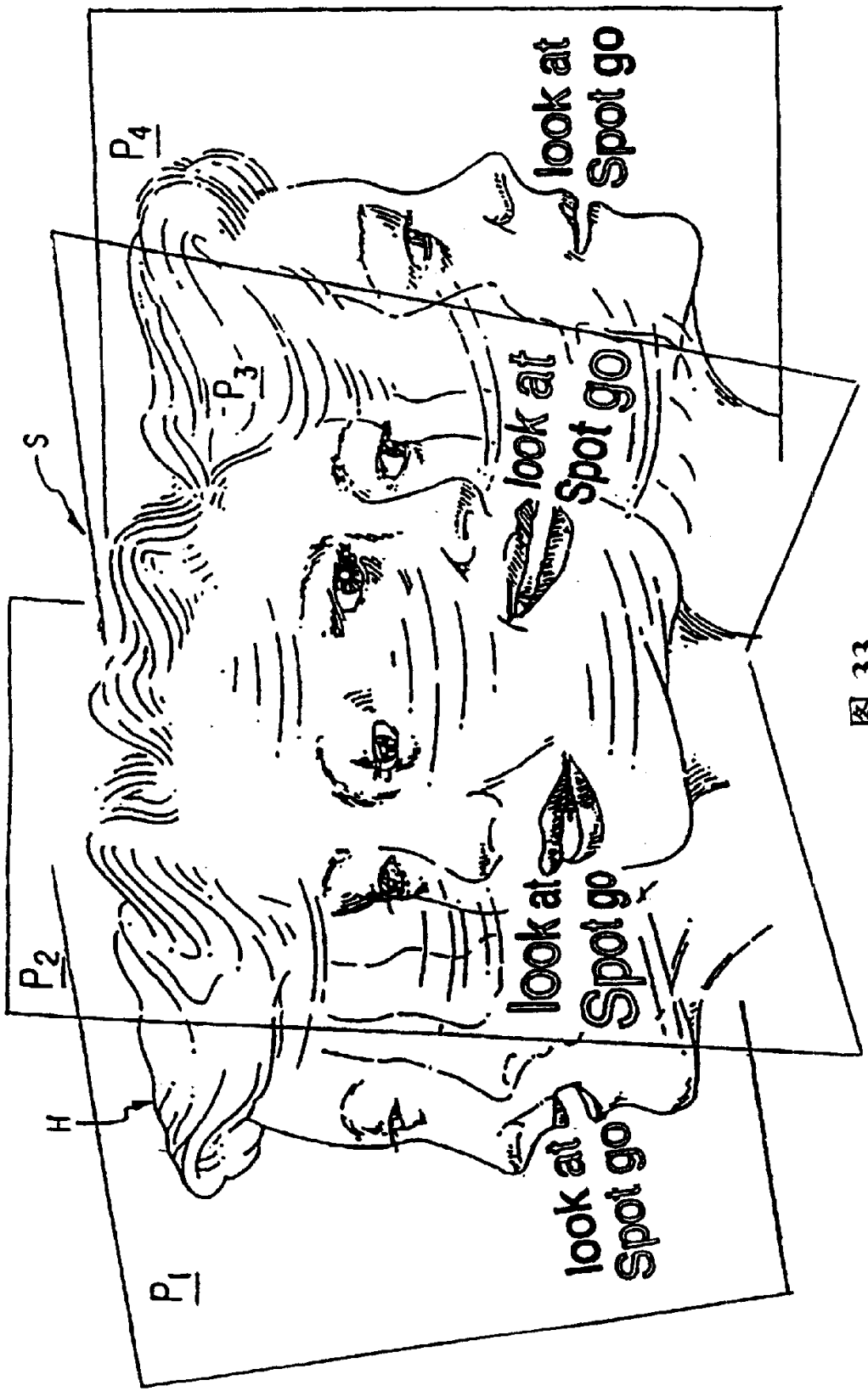

P4
P3
P2
P1
S
H
look at
Spot go
look at
Spot go
look at
Spot go
look at
Spot go

图 33

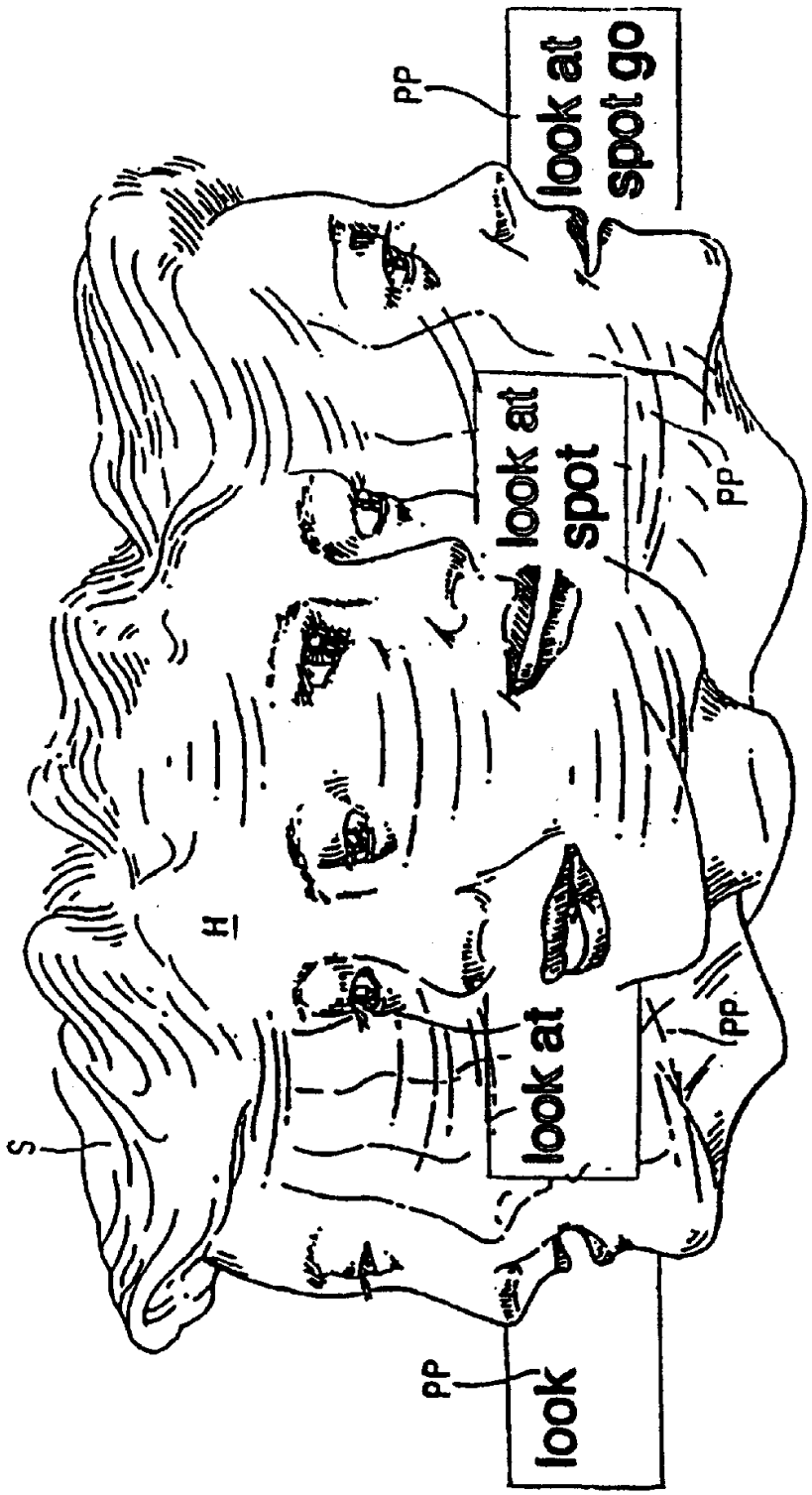


图 34

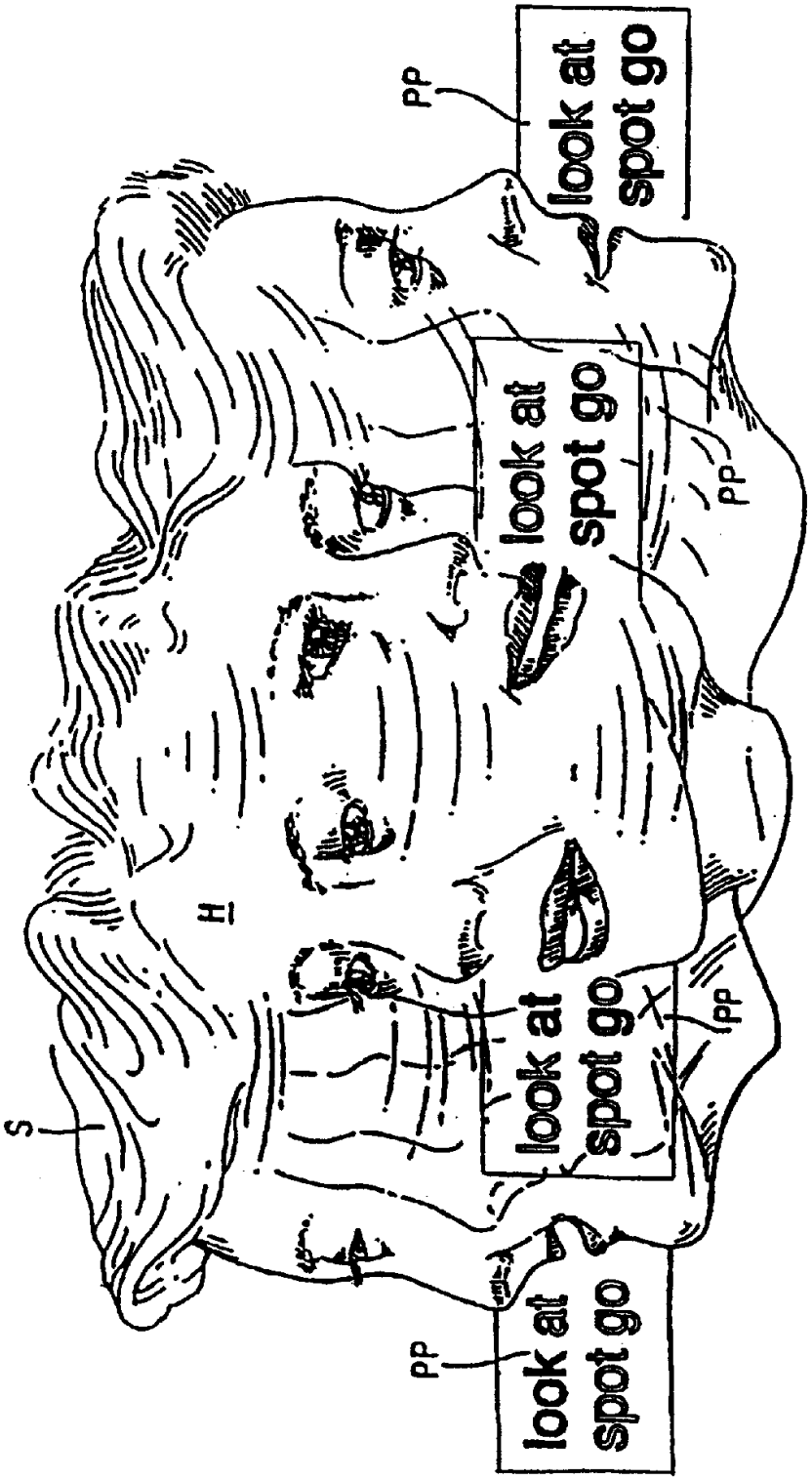

look at
spot go
look at
spot go
look at
spot go
look at
spot go
PP
PP
PP
PP
H
S

图 35

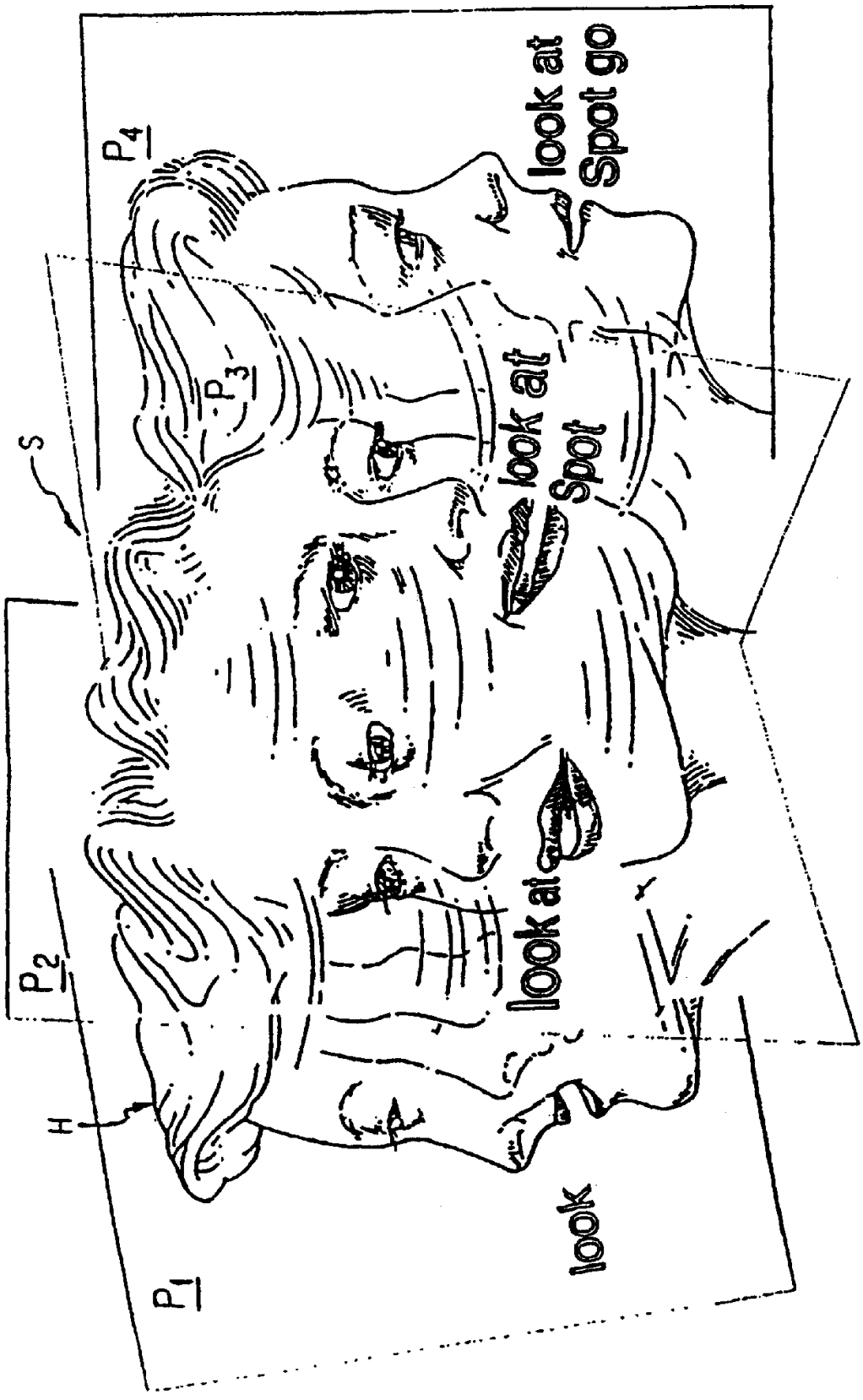


图 36

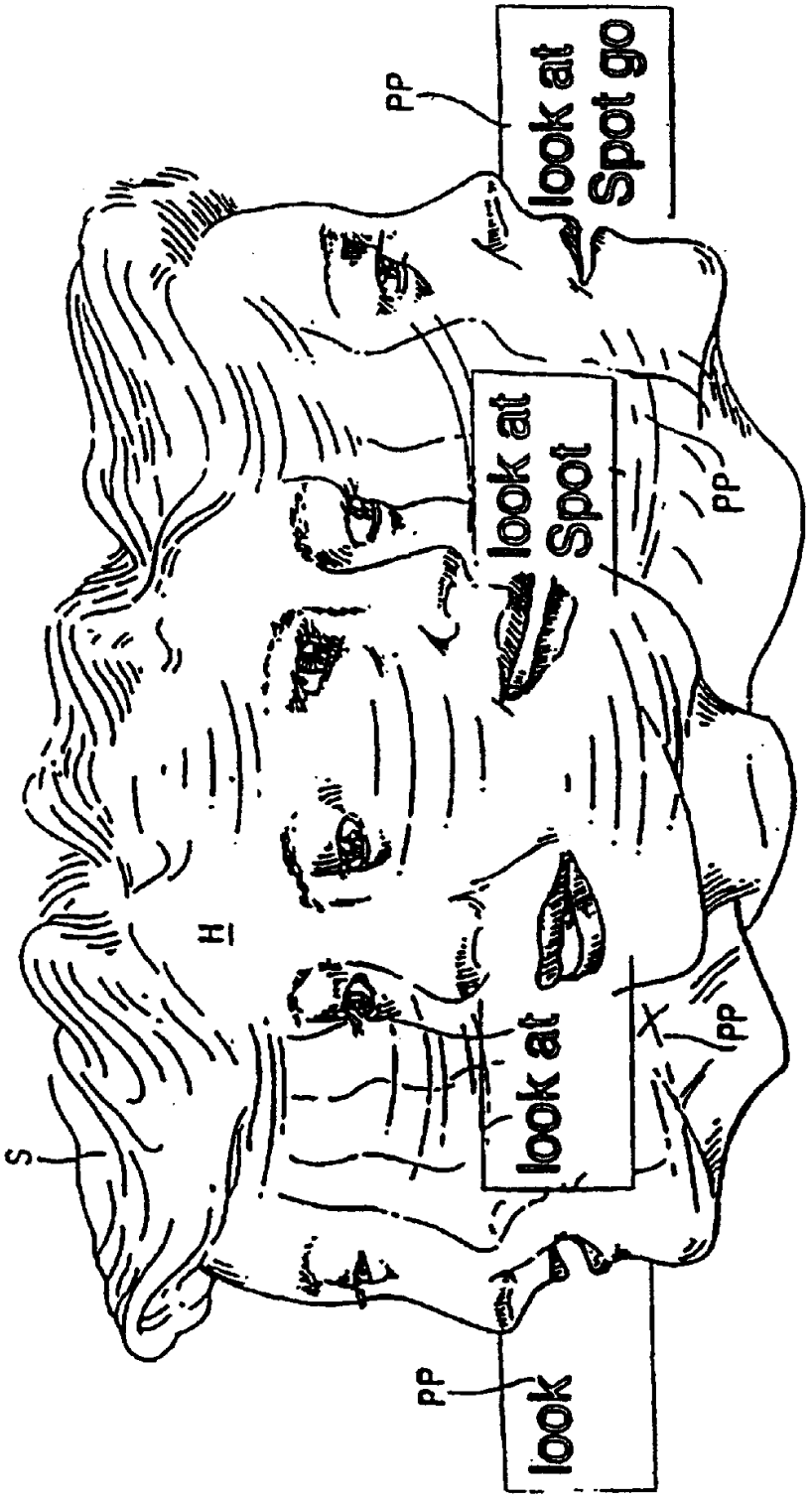


图 37

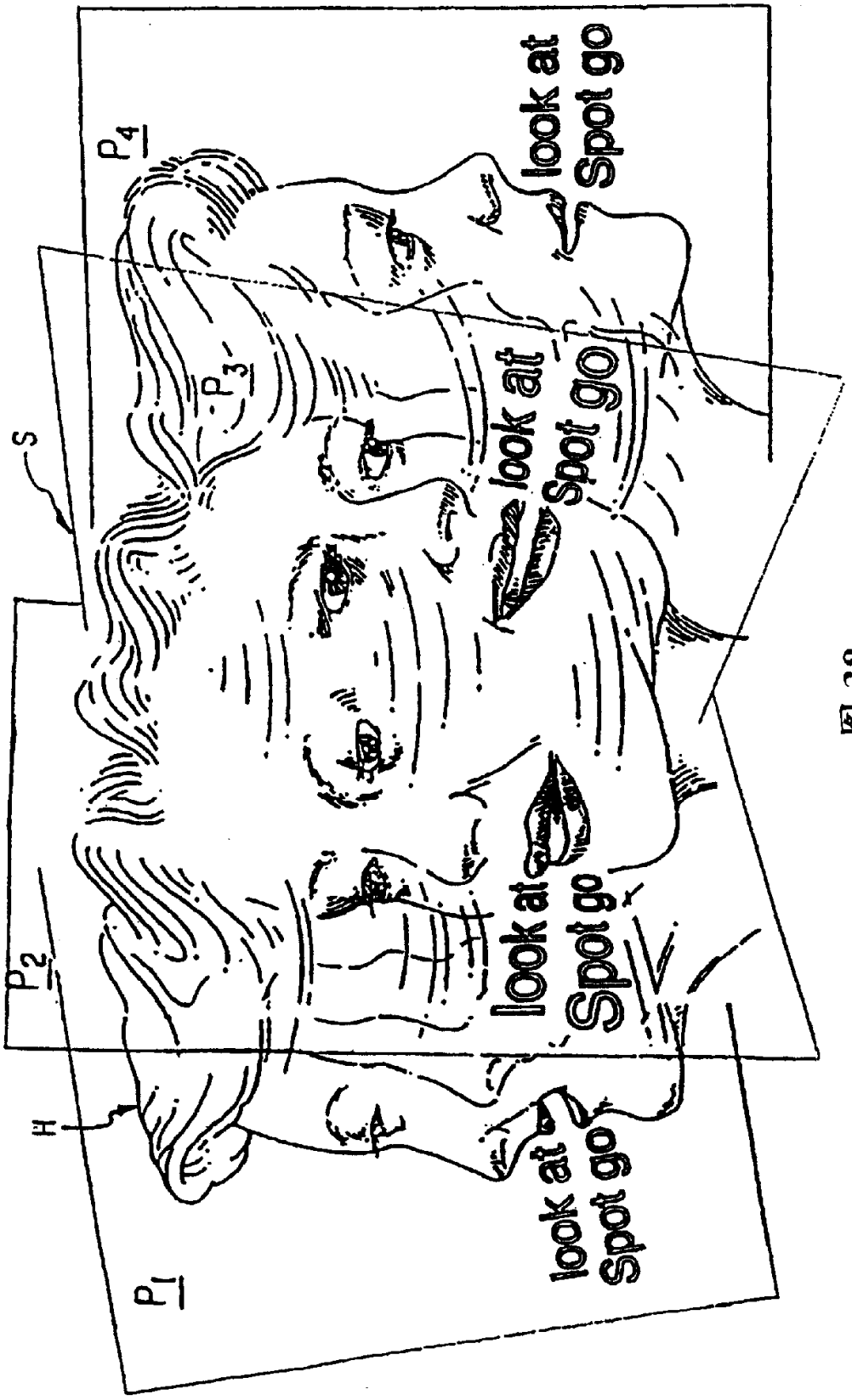

look at
Spot go
look at
Spot go
look at
Spot go
look at
Spot go
P1
P2
P3
P4
S
H

图 38

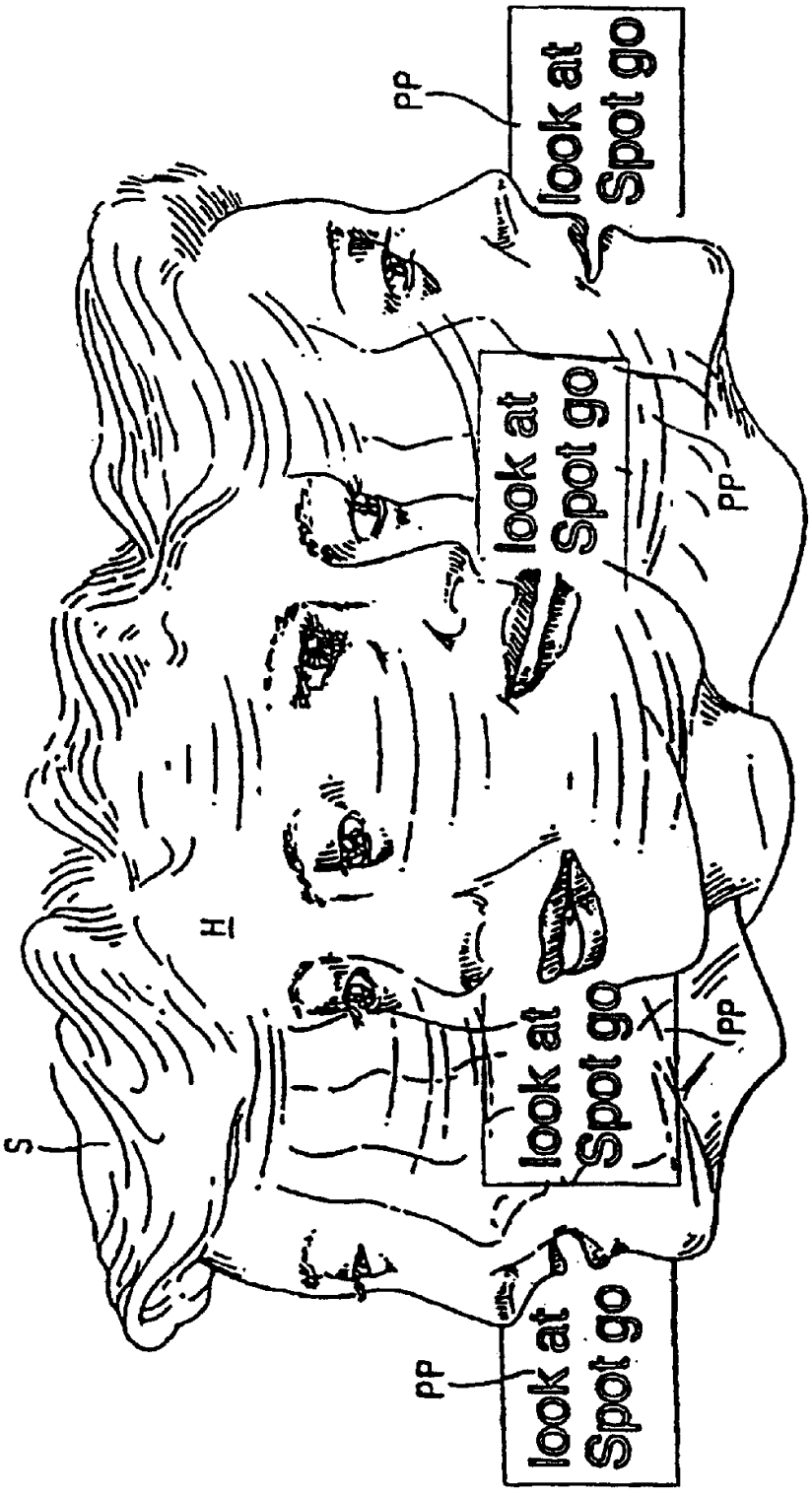

look at
Spot go
PP
look at
Spot go
PP
look at
Spot go
PP
look at
Spot go
PP
H
S

图 39