[19]中华人民共和国专利局

[11] 公开号 CN 1101446A

[12] 发明专利申请公开说明书

[21]申请号 94102645.0

[51]Int.Cl[5]

[43]公开日 1995 年 4 月 12 日

G09B 5 / 04

[22]申请日 94.1.20

[30]优先权

[32]93.1.21 [33]US[31]08 / 007,242

[71]申请人 DSP纵露神思国际公司

地址 美国加利福尼亚州

[72]发明人 齐夫·施皮罗

加布里埃尔·F·格罗尼尔

埃里克·奥登特里齐

[74]专利代理机构 永新专利商标代理有限公司

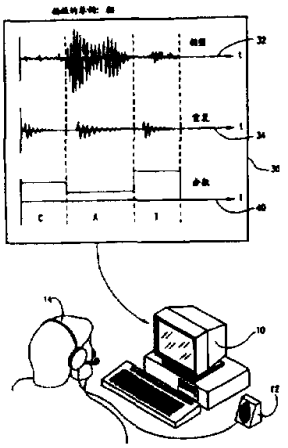
代理人 蹇 炜

说明书页数: 附图页数:

[54]发明名称 语音教学计算机化系统

[57]摘要

一种用于语音和发音训练的改进的计算机化系统，它将录制的参考语音样本提供给学生，并且将学生的重发与原来提供的参考语音样本之间的相似度的量化显示给使用者。



权 利 要 求 书

1、交互式语音训练装置，它包括：

一个声音样本发生器，它用于给使用者播放预录制的参考声音样本，从而使用者尝试着重发；

一个声音样本记分器，它用于对使用者的重发声音样本进行记分。

2、根据权利要求1所述的装置，其特征在于，所述声音样本记分器包括：

一个参考／响应比较单元，它用于将使用者的重发声音样本的至少一个特征与所述声音样本的至少一个特征进行比较；

一个相似度指示器，它用于提供一个至少一个重发声音样本的特征与至少一个参考声音样本的特征之间的相似度的输出指示。

3、根据权利要求2所述的装置，其特征在于，还包括一个使用者响应存储器，它用于存储参考声音样本的使用者的重发，所述参考／响应比较单元可以利用该使用者响应存储器。

4、根据权利要求2所述的装置，其特征在于，所述参考／响应比较单元包括一个音量／持续时间校正器，它用于将参考和重发声音样本的音量和持续时间校正。

5、根据权利要求2所述的装置，其特征在于，所述参考／响应比较单元包括一个参数化单元，它用于从所述参考和重发声音样本中抽取声音信号参数。

6、根据权利要求5所述的装置，其特征在于，所述参考／响应比较单元还包括用于将所述声音样本参数与所述重发声音样本参数进行比较的装置。

7、根据权利要求6所述的装置，其特征在于，所述用于比较的装置包括一个参数分数发生器，它提供一个表示所述参考与重发声音样本的声音信号参数之间的相似度的分数。

8、根据权利要求7所述的装置，其特征在于，所述输出指示包括所述分数的显示。

9、根据权利要求2所述的装置，其特征在于，所述输出指示包括至少一个声音波形的显示。

10、根据权利要求1所述的装置，其特征在于，还包括一个提示序列发生器，它给使用者产生提示序列。

11、根据权利要求1所述的交互式语音训练装置，其特征在于，还包括一个参考声音样本库，其中存储着参考声音样本，并且所述声音样本发生器可以利用该参考声音样本库。

12、根据权利要求11所述的装置，其特征在于，所述参考声音样本库包括由多个语音模型产生的声音样本的大量录音。

13、根据权利要求12所述的装置，其特征在于，所述多个语音模型以下特征中的至少一项中是相互不同的：

性别；

年龄；

方言。

14、交互式语音训练装置，它包括：

一个提示序列发生器，它给使用者产生提示序列并提示使用者产生相应的声音样本序列，

一个参考／响应比较单元，它将该使用者产生的每一个声音样本的至少一个特征与一个参考进行比较。

15、根据权利要求14所述装置，其特征在于，所述与单个使用者产生的声音样本比较的参考包括一个相应的存储的参考声音样本。

16、根据权利要求14所述的装置，其特征在于，所述提示序列根据使用者的表现转移。

17、根据权利要求14所述的装置，其特征在于，所述提示序列至少部分地由使用者指定的本国语确定。

18、根据权利要求14所述的装置，其特征在于，所述提示序列发生器包括一个多语言提示序列库，其中存储着以多种语言表示的多个提示序列，并且该提示序列发生器根据使用者指定的其本国语的单一语言操作产生多种语言中的一种单一语言的提示序列。

19、交互式语音训练装置，它包括：

一个声音样本录音机，它录制使用者产生的声音样本；

一个参考／响应比较单元，它将使用者产生的声音样本的至少一个特征与一个参考进行比较，该比较单元包括：

一个声音样本分段器，它将使用者产生的声音样本

分段成多个段；

一个段比较单元，它将多个段中的至少一个中的至少一个特征与一个参考进行比较。

20、根据权利要求**19**所述的装置，其特征在于，所述声音样本分段器包括一个语音分段器，它将使用者产生的声音样本分段成多个语音段。

21、根据权利要求**20**所述的装置，其特征在于，至少一个所述语音段包括一个语音。

22、根据权利要求**20**所述的装置，其特征在于，至少一个所述语音段包括一个音节。

23、根据权利要求**21**所述的装置，其特征在于，所述语音包括一个元音。

24、根据权利要求**21**所述的装置，其特征在于，所述语音包括一个辅音。

25、交互式语音训练装置，它包括：

一个声音样本录音机，它录制由使用者产生的声音样本；

一个非特定人声音样本记分器，它基于至少一个非特定人参数对使用者产生的声音样本进行记分。

26、根据权利要求**25**所述的装置，其特征在于，所述至少一个非特定人参数包括一个用于一个预定频率上的能量数值的阈值。

27、根据权利要求**1**所述的装置，其特征在于，还包括一个常规个人计算机。

说　　明　　书

语音教学计算机化系统

本发明一般地涉及教学系统并且更特别地涉及用于语音教学的计算机化系统。

近年来在计算机化语音教学领域已有了许多进展。将例如预录制的声音和单词的提示和插入信号提供给学生并录制或监听学生们的发音的语音实验室是众所周知的。

由IBM公司投放市场的语音观察器Ⅱ（Speech Viewer Ⅱ）是一种语音疗法（speech therapy）产品，它提供学生发声的视听反馈。

在下列出版物中描述了用于计算机化语音识别的公知方法和装置，其公开的内容在这里参考引用：

Flanagan，J．L．“听说计算机：通过话音的人机通信”，Proc IEEE，64卷，1976年，405—415页；

Itakura，F．“应用于语音识别的最小预测残差原理”，IEEE Trans．Acoustics，speech and Signal Processing，1975年2月—描述了一种时间校准算法和一种计算距离量度的方法；

Le Roux，J．和Gueguen，C．

一种部分相关系数的定点计算”，IEEE ASSP，1977年6月；

Peacocke，R. D. 和Graf，D. H，“一种语音及说话者识别介绍”，IEEE Computer，23（8）卷，1990年8月，26—33页；

L.R.Rabiner等，“采用群集技术的孤立单词的非特定人识别”，IEEE Trans. Acoustics，Speech and Signal Processing，ASSP-27卷，第4期，1979年8月，336—349页；

Rabiner，L. R，Levison，S. E. 和Sondhi，M. M，“向量量化和隐式马尔可夫模型应用于非特定人、孤立单词识别”，Bell Systems Tech J.，62（4）卷，1983年4月，1075—1105页；

Rabiner L. R.，和Sanbur M. R，“一种确定孤立发音终点的算法”，Bell Systems Tech J，1975年2月；

Rabiner，L. R和Wilpon，J. G.，“一种用于说话者训练的孤立单词识别系统的简化的鲁棒训练程序”，J. Acoustical Society of America，1980年11月。

所有上述出版物公开的内容在这里参考引用

本发明试图提供一种改进的用于语音和发音教学的计算机化系统，其中已录制的参考语音样本提供给学生并且将学生的重复与原来提供的参考语音样本之间的相似度的量化显示给使用者。

本发明也试图提供一种语音和发音教学系统，它特别适合于非特定人语音学习并且无需经过训练的人的语音和发音专家的参与。本发明的系统最好包括口头提示，它指导使用者通过一个教学系统而无需依靠一个教师而进行学习。最好对学生的表现进行监视并且口头提示序列的转移应考虑学生的表现。例如，预定类型的学生错误（例如一个特定语音的重复错误发音）可以从学生语音响应中抽取出来，并且口头提示序列可以转移到考虑每一种类型学生错误的出现或不出现的情形。

本发明也试图提供一种语音和发音教学系统，它特别适合于本国语说话者学习外语的优选发音的教学。最好，本发明的系统包括一个以多种语言和一个多语言信息提供的初始菜单，它提示使用者选择代表其本国语的菜单。根据使用者的本国语的选择，系统最好以其本国语操作目前的连续口头信息给使用者，和／或转移口头信息序列以考虑公知的使用者的本国语的说话者频繁发生的语言特征（例如发音错误）。例如，当以日语为本国语的说话者说英语时，常常混淆L和R声以及短I和长E声（例如在单词“ship”和“sheep”中）。以阿拉伯语和德语为本国语的说话者没有这些问题。

因此根据本发明的一个优选实施例，提供了交互式

语音训练的装置，它包括一个声音样本发生器和一个声音样本记分器，声音样本发生器给使用者播放预先录制好的参考声音样本从而让使用者尝试着重发，声音样本记分器对使用者的重发声音样本进行记分。

此外，根据本发明的一个优选实施例，声音样本记分器包括一个参考／响应比较单元和一个相似度指示器，参考／响应比较单元将使用者的重发声音样本的至少一个特征与参考声音样本的至少一个特征进行比较，相似度指示器提供重发声音样本的至少一个特征与参考声音样本的至少一个特征之间的相似度的输出指示。

此外，根据本发明的一个优选实施例，本装置还包括一个使用者响应存储器，它用于存储参考声音样本的使用者的重发，参考／响应比较单元可以利用使用者响应存储器。

另外，根据本发明的一个优选实施例，参考／响应比较单元包括一个音量／持续时间校正器，它将参考和重发声音样本的音量和持续时间校正。

此外，根据本发明的一个优选实施例，参考／响应比较单元包括一个参数化单元，它从参考和重发声音样品中抽取声音信号参数。

另外，根据本发明的一个优选实施例，参考／响应比较单元还包括将参考声音样本参数与重发声音样本参数进行比较的装置。

此外，根据本发明的一个优选实施例，用于比较的装置包括一个参数分数发生器，它提供一个表示参考与

重发声音样本的声音信号参数之间的相似度的分数。

此外，根据本发明的一个优选实施例，输出指示包括分数的显示。

根据本发明的另一个实施例，输出指示包括至少一个声音波形的显示。

此外，根据本发明的一个优选实施例，交互式语音训练装置包括一个提示序列发生器，它产生给使用者的提示序列。

此外，根据本发明的一个优选实施例，交互式语音训练装置还包括一个参考声音样本库，其中存储着参考声音样本并且声音样本发生器可以利用参考声音样本库。

另外，根据本发明的一个优选实施例，参考声音样本库包括由多个语音模型产生的许多声音样本的录音。

此外，根据本发明的一个优选实施例，多个语音模型在以下特征：性别、年龄和方言中的至少一项是相互不同的。

根据本发明的又一个优选实施例，还提供了用于交互式语音训练的装置，它包括一个提示序列发生器和一个参考／响应比较单元，提示序列发生器产生给使用者的提示序列并提示使用者产生相应的声音样本序列，参考／响应比较单元将由使用者产生的每一个声音样本序列的至少一个特征与一个参考信号进行比较。

此外，根据本发明的一个优选实施例，将单个使用者产生的声音样本与参考信号比较的参考信号包括一个

相应的存储参考声音样本。

此外，根据本发明的一个优选实施例，提示序列根据使用者的表现而转移。

另外，根据本发明的一个优选实施例，提示序列至少部分地由使用者指定的本国语确定。

此外，根据本发明的一个优选实施例，提示序列发生器包括一个多语言提示序列库，其中存储着以多种语言表示的多个提示序列，并且提示序列发生器根据使用者指定的其本国语的单一语言产生多种语言中的一种单一语言的提示序列。

根据本发明的另一个优选实施例，还提供了用于交互式语音训练的装置，它包括一个声音样本录音机和一个参考／响应比较单元，声音样本录音机用于录制由使用者产生的声音样本，参考／响应比较单元将使用者产生的声音样本的至少一个特征与一个参考信号进行比较。比较单元包括一个声音样本分段器和一个段比较单元，声音样本分段器用于将使用者产生的声音样本分成多个段，段比较单元用于将多个段中的至少一个中的至少一个特征与一个参考信号进行比较。

此外，根据本发明的一个优选实施例，声音样本分段器包括一个语音分段器，它将使用者产生的声音样本分成多个语音段。

另外，根据本发明的一个优选实施例，至少一个语音段包括一个语音（例如一个元音或辅音）。

根据本发明的另一个实施例，至少一个语音段可以

包括一个音节。

根据本发明的又一个实施例，提供了用于交互式语音训练的装置，它包括一个声音样本录音机和一个非特定人声音样本记分器，声音样本录音机用于录制使用者产生的声音样本，非特定人声音样本记分器根据至少一个非特定人参数对使用者产生的声音样本进行记分。

此外，根据本发明的一个优选实施例，至少一个非特定人参数包括一个用于一个预定频率上的能量数值的阈值。

此外，根据本发明的一个优选实施例，本装置还包括一个常规的个人计算机。

从以下结合附图的详细描述将理解和欣赏本发明，其中：

图1是根据本发明的一个优选实施例构造和操作的一个交互式语音教学系统的一般图示的示意图；

图2是图1系统的一个简化框图；

图3是图1系统中的一个部件的简化框图；

图4是显示用于本发明的预录制材料制备的一个简化流程图；

图5A与5B合起来是显示图1和图2装置的操作的一个简化流程图；

图6是一个语音模型重发单词“CAT”0.5秒的一个曲线图（声音幅度对时间（秒））；

图7是一个语音模型重发元音“A”0.128秒的由图6导出的一个曲线图（声音幅度对时间（秒））；

图8是一个学生尝试着重发单词“CAT”0.5秒的曲线图（声音幅度对时间（秒））；

图9是一个学生尝试着重发元音“A”0.128秒的由图8导出的曲线图（声音幅度对时间（秒））；

图10是一个学生尝试着重发单词“CAT”0.35秒的曲线图（声音幅度对时间（秒））；

图11是一个学生尝试着重发元音“A”0.128秒的由图10导出的曲线图（声音幅度对时间（秒））。

现在参看图1和图2，它们显示了根据本发明的一个优选实施例构造和操作的一个交互式语音教学系统。图1和图2的系统最好基于一个常规个人计算机10，例如一台IBM PC—AT，并且最好配备有一个辅助声音组件12。例如，一个合适的声音组件12是由美国加利福尼亚州帕洛阿尔托的Digispeech公司制造的DS201并且在商业上可从IBM教学系统中获得。一个耳机14最好与声音组件12相连接。

正如可以从图1看到，可选择地设有一台显示器30，它显示预录制的参考声音样本32和学生尝试重发34的校正声音波形。典型地显示有定量表示重发与参考声音样本之间的随时间的相似度的分数40，以给学生提供反馈。

可以采用任何合适的方法来产生相似度分数40，例如常规的相关法。在由Itakura所著的上述参考文献中描述了一种合适的方法，其公开的内容在这里参考引用。为了采用Itakura描述的距离量度，

从语音信号中抽取一阶线性预测系数。然后采用一种动态规划算法来计算学生的重复与一组模型之间的距离，即学生的重复与这些模型的相关程度。

最好，在图1的计算机10中装入合适的软件以执行图2的功能框图中提出的操作。另外，图2的结构也可以包括在一个常规的硬连线电路中。

现在参考图2的框图。图2的装置包括一个参考声音样本放音机100，可操作它给学生110播放参考声音样本。典型地通过多个语音模型的每一个预录制许多语音、单词和/或短语的每一个参考声音样本并且被存储在一个参考语音样本库120中。参考声音样本放音机100可以利用参考声音样本库120。

学生110尝试着重发每一个参考声音样本。他的口头尝试由学生响应样本接收机130接收并且最好由一个数字化转换器140数字化并存储在一个学生响应样本存储器150中。来自存储器150的每一个存储的学生响应在一个学生响应样本放音机154上可选择地放音给学生。当然，放音机100和154不必是分离的部件，图中所示的分离的方框只是为了清楚起见。

一个学生响应样本记分单元160通过利用学生响应样本接收机130用来评价参考声音样本。通过将学生的响应与由库120存取的相应参考声音样本进行比较来计算分数。

根据一个参考样本的学生响应来评价有时比最佳结果差一些，这是因为由一个单一语音模型产生的单一参

考样本不能精确地表示该样本的最佳发音。因此，可选择地或者另外，通过根据一个非特定人参考信号（例如存储在一个非特定人参数数据库170中的一组非特定人参数）评价学生响应，可以计算学生响应的分数。

根据本发明的一个优选实施例，数据库170中的非特定人参数对于说话者的年龄、性别和/或方言是特定的。换句话说，在每一个单独类型的特定的年龄、性别和/或方言的个人范围内，这些参数是与说话者无关的。

一个非特定人参数的例子是在一个取决于声音样本的特定频率上高能量的出现。例如在图6中，“猫”（CAT）波形包括第一和第三高频率、低能量部分和一个介于第一和第三部之间的且中频率、高能量的第二部分。第一和第三部分相应于CAT中的C和T声。第二部分相应于A声。

可以采用频率分析来评价响应样本。可以计算特定人参数（例如共振频率或线性预测系数），因而计算的数值可以与已知的正常范围进行比较。

学生响应样本记分单元160将参照图3进行更详细的描述。

由记分单元160导出的学生响应分数或评价在一个显示器（例如一个电视屏幕）180上显示给学生。最好，分数或评价也存储在一个学生跟踪数据库190中，数据库190累积有关每一个单独的学生为了跟踪目的的进展的信息。

系统与学生的接口最好由一个提示序列发生器200间接，可以操作提示序列发生器200给学生产生提示（例如语言提示），它既可以显示在显示器180上也可以可听地提供给学生。最好，提示序列发生器从记分单元160接收学生的分数，并可操作之将提示序列分支并将参考声音要样本提供给由其分数表明的相应的学生的进展。

根据本发明的一个优选实施例，提示序列发生器初始时给学生提供一个菜单，通过这个菜单学生可以指定其本国语。提示序列发生器最好以下列方式中的至少一种考虑学生的本国语：

（a）语言提示以其本国语提供给使用者。每一个提示以由系统支持的多个本国语的每一种语言存储在一个多语言提示库210中，提示序列发生器200可以利用多语言提示库210。

（b）提示序列和参考声音样本部分地由本国语指定而确定。例如，以希伯来语为本国语的说话者难以发英语的R声。因此，对于说希伯来语的人来说，提示序列和参考声音样本可能包括R声的基本训练。

现在参看图3，它是图2中的学生样本记分器160的一个优选实现的一个简化框图。

如上所述，作为输入记分单元160既可以直接从学生响应样本接收机130也可以间接地通过学生响应样本存储器150接收学生响应样本。响应的音量和持续时间最好由一个音量／持续时间校正器单元250用

常规方法校正。如果采用这里描述的参数抽取的线性预测编码方法，那么音量校正就不是必需的，因为在参数抽取期间音量是与其它参数相分离的。

可以采用由Itakura所著的上述参考文献中所描述的时间卷积方法校正持续时间。

如果希望只分析一个响应样本的一部分，或者希望分别分析响应样本的多个部分，那么一个分段单元260将每一个响应样本进行分段。每一段或每一部分可以包括一个语音单元（例如一个音节或语音）。例如，辅音C和T可以从一个学生的单词CAT的发音中除去，以允许单独地分析语音A。此外，每一段或每一部分可以包括一个记时单元。如果短的话，采用固定长度的段，那么持续时间校正就不是必需的了。

为了对一个响应样本进行分段，首先把静音边界（silence-speech boundary）识别为能量增高几倍于背景声级并保持高的点。可以采用任何合适的技术识别静音边界，例如在由Rabiner和Sambur所著的上述参考文献中所描述的技术，其公开的内容在这里参考引用。

接着，通过识别能量保持高但在主音频率降低至大约100至200赫兹的点，而识别辅音／元音边界。主频率可以由一个过零记数器进行测量，可操作过零记数器记数波形穿过横轴的次数。

此外，可以绕过或省去样本分段单元260，并且每一个响应样本可以作为一个单一的单元整体地进行分

析。

通过根据存储在图2中非特定人参数数据库170中的非特定人参数评价学生响应，可以操作参数比较单元280对学生响应进行记分。一个单个学生响应的分数最好代表由参数化单元270导出的单个学生响应的参数与存储在数据库170中的相应非特定人参数之间的相似度。

例如，系统可以将学生的响应样本与相应的多个存储参考样本进行比较，从而获得多个相似度数值，并且可以用这些相似度数值中的指示最大相似度的最高值作为学生响应的分数。

由参数比较单元280计算的学生响应分数最好提供给图1中的下列单元：

（a）显示器180，它用于显示给学生。可选择地，可以给学生提供一个指示分数的声音信息；

（b）学生跟踪数据库190，它用于存储；

（c）提示序列发生器200，以使提示序列发生器能适合于提示的连续顺序并使已录制的参考声音样本能适合于作为由分数表明的使用者的进展。

现在参看图4描述系统建立期间用于存储在参考声音样本库120中的预录制材料制备的一种优选方法。

如上所述，在系统建立期间，对要学习的每一个单词、语音或其它语音单元都要录制一个参考声音样本。在步骤300中，选择一组单词、语音、短语或其它声音样本。

最好，采用多个语音模型，以使之能代表多个性别、年龄和地方或民族的方言。例如，在一个设计用于英语发音教学系统中的多个语音模型可以包括以下六个语音模型：

男人——英国方言
女人——英国方言
儿童——英国方言
男人——美国方言
女人——美国方言
儿童——美国方言

在步骤310中，选择多个语音模型。由每一个语音模型产生在步骤300中选择的每一个声音样本。

在步骤320中，系统对每一个录制的声音样本进行录制、数字化并且存储在存储器中。

在步骤330中，对每一个录制的声音样本的幅度进行校正。

在步骤340中，最好将每一个录制的声音样本划分成时间段或语音段。

在步骤350中，通过从中抽取至少一个参数而将每一个录制的声音样本特征化。

现在参看照图5A—5B的流程图描述使用图1—3系统的一个典型使用者对话。

在步骤400中，给使用者提供一个语言菜单并提示他指定其本国语。此外，可以提示使用者用其本国语说一些单词，并且系统可以分析所说的单词并识别其本

国语。

在步骤405中，给使用者提供一个语音模型菜单，其选择对应于上述的多个语音模型，并且菜单提示使用者选择最适合他的语音模型。

在步骤410中，提示使用者选择一个初始参考声音样本（例如一个语音、单词或短语）以用于练习。此外，用于练习的样本可以由系统选择，最好部分地根据在步骤400中使用者指定的其本国语。

步骤420—给使用者播放参考声音样本，并且可选择地，其波形同时显示给使用者。

步骤430—使用者尝试着对参考声音样本的重发由系统接收、数字化并存储在存储器中。

步骤450—系统对声音电平和重发声音样本的持续时间进行校正。

步骤460—可选择地，重放重发声音样本并将重发声音样本的校正波形显示给使用者。

步骤490—系统通过样本的参数化从重发声音样本中抽取声音特征（例如线性预测系数）。合适的声音特征抽取方法在由Itakura听著的上述参考文献以及其中所引用的参考文献中进行了描述，其公开的内容在这里参考引用。

步骤500—系统将步骤490中抽取的参数与参考声音样本的存储特征进行比较并计算相似度分数。

步骤510—系统显示相似度分数。

步骤520—系统最好重放参考和重发样本，以供使用者进行声音比较。

步骤530—可选择地，系统存储相似度分数和/或重发样本本身，以用于随后的跟踪。

步骤540—除非系统或学生决定对话终止，否则系统返回步骤410。参考样本的系统选择最好考虑学生的表现。例如，如果对于一个特定的参考声音样本的相似度的分数低（表明学生的表现差），那么可以重复参考声音样本直到获得一个最低线为止。接着，可以采用一个相似的参考声音样本以确保获得的表现水平推广到相似的语音任务。

例如，如果使用者在重发CAT中的A时有困难，那么可以重复地出现样本CAT并且可以跟随着包括A的其它样本（例如BAD）。

图6—11是多个语音模型和学生产生的语音样本的波形曲线图。

图6表示一个语音模型重发单词“CAT”0.5秒的情形。因不是从图6所示的单词“CAT”的语音模型的重发中除去辅音获得一个语音模型重发元音“A”0.128秒的曲线图。如上所述，通过寻找“CAT”中的辅音元音边界，识别元音“A”的起点。根据本发明的实施例，每一个元音的持续时间预定的。已经发现0.128秒的预定元音持续时间能提供满意的结果，然而这个数值并不打算是限制性的。

根据本发明的另一个实施例，每一个元音的持续时

间不是预定的。取尔代之的是，通过对语音样本的合适的分析，识别元音／辅音边界。

图8是一个学生尝试着重发单词“CAT”0.5秒的曲线图。

图9是从图8所示的单词“CAT”的语音模型的重发中除去辅音获得的一个语音模型重发元音“A”0.128秒的曲线图。

图10是一个学生尝试着重发单词“CAT”0.35秒的曲线图。图11是从图9所示的单词“CAT”的语音模型的重发中除去辅音获得的一个语音模型重发元音“A”0.128秒的曲线图。

熟知本领域的技术人员应当理解的是：本发明并不限于以上所特别图示和描述的内容。相反地，本发明的范围只由随后的权利要求书所限定。

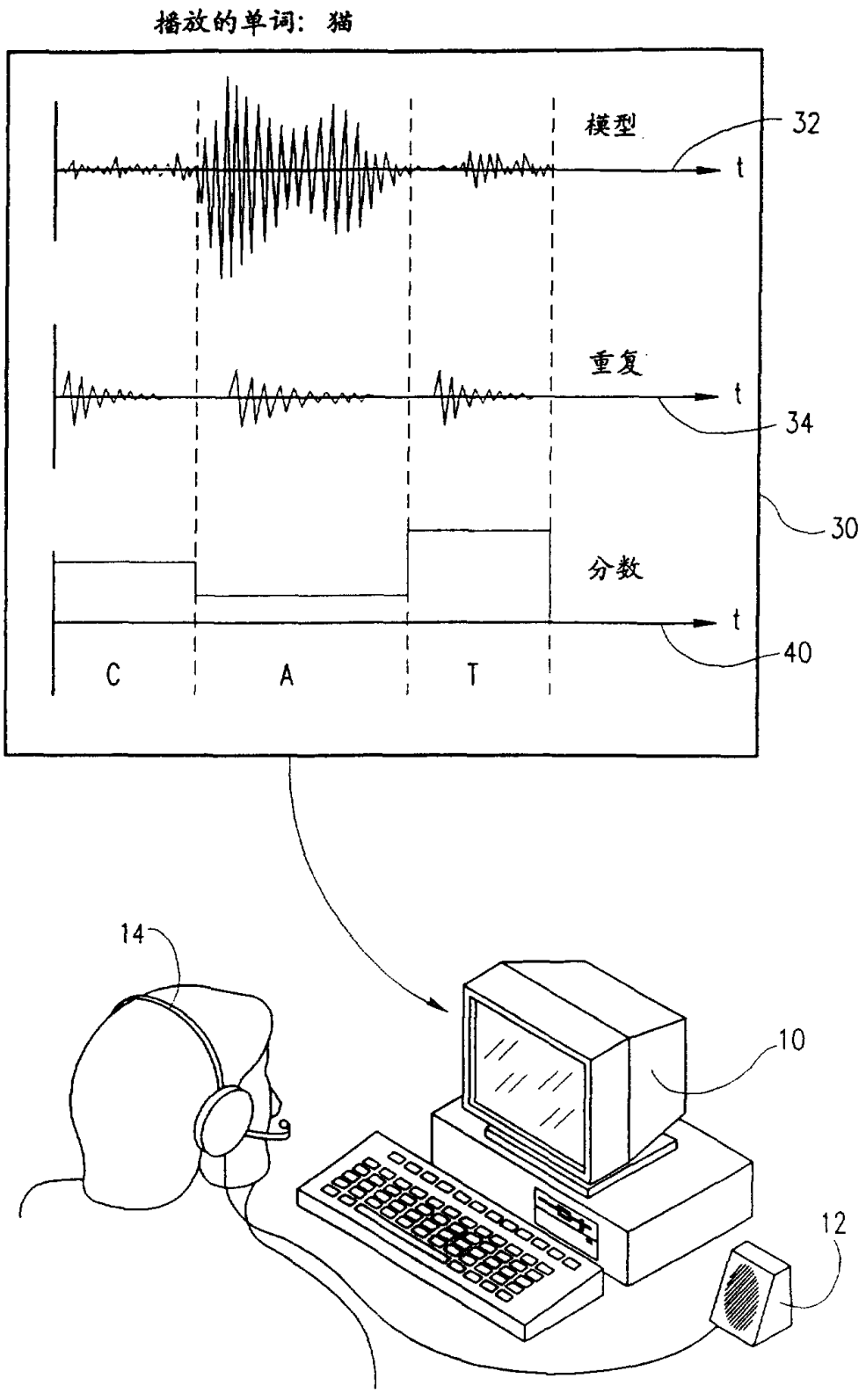


图 1

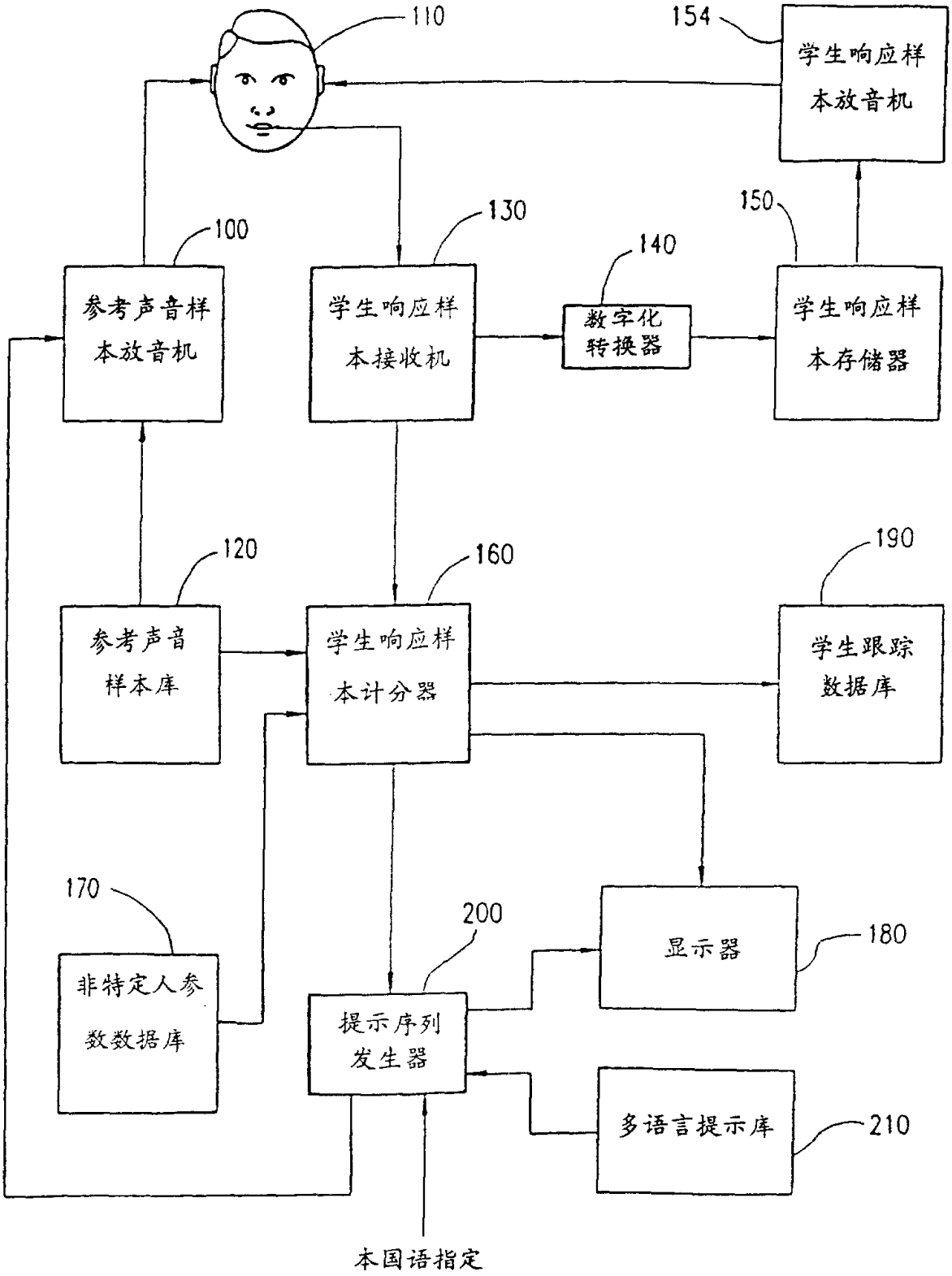

110
154
学生响应样本放音机
100
130
140
150
参考声音样本放音机
学生响应样本接收机
数字化转换器
学生响应样本存储器
120
160
190
参考声音样本库
学生响应样本计分器
学生跟踪数据库
170
200
180
非特定人参数数据库
提示序列发生器
显示器
多语言提示库
210
本国语指定

图 2

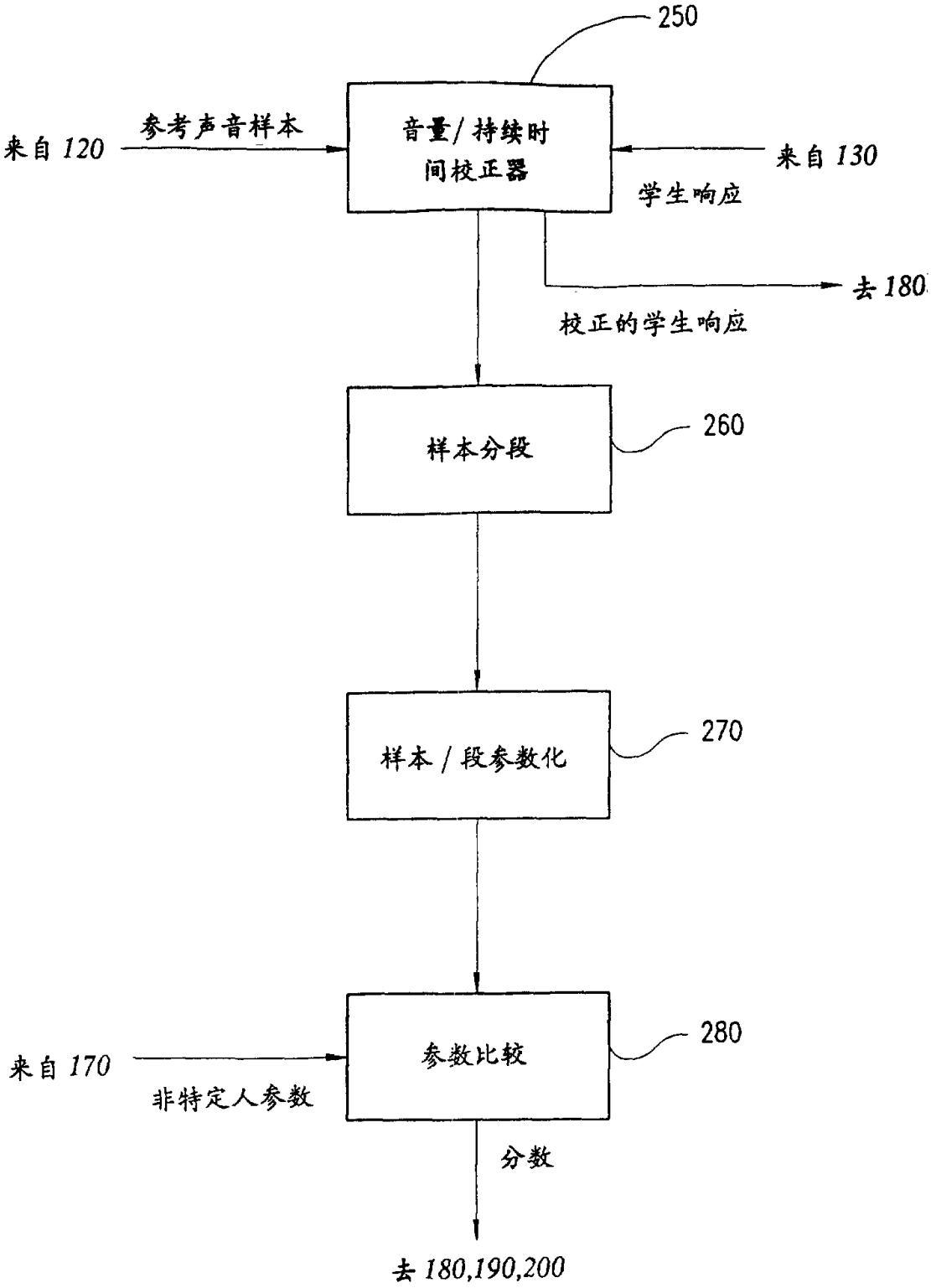

250
来自 120
参考声音样本
音量/持续时间校正器
来自 130
学生响应
去 180
校正的学生响应
样本分段
260
样本 / 段参数化
270
来自 170
非特定人参数
参数比较
280
分数
去 180,190,200

图 3

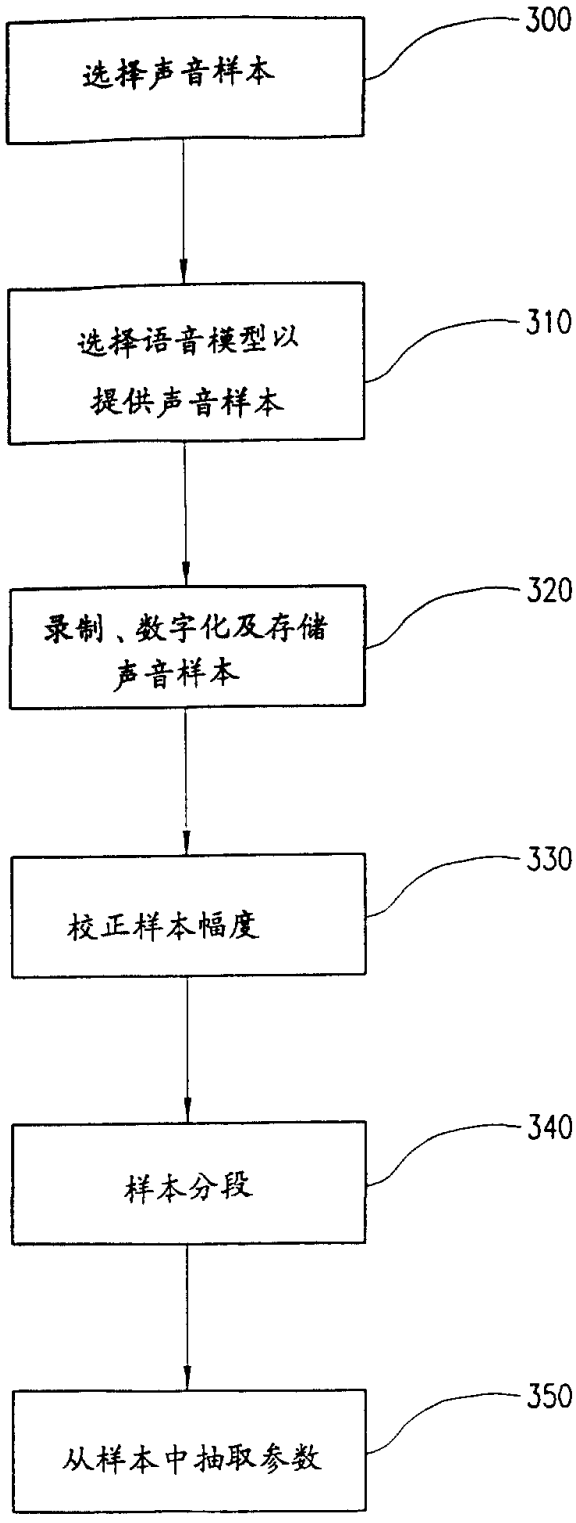

选择声音样本
300
选择语音模型以
提供声音样本
310
录制、数字化及存储
声音样本
320
校正样本幅度
330
样本分段
340
从样本中抽取参数
350

图 4

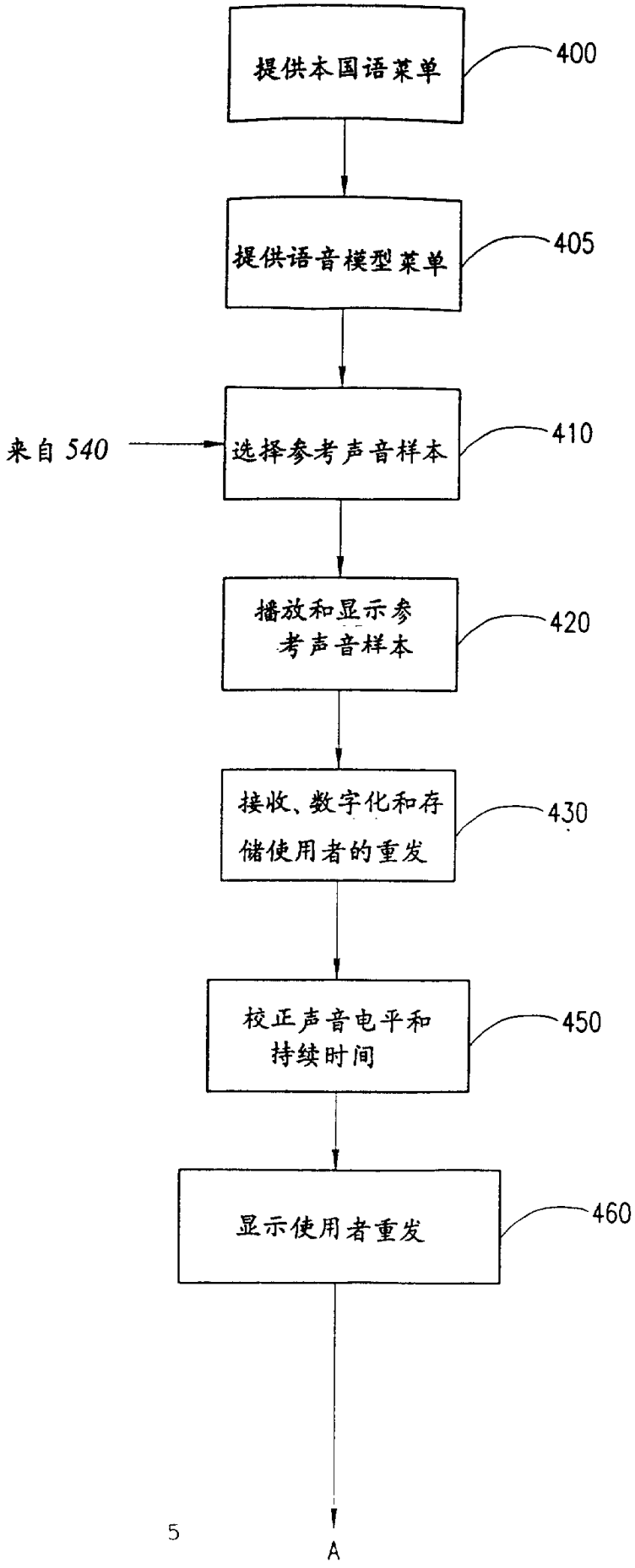

提供本国语菜单
400
提供语音模型菜单
405
来自 540
选择参考声音样本
410
播放和显示参考声音样本
420
接收、数字化和存储使用者的重发
430
校正声音电平和持续时间
450
显示使用者重发
460
A

图 5A

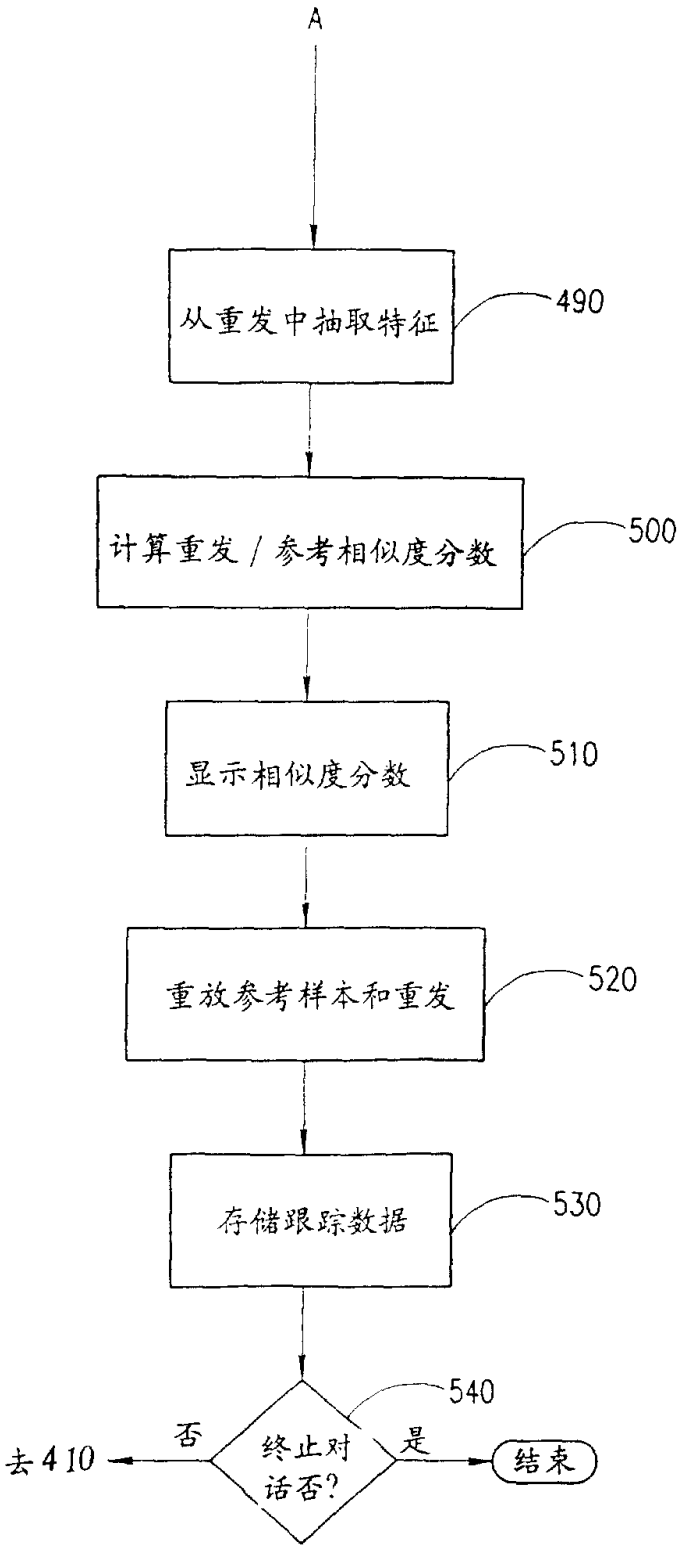

A
从重发中抽取特征
490
计算重发 / 参考相似度分数
500
显示相似度分数
510
重放参考样本和重发
520
存储跟踪数据
530
540
否
终止对话否?
是
结束
去410

图 5B

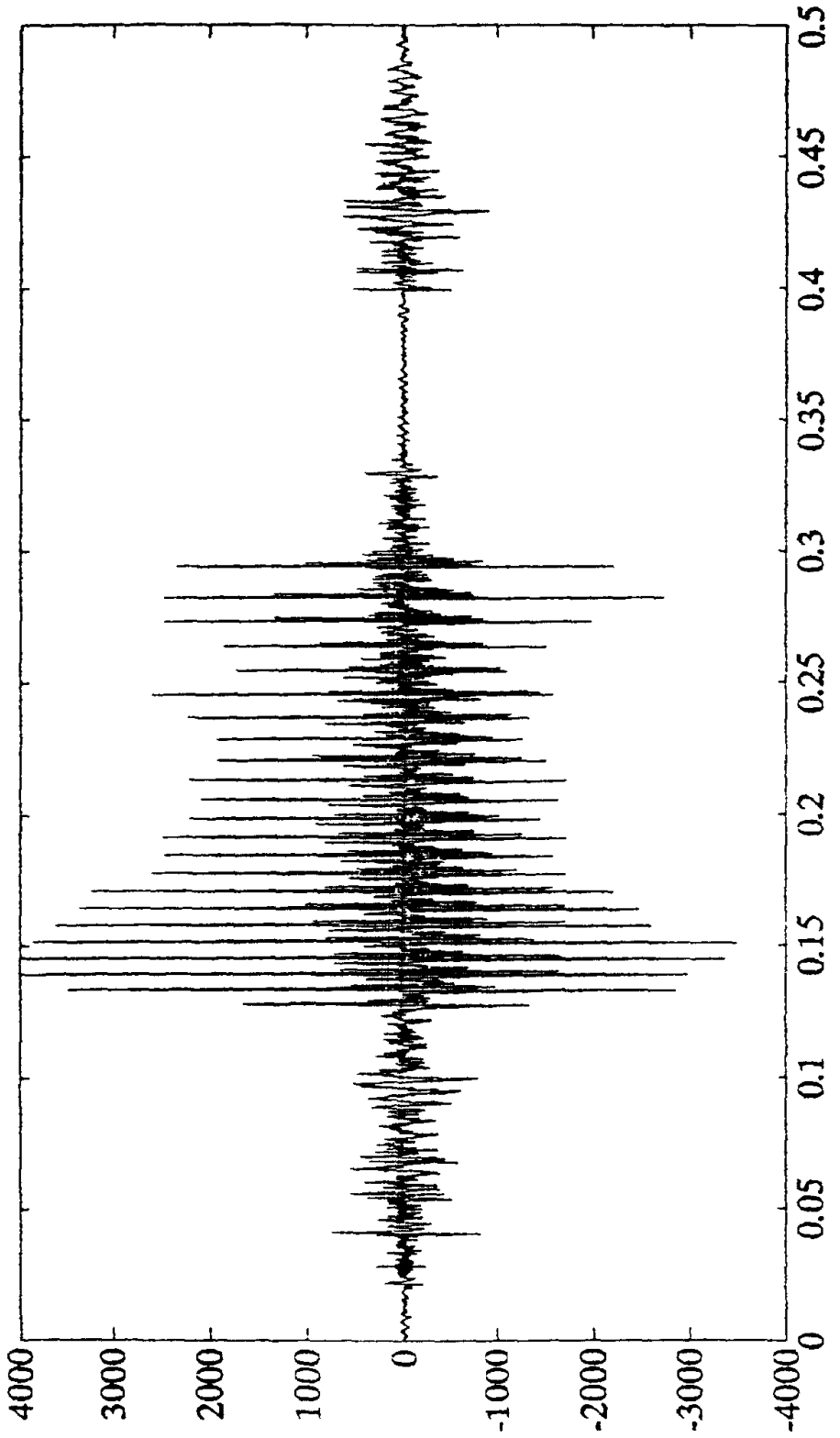

4000
3000
2000
1000
0
-1000
-2000
-3000
-4000
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5

图 6

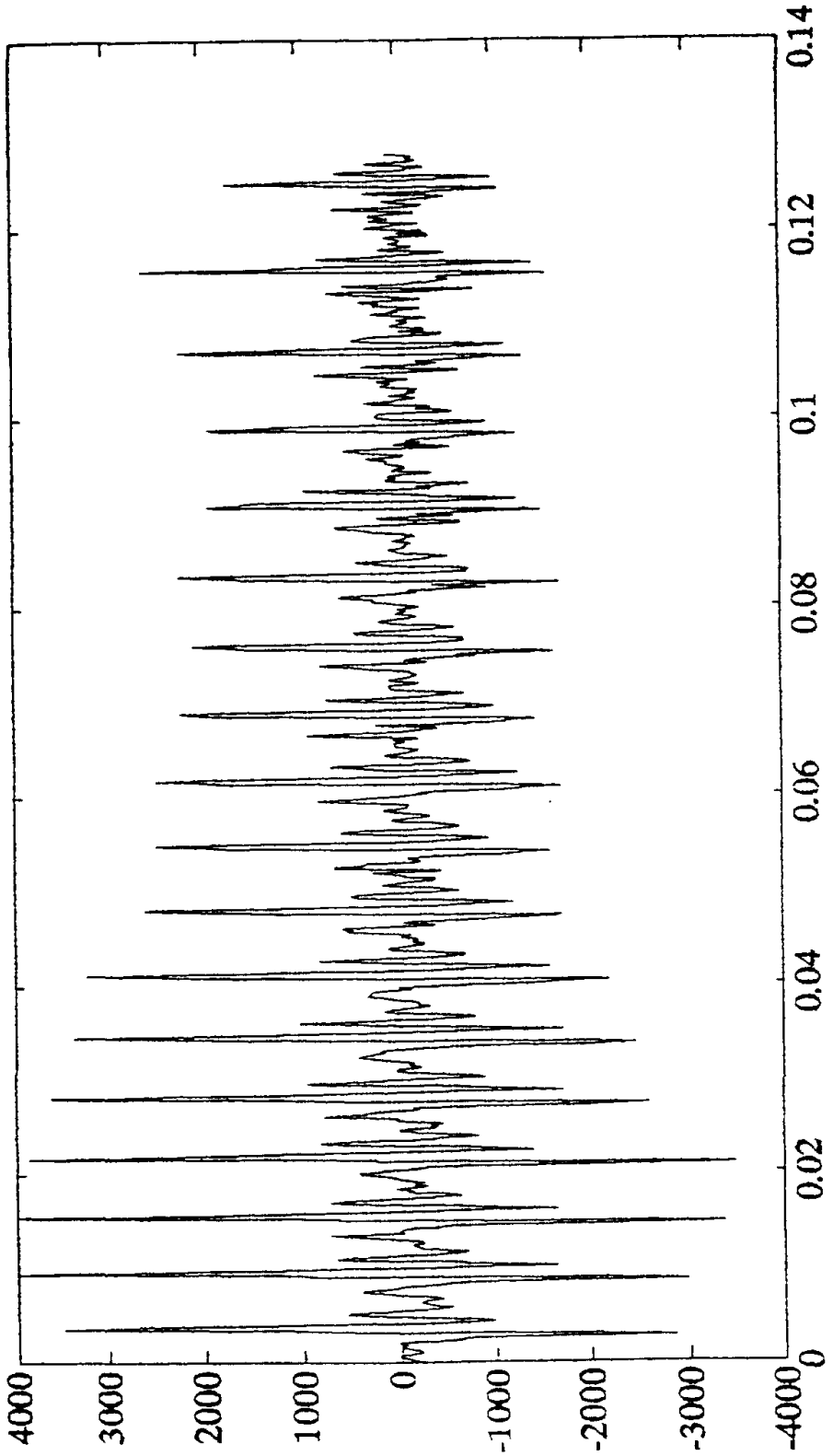

4000
3000
2000
1000
0
-1000
-2000
-3000
-4000
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14

图 7

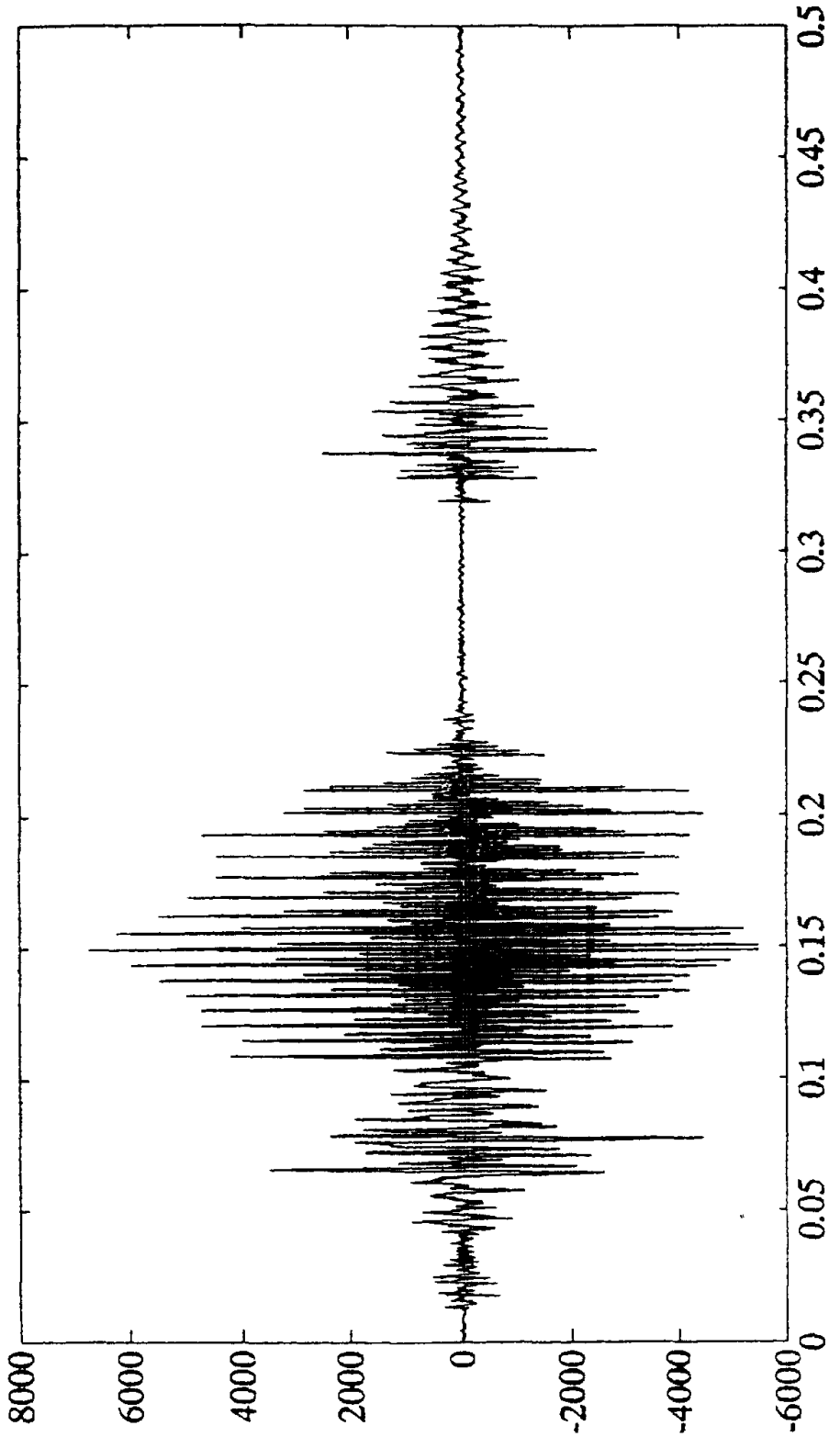

8000
6000
4000
2000
0
-2000
-4000
-6000
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5

图 8

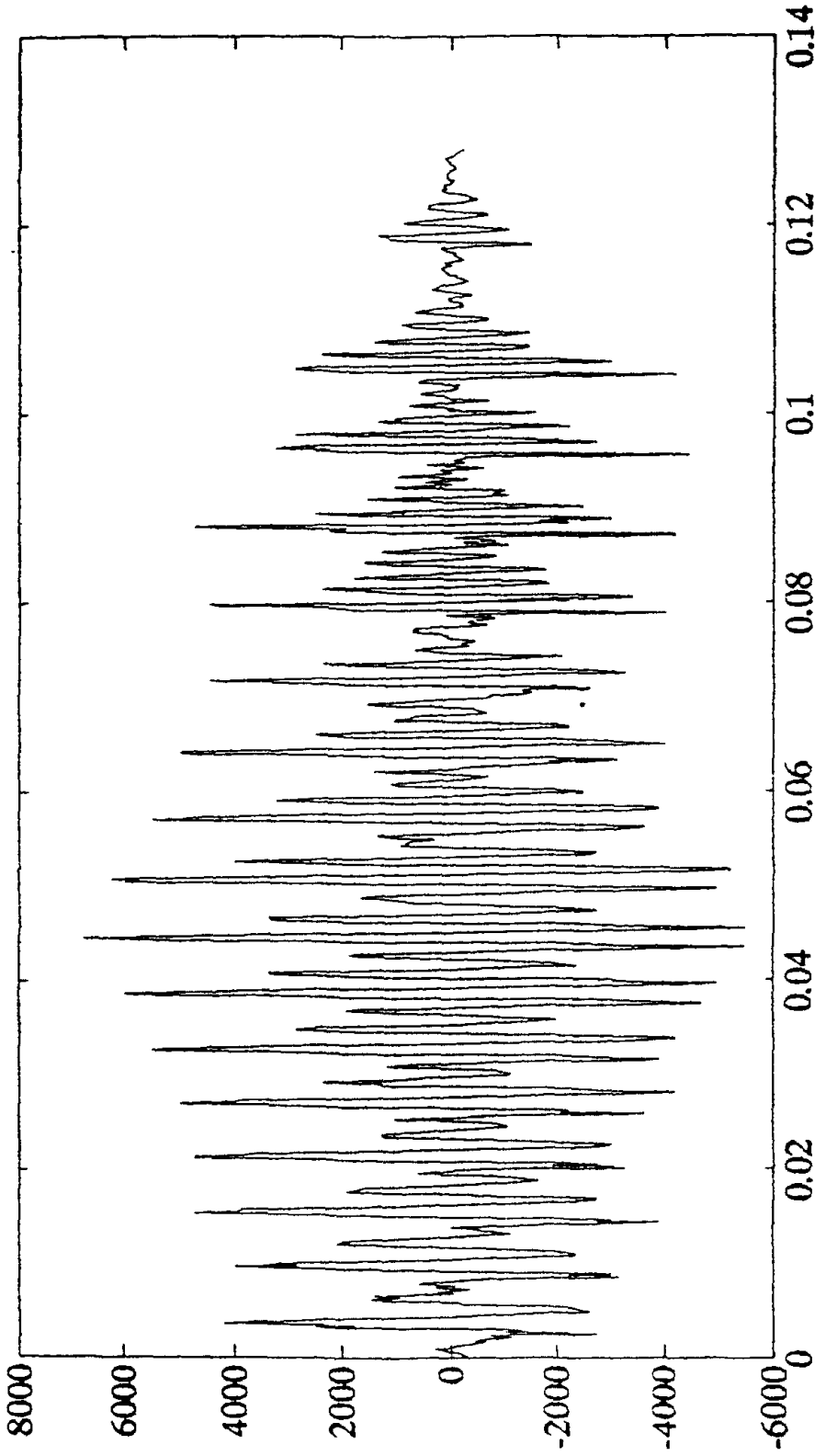

8000
6000
4000
2000
0
-2000
-4000
-6000
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14

图 9

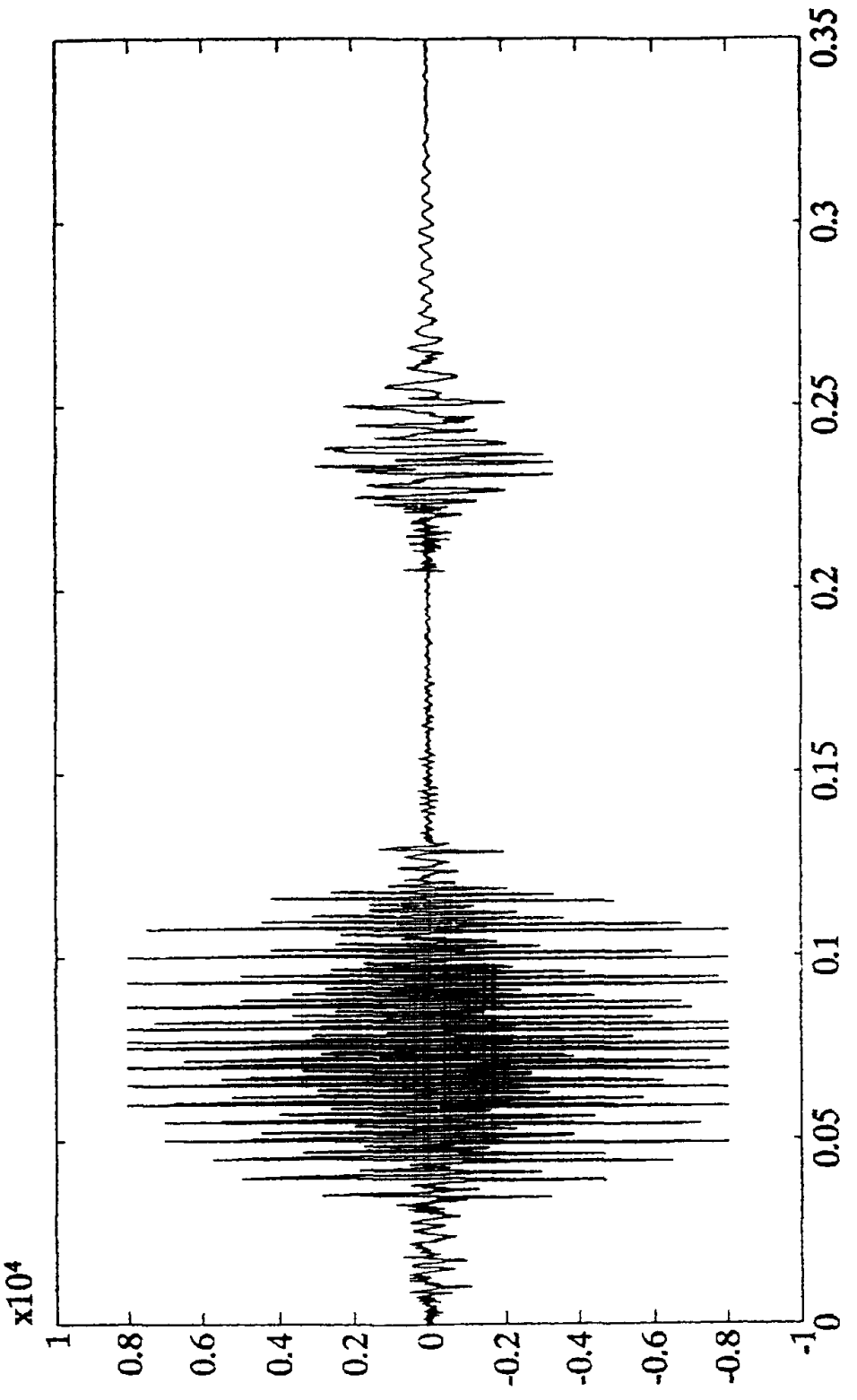


图 10

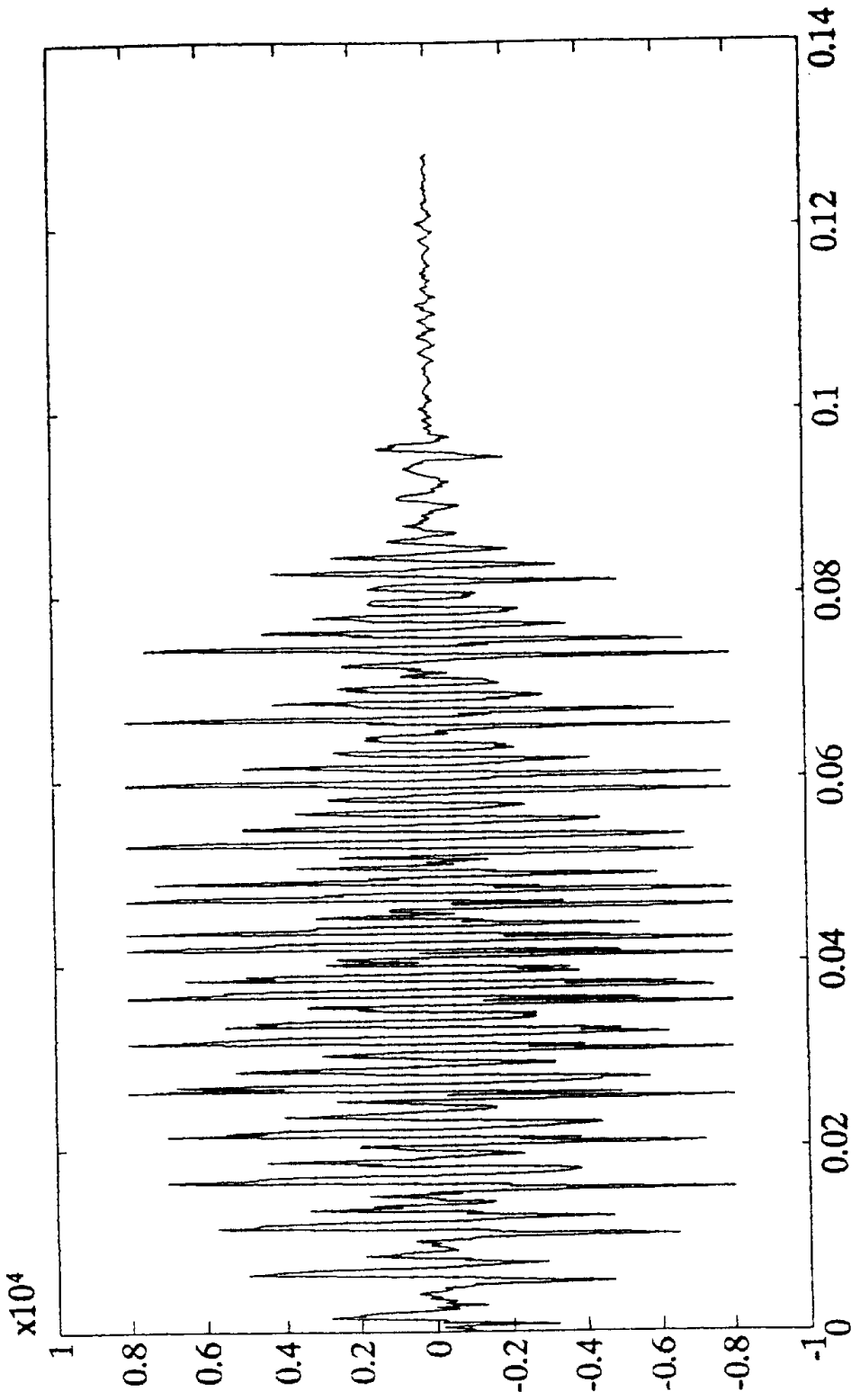


图 11