

6-1-2020

## Research of Chinese Diphthongs Pronunciation Based on DIVA Model

Shaobai Zhang

*Computer Department, Nanjing University of Posts and Telecommunications, Nanjing 210046, China;*

Yanli Chen

*Computer Department, Nanjing University of Posts and Telecommunications, Nanjing 210046, China;*

Liwen He

*Computer Department, Nanjing University of Posts and Telecommunications, Nanjing 210046, China;*

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the Artificial Intelligence and Robotics Commons, Computer Engineering Commons, Numerical Analysis and Scientific Computing Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Systems Science Commons

---

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

---

## Research of Chinese Diphthongs Pronunciation Based on DIVA Model

### Abstract

**Abstract:** DIVA (Directions Into Velocities of Articulators) is a kind of adaptive neural network model which is used to simulate and describe some associative functions in brain regions involved speech production and understanding. DIVA takes 29 essential English phonemes as its language background. Since the number of Chinese pronunciation phonemes is much larger than English and the pronunciation brain mechanisms of both also make a big difference, the adaptability of DIVA model for Chinese background has to be studied specially, in order that the model can read out the thinking processes in Chinese brain. *Based on DIVA, the Chinese pronunciation of diphthongs was explored and related issues on Chinese brain regions involved speech production and acquisition were discussed. The new modified model can distinguish Chinese vowels from English vowels clearly by adjusting formant and the parameters of the corresponding pronunciation organs in DIVA's simulative vocal tract.* This research lays a solid foundation for further comprehensive Chinese speech production and acquisition on DIVA model.

### Keywords

DIVA model, speech sound, Chinese compound vowel, LPMCC

### Recommended Citation

Zhang Shaobai, Chen Yanli, He Liwen. Research of Chinese Diphthongs Pronunciation Based on DIVA Model[J]. Journal of System Simulation, 2017, 29(2): 255-263.

# 基于 DIVA 模型的中文复合元音发音方法研究

张少白, 陈燕俐, 何利文

(南京邮电大学计算机学院, 江苏 南京 210046)

**摘要:** DIVA(Directions Into Velocities of Articulators)模型是一种被用来对涉及大脑中有关语音生成和理解区域的功能进行仿真和描述的自适应神经网络模型, 其依赖的语言背景是英文 29 个基本音素。由于汉语与英语发音区别很大, 且加工脑机制也颇为不同, 要想将汉语者大脑思维过程“阅读”出来, 需要对模型汉语背景的适应性进行专门研究。在 DIVA 模型的基础上研究汉语复合元音的发音方法, 探讨汉语者脑区语音生成与获取的相关问题。通过调节模型的共振峰以及模拟声道对应器官的参数, 新构建的模型能很好地辨识汉语与英语元音的区别。该研究为 DIVA 模型汉语语音生成与获取奠定了良好的基础。

**关键词:** DIVA 模型; 语音发音; 汉语复合元音; LPMCC

中图分类号: TP183      文献标识码: A      文章编号: 1004-731X (2017) 02-0255-09

DOI: 10.16182/j.issn1004731x.joss.201702004

## Research of Chinese Diphthongs Pronunciation Based on DIVA Model

Zhang Shaobai, Chen Yanli, He Liwen

(Computer Department, Nanjing University of Posts and Telecommunications, Nanjing 210046, China)

**Abstract:** DIVA (Directions Into Velocities of Articulators) is a kind of adaptive neural network model which is used to simulate and describe some associative functions in brain regions involved speech production and understanding. DIVA takes 29 essential English phonemes as its language background. Since the number of Chinese pronunciation phonemes is much larger than English and the pronunciation brain mechanisms of both also make a big difference, the adaptability of DIVA model for Chinese background has to be studied specially, in order that the model can “read out” the thinking processes in Chinese brain. *Based on DIVA, the Chinese pronunciation of diphthongs was explored and related issues on Chinese brain regions involved speech production and acquisition were discussed. The new modified model can distinguish Chinese vowels from English vowels clearly by adjusting formant and the parameters of the corresponding pronunciation organs in DIVA’s simulative vocal tract.* This research lays a solid foundation for further comprehensive Chinese speech production and acquisition on DIVA model.

**Keywords:** DIVA model; speech sound; Chinese compound vowel; LPMCC

## 引言

神经分析系统(Neuralynx System)是美国波士



收稿日期: 2016-03-15      修回日期: 2016-08-09;  
基金项目: 国家自然科学基金(61271334, 61373065);  
作者简介: 张少白(1953-), 男, 河北定县, 博士, 教授, 研究方向为智能系统与模式识别; 陈燕俐(1969-), 女, 湖北襄阳, 博士, 教授, 研究方向为智能系统与模式识别。

顿大学语音实验室冈瑟教授(Frank H. Guenther)所研制的一种新型的仪器<sup>[1]</sup>, 利用这种仪器, 可以将使用者头脑里所想象的东西, 用语音合成系统转换成语音来表述。系统主要由脑—机接口(Brain-Computer Interface, BCI)和 DIVA(Directions Into Velocities of Articulators)模型两部分所组成。其中, BCI 负责将检测到的脑电神经信号用来驱动

语音合成器的连续“运动”，为使用者提供实时语音输出；DIVA 模型则是一种在生物学意义上所构建的有关语音获取和生成的神经网络模型<sup>[2-3]</sup>，主要依据有关语音生成及感知心理物理学实验的行为数据、fMRI(functional magnetic resonance imaging)和 PET(Positron Emission computed Tomography)实验数据以及各种运动控制实验的神经生理学数据等而建立，目的是实施对声道运动的模拟以及对涉及到大脑中有关语音生成和理解区域的功能进行仿真和描述<sup>[2]</sup>，其原理如图 1 所示。

DIVA 模型的研究背景是英文发音的 29 个基本音素<sup>[3]</sup>，与最少 73 个发音音素的汉语相比，其大脑发音机制相差很大。要想探讨汉语者大脑里的思维过程，“阅读”汉语者大脑里所想象的东西，需要对 DIVA 模型汉语背景的应用结构进行专门研究。汉语元音分单元音和复合元音。对单元音的发音方法

我们已经有了初步的研究成果<sup>[4-7]</sup>，本文拟探讨应用 DIVA 模型来进行汉语复合元音的发音问题。

相对于单元音的发音，汉语复合元音的发音方法要复杂一些，有许多特殊的问题需要考虑，因而其生成与获取的方式也有很大不同。本文在 DIVA 模型的基础上对汉语复合元音特别是二合元音在模型中的发音机理和基本方法进行研究，主要工作如下：(1)探讨了应用 DIVA 模型处理汉语复合元音的可能性与适用性；(2)以汉语二合元音为例，在讨论汉英两种语言发音脑机制异同点的基础上，对 DIVA 模型进行了修改和重构；(3)应用修改后的 DIVA 模型对汉语二合元音的发音进行了仿真实验。

研究表明，改进后的 DIVA 模型能够正确生成汉语二合元音发音，且发音效果良好，这为以后继续研究汉语其它发音的生成与获取奠定了良好的基础。

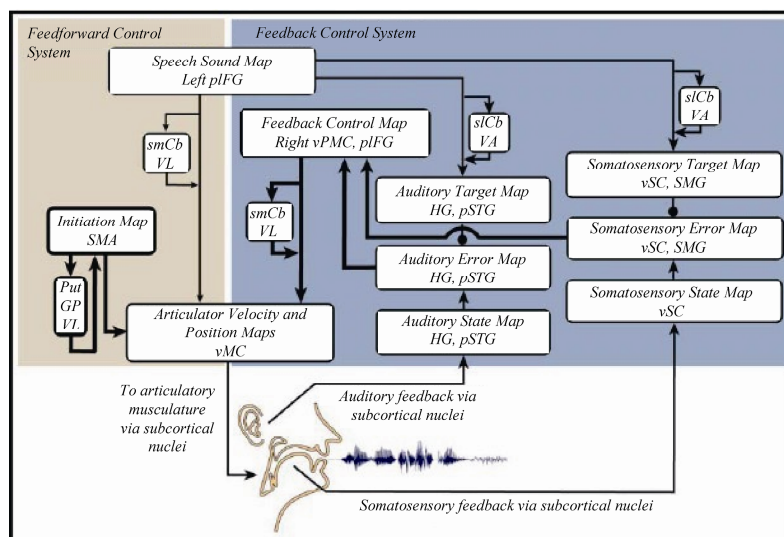


图 1 DIVA 模型示意图

Fig.1 A schematic diagram of the DIVA model

## 1 汉语复合元音发音机制与 DIVA 模型

### 1.1 汉语元音发音机制

从语言学角度来说，语音可以分为音素和音节。音素是语音的最小单位。音节是由一个或多个音素按照一定规律组合而成的、可以从听觉中直接

觉察的语言自然单位。

元音分为单元音和复合元音。无需与其它元音结合就能在音节中单独存在的元音称为单元音。单元音作为语言中的基本单位，其音质随着时间的延续基本保持不变。复合元音则相反，需要与其它元音结合才能在音节中存在，其音色一直处于变化状

态, 且这种变化是由调音器官连续运动产生的。复合元音有二合元音和三合元音之分。二合元音的两个组成成分在发音音质、响度以及稳定性等性能指标上均有所不同, 这实质上表明了单元音与二合元音甚至是三合元音的不同<sup>[8]</sup>。实际上, 不同类型的元音其共振峰的差异性在时间维度上的表现也是非常明显的, 这一点文献[8]中有详细说明。

对复合元音的解释与定义, 中文和英文不尽相同, 但基本认识一致, 即复合元音在音节内有滑移, 且发音时发音器官(例如舌、唇等)的相关部位会不断变化, 随之而来的是声道形态以及所发出声音的音色也会不断变化。因此可以说, 复合元音是一种动态变化的元音。

汉语普通话所包含的二合元音有[ai]、[ei]、[ao]、[ou]、[ia]、[ie]、[ua]、[uo]、[üe]等 9 个, 图 2 是这几个复合元音的声学空间图。图 2 中, 纵坐标由第一共振峰(F1)所组成, 横坐标则由第二共振峰(F2)所组成。

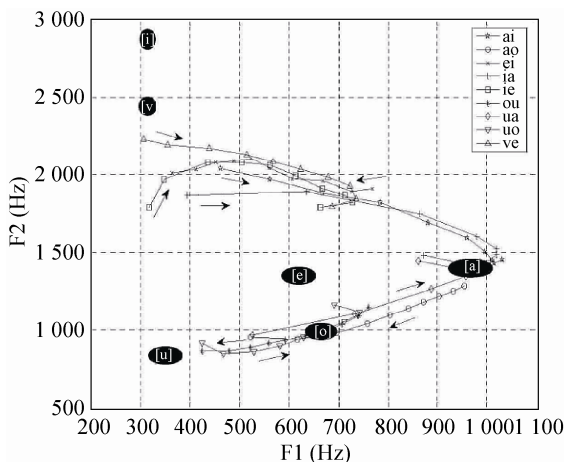


图 2 二合元音的 F1-F2 轨迹  
Fig.2 Trajectories of diphthongs in F1-F2

### 1.2 DIVA 模型中的汉语复合元音发音需考虑的问题

如前所述, 复合元音不是简单的两个或者三个元音合在一块读就可以的。相较于单元音, 它不仅是元音数量上的增多, 更是音质上的跨越。在一个音节里, 组合成汉语复合元音的单元音的目标值前

后不一致, 发音时嘴唇和舌位是从一个元音位置过渡到另一个元音位置的。因此, 在 DIVA 模型中要充分考虑到这样的因素, 例如: (1)同样是提取共振峰, 但单元音的提取方法不一定就适用于复合元音; (2)DIVA 中模拟声道参数的配置应与汉语复合元音发音的脑机制相一致; (3)单元音在多个发音之间不需要考虑相关的过渡关系, 但复合元音需要考虑这样的问题。也就是说, 元音间的线性拟合关系是复合元音发音所需要考虑的一个重要问题, 仅此才能实现多音素间的良好过渡。

## 2 共振峰频率及获取方法

语音发音过程中, 共振峰是一种用来描述声学共振现象反映声道特征的重要参数, 其基频为 F1, 接着为 F2, F3.....以此类推。文献[9]指出, F1 主要反映的是咽腔大小以及共鸣形态。医学常识表明, 咽腔大小由下颌运动方向所决定。当方向向下时, 口腔体积变大, 咽腔体积变小, 进而 F1 变大; 而当方向向上时, 口腔体积减小, 咽腔体积增大, F1 随之减少。也就是第一共振峰频率 F1 与下颌的开合度成正比<sup>[9]</sup>。

F2 反映的是口腔大小及共鸣状态。同样由医学常识表明, 口腔大小主要由舌的前后运动方向所决定。也就是说, 第二共振峰频率 F2 与舌体前后运动的方向直接相关。舌体向前, F2 增大, 舌体向后, F2 减小<sup>[9]</sup>。依据这些原理, 我们用 F2 作为横坐标, F1 作为纵坐标, 并将数值从大到小排列, 可以得到 F1-F2 动态舌位图, 如图 3 所示。

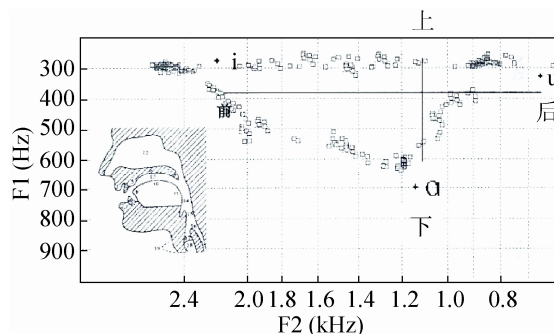


图 3 元音 F1-F2 舌位图  
Fig.3 Tongue bitmap of vowels in F1-F2

DIVA 模型的输入元素是共振峰频率。为适应汉语二合元音的发音,需要研究汉语复合元音的共振峰走势对模型的影响。文献[10]指出,复合元音的共振峰有着明显的动态特性,即当复合元音从一个音位向另一音位运动时会导致相关共振峰频率随时间变化而变化,从而使语音的每个音位或多或少地偏离各自的目标值,最终形成具有其特定形态

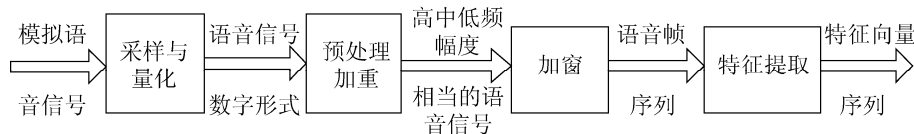


图 4 语音特征提取过程

Fig.4 Processes of speech feature extraction

2. 共振峰频率的计算方法常用的有: 基于线性预测(Linear Prediction Coefficient, LPC)的求根法<sup>[11]</sup>、倒谱法<sup>[12]</sup>和梅尔频率倒谱系数(Mel Frequency Cepstrum Coefficient, MFCC)法<sup>[13]</sup>。这几种方法中,与 LPC 等方法通过对发声机理的研究而得到声学特征不同, MFCC 是将语音获取机制与人耳的听觉特征结合起来,有一个将线性频谱转化为 Mel 对数频谱,最后再变换到倒谱域的过程。

本文拟采用线性预测和 Mel 倒谱相结合的方法,即 LPMCC (Linear Prediction and Mel Frequency Cepstrum Coefficient)法,用 MATLAB 仿真实现。基于该方法的指导思想是: Mel 频带划分考虑的对象主要是人耳听觉特征,在划分过程中,除了对音调高低要保持敏感外,还需要对声音响度予以感知,且这种感知与声频带宽是有密切关系的。为了更好地模拟听觉过程,将语音信号的频谱变换到感知频域中的方法无疑是一种较好的选择。因此,我们将常规 LPC 方法进行非线性变换,让其具有符合听觉特征的 Mel 尺度,从而在此基础上求得 LPMCC。

3. 采用 LPMCC 方法获取汉语复合元音共振峰频率的过程与我们先前所采用单元音获取过程相似(参见文献[7]),但重点考虑了双元音的特点,主要为: (1)对复合元音进行预处理,内容包括预

的共振峰动态模式。本文利用二合元音[ai]、[ao]、[ei]、[ou]、[ia]、[ie]、[ua]、[uo]、[üe]的这种特性,并结合适合这种特性的共振峰提取方法来获得符合 DIVA 模型要求的共振峰频率。

## 2.1 共振峰频率获取方法

1. 语音特征提取过程如图 4 所示。

滤波和消除噪音干扰等; (2)对处理后的复合元音分别进行分帧、预加重、加窗变换、端点检测,以确定起始端点和结束端点; (3)利用线性预测 Mel 倒谱系数对数据进行分析,获取双元音的共振峰频率; (4)针对所获取的双元音进行中值滤波和线性平滑等相关处理。

## 2.2 LPMCC 计算方法

LPMCC 的处理关键在于将常规 LPCC 方法进行非线性变换,让其符合人耳听觉特征的 Mel 尺度。在将 LPCC 通过 Mel 滤波器进行滤波时,特别要考虑带通滤波器的输出是 Mel 滤波器带宽内所有信号的加权幅值,依照这样的方法即可求出 LPMCC。过程见式(1)和式(2)。

$$H_m(k) = \begin{cases} 0 & k < f(m-1) \\ \frac{k-f(m-1)}{f(m)-f(m-1)} & f(m-1) < k \leq f(m) \quad (0 \leq m \leq M) \\ \frac{f(m+1)-k}{f(m+1)-f(m)} & f(m) < k \leq f(m+1) \\ 0 & k > f(m+1) \end{cases} \quad (1)$$

式中:  $M$  是滤波器个数,  $f(m)$  是 Mel 滤波器的中心频率。  $f(m)$  定义为:

$$f(m) = \left( \frac{N}{F_s} \right) F_{\text{mel}}^{-1} \left( F_{\text{mel}}(f_l) + m \frac{F_{\text{mel}}(f_h) - F_{\text{mel}}(f_l)}{M+1} \right)$$

其中:  $f_h$ ,  $f_l$  分别为滤波器的上下限频率;  $N$  表示帧长;  $F_s$  是采样频率;  $F_{mel}^{-1} = 700(e^{b/1125} - 1)$  是  $F_{mel}$  的逆函数。LPMCC 的递推式如式(2)。

$$MC_k(n) = \begin{cases} C_n + \alpha \cdot MC_0(n+1) & k=0 \\ (1-\alpha^2) \cdot MC_0(n+1) + \alpha \cdot MC_1(n+1) & k=1 \\ MC_{k-1}(n+1) + \alpha(MC_k(n+1) - MC_{k-1}(n)) & k>1 \end{cases} \quad (2)$$

其中:  $C_n$  和  $MC_k$  分别表示倒谱系数和 Mel 倒谱系数,  $n$  和  $k$  则表示迭代次数和倒谱阶数,  $n=k$ 。迭代时,  $n$  的取值范围为  $k, k-1, k-2, \dots, 0$ 。  $MC_0(0), MC_1(0), \dots, MC_k(0)$  里存放所求得的 Mel 倒谱系数结果。通常情况下,  $\alpha$  取值范围为 0.31~0.35, 目的是使其近似于 Mel 尺度。

### 2.3 二合元音共振峰频率的获取方法

为了深入了解复合元音中相邻音位如何相互影响, 可以通过使用单因素重复测量的方法对相似组共振峰的拟合斜率进行比较<sup>[14]</sup>。通过对比较结果的研究, 可以探测到二合元音/ai、ei、au、ou、ia、ua、ie、üe、uo/共振峰的走势<sup>[15-16]</sup>。例如应用这样的方法, 我们可以得到/ai/的发音走势, 见图 5。

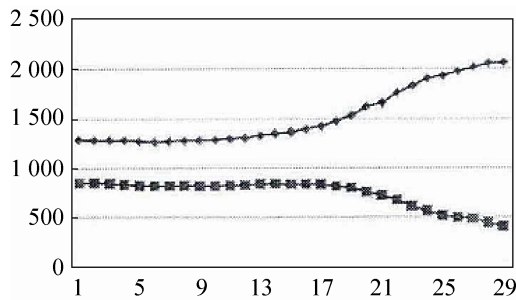


图 5 汉语复合元音/ai/的共振峰走势图  
Fig.5 Formant trend chart of Chinese diphthong /ai/

在此基础上, 再对类似图中的多个二合元音轨迹采用二次多项式分别进行拟合, 就可以得到如下结果:

$$\begin{aligned} ai: F2 &= -0.003F1^2 + 2.766F1 + 1429.3 \\ ao: F2 &= 0.006F1^2 - 9.95F1 + 4394.1 \\ ei: F2 &= -0.002F1^2 + 1.656F1 + 2322.6 \\ ou: F2 &= 0.008F1^2 - 9.87F1 + 3816 \end{aligned}$$

$$\begin{aligned} ia: F2 &= -1.71F1 + 3370.2 \\ ie: F2 &= -0.002F1^2 + 2.178F1 + 2296.3 \\ ua: F2 &= 0.003F1^2 - 3.554F1 + 2316.9 \\ uo: F2 &= 0.01F1^2 - 9.7F1 + 3189.6 \\ ue: F2 &= 0.001F1^2 - 1.604F1 + 2783.9 \end{aligned}$$

再将以上结果与 LPMCC 获取方法结合, 9 个二合元音的共振峰频率就可以被求出, 见表 1。

表 1 汉语二合元音共振峰频率表 /Hz

二合元音	F1	F2	F3
ai	654	1953	2970
ao	729	1724	2851
ei	641	2560	3014
ou	663	1216	3003
ia	645	2266	2666
ie	418	2827	2315
ua	524	1204	2431
uo	710	1362	2217
üe	370	2312	3120

## 3 汉语复合元音的声道配置与仿真

### 3.1 DIVA 模型接口

应用 DIVA 模型模拟汉语元音发音时, 要用到一个用户接口, 其功能为方便用户进行发音机制的控制, 如图 6 所示。该接口分为三个部分, 即: 1. DIVA 模型控制模块; 2. 声学表征空间模块; 3. 声道控制模块。

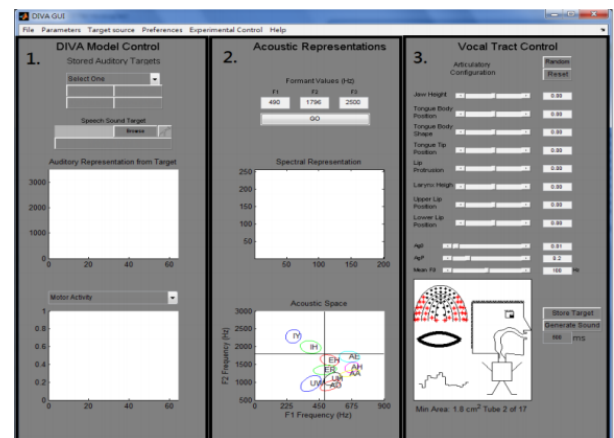


图 6 DIVA 模型用户接口  
Fig.6 User interface of DIVA

### 3.2 听觉感知功能模块的添加

DIVA 模型接口中虽然设有虚拟听觉声学空间,目的是给用户提供一个对想要发出的元音种类进行控制。但限于各种原因,模型中元音的分类是通过人为设置的方法而生成的,这是 DIVA 模型的一个遗憾。由于涉及到模型结构,我们在文献[17]中对此作了修改。修改后的模型接口如图 7 所示,图中黑框部分为添加的感知模型。依据感知磁效应(perceptual magnet effect)原理<sup>[18]</sup>,后天语音环境对人类听觉感知空间会有较大的影响,应用文献<sup>[17]</sup>所介绍的训练方式,我们修改过的感知空间能自主形成不同的元音分类。



图 7 经过修改的 DIVA 模型接口  
Fig.7 Modified user interface of DIVA

### 3.3 声道配置

做好以上工作后,就可以进行复合元音的声道配置了。首先在 MATLAB 运行环境中将 DIVA 模型源程序文件夹设置为当前文件夹,然后在命令窗中输入命令“DIVA”,此时界面显示如图 8。通过该界面,用户可以让模型学习新的发音,也可以学习已经学过的发音。“参数”菜单里可以更改模型的一个或者多个参数,也可以更改每次训练的次数。模型中,每个音节的听觉空间都有唯一的声道配置与之对应。DIVA 模型中的体觉方向向量(Orosensory Direction Vector)经过 22 维的感觉表征转换,可以将语音映射集(Speech Sound Map)中的每个音节映

射成唯一与之对应的前田声道(Maeda channel)配置。图 8 也是复合元音/ai/对应的声道配置图,其中声道参数的数值在-4~4 之间。

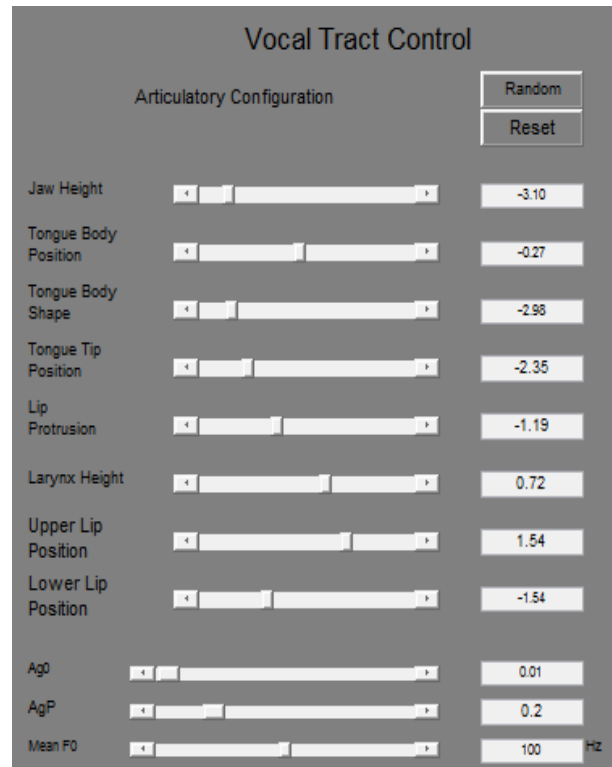


图 8 汉语复合元音/ai/对应的声道配置  
Fig.8 Track configuration of Chinese diphthong /ai/

### 3.4 /ai/发音过程仿真

(1) 应用第三节所述方法获取复合元音/ai/的前三个共振峰频率值(参见表 1);

(2) 利用声韵母原则建立音节单元集,然后映射到 DIVA 模型的语音映射集(Speech Sound Map)中。所有单元集中音节单元的初始状况都为未激活状态;

(3) 将已计算出的/ai/共振峰频率输入到 DIVA 模型中,并通过对声道各参数的调节,对/ai/实施训练和学习<sup>[9]</sup>。具体训练过程为:判断单元集中是否存在与/ai/共振峰频率相同的单元,如果存在,将已经映射到模型中的对应单元激活,否则在对应单元中建立相应的发音;

(4) 随机输入任意一个复合元音的共振峰频



率, 判断是否为 /ai/, 如果是, 跳出此过程, 直接接通扬声器发出 /ai/ 的声音; 否则, 回到步骤 3 重新开始。以上过程循环往复, 直至将所有的复合元音都训练学习完毕为止;

(5) 将所有训练完毕的复合元音通过扬声器加以输出。

其它复合元音的发音方法与此类似, 在此不再赘述。

### 3.5 仿真结果

通过在图 9 中设置经由第 3 节描述的方法所获取的共振峰频率, 并在图 8 中配置如前所述的各模拟参数, 就可以获取 /ai/ 的发音仿真结果, 如图 10 所示。

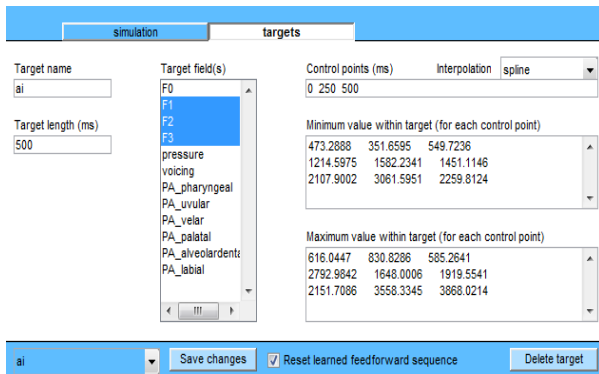
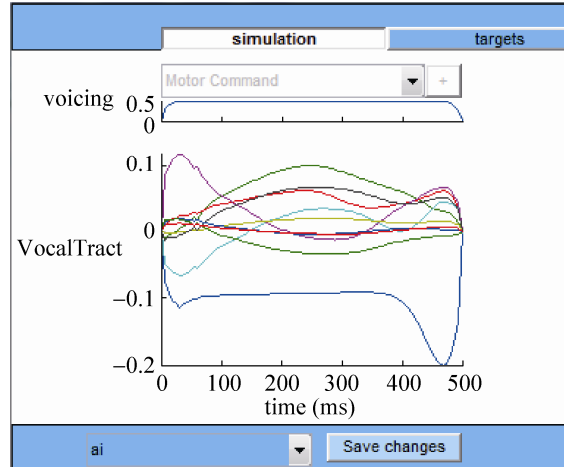


图 9 模型中汉语二合元音 /ai/ 的参数

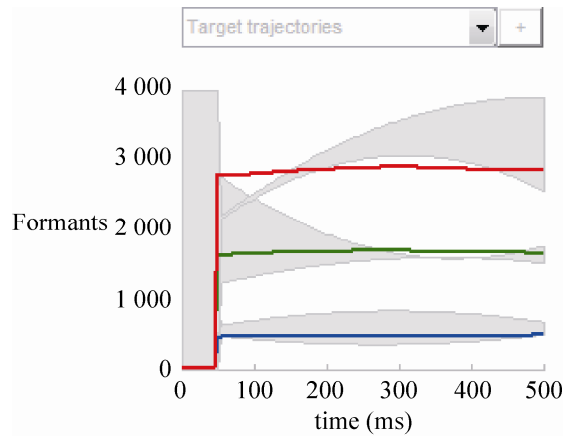
Fig.9 Parameters of Chinese diphthong /ai/ in the model

图 10(a)和图 10(b)分别代表发音运动控制时的声道变化情况以及共振峰频率变化示意图。图 10(b)中蓝线代表 F1, 绿线代表 F2, 红线代表 F3。反馈控制运动训练结果见图 10(c)。

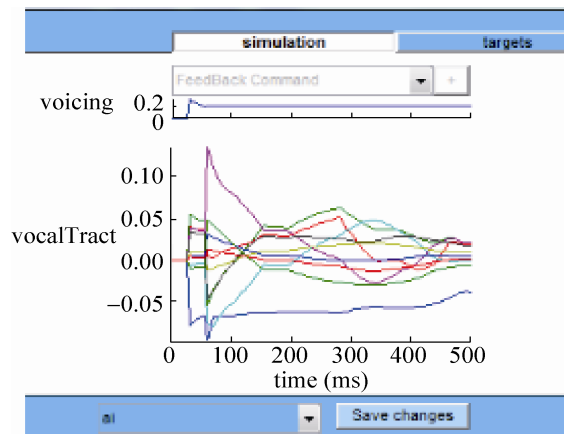
为了检验 DIVA 模型在感知空间中模拟二合元音发音的正确性, 本文参照文献[17]给出的方法, 在 DIVA 模型中取出模拟 /ai/ 发音时的输出波形(基频 197 Hz, 取样频率 10 kHz, 时间 500 ms), 如图 11 所示。通过对波形图中的一小段语音信号进行分析, 然后将分析结果与真人元音 /ai/ 的发音波形图(图 12)加以比较, 可以观察到相同时间段内的周期是否一致。



(a) /ai/声道变化示意图  
(a) Track variation diagram of /ai/



(b) /ai/的共振峰示意图  
(b) Formant diagram of /ai/



(c) /ai/反馈控制训练结果  
(c) Feedback control training results of /ai/

图 10 /ai/的发音仿真结果  
Fig.10 Simulation results of /ai/ pronunciation

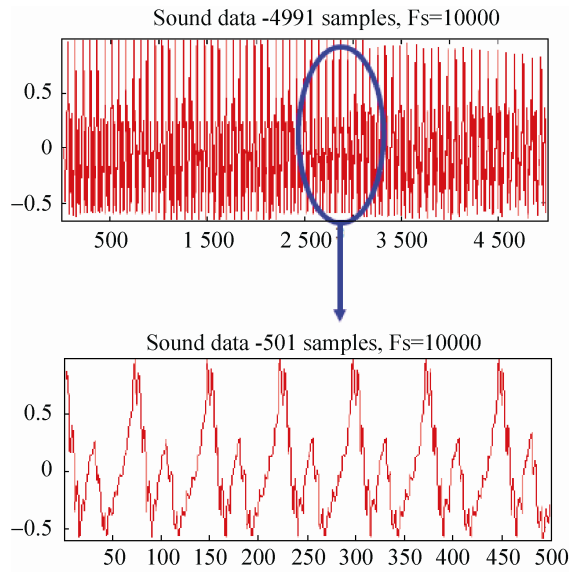


图 11 DIVA 模型输出的 /ai/ 波形图  
Fig.11 Output waveform of /ai/ in DIVA

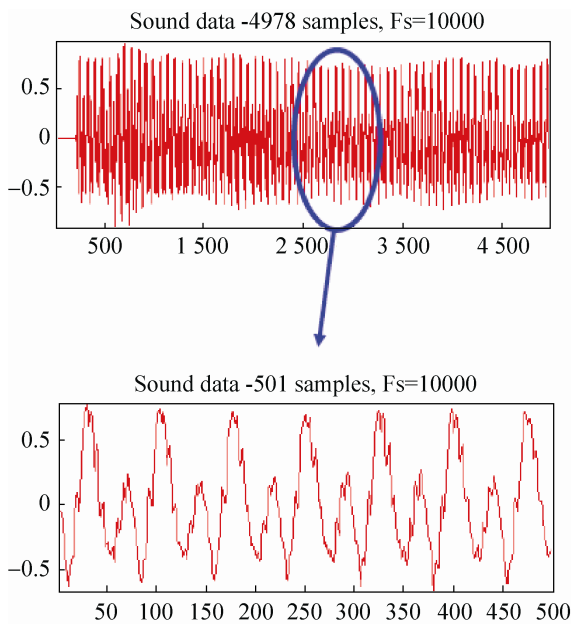


图 12 真人发出的 /ai/ 波形图  
Fig.12 Waveform of /ai/ by a real person

通过比较可以看到, DIVA 模型所产生的语音波形在周期上与真人没有太大差异。由此, 利用 DIVA 模型学习汉语二元元音发音的可行性得到了进一步证明。

## 4 结论

为了应用 DIVA 模型对汉语者大脑中的思维过程进行语音合成, 使模型能够对汉语复合元音进

行有效模拟, 本文在现有模型的基础上, 通过添加听觉感知空间, 分析复合元音的共振峰走势, 比较汉语单元音和复合元音发音方法的差异, 利用 LPMCC 方法和二次多项式对复合元音的共振峰进行拟合, 以及对 DIVA 模型中共振峰和声道对应发音器官的相关参数进行重新配置等方法, 重点对适用于 DIVA 模型的汉语复合元音发音方法进行了理论探讨和实验仿真。当然, 这些都是在我们前期工作中对 DIVA 模型硬件进行修改, 也就是在 DIVA 模型中添加了语音感知部分的基础上所完成的(参见文献[17])。仿真结果表明, 采用新的方法以后, DIVA 模型能够很好的识别汉语复合元音的发音。以此可以断定, 依据本文所提供的研究方法对 DIVA 模型加以改进或重构, 一定能使模型获得更好的符合汉语语音发声规律的效果, 从而为构造具有中国人思维特征的汉语神经分析系统(Chinese Neuralynx System, CNS)奠定理论和实践基础。

## 参考文献:

- [1] Tourville J T, Guenther F H. The DIVA Model: A Neural Theory of Speech Acquisition and Production [J]. *Language and Cognitive Processes* (S0169-0965), 2011, 25(7): 952-981.
- [2] Ghosh S S. *Understanding Cortical Contributions to Speech Production through Modeling and Functional Imaging* [D]. USA: Boston University, 2005.
- [3] Guenther F H. A neural network model of speech acquisition and motor equivalent speech production [J]. *Biological Cybernetics* (S0340-1200), 1994, 72(1): 43-53.
- [4] Shaobai Zhang, Zheng Zhang, Xin Liu. Research on the Chinese vowels Pronunciation method Based on DIVA model [J]. *WIT Transactions on Modelling and Simulation* (S1743-355X), 2014, 60(1):809-818.
- [5] Zhang S B, Ji Y C, He L W. Research on the Mechanism for Phonating Stressed English Syllables Based on DIVA Model [J]. *Neurocomputing* (S0925-2312), 2015, 152(3): 11-18.
- [6] 张少白, 王勇, 何利文. 汉语神经分析系统研究现状与展望 [J]. *中国科学:信息科学*, 2015, 45(7): 849-868. (Zhang Shaobai, Wang Yong, He Liwen. Research status and prospect of Chinese Neuralynx System [J]. *Scientia Sinica (Informationis)*, 2015, 45(7): 849-868.)

- [7] 刘欣, 基于 DIVA 模型的汉语元音发音方法及口吃成因研究 [D]. 南京: 南京邮电大学, 2013.(Liu Xin. Research on the Method of Chinese vowels Pronunciations and the Causes of Stuttering Based on DIVA model [D]. Nanjing, China: Nanjing University of Posts and Telecommunications, 2013.)
- [8] He Feng, Chen Xiaoqing, Li Guosuo, et al. A New Calculation Method of Extracting Formant [J]. Signal Processing (S0165-1684), 2007, 23(4): 618-621.
- [9] 金星, 万萍, 吴绪波, 等. 痉挛型与手足徐动型脑性瘫痪患儿口腔共鸣功能的比较研究 [J]. 中国康复理论与实践, 2012, 18(10): 913-915. (Jin Xing, Wan Ping, Wu Xubo, et al. Comparison of Oral-pharyngeal Resonance Function between Spastic and Athetoid Cerebral Palsy Children [J]. Chinese Journal of Rehabilitation Theory and Practice, 2012, 18(10): 913-915.)
- [10] Bush B O, Kain A. Estimating Phoneme Formant Targets and Coarticulation Parameters of Conversational and Clear Speech [C]// Acoustics, Speech and Signal Processing (ICASSP). Canada: IEEE, 2013: 8017-8021.
- [11] Lee S, Potamianos A, Narayanan S. Developmental Aspects of American English Diphthong Trajectories in the Formant Space [C]// Proceedings of Meetings on Acoustics. USA: Acoustical Society of America, 2013, 19(1): 60-67.
- [12] M Zbancioc, Marius, Mihaela Costin. Using Neural Networks and LPCC to Improve Speech Recognition [C]// International Symposium on Signals, Circuits and Systems. USA: IEEE, 2003, 3(3): 256-274.
- [13] Yujin Yuan, Zhao Peihua, Zhou Qun. Research of speaker recognition based on combination of LPCC and MFCC [C]// Intelligent Computing and Intelligent Systems (ICIS), 2010 IEEE International Conference on. USA: IEEE, 2010.
- [14] Wang Wei. One Factor Repeated Measures ANOVA Analysis of Gene Expression Profile Data Based on Database Language [J]. China Medical Devices (S1674-1633), 2013, 28(11): 34-36.
- [15] Yu Hang, Zhou Qiang. Intra-Chunk Relationship Analyse for Chinese Base Chunk Labeling Systems [J]. Journal of Tsinghua University (Sci & Tech) (S1000-0054), 2009, 49(10): 1708-1711.
- [16] Cruys T V, Moir'ona B V. Lexico-Semantic Multiword Expression Extraction [C]// Proceedings of the 17th Meeting of Computational Linguistics, Netherlands: Computational Linguistics. Belgium: Lot Occasional, 2007: 175-190.
- [17] 徐歆冰. 基于DIVA神经网络模型模拟语音感知神经机制研究 [D]. 南京: 南京邮电大学, 2014.(Xu Xinbing. Simulation of Neural Mechanisms for Speech Perception with DIVA Neural Network Model [D]. Nanjing, China: Nanjing University of Posts and Telecommunications, 2014.)
- [18] Guenther F H, Gjaja M N. The Perceptual Magnet Effect as an Emergent Property of Neural Map Formation [J]. Journal of the Acoustical Society of America (S1520-8524), 1996, 100(2): 1111-1121.

(上接第 254 页)

- [17] Wei Ding. A New Method for Image Noise Removal Using Chaos PSO and Nonlinear ICA [J]. Procedia Engineering (S1877-7058), 2011, 24(1): 111-115.
- [18] Yu Z, Li J, Liu K. Radar Emitter Recognition based on PSO-BP Network [J]. AASRI Procedia (S2212-6716), 2012, 1(4): 213-219.
- [19] Hu Y, Wang X. PSO-based Energy-balanced Double Cluster-heads Clustering Routing for Wireless Sensor Networks [J]. Procedia Engineering (S1877-7058), 2011, 24(15): 3073-3077.
- [20] 刘俊杰, 巴海涛. 基于粒子群优化的观测站部署算法 [J]. 指挥控制与仿真, 2014, 36(3): 40-43. (Liu Junjie, Ba Haitao. Management of Multi-sensor Based on Particle Swarm Optimization [J]. Command Control & Simulation, 2014, 36(3): 40-43.)
- [21] 史峰, 王辉, 郁磊, 等. 智能算法30个案例分析 [M]. 北京: 北京航空航天大学出版社, 2010. (Shi Feng, Wang Hui, Yu Lei, et al. Intelligent Algorithms 30 Case Analysis [M]. Beijing, China: Beihang University Press, 2010.)
- [22] 高尚, 杨静宇, 吴小俊, 等. 圆排列问题的蚁群模拟退火算法 [J]. 系统工程理论与实践, 2004, 8(8): 102-106. (Gao shang, Yang Jingyu, Wu Xiaojun, et al. Solving Circle Permutation Problem with Ant Colony Simulated Annealing Algorithm [J]. System Engineering Theory & Practice, 2004, 8(8): 102-106.)
- [23] 章义刚, 王会颖. 改进蚁群算法求解圆排列问题 [J]. 机电工程, 2008, 25(5): 92-95. (Zhang Yigang, Wang Huiying. An Improved Ant Colony Algorithm of Solving Circle Permutation Problem [J]. Mechanical & Electrical Engineering Magazine, 2008, 25(5): 92-95.)