

DOI: 10.3969/j.issn.1006-9771.2013.10.027

· 康复工程 ·

虚拟人舌运动可视化在发声中的研究

陈志翔^{1,2}, 信琴琴², 朱月秀^{1,2}, 林姿琮^{1,2}, 王琳³

[摘要] 发音障碍是听力残疾者和言语残疾者的显著特征之一。舌部作为发音的主要器官,在发音训练中起重要作用。如果将发音可视化引入发音障碍康复训练中,可以让聋人和言语残疾者直观地看到发音过程中舌位及舌形的变化,则康复训练时将起到事半功倍的效果。在研究舌部解剖结构及其运动特征的基础上,通过建立三维舌体肌肉模型,结合发音中舌的X光图片和电子颞位图仪等相关数据,实现舌体模型仿真发音中常见的运动及形状,进而利用这种可视化技术,帮助听力残疾者和言语残疾者达到矫正和康复的目的。

[关键词] 发音可视化;康复训练;发音障碍;虚拟舌部建模;运动仿真

Visualization Study of Virtual Human Tongue in Speech Production CHEN Zhi-xiang, XIN Qin-qin, ZHU Yue-xiu, et al. Department of Computer Science and Engineering, Minnan Normal University, Zhangzhou 363000, Fujian, China

Abstract: Pronunciation obstacle is one of the characteristics of hearing and speech disabled persons. As the main organ of pronunciation, tongue plays an important role in pronunciation training. If the pronunciation visualization is applied to rehabilitation training, this can make the hearing and speech disabled persons intuitively watch the change of the tongue in the process of pronunciation, which may promote the rehabilitation training. On the basis of tongue anatomical structure and movement characteristics, the common movements of tongue in pronunciation are realized after the establishment of three-dimensional tongue muscle model and the relevant data of the tongue X-ray images and electropalatography. Using this kind of visualization technology, we can help correction and rehabilitation for the hearing and speech disabled persons.

Key words: pronunciation visualization; rehabilitation training; pronunciation obstacle; virtual tongue modeling; motion simulation

[中图分类号] R493 **[文献标识码]** A **[文章编号]** 1006-9771(2013)10-0993-05

[本文著录格式] 陈志翔, 信琴琴, 朱月秀, 等. 虚拟人舌运动可视化在发声中的研究[J]. 中国康复理论与实践, 2013, 19(10): 993-997.

据2012年3月5日中国残疾人联合会发布的“残联(2012)25号文件”,目前全国残疾人总数为8502万人,其中听力残疾人数占残疾人总数的24.16%,言语残疾的人数占残疾人总数的1.53%^[1]。

听力残疾者因为听不见或听力损失,与他人交流的时候多有不便。特别是由于听力原因导致的言语障碍,将使得言语残疾者不能进行正常的言语交往活动。目前已有的康复技术主要集中在补听和助听方面^[2-3],试图提高他们的听力理解能力,间接地为语言功能的恢复提供帮助。然而由于生理上的缺陷,使得他们对于那些发音时外观比较接近的音无法准确掌握。为了解决这一难题,国内外许多学者致力于发音方式和发音位置的研究。例如,2005年~2012年期间,Wayland和Li做了视觉信息与语言感知之间关系的研究,发现/l/和/n/有相似的发音位置,但视觉上存在明显差异,最终通过这一视觉信息帮助人们

提高了对于英语辅音的正确感知率^[4-6]。2012年夏静宇等发现,对听力正常儿童的发音方式和发音位置的研究,对于有听力障碍问题的儿童的训练有重要意义^[7]。张磊等也发现,发音的可视性差是造成学龄前聋儿发音难的一个主要原因^[8]。故在发音障碍康复训练中,准确把握发音方式及发音位置,引入发音可视化技术,可以极大地帮助听力残疾者及言语残疾者。

构成语言器官的一个重要部分是舌,舌部每一段都参与发音,诸如舌尖音、舌面音及舌根音等,它的动作很灵敏,可以前伸后缩,也可以上升下降。听力残疾者由于听力缺陷,言语发展受到障碍,使舌部得不到很好的锻炼,舌肌僵硬,有的字音发不出来,或者发得残缺不全。而言语残疾者也存在这方面的缺陷。因此,对听力残疾者和言语残疾者的舌部进行锻炼十分重要。但听力残疾者的听力缺陷及发音时口腔中舌体的难见性,给康复训练师带来了极大的难题。

基金项目: 1.国家自然科学基金(No. 61170128); 2.福建省自然科学基金(No. 2012J05128); 3.福建省教育厅资助科技项目(No. JA11169; No. JA12222)。

作者单位: 1.闽南师范大学计算机科学与工程系,福建漳州市 363000; 2.福建省粒计算及其应用重点实验室,福建漳州市 363000; 3.闽南师范大学医院,福建漳州市 363000。作者简介: 陈志翔(1982-),男,福建福州市人,博士,讲师,CCF会员,主要研究方向:计算机视觉、虚拟现实技术、可视化方法等。

听力残疾者和言语残疾者虽在听力上存在缺陷,可是其在视觉上并不存在障碍。若果能够使他们清晰直观地观察到发音中口腔内舌的运动及位置的变化,即将发音可视化,无疑对听力残疾者和言语残疾者的发音康复训练带来极大帮助。然而,由于舌部肌肉的复杂性难于采集实验数据,目前对于舌部模型的建模及可视化技术研究较少,已有的少数模型运动也较为机械。Stone 等建立了第一个 3D 舌体模型^[9],该模型将舌体沿冠状面和径向方向各分成 5 段,但只能实现舌体简单的形变。Pelachaud 等^[10]在基于 Wyvill 等^[11]提出的点状模型的基础上,用 9 个三角形表示舌部,并分析发音中舌体的初步运动情况,由于其几何结构简单,无法真实再现舌部灵活的运动特征。

本文通过对舌部解剖结构^[12]以及发音中舌体位置和舌形变化的研究^[13],将发音中的几种主要的舌形进行分类,通过建立相应的肌肉模型,控制舌部肌肉进行相应变化,并结合三维建模技术及真实纹理映射,进而实现发音的可视化。实验表明,该方法能较好仿真及控制舌部发音过程中的多种运动,实现发

音的可视化,在实际康复训练中能够起到重要的辅助作用。

1 舌部结构及建模方法

1.1 舌部的解剖结构

舌肌为横纹肌,可分为舌固有肌和舌外肌两类。舌固有肌指组成舌体本身的肌肉,起止均在舌内,其肌纤维走向分纵行、横向和垂直三种,收缩时,分别可使舌缩短、变窄或变薄。舌外肌起自舌外,止于舌内,根据其解剖结构对应的生理运动功能,颊舌肌、茎突舌肌和舌骨舌肌较为重要。颊舌肌是一对强有力的肌肉,起自下颌体后面的颊棘,肌纤维呈扇形向后上方分散,止于舌中线两侧。两侧颊舌肌同时收缩时,拉舌向前下方,即实现伸舌运动,单侧收缩时,使舌伸向对侧,左右交替收缩时可以实现舌体左右运动。茎突舌肌起自茎突,斜向前下,止于舌旁和舌底,其功能为引舌向后上方。舌骨舌肌为薄四边形,起于舌骨大角,直向上分布于舌体底部,由舌下神经支配,作用是牵舌向下,使舌部中线下陷,形成谷形。见表 1。

表 1 舌体肌肉功能表

肌肉	舌上纵肌	舌下纵肌	茎突舌肌	舌横肌	舌垂直肌	舌骨舌肌	颊舌肌
作用	收缩时舌体变短,向上弯曲	收缩时舌体变短,向下弯曲	收缩时引舌体后部升高	收缩时舌体变厚、变窄	收缩时舌体变宽、变薄	收缩时牵引舌中线下陷,形成谷形	左侧颊舌肌收缩,引舌向右侧运动;右侧颊舌肌收缩,引舌向左侧运动;双侧同时收缩引舌向前下

1.2 舌部发音的功能模型

舌部是造成共鸣腔形状变化的主要发音器官,以元音为例,依据舌部起作用的具体部位不同,可以将其分为舌面元音、舌尖元音与卷舌元音三类。例如,舌面元音主要由舌面调节共鸣腔形状而发出的元音,主要有三种方式:①把舌面抬得高一点或低一点;②把舌面伸得前一点或后一点;③嘴唇呈圆形或者展平。可将这三种方法归纳为舌位的高低、舌位的前后和唇型的圆展。这 3 个要素的结合决定着每个舌面元音的音质,即每个舌面元音都是由这 3 个参数的变化所控制的,每给定一组 3 个参数,就能得到一个舌面元音。语音学家绘制了元音舌位图(见图 1)来对舌面元音的 3 个参数进行分析,从而对舌面元音进行分类和描述^[13]。

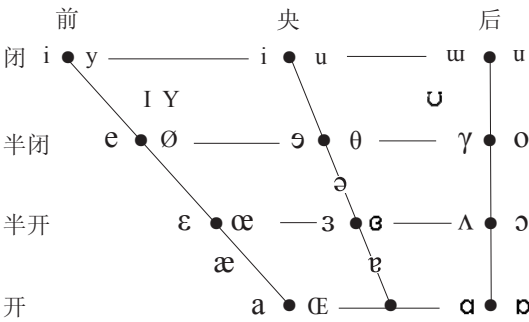


图 1 元音舌位图

随着实验语音学的快速发展,电子颞位图仪、超声影像技术、X 光摄像等先进的设备和技术被引入到发音研究,为进一

步了解口腔内各组织在元音及辅音发音过程中的动作提供了依据。电子颞位图仪即在舌和上颞上安装上一定数量(一般是 96 个)的电极,通过标准发音者发某些音标,记录舌和颞的接触情况,可以分析舌位及口腔开合的情况。图 2 为舌体在 X 光下拍摄的图片及电子颞位图^[14]。

根据电子颞位图仪的记录,可以按发音时舌的形状将音标分为 3 类:前升高、后升高和谷形^[14]。

考虑到发音过程中舌体的运动存在连贯性,期间的连贯动作多涉及左右及前后的衔接,在通过对舌部解剖结构及生理运动特征研究的基础上,本文将舌部参与发音的主要肌肉分成 4 类:①控制舌体前升高的舌纵肌;②控制舌体后升高的茎突舌肌;③控制舌体谷形运动的舌骨舌肌、舌横肌及舌垂直肌;④控制舌体左右运动及前下运动的颊舌肌。

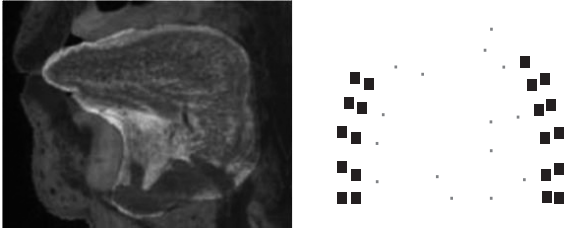


图 2 舌 X 射线图及电子颞位图

1.2.1 舌前升高建模 舌部在发某些音时,其电子颞位图如图 3 所示,与上颞主要接触点集中在舌尖部分,这即为前升高。在舌部上纵肌的控制下,舌体向上卷曲,形成前升高。根据本文的前期工作^[15-16],可以通过控制上纵肌的收缩量,实现不同程

度的前升高,具体如图4所示。舌体发某些音时,舌体前升高,通过控制方法,可以达到不同程度的前升高,满足发音过程中的要求。

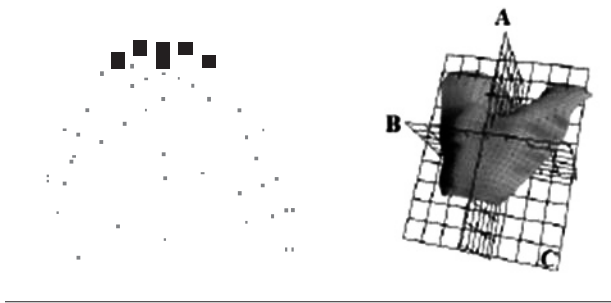


图3 前升高图

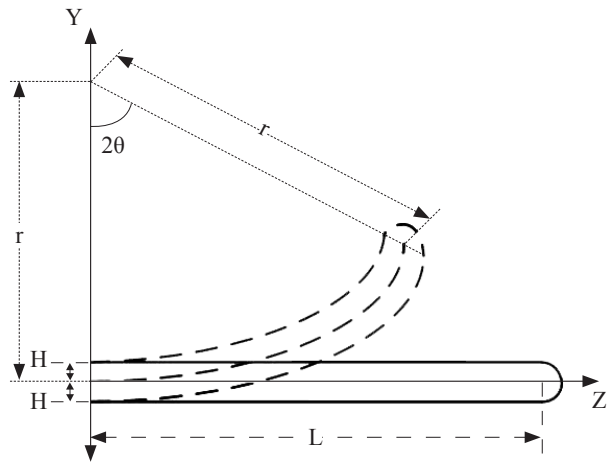


图4 舌前升高模型

1.2.2 舌后升高建模 舌体在发一些音时,舌体后部有升高趋势,其电子颞位图如图5所示。通过对舌部解剖结构的研究,知道其主要是通过茎突舌肌控制的。由前期工作可知,控制茎突舌肌的收缩量,可以实现不同程度的后升高^[17],图6展示了不同收缩量下舌体侧中线所呈现的多种后升高状态。在发音可视化过程中,若某些音有后升高趋势,则通过控制模型中相关参数,可达到对应发音的舌位要求。

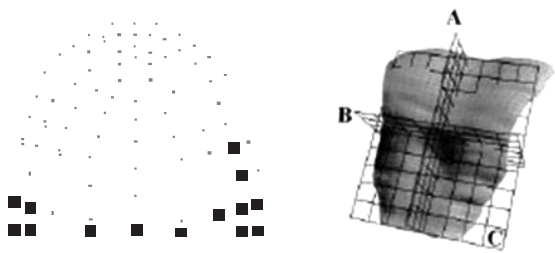


图5 后升高图

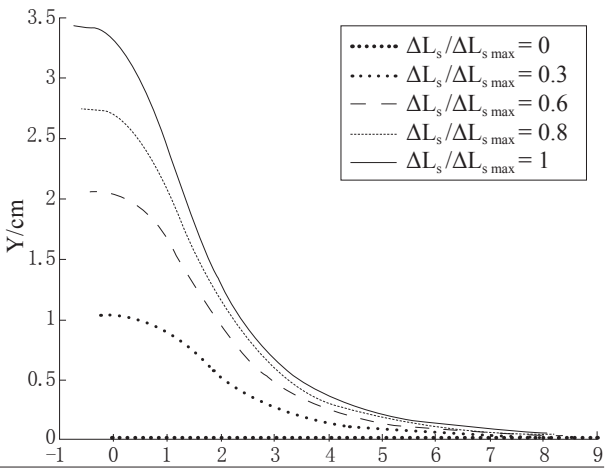


图6 舌后升高示意图

1.2.3 舌谷形建模 舌体谷形形成的机理主要为舌中线下陷,舌部形成类似山谷的形状,其电子颞位图如图7所示。该形状主要是由舌骨舌肌、舌横肌及舌垂直肌控制形成。其控制模型由本文前期工作^[18]可得,图8a给出了相同形状因子下不同收缩量的舌体谷形效果,图8b考虑到不同人种在形成谷形时的区别,利用各自的形状因子来实现不同形状的谷形。在发音可视化过程中,可通过控制模型中的参数,以达到使舌体形成预期谷形的目的。

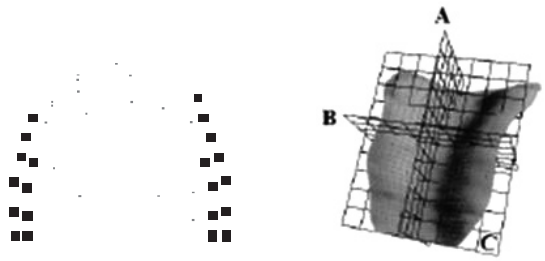


图7 谷形图

1.2.4 舌左右运动建模 舌体的另一种主要运动为左右运动。此运动主要通过颏舌肌控制。颏舌肌为一对扇形肌,单侧收缩,可引舌伸向对侧,实现舌体的左右运动。由于其结构对称,根据本文的前期工作^[18],图9给出了单侧颏舌肌的几何模型,在确定三个角度符合生理结构的前提下,通过控制相关肌肉收缩量,可实现舌体左右运动。

2 实验结果

2.1 舌部的的基本运动

进一步,通过实验获取真实舌部纹理图案,在三维网格标定及纹理映射下,可以实现舌部前后升高、左右运动以及谷形运动等常见运动仿真。为方便显示,从实验仿真出的发音运动视频中等时截取了四幅序列图。由于建立的是三维模型,故可以从任意角度观察舌体运动形态,图10a~c分别从左上方观察舌体从自然状态运动到前升高、后升高及谷形的过程,图10d为从正上方观察舌左右运动的实验结果。可以看出,其效果较为逼真,运动较自然。

2.2 舌部的发音运动

将上述方法运用于舌体发音控制。根据控制模型，控制对应的肌肉收缩量，可得到其发音过程(图 11)，图中可以清晰观察到谷形运动的整个过程，较好地实现了舌体的可视化。特别地，在图 11 中舌体前、后升高及前下降图中我们加入了口腔部分，这样可以清晰的看到舌体前升高及后升高在口腔中的变化，为发音训练的可视化研究奠定基础。

3 对发音障碍康复训练的意义

舌部是人发音中涉及到的最重要的一个部分，实现虚拟舌体灵活便捷地控制，对于发音教学、可视化仿真、发音障碍康

复训练等有着及其重要的意义。本文提出了一种基于肌肉控制的舌部模型。根据舌生理解剖结构和舌体发音原理，将舌体肌肉按发音功能分为四类，分别建模控制。实验结果表明，针对日常发音过程中几种主要的舌部运动，该方法能够较好地仿真相应的运动，且所需数据量小，驱动方便。舌体发音过程仿真图中舌形及舌位均可以清楚地展示，这表明发音的可视化是完全可以做到的。如果我们将发音的可视化运用于发音障碍康复训练、虚拟发音教学、聋哑人发音矫正中，将会给听力残疾者和言语残疾者的康复训练带来极大的帮助。

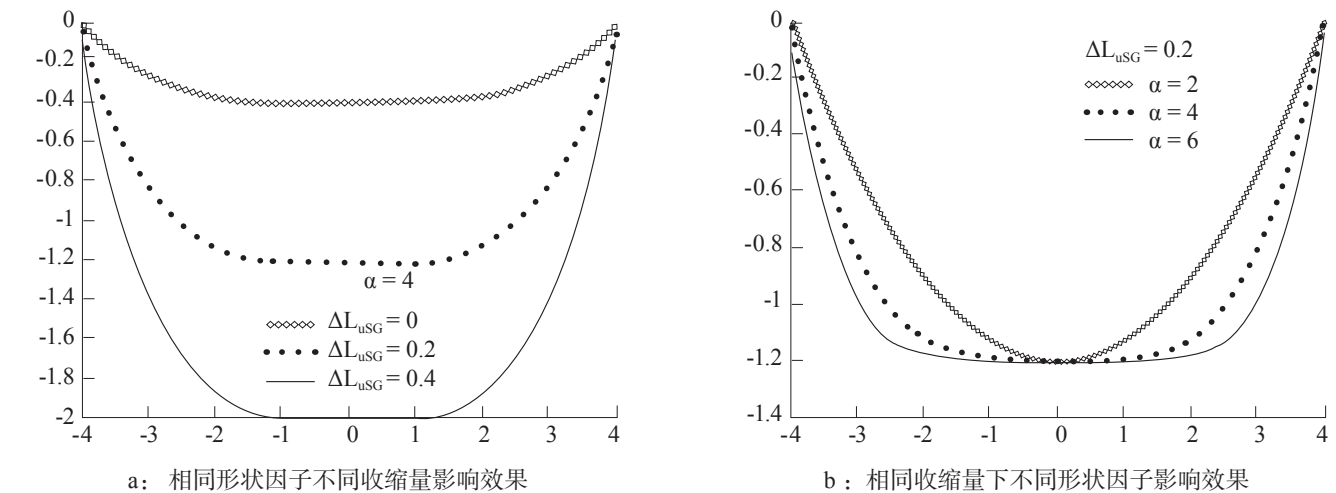


图 8 谷形模型示意图

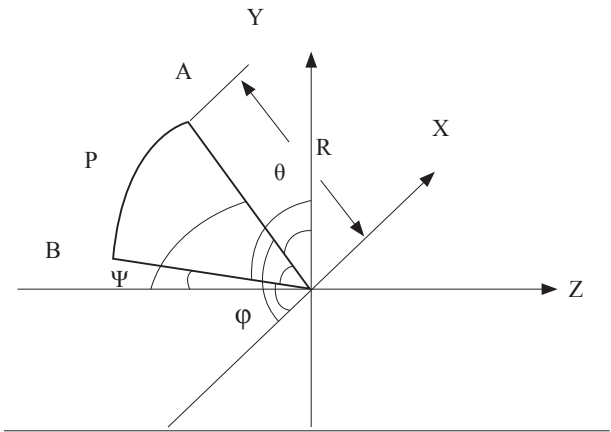


图 9 舌舌肌几何模型

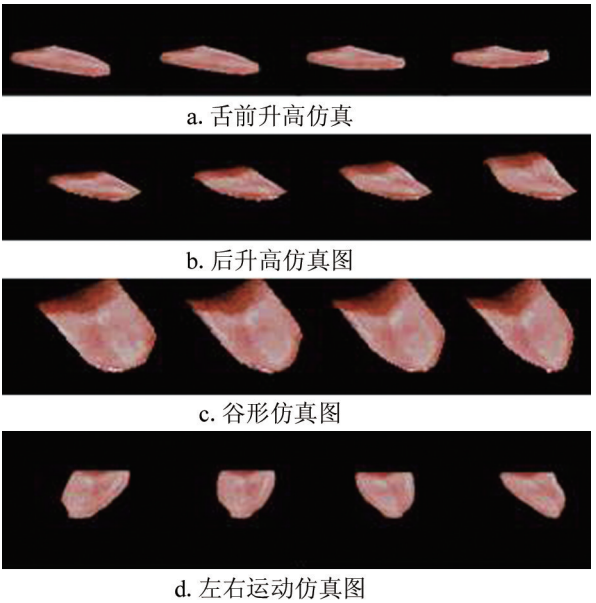


图 10 舌体前、后升高、谷形及左右运动仿真图

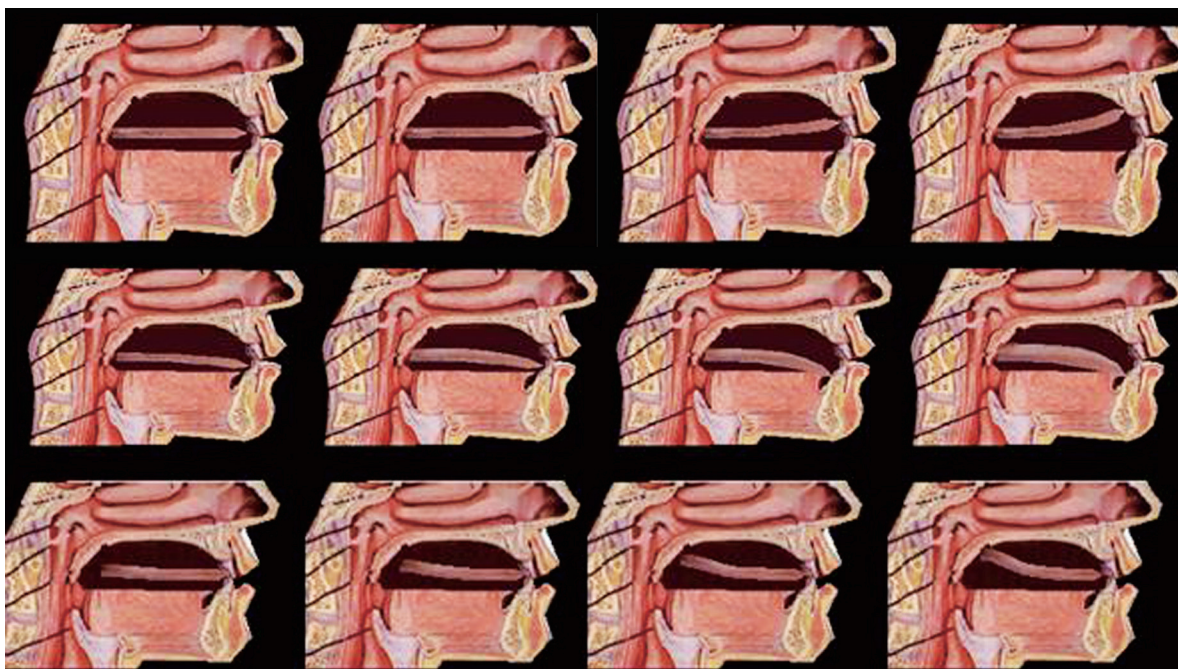


图 11 口腔中舌体的前升高, 前下降, 后升高发音过程可视化序列

[参考文献]

- [1] 中国残疾人联合会. 关于使用 2010 年末全国残疾人总数及各类、不同残疾等级人数的通知残联[2012]25 号[EB/OL]. [http:// www.cdprf.org.cn/wxzx/content/2012- 03/12/content_30383765. htm](http://www.cdprf.org.cn/wxzx/content/2012-03/12/content_30383765.htm). [2012-03-05]
- [2] 李郁明, 梁勇, 谭少珍. 语前聋儿童康复训练后听觉能力的动态评估[J]. 听力学及言语疾病杂志, 2013, 21(2): 174-177.
- [3] 李靖, 陈雪清, 吴燕君, 等. 耳聋程度对选配助听器婴幼儿言语产出能力的影响[J/OL]. 听力学及言语疾病杂志, 2013, 21(4): 391-394.
- [4] Wayland R, Li B. Training native Chinese and native English listeners to perceive Thai tones [R]. London, UK: ISCA Workshop on Plasticity in Speech Perception, Senate House, 2005: 62-65.
- [5] Wayland R, Li B. Effects of two training procedures in cross-language perception of tones [J]. J Phonetics, 2008, 36(2): 250-267.
- [6] 李彬. 视觉信息与辅音感知[J]. 语言科学, 2012, 11(4): 403-411.
- [7] 夏静宇, 管燕平, 薛永强. 4.4-5.5 岁聋儿与正常儿童语音清晰度的比较[J]. 中国康复理论与实践, 2012, 18(8): 707-709.
- [8] 张磊, 朱群怡, 黄邵鸣, 等. 学龄前聋儿声母发音难度研究[J]. 听力学及言语疾病杂志, 2012, 20(2): 102-104.
- [9] Stone M. Toward a model of three-dimensional tongue movement [J]. J Phonetics, 1991, 19: 309-320.
- [10] Pelachaud C, Overveld, Seah C. Modeling and Animating the Human Tongue During Speech Production [C]. Proceedings of Computer Animation, 1994: 40-49.
- [11] Wyvill G, McPheeters C, Wyvill B. Data structure for soft objects [J]. The Visual Computer, 1986, 2(4): 227-234.
- [12] 曾智成. 新编人体解剖学图谱[M]. 西安: 世界图书出版社, 2006: 55-58.
- [13] 周殿福, 吴宗济. 普通话发音图谱[M]. 北京: 商务印书馆, 1963: 6-18.
- [14] Stone M, Lundberg A. Three-dimensional tongue surface shapes of English consonants and vowels [J]. J Acoust Soc Am, 1996, 99(6): 3728-3737.
- [15] 陈志翔, 程义民, 曾丹, 等. 人舌及嘴部的 3D 控制模型[J]. 中国科学院研究生院学报, 2008, 25(3): 372-378.
- [16] 陈志翔, 程义民, 曾丹, 等. 人舌运动与嘴部表情的多线谱仿真[J]. 系统仿真学报, 2009, 21(23): 7518-7521.
- [17] Chen ZX, Zhang XJ, Wu ZR. A new tongue model based on muscle-control [C]. IEEE International Conference on Granular Computing, 2011: 132-137.
- [18] 陈志翔. 虚拟人舌运动与嘴部表情的研究[D]. 合肥: 中国科学技术大学, 2010: 59-70.

(收稿日期: 2013-03-30 修回日期: 2013-04-12)