

文章编号 1004-924X(2015)02-0582-07

基于舌部可视化模型的听障儿童发音训练

史丽娟^{1,2}, 赵 剑³, 王丽荣⁴, 杜钦生³, 安志勇^{1*}, 车 娜³

(1. 长春理工大学 光电工程学院, 吉林 长春 130022;

2. 长春大学 电子信息工程学院, 吉林 长春 130022;

3. 长春大学 计算机科学技术学院, 吉林 长春 130022;

4. 苏州大学 电子信息学院, 江苏 苏州 215000)

摘要: 基于可视化模型提出了一种听力障碍儿童的舌部训练方法。通过对听障儿童舌部发音位置的训练, 辅助其获取正确的发音位置和发音方式。该方法采用三维建模技术和传感器技术相结合的舌部康复训练方法, 建立舌部三维模型; 用腭位传感器获取发音过程中舌部位置信息; 选取 23 个声母作为样本, 建立健听儿童声母发音舌位置信息样本数据库; 最后, 驱动舌部模型, 引导听障儿童进行舌部发音位置训练。该方法能够显示健听儿童发音时舌部的运动过程, 克服了目前听障儿童语言训练中舌部训练过程中不可视的问题。实验结果表明: 该方法提高了听障儿童发音训练的效果, 缩短了康复训练时间, 尤其对于舌、下腭之间配合较为复杂, 发音位置较为隐蔽的发音, 提高发音训练的准确度超过 15%。

关键词: 听障儿童; 三维建模; 舌部可视化模型; 发音位置; 康复训练

中图分类号: TP391; R764.44 **文献标识码:** A **doi:** 10.3788/OPE.20152302.0582

Articulation training for children with hearing disorder based on visualization model of tongue

SHI Li-juan^{1,2}, ZHAO Jian³, WANG Li-rong⁴,

DU Qin-sheng³, AN Zhi-yong^{1*}, CHE Na³

(1. College of Opto-electronic Engineering, Changchun University of Science and Technology, Changchun 130022, China;

2. College of Electronic Information and Engineering Changchun University, Changchun 130022, China;

3. College of Computer Science and Technology, Changchun University, Changchun 130022, China;

4. School of Electronic and Information Engineering, Suzhou University, Suzhou 215000, China)

* Corresponding author, E-mail: an_zhiyong@126.com

Abstract: An articulation training method for children with hearing disorder was proposed based on a visualization tongue model. By training on the position of articulation for children who have residual hearing and hearing compensation, these children could acquire the accurate place and way of articulation. A tongue three dimensional model was established based on the tongue rehabilitation training

收稿日期: 2014-10-16; **修订日期:** 2014-11-11.

基金项目: 中国博士后科学基金面上项目 (No. 2013M541298); 教育部“春晖计划”合作项目 (No. Z2012062); 吉林省科技厅科技发展计划资助项目 (No. 20130522162JH); 吉林省教育厅资助项目 (No. 2013285, No. 2014281, No. 2013278); 长春市科技计划资助项目 (No. 13GH03)

with the combination of three-dimensional modeling technology and sensor technology. The three-dimensional positions of tongue during the process of articulation was acquired by a palate sensor. Then, 23 initial consonants were selected as samples for establishing a sample database of tongue positions of children with normal hearing during articulation. Such the tongue model could be activated to provide a guidance for children with hearing disorder during training on the position of articulation. This method displays the movement of tongues of children with normal hearing during articulation and overcomes the shortcomings that has been neglected in tongue training and language training. Experimental result indicates that this method improves the effect of articulation training for children with hearing disorder and shortens training time. Especially for those articulation processes involving complicated coordination between tongue, lower jaw and unobvious position of articulation, the method increases the accuracy of articulation training by 15%.

Key words: children with hearing disorder; three-dimensional modeling; tongue visualization model; articulation position; rehabilitation training

1 引言

2012年3月12日中国残疾人联合会发布的“残联(2012)25号文件”推算2010年末全国残疾人总数为8502万,其中听力残疾人为2054万,约占24.16%,言语残疾人为130万,约占1.53%^[1],其中听力障碍儿童约为83.5万。由于听力障碍,这些儿童的语言能力远远落后于同龄健全儿童^[2]。听障儿童发音器官是正常的,但听障儿童说出的话一般较难听懂,主要是因为他们在发音方式和发音位置上会出现许多错误^[3]。在听力障碍儿童群体中有残余听力的约占85%以上^[4]。因此,只要听力能得到及时补偿和重建,再进行科学的康复训练,大部分听障儿童是可以掌握有声语言的。

听障儿童的语言康复训练一般包括呼吸训练、舌部训练、唇部训练、鼻音训练、嗓音训练等。舌部是发音器官的一个重要部分,也是口腔器官中最灵活的部件,在发音中起着最重要的作用。听障儿童由于长期不用舌部或者利用不全面,使得舌头运动幅度不大,从而导致舌肌僵硬,转动不灵活,说话时跟不上语言节拍的速度,有的字音发不出来,或是发的不准确^[5]。传统的舌部训练方式是让听障儿童进行舌操训练,即安排时间组织听障儿童做舌操,如伸缩舌头、顶腮、翻转舌头等,以期望锻炼孩子舌头的灵活性,柔软性及正确性^[6]。但是这种训练方式往往比较枯燥、乏味,容易让孩子厌烦,从而失去训练耐心。随着信息

科技的飞速发展,许多学者将唇读技术^[7]、三维重建技术^[8-9]、多媒体技术等引入听障儿童康复训练过程,进行辅助教学,因此涌现了一批视听结合的聋儿言语康复系统。Rathinavelu等人设计了一款计算机辅助聋儿康复系统,该系统使用磁共振成像技术(MRI)探知发音器官在发音过程中的运动轨迹,并生成对应的模拟发音图片序列,用于聋儿的发音教学^[10]。Olle Bälter等人开发出了一款ARTUR(发音向导)系统,该系统包含一个用于发音模拟和教学的三维虚拟发音人头像,使教学更加形象、生动,易于激发听力障碍儿童的学习动力^[11],但是这种方法可视性差^[12],仍然不能为听障儿童直观地显示发音过程中内部发音器官的动作,未能达到康复训练的理想效果^[13]。

考虑到听障儿童的视觉感知能力超出正常儿童,因此,本文基于可视化模型提出了一种听力障碍儿童的舌部训练方法。该方法将三维建模技术与传感器技术相结合,建立舌部在口腔中运动的真实感模型来模拟正常人发音时舌的运动状态,使听障儿童可以获取正确的发音位置和发音方式。这种方法有助于克服目前听障儿童语言训练中舌部训练过程不可视的问题,可以提高听障儿童发音训练的效率,以及发音训练的准确度。

2 系统整体设计方案

首先,建立舌部模型;接着,通过采集正常儿童发音时的发音位置信息建立标准样本库;之后,通过运动特征参数变换驱动舌部模型运动,采集

听障儿童发音位置信息,并对其进行运动特征参数变换,驱动模型运动;最后,将听障儿童发音模

型与正常儿童的发音模型进行分析比较。系统整体设计如图 1 所示。

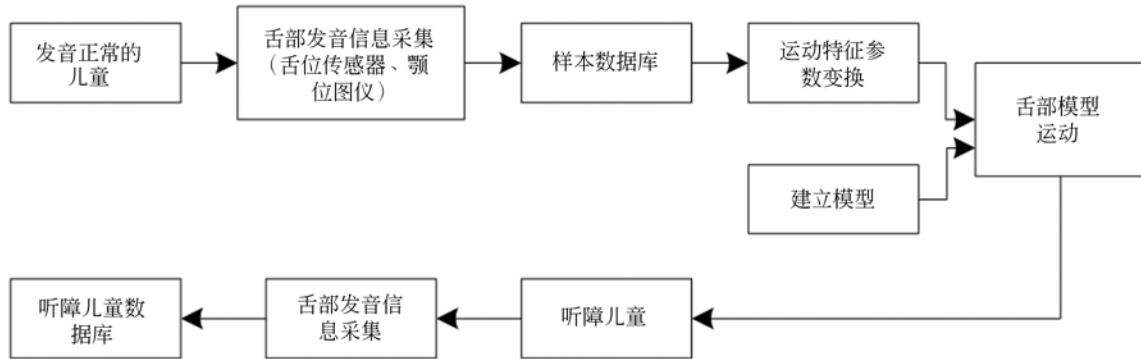


图 1 系统整体设计

Fig. 1 Overall system design

2.1 舌部模型的建立

舌部训练的目的在于通过一系列发音的强化训练,增强舌内肌的运动感受性,提高舌头的灵活性,如舌部的伸、缩、顶、卷、弹等。因此,需要建立一个高自然度的舌部模型,用以模拟灵活的舌部运动。陈志祥博士根据舌部解剖结构,提出了一种基于肌肉运动特征的三维舌体模型^[14],如图 2

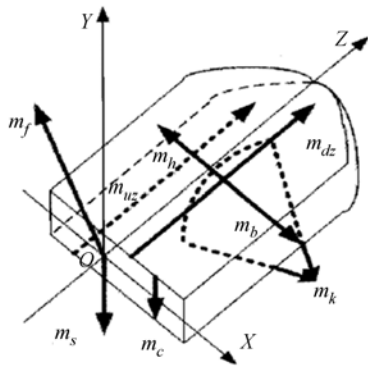


图 2 舌部几何模型

Fig. 2 Geometric model of tongue



(a)舌部三维模型平面图 (b)舌部三维模型侧面图
(a)3D model ichnography (b)3D model side elevation
of tongue of tongue

图 3 舌部三维模型

Fig. 3 3D model of tongue

所示。该模型能够较为灵活地控制舌部发音过程的多运动,因此本系统选择该模型进行分析。图 3 为本文基于舌部几何模型建立的舌部三维模型的平面图及侧面图。

2.2 舌部发音位置的参数采集

发音时采用腭传感器获取舌与腭的信息,该传感器中 126 个镀金腭传感器平均分布于网架状组织中,图 4 为腭传感器的电极分布图。当舌与这些电极接触时,传感器电极能够迅速地采集舌与假腭的接触信息,黑色圆点表示舌一腭接触的位置,(该数据显示讲话时舌与腭部贴合的部位)。图 5 为发音过程中采集的发音序列关键帧,用于提取舌腭部发音时的动态参数信息。

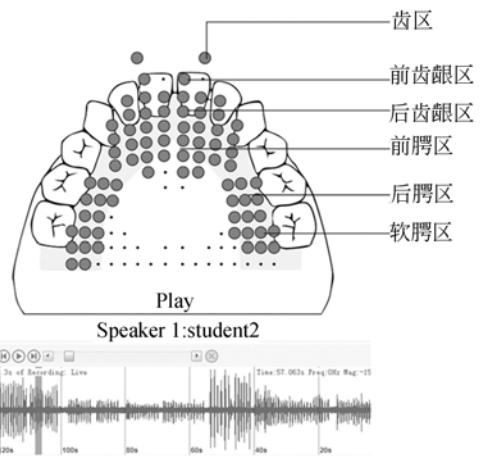


图 4 腭传感器 126 个电极分布图

Fig. 4 126 electrode distribution map of palatal sensor

将舌腭发音信息数据与舌发音位置三维数据相结合来驱动舌部模型,并且在此过程中与语音同步输出一个具有真实感的模拟正常人发音动作的舌部模型。

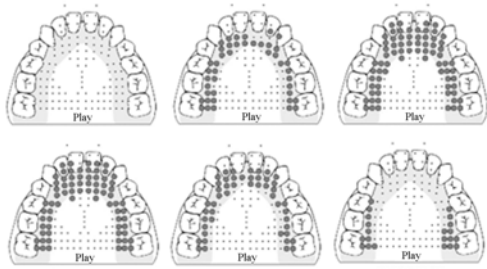


图 5 舌部发音参数的采集

Fig. 5 Parameter acquisition of tongue pronunciation

3 实验过程及结果分析

选择一名 6 岁,说普通话,且发音稳定的健听儿童作为参考对象。信号采样频率为 100 Hz,语音信号采样频率为 1 600 Hz,选择 4 名听障儿童作为实验对象。其中 2 名为 3 岁,已进行 2 年听力补偿,听力达到合适范围,之后,又进行 1 年语言康复;另外 2 名听障儿童均为 5 岁,进行了 3 年听力补偿,2 年语言康复。将 4 名听障儿童分成两队,每队由一名 3 岁听障儿童和一名 5 岁听障儿童组成。第一队由语训老师进行康复训练,第二队由舌部康复训练系统进行训练,时间长度均为 3 个月。语训材料从声母开始,声母因为时长短、能量小,且需要不同发音部位形成不同方式的阻塞,是发音康复训练的重点和难点。在听障儿童口语朗读训练资料及孙喜斌主编的《聋儿听力语言康复评估题库》的声母及韵母测试表中选取 4 组发音进行测试。第一组为双唇音、唇齿音(b, p, m, f);第二组为舌尖音(d, t, n, l);第三组为舌根音(g, k, h);第四组为舌面音(j, q, x)。

3.1 健听儿童发音时舌部信息采集及模型运动

首先采集健听儿童的舌部发音信息,并建立数据库,驱动三维模型运动。图 6~图 8 为健听儿童发标准音时的信息采集及模型动作情况。

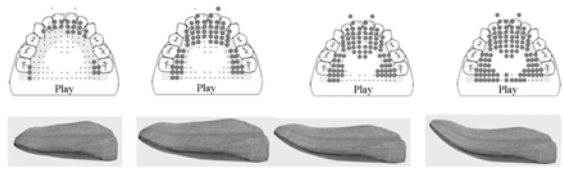


图 6 发舌尖音 d 时参数采集及舌模型运动的关键帧
Fig. 6 Key frames of parameter acquisition and model motion when normal children pronounce apical tone “d”

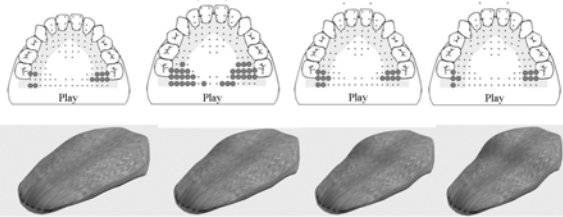


图 7 发舌根音 g 时参数采集及舌部模型运动的关键帧
Fig. 7 Key frames of parameter acquisition and model motion when normal children pronounce velar tone “g”

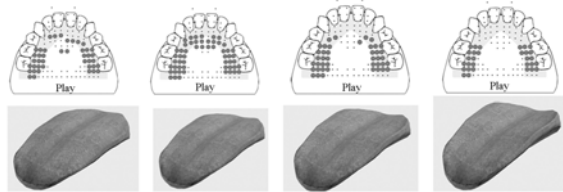


图 8 发舌面音 q 时参数采集及舌部模型动作的关键帧
Fig. 8 Key frames of parameter acquisition and model motion when normal children pronounce palatal tone “q”

3.2 第二组听障儿童未受训之前发音时舌部信息

图 9~图 11 为听障儿童未受训之前的发音状况。



图 9 未受训儿童发舌尖音 d 时参数采集的关键帧
Fig. 9 Key frames of parameter acquisition when untrained children pronounce apical tone “d”

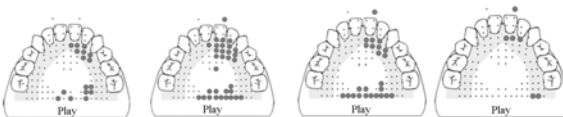


图 10 未受训儿童发舌根音 g 时参数采集的关键帧
Fig. 10 Key frames of parameter acquisition when untrained children pronounce velar tone “g”



图 11 未受训儿童发舌面音 q 时参数采集的关键帧
Fig. 11 Key frames of parameter acquisition when untrained children pronounce palatal tone “q”

3.3 第二组听障儿童受训之后的舌部发音信息采集

由听障儿童选择一个训练环节,如第二组舌尖音(d,t,n,l),或第三组舌根音(g,k,h)等。系统输入相应训练单元,并从参数库中调取参数驱动舌部模型,此时,屏幕上会播放舌部模型的发音过程,同时播放该信息的语音,该发音过程经过参数驱动三维舌部模型完成。听障儿童根据屏幕显示的发音过程进行发音练习,同时系统收集听障儿童练习过程的音频信息和传感器信息,并通过传感器技术和语音识别技术进行分析,通过判断模块给听障儿童发音进行评价。如果发音正确,系统会给予鼓励和赞扬,如果发音错误,可以选择重新播放正确的三维舌部模型发音过程。通过视频分析提示其在刚刚发音过程中的错误之处。整个学习过程,系统会建立受训者资料数据库,记录受训者的学习情况,图 12~图 14 为听障儿童经过 3 个月舌部训练系统康复之后的发音状况。表 1 为听障儿童受训之前的发音清晰度测试表。

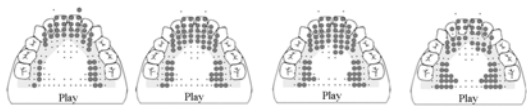


图 12 受训儿童发舌尖音 d 时参数采集的关键帧
Fig. 12 Key frames of parameter acquisition when training children pronounce apical tone “d”

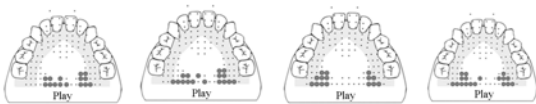


图 13 受训儿童发舌根音 g 时参数采集的关键帧
Fig. 13 Key frames of parameter acquisition when training children pronounce velar tone “g”

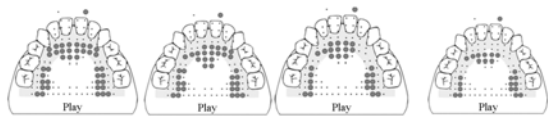


图 14 受训儿童发舌面音 q 时参数采集的关键帧
Fig. 14 Key frames of parameter acquisition when training children pronounce palatal tone “q”

表 1 未受训听障儿童不同发音部位声母发音清晰度比较(每个音发音 50 次的正确率)

Tab. 1 Definition comparison of consonant pronunciation in different places for untrained children (The correct rate with each tone pronouncing 50 times) (%)

		双唇音 b	舌尖音 d	舌跟音 g	舌面音 q
第一组 (语训老 师组)	3 岁听 障儿童	70.7	54.2	40.7	30.1
	5 岁听 障儿童	78.5	60.2	45.2	40.2
第二组 (舌部康 复训练组)	3 岁听 障儿童	70.3	52.1	42.2	33.1
	5 岁听 障儿童	77.1	63.5	43.7	42.2

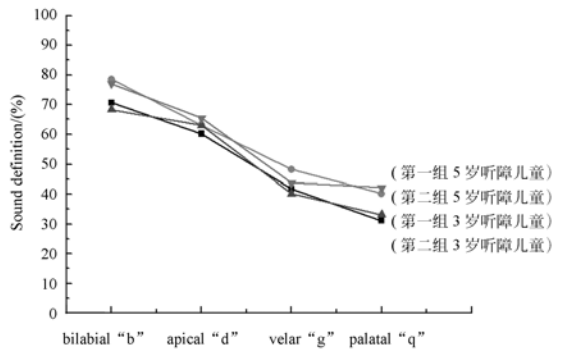


图 15 不同发音部位声母发音难度比较
Fig. 15 Difficulty comparison of initial consonant pronunciation in different places

由图 15 可知,发音清晰度最高的声母为双唇音 b 和舌尖音 d,其次为舌跟音 g,之后是舌面音 q,这说明发音清晰度与发音的复杂程度有关。通过分析可知,发双唇音和舌尖音时,口型很容易看清楚,而发舌跟音及舌面音时所需唇、舌、齿以及上腭之间的配合,且发音位置较为隐蔽,所以听障儿童学习起来很困难。

表2 由语训老师训练的听障儿童声母发音清晰度比较
(每个音发音50次的正确率)

Tab.2 Definition comparison of initial consonant pronunciation of hearing-impaired children trained by language training teacher

(The correct rate with each tone pronouncing 50 times)
(%)

	双唇音 b	舌尖音 d	舌跟音 g	舌面音 q
受训儿童 1	78.7	69.2	48.9	37.1
受训儿童 2	83.5	70.2	56.2	45.2

表3 由舌部康复训练系统训练的听障儿童声母清晰度的测试表(每个音发音50次的正确率)

Tab.3 Definition comparison of initial consonant pronunciation of hearing-impaired children trained by the rehabilitation training system

(The correct rate with each tone pronouncing 50 times)
(%)

	双唇音 b	舌尖音 d	舌跟音 g	舌面音 q
受训儿童 1	80.3	73.3	53.5	50.1
受训儿童 2	85.5	77.1	58.8	56.8

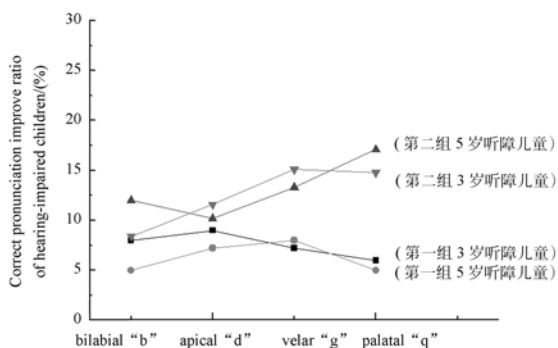


图16 语训老师及训练系统的康复效果比较

Fig.16 Comparison of rehabilitation effects between language training teacher and rehabilitation training system

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]. [2012-03-12] http://www.cdpf.org.cn/wxzx/content/2012-03/12/content_30383765.htm
- [2] 张芳.听障与健听儿童连续语音切换能力的比较及应用研究[D].上海:华东师范大学,2009.

表2为经过语训老师的康复训练之后,听障儿童的发音清晰度正确率,表3为听障儿童经过系统训练之后的发音清晰度正确率。将表2、表3的数据与表1进行比较得到图16。图16为经过语训老师训练的听障儿童和经过训练系统康复训练的听障儿童训练效果对比。由图16可见:由于双唇音及舌尖音口型容易看清楚,所以对于双唇音和舌尖音,语训老师和可视模型的舌部系统康复训练效果的区别不是很大;但对于舌根音及舌面音,发音难度大,并且发音位置隐蔽,语训老师无法让受训儿童观察到内部口型,因此康复训练效果有所下降,而利用可视模型的舌部康复训练系统由于能让受训儿童直接观察舌位发音器官的动作,故其康复训练效果明显优于语训教师。

4 结论

本文建立了三维可视化舌部模型,基于该模型提出了听障儿童的舌部康复训练方法,即应用逼真的三维模型充分指导听障儿童进行舌部训练,使听障儿童可以从各个方向了解发音时舌部的位置和状态。该方法还通过奖励和鼓励措施调动听障儿童训练兴趣,并且能够对学习效果进行评估和反馈,故可以实时监测听障儿童的康复训练程度。选择发音位置较为典型的4组声母作为语训材料,对2组听障儿童进行实验。实验结果表明,采用该方法能够提高康复发音的训练效果,并且能够缩短听障儿童发音训练用时。

ZHANG F. *Speech between hearing-impaired and normal hearing Children and its application* [D]. Shanghai: East China Normal University, 2009. (in Chinese)

- [3] 咎飞,刘春玲.儿童语音发展的研究方法、对象及规律[J].中国特殊教育,2000,3:15-16.
- JIU F, LIU CH L. Patterns in phonological development of Children with special needs[J]. *Chinese Journal of Special Education*, 2000, 3: 15-16. (in Chinese)

- [4] 夏静宇,管燕平,薛永强. 4.5~5.5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.
XIA J Y, GUAN Y P, XUE Y Q. Comparison of phonetic articulation between hearing-impaired and normal children in 4.5-5.5 years old groups [J]. *Chinese Journal of Rehabilitation Theory and Practice*, 2012, 18(8): 707-709. (in Chinese)
- [5] 黄振美. 聋儿语训初探[J]. 潍坊教育学院学报, 2005, 18(4): 80-81.
HUANG ZH M. Discussion on the language training of deaf children [J]. *Journal of Weifang Educational College*, 2005, 18(4): 80-81.
- [6] 张莉莉. 听力障碍儿童语言康复训练的构建与实施[J]. 中国教育技术装备, 2012, 1: 20-23.
ZHANG L L. Implementation and construction of hearing-impaired Child rehabilitation training [J]. *Chinese Educational Technology and Equipment*, 2012, 1: 20-23. (in Chinese)
- [7] 王丽荣, 王建蕾. 基于主成份分析的唇部轮廓建模[J]. 光学精密工程, 2012, 20(12): 2276-2279.
WANG L R, WANG J L. Lip contour modeling based on PCA [J]. *Opt. Precision Eng.*, 2012, 20(12): 2276-2279. (in Chinese)
- [8] 王欣, 袁坤, 于晓, 等. 基于运动恢复的双目视觉三维重建系统设计[J]. 光学精密工程, 2014, 22(5): 1380-1387.
WANG X, YUAN K, YU X, et al.. Design of binocular vision 3D reconstruction system based on motion recovery [J]. *Opt. Precision Eng.*, 2014, 22(5): 1380-1387. (in Chinese)
- [9] RATHINAVELU A, THIAGRAJAN H, RAJKUMAR A. Three dimensional articulator model for speech acquisition by children with hearing loss [C]. *Proceedings of 4th International Conference on Universal Access in Human Computer Interaction, Beijing, China UAHCI 2007*: 786-794.
- [10] RATHINAVELU A, YUVARAJ G. Data visualization model for speech articulators [C]. *Proceedings of Annual International Conference on Emerging Research Areas, AICERA*, 2011: 155-159.
- [11] 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.
ZHANG L, ZHU Q Y, HUANG SH M, et al.. A study of consonant production difficulty level in preschool children with hearing impairments [J]. *Journal of Audiology and Speech Pathology*, 2012, 20(2): 102-104. (in Chinese)
- [12] ZHAO J, SHI L J, DU Q S, et al.. Special human-computer interaction model oriented to hearing impaired children [J]. *Journal of Convergence Information Technology*, 2013, 8(3): 503-511.
- [13] ZHAO J, SHI L J, FAN Q Y, et al.. Breath training for hearing impaired hearing children based on computational fluid dynamics [C]. In *Proceeding of the 2013 International Workshop on Assistive Engineering and Information Technologies, Dalian, China: AEIT*, 2013: 506-509.
- [14] 陈志翔, 信琴琴, 朱月秀, 等. 基于发音运动特征的舌部肌肉建模及仿真[J]. 计算机工程与应用, 2013, 49(20): 255-258, 262.
CHEN ZH X, XIN Q Q, ZHU Y X, et al.. Modeling and simulation of tongue muscle based on pronunciation characteristics [J]. *Computer Engineering and Applications*, 2013, 49(20): 255-258, 262. (in Chinese)

作者简介:



史丽娟(1978—),女,吉林榆树人,博士研究生,长春大学电子信息工程学院讲师,2009年于长春理工大学获得硕士学位,主要从事仪器科学与技术及康复工程等方面的研究。E-mail: shilijuan.ccu@gmail.com

导师简介:



安志勇(1943—)男,教授,博士生导师,河北滦县人,1966年毕业于天津大学精密仪器工程系光学仪器专业,1978年~1988年作为访问学者在德国柏林工业大学光技术研究所学深造,“光学工程”国家重点学科带头人,曾获国务院政府特殊津贴,主要从事光学工程、测试计量技术及仪器等方面的研究。E-mail: an_zhiyong@126.com