

频率跟随响应对音调跟随准确性的反映

宫 琴^{1,2}, 孙文生¹

(1. 清华大学 生物医学工程系, 北京 100084; 2. 清华大学 深圳研究生院, 生物医学研究中心, 深圳 518055)

摘 要: 频率跟随响应(FFR)能够反映受试者的音调感知能力。为了研究 FFR 对具有相同听力特征人群的个体间音调感知能力差异的表征,该文测试了母语为汉语普通话的听力正常受试者对频率为成年男性语音基频典型值的纯音的 FFR 响应,分别采用短时自相关及谱图方法从时域及频域 2 个角度对 FFR 信号音调跟随准确性特征参数进行了提取。通过与受试者频率分辨阈值的心理声学测试结果的相关分析,得出短时自相关及谱图方法与行为学频率分辨阈值之间具有显著相关性的结论。结果表明:使用 FFR 能对同一听力特征群体中个体之间音调感知能力的差异进行表征;另外,基于短时自相关的 FFR 音调跟随准确性特征比谱图方法更适合用于提取个体的音调感知能力差异。

关键词: 频率跟随响应; 音调感知; 频率分辨阈值

中图分类号: R 318.03

文献标志码: A

文章编号: 1000-0054(2013)01-0133-06

Reflection of pitch following accuracy with frequency following response

GONG Qin^{1,2}, SUN Wensheng¹

(1. Department of Biomedical Engineering, Tsinghua University,
Beijing 100084, China;

2. Research Center of Biomedical Engineering, Graduate School
at Shenzhen, Tsinghua University, Shenzhen 518055, China)

Abstract: Frequency following response (FFR) reflects the subjects' pitch sensation abilities and it is used to study individual differences in pitch sensation ability among populations with the same hearing characteristics. The FFR responses to pure tones with the typical adult male speech fundamental frequency were tested with normal hearing native Chinese mandarin speakers. The characteristic parameters of their pitch following accuracy were extracted from the FFR signals using short-time auto-correlation and spectrogram methods in the time and frequency domains. Correlation analysis showed that FFR pitch tracking accuracy by short-time auto-correlation and spectrogram methods significantly correlated with the subjects' frequency discrimination limens from psychoacoustic tests. The results show that FFR is able to reflect the differences in pitch sensation ability among individuals from populations with the same hearing characteristics. In addition,

analysis of the FFR pitch following accuracy characteristics using short-time auto-correlation is more suitable for extracting the individual pitch sensation ability differences than the spectrogram method.

Key words: frequency following response; pitch sensation; frequency discrimination limen

音调是声音的一个非常重要的感知属性,它是音乐感知的基础,并且在有调语言(如汉语)中起着重要的表意作用。目前广泛使用行为学表现进行音调感知能力的评估,此类方法易受主观因素的影响,可靠性较差,实验过程较为复杂。因此,研究音调感知的行为学表现与听觉神经生理信号的关系对理解音调感知的神经机理有重要的意义。听觉电生理信号测试将提供一种更加客观、可靠、操作简便、能够反映内在神经活动情况的测量方法,从而更好地促进电子耳蜗音调感知新算法的研究和开发^[1-2],以及临床对电子耳蜗佩戴者听力康复训练效果的客观评估与方案设计。

频率跟随响应(frequency following response, FFR)是在头皮上记录到的脑干对声音信号的稳态响应。FFR 的潜伏期在 5~7 ms,它是脑干神经元对声音中周期性成分锁相响应的集总表现,与脑对声音音调的感知有着密切的联系^[3]。FFR 研究和应用的一个重要方面是脑对音调的编码和感知^[3-4]。Krishnan 等 2004 年用汉语音节/yi/的 4 个声调分别做刺激声诱发产生 FFR,并用基于自相关的方法从 FFR 信号中提取音调高度信息,得到的音调高度变化曲线与刺激声的音调高度非常一致^[5]。这说明 FFR 信号中携带了语音音调高度

收稿日期: 2012-03-08

基金项目: 国家自然科学基金资助项目(61271133);

深圳基础研究基金重点资助项目(JC201105180808A)

通信作者: 宫琴(1968—),女(汉),山东,副教授。

E-mail: gongqin@mail.tsinghua.edu.cn

相关的信息,反映了脑干对音调高度信息的编码功能。

在对频率跟随响应的研究中,已经证明了不同听力经历引起的音调感知的差异^[6-7];而且,在母语为非有调语言的人群中,长期听力训练引起2个人群的音调感知能力的差异^[8];而且还发现,在对母语为英语的成人进行使用汉语声调的短期训练之后,他们的频率跟随响应对汉语声调的音调跟随变得更加准确^[9]。这一方面说明了听觉脑干的可塑性,另一方面也说明了FFR信号能够反映受试者短期内音调感知能力的变化。

已有的对频率跟随响应与人的音调感知表现关系的研究侧重于以人群特征为研究对象,在具有不同听力特征的群体之间对频率跟随响应及音调感知表现进行比较。但对于具有相同听力经历及特征的人群,如母语同为汉语普通话的人群,群体中不同个体之间对音调的感知的差异是否能够通过电生理的方式体现出来的问题目前研究很少。因此,本文通过提取FFR中的特征参数验证其在具有相同听力特征人群中能够识别个体间音调感知的差异。从反映群体的听力特征到反映个体的听力情况,是在更加严格的尺度上观察频率跟随响应与音调感知能力之间的关系,这对认识脑干在音调感知中的作用有重要的意义。此外,对个体听力情况的反映是临床听力研究的核心课题,从这一角度对频率跟随响应进行观察是将其应用于临床音调感知能力测量的基础。

本文以母语为汉语普通话的正常听力受试者为研究对象,分别从时域及频域2个角度,提取FFR中关于受试者音调跟随准确性的信息,探讨用频率跟随响应对这一相同听力特征群体中音调感知能力的个体差异进行表征的可能性及合适的表征参数。

1 实验方法

为避免声音复杂度对实验结果带来影响,使用纯音为声学材料,测试受试者对频率为成年男性语音基频典型值140 Hz纯音的FFR响应。受试者对音调的跟随越准确,其行为学上表现出的频率分辨阈值,即受试者能够区分出其音调高低的纯音间最小的频率差异越小。为了与FFR信号的度量进行对比,用心理声学方法对受试者在140 Hz处对纯音音调高低的频率分辨阈值进行测量,进而考察用FFR音调跟随准确性特征参数表

征同一听力特征群体内个体音调感知差异的可能性,以及能较好地反映FFR音调跟随性特征的方法。

1.1 受试者

实验对16名母语为汉语的听力正常青年受试者的右耳进行了测试,受试者的平均年龄为22岁。所有受试者听力正常,没有听力疾病病史。实验前与他们签订了知情同意书,所有测试遵守清华大学对实验受试者的要求及规定。

1.2 频率分辨阈值的心理声学测量

对频率分辨阈值的心理声学测试采用经典的“3选1必须作答”的“2上1下自适应测试”方法^[10-11]。测试系统以MATLAB科学计算软件为平台、计算机声卡等硬件为基础搭建,测试软件的主界面如图1所示。实验在电声屏蔽室中进行,受试者通过佩戴单耳入耳式耳机接收刺激声。1次测试(test)由数次试验(trial)组成。每次试验中,向受试者随机播放3个声音,其中2个是频率为测试点频率140 Hz的参考声音,另1个是频率高于测试点频率的对比声音。3个声音都是时长为250 ms的正弦信号,其中开头和结尾10 ms分别加有Blackman窗的上升和下降段,声音强度均为83 dB SPL(sound pressure level)。相邻2个声音之间的静音段长度为800 ms,要求受试者在听到3个声音后选出其中音调高的1个。选定第1次试验中对比声音的频率后,每当受试者连续回答正确2次试验,对比声音的频率就会下降1个步长,缩小对比声音与参考声音之间的频率差;每当受试者回答错误1次试验,对比声音的频率就会上升1个步长,增大对比声音与参考声音之间的频率差。测试中,对比声音频率变化方向发生改变的试验叫做转折点。对比声音频率的变化步长在每个转折点处以1.414的系数缩小,但最小不小于0.1 Hz,若计算结果小于0.1 Hz,则之后步长维持在0.1 Hz。整个实验由1次练习、2次正式测试组成。练习在测试达到8个转折点时结束,正式测试在测试达到14个转折点时结束。每次测试后,将从第7个转折点起向后所有转折点试验中对比声音与参考声音的频率差的几何平均值作为当次测试测得的频率辨别阈值,将2次正式测试结果的平均值作为最终的测试结果。



图1 心理声学测试界面及参数选择界面

1.3 FFR 信号的记录

FFR 信号采集系统框图如图 2 所示。

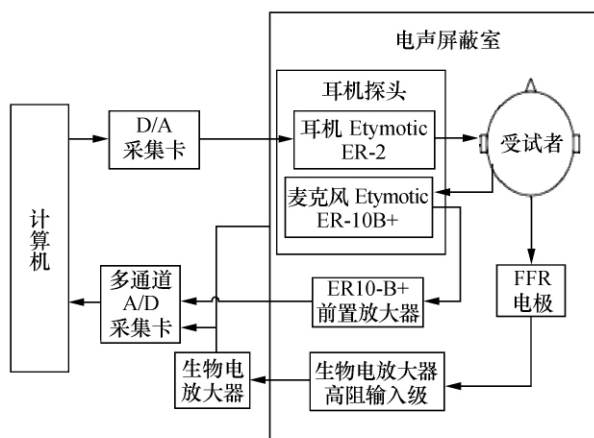


图2 FFR 信号的采集系统框图

FFR 信号的记录在电声屏蔽室中进行,受试者舒服地靠坐在 1 张大椅子上,通过佩戴听力专用单耳入耳式耳机 Etymotic ER-2 接收刺激声。FFR 信号由贴于受试者头部的 3 个电极拾取^[4]: 额中发际 Fz 处为记录电极,刺激声接收同侧耳(右耳)乳突处为参考电极,公共地电极位于眉心 Fpz 处。电极拾取到的 FFR 信号顺次经过 50 000 倍放大,以及 1~3 000 Hz 的硬件带通滤波,进入 A/D 采集卡和计算机。为了避免刺激伪迹对 FFR 信号记录的影响,对 ER-2 耳机的音频传输线以及传感器用锡纸进行了封闭包装,形成了一个用于防止电磁泄露的 Farady 笼^[12]。ER-2 耳机的入耳式探头中同时含有一个麦克风(Etymotic ER-10),测试过程中同时记录麦克风的输出信号,以监测测试时的耳道内声压。

实验所用刺激声的频率为 140 Hz,强度为 83 dB SPL,长度为 144 ms。为减少脑干对声音“出现”和“消失”响应的瞬态成分,并减小刺激声由于有限长度带来的“频谱飞溅”效应,对刺激声开始和结

束的 7 ms 分别加了 Blackman 窗的上升和下降部分进行平滑^[3],即刺激声由 7 ms 上升期+130 ms 平台期+7 ms 下降期组成。相邻刺激声的给声间隔为 416.7 ms。信号记录时窗为从刺激声给声前 77.7 ms 至刺激声结束后 61.6 ms。在一次不间断的测试段中,重复记录 1 000 次刺激声诱发的 FFR 信号,整个测试包含 2~3 个测试段。每个测试段完成后,受试者休息片刻再进行下个测试段的记录。

1.4 FFR 信号的分析方法

1.4.1 数据筛选和预处理

信号的筛选和处理在数据采集完成后离线进行。由于在测试过程中,受试者头部可能会发生微小的运动,从而导致耳机在耳道中发生微小的位移,或者对电极与皮肤的良好接触产生影响,因此,剔除重复测量中下列不符合条件的测量数据:

- 1) 耳道内声压最大值和最小值的差在所用刺激声强度(83 dB SPL)±6 dB SPL 对应的理论声压最大值和最小值的差的范围之外;
- 2) 耳道内声压最大值大于刺激声强度 83 dB+6 dB SPL 对应的声压最大值;
- 3) 将刺激声给声时记录到的耳道内声压按时间平均分为 5 段,某一段的声压最大值小于 5 段整体声压最大值的 75%,或者某一段的声压最小值大于 5 段整体声压最小值的 75%;
- 4) 测得的 FFR 信号的最大值或最小值在测量设备量程的 95%以外。

在测量筛选步骤完成后,对每名受试者各个测试段的 FFR 信号分别进行 Wiener 滤波、相干平均、带通滤波,继而将所有测试段的信号进行平均,并通过信噪比(SNR)的估计,剔除信噪比过低的受试者的数据。本文使用 SNR=10 作为筛选受试者的依据。

1.4.2 短时自相关音调跟随准确性特征提取

为提取 FFR 信号周期性成分的时域特征,使用基于滑窗和信噪比筛选的短时自相关方法提取 FFR 信号的音调跟随准确性特征参数。特征提取时使用对应于刺激声平台期的 FFR 信号,考虑潜伏期及耳道内回声的影响,采用 89.7~223.3 ms 的 FFR 信号。将 FFR 信号用时间窗截取出一系列 30 ms 长的片段,相邻 2 个时间窗的时移为 1 ms,则共产生 104 个信号片段。为了更加准确地提取信号特征,对上述全部 104 段时间信号施加了信噪比筛

选,使用其中符合信噪比筛选条件的片段进行下一步计算。筛选方法如下:使用实验中刺激声给声前记录的信号的时域均方根值作为噪声水平的估计,选取待分析 FFR 各信号片段中时域均方根大于噪声均方根的片段用于下一步分析。记通过信噪比筛选的第 i 个信号片段为 $s_i(n)$, $i=1\sim Q$, Q 为通过筛选的片段个数。对每个时间片段的信号,计算其自相关函数 $r_i(m)$,即信号与其自身延时后得到的信号的点积,如式(1)所示。

$$r_i(m) = \frac{\sum_{n=1}^{N_s} s_i(n)s_i(n-m)}{\sum_{n=1}^{N_s} s_i^2(n)},$$
$$m = 0 \sim (N_s - 1).$$
(1)

自相关函数的峰对应着信号中具有较强周期性的成分,其除零延时点外的最大峰值对应的时间延时即为 FFR 信号中最强的频率成分对应的时间周期,表明了此片段 FFR 信号频率跟随的位置,记第 i 段信号片段此最大峰值对应的延时为 τ_i 。求出所有 Q 个时间片段的延时与刺激声周期的差的均方根,即可表征整段 FFR 信号的主要周期与刺激声周期的偏差程度,从而作为 FFR 音调跟随准确性的度量,如式(2)所示。记此度量为 Δa ,则 Δa 值越大,表明该 FFR 信号的音调跟随准确度越差。

$$\Delta a = \sqrt{\left(\sum_{i=1}^Q \left(\tau_i - \frac{1}{f_0}\right)^2\right)/Q}.$$
(2)

1.4.3 基于谱图的音调跟随准确性特征提取

短时自相关方法是在时域对受试者的 FFR 音调跟随准确度进行的分析,谱图方法则从频域对 FFR 信号音调跟随准确性特征进行提取。其进行音调跟随准确性特征参数提取的方法是:首先使用同样的滑窗方法截取一系列信号片段并进行信噪比筛选。对于通过筛选的信号片段,分析其频谱成分,记其频谱中具有最大幅值的频率成分为该段信号的特征频率,表征该片段内受试者感知到的音调值。然后,将各片段的特征频率与刺激声频率作差,此差值的均方根即为对受试者音调跟随准确性的估计。

2 实验结果

表 1 给出了通过信噪比筛选的 11 名受试者的频率分辨阈值心理声学测试结果。受试者频率分辨阈值的平均值为 3.12 Hz,标准差为 1.33 Hz。

表 1 频率分辨阈值的测试结果

受试者序号	测试日期	姓名	对比声音的频率阈值/Hz	频率分辨阈值/Hz
1	2011-06-22	Liu T	142.96	2.96
2	2011-06-25	Zhang S	142.67	2.67
3	2011-06-25	Dong XX	145.54	5.54
4	2011-06-25	Liu Y	143.24	3.24
5	2011-06-27	Tong M	144.31	4.31
6	2011-06-29	Li J	141.36	1.36
7	2011-06-30	Zheng XC	144.53	4.53
8	2011-07-04	Zhang Y	142.96	2.96
9	2011-06-29	Wang X	143.58	3.58
10	2011-07-05	Zhang YY	141.44	1.44
11	2011-07-06	Liu X	141.72	1.72

表 2 给出了 11 名受试者基于短时自相关的 FFR 音调跟随准确性指标计算结果。

表 2 短时自相关音调跟随准确性指标的计算结果

受试者序号	行为学分辨阈值/Hz	短时自相关频率跟随准确度指标/(ms)
1	2.96	0.20
2	2.67	0.11
3	5.54	0.29
4	3.24	0.12
5	4.31	0.32
6	1.36	0.16
7	4.53	0.23
8	2.96	0.15
9	3.58	0.20
10	1.44	0.12
11	1.72	0.22

为了检验 FFR 短时自相关音调跟随准确性指标对音调感知能力的表征情况,对 11 名受试者的 FFR 音调跟随准确性指标与其心理声学行为测试得到的频率分辨阈值做相关分析,结果如图 3 所示。

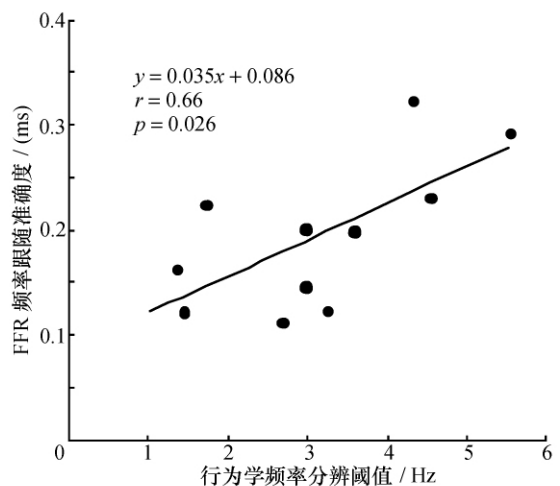


图 3 短时自相关音调跟随准确度与频率分辨阈值的相关分析

可以看到,受试者的频率分辨阈值越大,即频率分辨能力越差,相应的短时自相关 FFR 音调跟随准确度指标值越大,即 FFR 表现出的音调跟随特性越不准确。分析得到二者的相关系数为 $r=0.66$, $p=0.026$, 相应的最小二乘拟合得到的拟合关系式为 $x=0.035x+0.086$, 其中: x 表示受试者的频率分辨阈值, y 表示 FFR 音调跟随准确度。此结果表明用短时自相关延时偏差表征的 FFR 音调跟随准确度与受试者的频率分辨阈值行为学测试结果之间具有显著的相关性,这种相关性远高于随机水平。这一结果一方面说明采用 FFR 短时自相关方法提取的音调跟随准确性指标能够对同一听力特征人群中个体的音调高度感知能力的差异进行表征;从另一方面,上述相关性结果显示基于此方法可用 FFR 对行为学频率分辨阈值进行估计,估计结果与实测值之间还存在一定程度的偏差。表 3 是基于谱图方法提取的 FFR 音调跟随准确度指标的结果。

表 3 基于谱图方法的 FFR 音调跟随准确度指标计算结果

受试者序号	行为学分辨 阈值/Hz	基于谱图的 FFR 音调 跟随准确度指标/Hz
1	2.96	8.92
2	2.67	1.80
3	5.54	6.12
4	3.24	3.41
5	4.31	7.29
6	1.36	2.48
7	4.53	5.77
8	2.96	3.47
9	3.58	4.74
10	1.44	2.97
11	1.72	5.19

为了验证谱图方法得到的音调跟随准确性指标对频率分辨阈值的估计情况,对二者进行相关分析。图 4 为相关分析的结果,得到二者的相关系数为 0.51, 统计分析的 p 值为 0.11。用最小二乘拟合得到的二者的定量关系在图中进行了标注。

从图中可见,谱图方法得到的频率跟随准确性与行为学频率分辨阈值之间仍具有一定的相关性。但是二者的相关性比短时自相关方法分析中得到的相关性差。实验结果说明采用短时自相关的方法比谱图的方法能更好地提取 FFR 信号中的音调感知能力信息。

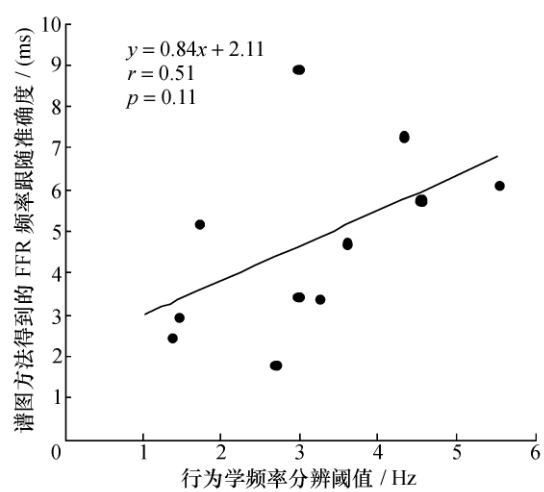


图 4 谱图方法音调跟随准确性与行为学频率分辨阈值的相关分析

3 结 论

本研究在母语为汉语普通话的正常听力受试者中,探讨了频率跟随响应对具有相同听力特征群体的音调感知能力的个体差异进行表征的能力。通过使用短时自相关及谱图方法,分别从时域及频域 2 个角度提取了 FFR 信号的音调跟随准确性特征参数,并将 FFR 信号的音调跟随准确性特征参数与受试者的频率分辨阈值心理学行为测试结果进行了相关分析,结果显示短时自相关及谱图的方法与行为学频率分辨阈值之间均具有相关性;并且,短时自相关的方法比谱图方法分析中得到的相关性更好,即 FFR 信号反映出的脑对音调高度信息编码的准确性能够较好地体现受试者的音调高度感知能力。由实验结果可得出:FFR 能对具有相同听力特征群体中音调感知能力的个体差异进行表征;由短时自相关方法得到的 FFR 信号音调跟随准确性特征参数比谱图的方法提取的特征参数与受试者的行为学频率分辨阈值有着更显著的相关性,即与谱图方法相比,短时自相关方法更适合用以提取 FFR 信号的音调跟随准确性特征。

参考文献 (References)

[1] Wilson B S, Dorman M F. Cochlear implants: A remarkable past and a brilliant future [J]. *Hearing Research*, 2008, 242(1-2): 3-21.

[2] Lan N, Nie K B, Gao S K, et al. A novel speech-processing strategy incorporating tonal information for cochlear implants [J]. *IEEE Transactions on Biomedical Engineering*, 2004, 51(5): 752-760.

- [3] Burkard R F, Eggermont J J, Don M. Auditory Evoked Potentials: Basic Principles and Clinical Application [M]. Philadelphia, USA: Lippincott Williams & Wilkins, 2007.
- [4] Skoe E, Kraus N. Auditory brain stem response to complex sounds: A tutorial [J]. *Ear Hear*, 2010, **31**(3): 302-324.
- [5] Krishnan A, Xu Y, Gandour J T, et al. Human frequency-following response: Representation of pitch contours in Chinese tones [J]. *Hearing Research*, 2004, **189**(1-2): 1-12.
- [6] Krishnan A, Gandour J T. The role of the auditory brainstem in processing linguistically-relevant pitch patterns [J]. *Brain Language*, 2009, **110**(3): 135-148.
- [7] Krishnan A, Xu Y, Gandour J T, et al. Encoding of pitch in the human brainstem is sensitive to language experience [J]. *Cognitive Brain Research*, 2005, **25** (1): 161-168.
- [8] Wong P C M, Skoe E, Russo N M, et al. Musical experience shapes human brainstem encoding of linguistic pitch patterns [J]. *Nature Neuroscience*, 2007, **10**(4): 420-422.
- [9] Song J H, Skoe E, Wong P C M, et al. Plasticity in the adult human auditory brainstem following short-term linguistic training [J]. *Journal of Cognitive Neuroscience*, 2008, **20**(10): 1892-1902.
- [10] Levitt H. Transformed up-down methods in psychoacoustics [J]. *Journal of the Acoustical Society of America*, 1971, **49** (2): 467-477.
- [11] Krishnan A, Bidelman G M, Gandour J T. Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices [J]. *Hearing Research*, 2010, **268**(1-2): 60-66.
- [12] Akhoun I, Moulin A, Jeanvoine A, et al. Speech auditory brainstem response (speech ABR) characteristics depending on recording conditions, and hearing status: An experimental parametric study [J]. *Journal of Neuroscience Methods*, 2008, **175**(2): 196-205.

(上接第 128 页)

- [12] 向春. 大学生创业倾向影响因素研究——以清华大学学生为对象 [D]. 北京: 清华大学, 2010.
XIANG Chun. Influencing Factors on Entrepreneurial Intention of College Students: Based on Students at Tsinghua University [D]. Beijing: Tsinghua University, 2010. (in Chinese)
- [13] 雷家骕, 王兆华. 高技术创业管理: 创业与企业成长 [M]. 北京: 清华大学出版社, 2008.
LEI Jiasu, WANG Zhaohua. High-Tech Entrepreneurship Management [M]. Beijing: Tsinghua University Press, 2008. (in Chinese)
- [14] 钱永红. 创业意向影响因素研究 [J]. 浙江大学学报: 人文社会科学版, 2007, **37**(4): 144-152.
QIAN Yonghong. On the determinants of entrepreneurial intentions [J]. *Journal of Zhejiang University: Humanities and Social Sciences*, 2007, **37**(4): 144-152. (in Chinese)
- [15] 邱皓政. 结构方程模式: LISREL 的理论、技术与应用 [M]. 台北: 双叶书廊有限公司, 2005.
QIU Haozheng. Principles and Practice of Structural Equation Modeling with LISREL [M]. Taipei: Yeh Yeh Book Gallery, 2005. (in Chinese)
- [16] 黄芳铭. 结构方程模型: 理论与应用 [M]. 北京: 中国税务出版社, 2005.
HUANG Fangming. Structural Equation Modeling: Theory and Practice [M]. Beijing: China Tax Publishing House, 2005. (in Chinese)
- [17] 林嵩. 结构方程模型原理及 AMOS 应用 [M]. 武汉: 华中师范大学出版社, 2008.
LIN Song. Theory on Structural Equation Modeling and Application of AMOS [M]. Wuhan: Huazhong Normal University Press, 2008. (in Chinese)