

## 韵律块基频曲线的优化及规则

刘浩杰<sup>①②</sup> 杜利民<sup>①</sup> 付跃文<sup>③</sup>

<sup>①</sup>(中国科学院声学研究所 北京 100080)

<sup>②</sup>(胜利油田物探研究院 东营 257000)

<sup>③</sup>(南京工业大学信息科学与工程学院 南京 210009)

**摘要:** 汉语规则合成系统中, 连续语流基频曲线(F0 曲线)的生成并不是各个合成单元 F0 曲线的简单拼接, 而是语音学功能单元的综合作用。该文提出了汉语韵律块基频曲线优化的正演思想, 使韵律块内的基频曲线融入重音强度、形状失真度以及发音速度等语境和发音的个体信息, 提高合成语音的自然度。基于这种优化思想, 该文针对聚类后的单音节、二音节和三音节韵律块的基频曲线, 利用最小均方差准则通过反演提取了各个单元的优化相关参数(高音线、低音线、平滑因子、形状失真度、重音强度)。对音节在韵律块中的位置因素和声调因素对优化相关参数的影响的统计分析表明了参数提取结果的可靠性和基频曲线优化的合理性, 得到了优化控制参数在规则合成系统中具体的应用规则。实际的听测实验表明, 韵律块基频曲线进行优化前后, 合成系统的清晰度分别为 3.25 和 3.35, 自然度分别为 2.9 和 3.31。

**关键词:** 规则语音合成; 基频(F0)曲线; 优化; 自然度

中图分类号: TN912.3

文献标识码: A

文章编号: 1009-5896(2007)01-0071-05

## F0 Contour Optimization and Its Rules in Chinese

Liu Hao-jie<sup>①②</sup> Du Li-min<sup>①</sup> Fu Yue-wen<sup>③</sup>

<sup>①</sup>(Institute of Acoustics, Chinese Academy of Sciences, Beijing 100080, China)

<sup>②</sup>(Geophysical Research Institute, Shengli Oil Field, Dongying 257000, China)

<sup>③</sup>(College of Information Science and Engineering, Nanjing University of Technology, Nanjing 210009, China)

**Abstract:** The fundamental frequency contour (F0 contour) for utterance in rule-based speech synthesis system, is shaped by many functional unit in phonetics, not only the simple concatenation of F0 contour among the nearby syllables. In order to improve the naturalness of synthesized speech, this paper proposes a new forward idea of F0 contour optimization in Chinese prosodic chunk, which can integrate the environmental factors (such as, the stress, the distortion of syllable, the articulation velocity, etc.) into the F0 contour. And based on the idea of optimization, this paper inversely extracts the parameters associated with optimization (namely the top-line, the bottom-line, the smoothness, the distortion, the stress) from the clustered F0 contour using the MMSE principle for the monosyllable, the disyllable, the trisyllable chunks. Further, this paper analyzes the influence of position and tone to the parameters associated with optimization. The analyzed result shows the reliability of the extracted parameters and the rationality of the optimization theory on the whole, so the rules of the parameters associated with optimization can be got for the different prosodic chunk in speech synthesis system. The actual listening test shows that, the scores of intelligibility are 3.25 and 3.35 before and after the optimization, and the scores of naturalness are 2.9 and 3.31.

**Key words:** Rule-based speech synthesis system; F0 contour; Optimization; Naturalness

### 1 前言

人类自然言语的基频曲线融入了语法、语音、情感及发音实体等多方面的因素, 因而发出的语音自然流畅。按照这些因素对基频曲线作用方式的不同, 可以分为 3 种: (1) 本征作用: 它对应于自然语流中基本的声调模式(对于普通话有 4 种基本声调: 阴平、阳平、上声、去声); (2) 心理作用: 依据所要表达的内容及方式(即心理作用), 发音人通过控制音域及重音的方式实现发音的高低强弱; (3) 生理作用: 语音的生成是通过

发音器官的运动产生的, 这些运动受到发音器官各种限制性因素(运动的连续性和平滑性)的作用<sup>[1-3]</sup>。因此在规则合成系统中, 寻求这种综合作用的实现方式, 对合成系统自然度的提高有重要意义。

以往的研究着重于对各种作用的单独分析, 而对其综合作用的形式及效果则涉及很少。参考 Stem-ML 标注体系<sup>[1]</sup>, 本文提出了对汉语韵律块的 F0 曲线进行优化的正演思想。在规则合成系统的音域调整的基础上, 这种优化能够在指定平滑因子、音节重音强度、音节形状失真度的条件下, 实现音域调整基础上的基频(F0)曲线的优化, 在很大程度上消除了拼接处的

基频间断,使F0曲线的生成更合乎生理机制的特点。

进而,如果已知优化后的F0曲线(即理想基频曲线)和模板基频,则利用最小均方误差(MMSE)准则通过反演的方式可得到理想的优化相关韵律参数,唯一的区别就是问题的求解由线性转变成了非线性。由于发音的随意性和灵活性,为了得到优化控制参数的一般规则,问题的关键是优化后的理想的F0曲线的确定、优化搜索算法的选择以及参数提取后的分析。本文基于聚类后的F0曲线,利用内部映射牛顿法的子空间置信域法进行了优化相关参数的提取,并对这些参数所受的音节在韵律块中的位置因素和声调因素的影响分别进行了统计分析和比较,从而检验韵律参数提取方法的可靠性,并得到其在合成系统中的具体应用规律。

本文分为以下几个部分。第2节阐述了汉语韵律块基频曲线优化的基本思想;第3节说明了优化相关参数具体提取的实验设计;第4节则对优化相关参数提取的结果进行了统计和方差分析,对这些参数在合成系统中的具体应用规则做了总结;第5节对64种声调组合的三音节韵律块进行了听测实验,判断基频曲线优化对提高合成系统性能的效果;最后对全文进行了总结。

## 2 韵律块基频曲线的优化思想

韵律块的F0曲线的优化是在对发音的生理和心理机制进行初步认识的基础上,使发音的费力程度和准确程度达到相对的平衡,即实现F0曲线生成的优化。其中,准确度用合成单元的F0曲线与其稳定模式下的F0曲线的差别程度表示(包括F0曲线的形状差别和平均值差别);费力程度对应于F0曲线的平滑和连续性(前者用F0曲线的二阶导数的累加描述;后者则用F0曲线斜率的累加来描述)。这种优化实现了重音作用下F0曲线的优化,在很大程度上消除了由于基频拼接的不连续所造成的发音间断。

人体已经建立起了本能的趋势在自然发音的准确度和费力程度之间实现最优化,而在语音合成系统中则需要通过一定的数学模型来实现。在Stem-MI标注体系所建立的数学模型的基础上,本项研究利用基于 $\chi^2$ 估计的方法<sup>[4]</sup>,使非线性问题转化为简单线性问题的求解。假定韵律块所包含的音节个数为 $M$ (即模板个数为 $M$ ),基频的采样周期为 $\Delta T$ ,与发音人特征有关的平滑因子为smooth,强度(stress( $i$ ), $i=1,\dots,M$ ),音节形状失真度,则各个模板的基频采样点数目为(sample( $i$ ), $i=1,\dots,M$ ),总的基频采样点为 $N = \sum_{i=1}^M \text{tag}(i)$ 。待求解的最优化问题可描述为

$$G = \sum_{i=1}^{N-1} \left( w_l \cdot \frac{-p_i + p_{i+1}}{\Delta T} \right)^2 + \sum_{i=1}^{N-2} \left( w_s \cdot \frac{-p_i + 2 \cdot p_{i+1} - p_{i+2}}{\Delta T^2} \right)^2$$

$$R = \sum_{m=1}^M \sum_{j=1}^{\text{sample}(m)} \{w_c(m) \cdot [p'(m,j) - y'(m,j)]\}^2$$

$$+ \sum_{m=1}^M \{w_a(m) \cdot [\bar{p}'(m) - \bar{y}'(m)]\}^2 \quad (1)$$

$$E = G_{(\text{费力程度})} + R_{(\text{准确程度})}$$

其中, $w_l, w_s$ 分别为韵律块内F0曲线的连续性和平滑性特征的权重; $w_c(m), w_a(m)$ 分别为第 $m$ 个模板的F0曲线的形状特征和平均值特征权重; $p_i$ 为韵律块内第 $i$ 个基频采样值; $p'(m,j)$ 为第 $m$ 个模板的第 $j$ 个优化后的基频采样值; $y'(m,j)$ 为第 $m$ 个模板的第 $j$ 个原始采样的基频值; $\bar{p}'(m), \bar{y}'(m)$ 分别为第 $m$ 个模板优化后和原始的基频平均值。

该模型能够在指定平滑因子、各个音节重音强度以及形状失真度等优化控制参数的条件下,使韵律块内F0曲线的生成以最优的方式(最省力,最准确)表达特定的目的。

进而,如果已知了最优的F0曲线(即理想F0曲线,如假定真实语料的基频为优化后的最优F0曲线),则韵律块F0曲线优化的结果已知,此时调整平滑因子、重音强度、形状失真度等相关韵律控制参数的值,使在该条件下优化后的F0曲线与已知的最优F0曲线的吻合程度(用两者的最小均方误差衡量)最高,此时的最优控制参数值可以认为是最佳的参数值。另外,韵律块F0曲线的优化建立在音域调整的基础上,且模板基频已经进行了线性标准化<sup>[5]</sup>,因而在最优控制参数的搜索中,也需要确定最优F0曲线的音域(高音线和低音线)。该问题的求解可以用如下的数学模型表示:

假设最优基频曲线 $P_i$ 标准化后的模板基频曲线 $P_T$ ,音域调整后的模板基频曲线 $P_i$ ,高音线为 $F_h$ ,低音线为 $F_l$ ,调中值为 $F_0$ ,调域为 $F_d$ ,基频曲线优化函数为( $P_i = \text{Opti}F_p(P_t, \text{smooth}, \text{Stress}, \text{Distortion})$ ,

$$F_0 = (F_h + F_l) / 2; F_d = (F_h - F_l); P_i = P_T \cdot F_d + F_0$$

则待求解问题为

已知变量:  $P_i$  和  $P_T$

未知变量:  $F_h, F_l, \text{smooth}, \text{Stress}, \text{Distortion}$

求解准则: 使  $\text{Err} = \sum (P_i - P_0)^2$  最小

因此,在对最优控制参数的求解过程中,也得到了已知最优F0曲线所对应各个音节的高低音线。这对于所求取参数的可靠性判断以及相关的语调研究都提供了基础。该问题的求解已经超出了线性问题的范畴,是非线性复杂问题的优化求解。本文利用基于内部映射牛顿法的子空间置信域法<sup>[6]</sup>进行求解优化相关的参数。该方法比较适合于大型复杂优化问题。

## 3 基频曲线优化的实验设计

本文所有的实验语料都是建立在“863”语音合成语料库的基础上。所研究的对象主要是浊辅音声母(首音节除外)组合的单音节、双音节、三音节韵律块(保证基频的连续性),从中寻找相关韵律参数在韵律块的不同位置、不同声调等语境时的变化规律。

考虑到人类发音的灵活性所导致的基频变化的复杂性,本文利用李香春<sup>[7]</sup>对“863”语料库两位男发音人的基频数据进行聚类分析得到的单音节(4)、双音节(16)、三音节(64)组合的基频数据为研究对象。聚类后的各个韵律块组合的基频数据在很

大程度上剔除了发音的灵活性所造成的随机误差。

分别以聚类后的各个韵律块的基频数据作为已知的理想 F0 曲线, 逐个单元提取优化相关参数, 按照不同的韵律组合分别进行统计分析, 得到不同韵律块组合的、分别随位置因素和声调因素变化的各优化相关参数的平均值。通过对各个韵律块的优化相关参数的统计平均值的分析, 对优化问题求解的可靠性进行判断, 并寻求其在合成系统中的具体应用规则。

### 4 实验结果及分析

本文在优化相关参数的提取中, 利用的是迭代优化算法, 因此, 在求解过程中, 求取的结果有可能陷入局部极小点而使参数求取的结果出现偏差, 甚或出现很大的异常。为了减小局部极小点对参数提取结果的影响, 本文在搜索过程中限定了各种参数的取值范围, 但这种限制仍然不能解决所有的问题。在优化相关参数的提取中, 仍然出现了参数的值在各自的限定极值的情形, 而且有的音节优化后的基频与理想基频的相关系数甚至为负值, 这些情况很有可能干扰对参数正确规律的把握。特别地, 高低音线的正确求取对于优化控制参数的正确获得是至关重要的, 因为只有音域调整合适, 才能在韵律块基频优化理论的基础上对优化控制参数进行正确提取。为了剔除这种音域调整的可能不合适给优化控制参数规律的研究的干扰, 本文在以下的分析中剔除了高低音线的值在音域极值时和优化后的基频与理想基频的相关系数为负值时的音节的相关参数。

#### 4.1 实验结果的误差及平滑因子分析

首先统计了各个韵律块组合内各个单元的平滑因子、平均迭代次数以及其理想基频和优化后基频的均方差, 如表 1 所示。优化后基频与理想基频的均方差表示的是两者的差别程度, 因此与原发音人的基音大小是不相关的。各韵律块组合的平均均方差小则表明, 各种参数对模板基频的综合作用能够实现理想基频的效果。而且随着韵律块音节数目的增多, 平滑因子逐渐增大, 这与发音的一般认识相一致。因为音节数目越多, 相互影响的程度就强, 因此各个音节的发音准确度相对降低, 所以平滑因子增大。

表 1 各韵律块组合的平均平滑因子、均方差及迭代次数

	平滑因子	均方差(Hz)	迭代次数
单音节	0.01	12.2	98
双音节	0.03	4.2	517
三音节	0.05	2.6	209

#### 4.2 位置因素对优化相关参数的影响

图 1, 图 2, 图 3 分别为各韵律块组合的高低音线、重音强度和形状失真度的平均值的示意图。从图 1 中可以明显的看出: (1)各音节组合的高低音线均呈明显的下倾趋势, 低音线的变化大于高音线的变化。这种音高下倾效应与沈炯<sup>[8]</sup>、王安红等<sup>[9]</sup>对语调的研究结论相一致。(2)各韵律块的高音线随着韵律块内音节数目的增加, 均有显著的增大, 这与人发音的“预置效应(pre-planning)”<sup>[3]</sup>相一致。(3)各韵律块组合的低音线出现了重

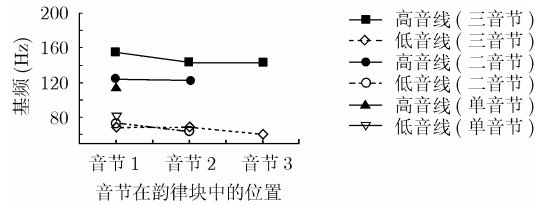


图 1 各韵律块组合的高低音线

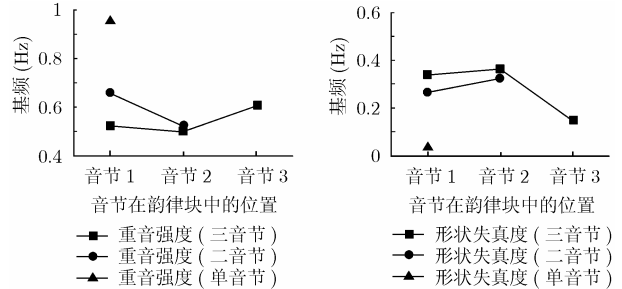


图 2 各韵律块组合的重音强度 图 3 各韵律块组合的形状失真度

叠的下倾趋势, 说明了低音线主要与韵律节奏的完整性相关。这些现象不但从量化的角度证实了抽象的语音现象, 也初步说明了所提取的高低音线的可靠性。

图 2 中, 双音节韵律块的重音强度为“前重”模式; 三音节韵律块的重音强度呈现“中轻重”模式。对于双音节韵律块, 不同的人对不同语料的统计结果并不一致<sup>[10]</sup>, 所以无法确定哪种模式更合乎人发音的自然规律; 对于三音节韵律块, “中轻重”的模式似乎得到了普遍的承认<sup>[11]</sup>。因此从这个意义上, 也表明了优化相关参数提取的可靠。

图 3 中说明, 各韵律块组合的形状失真度的变化与图 2 对应组合的重音强度的变化相对应, 即韵律块组合中重音强度大, 其形状失真度就小。另外, 随着韵律块内音节数目的增加, 音节相互间的影响增大, 因而形状失真度增大(双音节韵律块和三音节韵律块的前两个音节), 但三音节的末音节形状失真度很小。

#### 4.3 声调因素对高低音线的影响

图 4 为各韵律块组合的高低音线随声调因素的变化。从图 4 中可以看出, 随着韵律块音节数目的增加, 音域出现加宽(高音线上升, 低音线下降), 且低音线变化的幅度明显的低于高音线的变化, 这与“预置”效应<sup>[3]</sup>和低音线所承担的语言学功能相对应。这些现象也说明了优化参数提取的结果与一般的语音学理论一致。

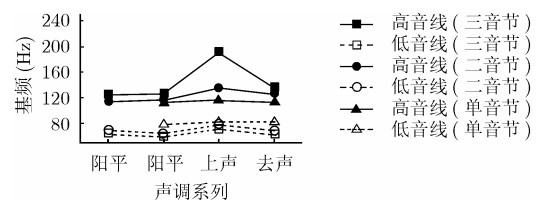


图 4 各韵律块组合的高低音线随声调因素的变化

#### 4.4 优化相关参数的单因素及多因素方差分析

本文进一步对三音节韵律块的优化参数的提取结果分别按位置因素(三种水平, 即首音节、中音节和末音节)和声调因素(四种水平, 即阴平, 阳平, 上声和去声)进行了方差分析(如表

2 所示)。

表 2 两种因素对优化相关参数影响的单因素方差分析  
(显著性水平为 0.05)

	高音线	低音线	重音强度	形状失真度
音节在韵律块中的位置因素	0.3246	0.011	0.2663	0.0056
音节声调因素	0	0.002	0.1782	0.286

从表 2 中可以得到以下认识: (1)音节在韵律块中的位置对于低音线、形状失真度都有显著的影响( $p < 0.05$ ),而对高音线和重音强度的影响有限( $p > 0.05$ )。这与这些参数所承担的语言学功能一致。低音线所受的显著影响与低音线所承担的“节奏的完整性”语言学功能一致,形状失真度表示音节在具体语境中的基频的形状失真度,所受显著影响也表明其在韵律块中也遵循相应的模式,即首音节及末尾音节保持其声调特征的程度强,中间音节保持其声调特征的程度就弱;高音线和重音强度由于是与发音的强弱有关,因此受位置因素的影响就小。(2)从表 2 可以看出,重音强度以及形状失真度不受音节声调变化的显著影响( $p > 0.05$ ),而高低音线则受音节声调的显著影响( $p < 0.05$ )。这是一个奇怪的现象,从图 4 声调因素对高低音线影响的示意图中,也可以明显地看到此现象,且上声音节的高低音线普遍抬高,导致音域上升。这是什么原因呢?一方面有可能是统计的方法和研究对象导致的,另一方面可能是音节本身的特殊性质,因为不同声调的音节,发音时声带的紧张程度等力量因素不同。我们发不同声调的单音节时可以明显的体会到这种差异。

为了能够确定位置因素、声调因素对优化相关参数的综合作用的程度及主要因素,进行了两者的  $3 \times 4$  水平的双因素方差分析(表 3)。从表 3 可以进一步看出,位置因素对低音线、形状失真度都有显著的影响,而对高音线和重音强度则没有显著影响;声调因素对高低音线有显著的影响,而对重音强度和形状失真度则没有显著影响。这与位置因素及声调因素的单因素方差分析的结论一致。双因素方差分析的结果也表明,位置因素与声调因素对形状失真度和重音强度存在显著的交互效应,而对高低音线则没有显著交互影响。这为在音域调整基础上实现重音和形状失真度联合作用下的韵律块基频曲线的优化提供了进一步的理论基础。

表 3 两种因素对优化相关参数的双因素方差分析(显著性水平为 0.05)

	高音线	低音线	重音强度	形状失真度
音节在韵律块中的位置因素	0.4474	0.0401	0.1344	0.0348
音节声调因素	0	0.0046	0.1947	0.6724
两者的交互作用	0.2349	0.9371	0.0048	0.0002

#### 4.5 优化控制参数在合成系统中的应用规则

上述的优化相关参数的各种分析表明,基于基频曲线优化

所提取的各种参数的规律与相关的语音学理论和发音的生理和心理机制基本一致。这既说明了韵律块基频曲线优化的合理性,也为各种参数提供了具体的优化规则。具体而言,在实现韵律块基频曲线的优化时,可以为不同的韵律块设置基本的优化控制规则,然后根据语境进行适当的调整。

对于平滑因子,韵律块内音节数目越多,语速越快,其作用的程度越强,值就越大。可以设置单音节、双音节、三音节韵律块的平滑因子分别为 0.01, 0.03, 0.05。更多音节数目的韵律块可适当增大平滑因子。

对于形状失真度,尽管位置因素和声调因素对其有显著的影响,但这种影响很难区分,因此在具体应用时,可以只考虑位置因素对其影响。根据不同的韵律块,可以进行如下设置:单音节(0.03);二音节组合(0.27, 0.33);三音节组合(0.34, 0.36, 0.15)。

对于重音强度,由于其所受的影响因素十分复杂,因此可以根据具体的合成系统进行适当的调整。一般遵循的模式如下:单音节为“重”模式;双音节为“前重模式”;三音节韵律块为“中轻重”模式。

## 5 韵律块基频曲线优化的听测试验

上述几节分别论述了韵律块基频曲线优化的思想,优化控制参数的反演提取以及相关参数的统计分析。结果初步表明了韵律块基频曲线优化的合理性。最后得到了优化控制参数在合成系统中的基本应用规则。所以实验的目的就是验证韵律块基频曲线的优化对合成系统性能的影响。

听测实验的平台为商用 VSS 合成系统。该系统采用了谐波噪声的合成内核,能够在较大程度改变合成单元超音段特征参数的条件下仍能保持较高的可懂度。实验中,听音人对合成系统的原始生成语音、韵律块基频曲线优化后的生成语音与真人发音进行比较,并对前两者进行主观评分(Mean Opinion Score, MOS)。评分的项目包括发音的清晰度和自然度。

15 名听音人听力正常且熟悉普通话。进行听测实验时,事前并不告诉听音人所发音的类别,以避免主观的影响,但告知听音人着重对合成语音的韵律状况进行评分。考虑到不同的听音人对事物的判断标准不一样,因而对合成语音的评分标准也是不一致的。因此统计时需要首先对每个人的所有打分进行归一化,然后再对所有人的打分进行平均。

实验的语料全部来自“863”语音合成语料库。包括 64 个浊辅音声母组合(首音节除外)的三音节韵律块,包括了所有声调组合的三音节韵律块,因此有较大的代表性。

听测实验的结果表明,进行韵律块基频曲线优化后,合成语音的清晰度得分由 3.25 提高到 3.35;自然度得分由 2.9 提高到 3.31。

## 6 结束语

在对发音的生理和心理机制认识的基础上,本文分别从正演和反演的角度对汉语的基频曲线进行了研究。首先,基于基

频曲线的正演, 本文提出了对韵律块基频曲线进行优化的思想, 使合成单元的基频曲线融入了重音强度等语境信息和发音的生理限制, 使规则合成系统基频曲线的生成更合乎自然语音的规律。然后, 本文又从反演的角度, 针对汉语自然语流基频曲线的特点, 利用优化算法量化提取了高音线、低音线、重音强度、形状失真度等相关韵律参数, 并结合相关的语音学理论对这些参数的可靠性进行了统计和方差分析, 得到了优化控制参数在合成系统中的具体应用规则。

研究结果从量化的角度证实了以下语音学认识: (1)音高下倾; (2)低音线主要与韵律节奏的完整性有关; (3)预置效应; (4)韵律块的末音节保持其声调特征的能力较强, 即末音节的形状失真度低, 重音强度一般较大; (5)各韵律块的重音强度遵循一定的稳定模式。这些一致的认识初步地说明了韵律块基频曲线优化的合理性和优化控制参数提取的可靠性。

实际的听测实验表明, 对规则合成系统的韵律块基频曲线进行优化处理后, 合成语音的清晰度提高了 0.1, 自然度提高了 0.41。

### 参 考 文 献

- [1] Greg Kochanski and Chin Shih. Stem-ML: Language-independent prosody description. in Proceedings of the International Conference on Spoken Language Processing 2000, Beijing, China, 2000, Vol.3: 239–242.
  - [2] Fujisaki Hiroya. The fundamental frequency contour of speech: Its modeling, underlying mechanisms, and application to multilingual speech synthesis. In Proceedings of ICSP'99, Seoul Korea, 1999: 19–26.
  - [3] Xu Yi. Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech Communication*, 2001, 33(4): 319–337.
  - [4] Press H and Teukolsky A, *et al.* Numerical recipes in c. New York: Cambridge University Press, 1992: 657–661.
  - [5] 杨顺安. 浊声源动态特性对合成音质的影响. *中国语文*, 1986, (3): 173–181.
  - [6] Coleman T F and Li Y. An interior, trust region approach for nonlinear minimization subject to bounds. *STAM Journal on Optimization*, 1996, (2): 418–445.
  - [7] 李香春. 汉语单音节、两音节组和三音节组基频曲线建模方法研究. [博士论文], 中国科学院声学研究所, 2002.
  - [8] 沈炯. 北京话声调的音域和语调. *北京语音实验录*, 北京: 北京大学出版社, 1985: 73–130.
  - [9] 王安红, 陈明, 吕士楠. 基于言语数据库的汉语高音下倾现象研究. *声学学报*, 2004, 29(4): 353–358.  
Wang Anhong, Chen Ming, and Lu Shinan. The study of declination in speech database in standard Chinese. *Acta Acustica*, 2004, 29(4): 353–358.
  - [10] 王韞佳, 初敏等. 连续话语中双音节韵律词的重音感知. *声学学报*, 2003, 28(6): 534–539.  
Wang Yunjia and Chu Min, *et al.* The perception of disyllabic word stress of Chinese speech in utterance. *Acta Acustica*, 2003, 28(6): 534–539.
  - [11] 颜景助, 林茂灿. 北京话三字组重音的声学表现. *方言*, 1988, (3): 227–237.
- 刘浩杰: 男, 1976 年生, 博士, 从事声学信号处理、语音合成方面的研究。  
杜利民: 男, 1957 年生, 研究员, 博士生导师, IEEE 高级会员、中国电子学会理事, 中国声学学会会士, 主要从事语音信号处理、自然语言理解、语音交互信息技术方面的研究。  
付跃文: 男, 1968 年生, 博士, 从事语音信号处理研究。