

文章编号:1001-9081(2009)05-1433-04

基于自相关平方函数与小波变换的基音检测

林 琴,郭玉堂,刘亚楠

(合肥师范学院 计算机科学与技术系,合肥 230061)

(linqin1201@yahoo.com.cn)

摘 要:在背景噪声干扰条件下,研究语音信号的基音周期,提出了一种基于自相关平方函数与小波变换结合的基音检测算法。该算法先用小波变换对带噪语音去噪,然后再求语音的自相关平方函数以突出真实基音周期的峰值,以获取较精确的基音周期。实验结果表明,与传统的自相关法相比,该算法鲁棒性好,具有更高的准确性,且计算复杂度低,利于语音合成和编码的实时处理。

关键词:语音信号;自相关平方函数;小波变换;基音检测

中图分类号:TP391.42 **文献标志码:**A

Pitch detection based on autocorrelation square function and wavelet transform

LIN Qin, GUO Yu-tang, LIU Ya-nan

(Department of Computer Science and Technology, Hefei Normal College, Hefei Anhui 230061, China)

Abstract: A robust pitch detection algorithm based on autocorrelation square function and wavelet transform was proposed for speech signals severely corrupted by noise. Noisy speech signal was first preprocessed by wavelet transform. The autocorrelation square function had been applied to emphasize the peak of the true pitch period. This algorithm can get the exact pitch of the speech signals in strong noisy environment. Experimental results indicate that, compared with classical pure autocorrelation method, the proposed algorithm has better robustness and higher accuracy with lower computational complexity, which is good for speech synthesizing and coding in real-time.

Key words: speech signal; autocorrelation square function; wavelet transform; pitch detection

0 引言

人在发语音时,声门周期性地打开与闭合,使来自肺部的空气流形成一串周期性的脉冲气流进入声道,这串周期性脉冲的周期称为基音周期。基音周期是语音信号处理中的重要参数之一,准确而快速的基音周期提取对语音编码、合成、识别等都具有重要意义。近几十年来,提出了多种基音检测方法,最具代表性是自相关法^[1]、平均幅度差函数法、倒谱法及小波变换法。理论上,它们都可以准确地提取无背景噪声的语音基音周期。但通常得到的语音信号有各种背景噪声,所以噪声环境下的基音检测是必要的。

因此,为了在噪声环境下精确提取基音周期,本文提出了一种基于自相关平方函数和小波变换相结合的基音检测算法。即首先利用小波变换,滤除非锐变信息和噪声的影响,以加强声门闭合瞬间的锐变信息;然后把小波变换后的结果作为被处理信号,计算其自相关平方函数;最后利用自相关平方函数在基音周期的各整数倍点上具有较高的峰值,就可以检测出基音周期。该算法在信噪比很低的情况下能准确地检测出基音周期,大大提高了算法的鲁棒性,并且容易实现。

1 原理

1.1 自相关函数

自相关分析是语音信号时域分析的一种常用方法,它由相关函数^[2]定义。相关函数可用来测定信号间的时域相似性。如果两个信号完全不同,相关函数接近于零;如果两个信

号波形相同,就会在超前、滞后处出现峰值。自相关函数可用于研究信号本身,如波形的同步性,周期性等。

对于确定性信号序列,自相关函数定义如下:

$$R(k) = \sum_{m=-\infty}^{\infty} x(m)x(m+k) \quad (1)$$

对于随机性信号序列或周期信号序列,自相关函数定义为:

$$R(k) = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{m=-N}^N x(m)x(m+k) \quad (2)$$

如果序列 $x(n)$ 具有周期 N_p , 即 $x(n) = x(n + N_p)$, 则:

$$R(k + N_p) = \sum_{m=-\infty}^{\infty} x(m)x(m+k+N_p) = \sum_{m=-\infty}^{\infty} x(m)x(m+k) = R(k) \quad (3)$$

$$R(k + N_p) = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{m=-N}^N x(m)x(m+k+N_p) = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{m=-N}^N x(m)x(m+k) = R(k) \quad (4)$$

由此可得,自相关函数 $R(k)$ 与周期序列 $\{x(n)\}$ 具有相同的周期 N_p 。

语音信号中清音没有周期性,它的自相关函数 $R(k)$ 也没有周期, $R(k)$ 会随着 k 的增大迅速衰减。浊音信号具有准周期性,它的自相关函数 $R(k)$ 具有与 $\{x(n)\}$ 具有相同的周期。自相关法的基音检测就是利用 $R(k)$ 的这一性质对语音信号进行基音检测的。

收稿日期:2008-11-24;修回日期:2009-02-18。 基金项目:安徽省教育厅教育科研基金资助项目(2007JYXM547)。

作者简介:林琴(1982-),女,安徽六安人,硕士研究生,主要研究方向:语音信号处理、人工智能、模式识别;郭玉堂(1962-),男,安徽安庆人,教授,主要研究方向:模式识别、图像处理;刘亚楠(1984-),女,山东济宁人,硕士研究生,主要研究方向:模式识别、图像处理。

1.2 小波变换

设 $\varphi(t) \in L^2(R)$ ($L^2(R)$ 表示均方可积一维函数的 Hilbert 空间), 其傅里叶变换为 $\hat{\varphi}(w)$ 。

当 $\hat{\varphi}(w)$ 满足完全重构条件:

$$C_\varphi = \int_R \frac{|\hat{\varphi}(w)|^2}{|w|} dw < \infty \quad (5)$$

称 $\varphi(t)$ 为母小波。将母小波 $\varphi(t)$ 经二进伸缩和平移后, 就可以得到一个小波序列:

$$\varphi_{j,k}(t) = 2^{-j/2} \varphi(2^{-j}t - k); j, k \in \mathbf{Z} \quad (6)$$

显然 $\varphi_{j,k}(t) \in L^2(R)$, 若存在 $0 < A < B < \infty$, 使 $A \leq \sum_{j \in \mathbf{Z}} |\hat{\varphi}(w)|^2 \leq B$, 则称 $\varphi_{j,k}(t)$ 为一个二进小波。对任意函数 $f(t) \in L^2(R)$ 的二进小波变换为:

$$W_\varphi f(j, k) = \langle f, \varphi_{j,k} \rangle = 2^{-j/2} \int_{-\infty}^{\infty} f(t) \varphi(2^{-j}t - k) dt \quad (7)$$

小波变换一般由符合条件的有限长脉冲响应滤波器 (Finite Impulse Response, FIR) 实现, 取 $t = n, n \in \mathbf{Z}$ 时得到离散信号的二进小波变换, 其实现算法^[3]为:

$$\begin{cases} S_{2^j} f(n) = \sum_{l \in \mathbf{Z}} h_l S_{2^{j-1}} f(n - 2^{j-1}l) \\ W_{2^j} f(n) = \sum_{l \in \mathbf{Z}} g_l S_{2^{j-1}} f(n - 2^{j-1}l) \end{cases}; j = 1, 2, \dots, J \quad (8)$$

本文选取具有紧支集的 2 次样条小波, 相应的滤波器传递函数^[4]为:

$$\begin{cases} H(w) = e^{iw/2} [\cos(w/2)]^3 \\ G(w) = 4ie^{iw/2} \sin(w/2) \end{cases} \quad (9)$$

可得理想半带低通滤波器 $H(n)$ 和理想半带高通滤波器 $G(n)$ 的有限脉冲响应如表 1。

表 1 $H(n), G(n)$ 的有限脉冲响应

滤波器	n			
	-1	0	1	2
$H(n)$	0.125	0.375	0.375	0.125
$G(n)$	0.000	-2.000	2.000	0.000

在每一尺度 $s = 2^j$ 上, 信号 $S_{2^j} f$ 通过滤波器 $H(n), G(n)$ 后, 分解为低半带 $S_{2^j} f$ 和高半带 $W_{2^j} f$ 。

$$\begin{cases} S_{2^j} f = S_{2^j} f \cdot G(n) \\ W_{2^j} f = S_{2^j} f \cdot H(n) \end{cases} \quad (10)$$

当 $j = 0$ 时, $S_{2^j} f = S_1 f$, 代表语音信号的采样值。

1.3 自相关平方函数与小波变换的结合

自相关函数在语音基音周期的各整数倍点上具有较高的峰值, 因此只要找到 $R(k)$ 的第一峰值 (除 $R(0)$ 外) 点^[5]的位置与 $k = 0$ 点的间隔, 便能估计出基音周期。可见, 第一峰值点的位置非常重要, 峰值越尖锐则确定的位置就越准确。为此本文在短时自相关函数基础上寻求一种更好的解决方法, 即求自相关函数 $R(k)$ 的平方, 用符号 $SR(k)$ 表示, $SR(k)$ 与 $R(k)$ 一样都呈周期性变化, 且周期一致, 可以发现 $SR(k)$ 的波形要比 $R(k)$ 的波形尖锐得多, 在峰值处尤为明显。因此用短时自相关的平方函数进行基音周期提取要准确得多。

自相关函数能准确地提取无背景噪声的语音基音周期, 但实际上得到的语音信号都含有各种背景噪声, 这时用自

相关函数法已经无法求出语音的基音周期, 并且把浊音错误地判别为清音。因此, 为了在噪声环境下精确提取基音周期, 本文提出了一种基于自相关平方函数和小波变换相结合的基音检测算法。即首先利用小波变换, 滤除非锐变信息和噪声的影响, 以加强声门闭合瞬间的锐变信息; 然后把小波变换后的结果作为被处理信号, 计算其自相关平方函数; 最后利用自相关平方函数在基音周期的各整数倍点上具有较高的峰值, 就可以检测出基音周期。

2 算法实现

自相关平方函数与小波变换结合的算法大致分为以下几个功能模块: 取样、预处理、分帧、二进小波变换、自相关平方运算、清浊音判决、基音周期提取和中值平滑滤波。

2.1 预处理

2.1.1 低通滤波

语音信号中包含了十分丰富的谐波分量, 所以语音信号中的基频分量往往不是最强的分量, 其中多次谐波成分有时甚至比基波分量还要强。然而, 语音的基音周期的变化范围是有限的, 变化范围为 20 ~ 150 样值 (8 kHz 采样), 与其对应的基音频率变化范围为 60 ~ 450 Hz。因此, 通过选择截止频率合适的前置低通滤波器, 可以降低语音信号的共振峰的影响, 又不削弱基音频率的信息。

低通滤波函数有很多, 考虑到计算量和实时处理的要求, 本文根据双线性变换法, 设计 800 Hz 5 阶椭圆低通滤波器, 通带的波纹不超过 3 dB, 阻带的衰减不小于 50 dB, 该低通滤波器的传递函数为:

$$\begin{aligned} H(z) = & (0.0063642 - 0.0079932z^{-1} + 0.0060935z^{-2} + \\ & 0.0060935z^{-3} - 0.0079932z^{-4} + 0.0063642z^{-5}) / \\ & (1 - 4.1226z^{-1} + 7.272z^{-2} - 6.7917z^{-3} + \\ & 3.3507z^{-4} - 0.69941z^{-5}) \end{aligned} \quad (11)$$

2.1.2 数值滤波

经 800 Hz 低通滤波的语音信号, 主要去除了第三和第四个高频共振峰的影响, 但第一个和第二个共振峰有时依然存在, 使浊音段语音信号的周期性模糊, 产生了错误的基音估计。

大量实验发现, 若在 800 Hz 低通滤波器之后级联一个宽度 $N = 9$ 的数值滤波器可有效去除这一影响, 突出浊音段语音信号的周期性, 该数值滤波器的传递函数^[7]为:

$$H(z) = \frac{1}{9} \sum_{n=0}^8 z^{-n} \quad (12)$$

2.2 算法步骤

设 $x(n)$ 为一帧带噪语音数据, 本文算法步骤如下。

步骤 1 对 $x(n)$ 作预处理 (预处理后的语音数据仍记为 $x(n)$)。

步骤 2 对 $x(n)$ 进行二进制小波变换, 则变换的最佳尺度^[8] $J = \text{int}(\text{lb}(f_s/f_o))$ 。其中 f_s 为信号采样频率, f_o 为基音频率上限。多数情况下基音频率在 100 ~ 250 Hz 之间, 选择其上限为 250 Hz)。

步骤 3 对小波处理后的信号求自相关平方函数:

$$R(k) = \sum_{m=1}^{len} x(m)x(m+k); k = 0, 20, 21, 22, 23, \dots, 150 \quad (13)$$

$$SR(k) = R^2(k) \quad (14)$$

式 (13) 中 len 为语音信号的帧长。

步骤 4 对 $SR(k)$, 采取以下判决准则:

$$MAX \begin{cases} \geq 0.25SR(0), \text{判为浊音} \\ < 0.25SR(0), \text{判为清音} \end{cases} \quad (15)$$

其中 $SR(0) \geq |SR(k)| MAX = \max SR(k) = SR(N_p)$; 基音周期 $T = 1000 \times N_p / f_s$ (ms) 或基音频率 $f = f_s / N_p$ (Hz)。

步骤 5 求出所有语音帧的基音周期后, 对基音进行中值平滑处理, 纠正个别估值的错误。

3 实验结果与分析

本实验所用的语音信号是在实验室的环境下录制, 采样频率为 8 kHz, 16 bit 量化。将语音信号分为若干帧, 每帧长 240 点, 相邻帧交叠 80 点, 利用 Hamming 窗进行加窗处理, 为了验证本文算法的有效性和准确性, 对语音信号加入高斯白噪声。

对单帧语音数据实验, 语音信号为一女性读英文元音 [a:]。

图 1 是用自相关方法对纯净语音信号进行基音检测的实验结果, 图 (a) 为一帧纯净语音信号, 图 (b) 为其自相关函数的波形, 图 (c) 为其自相关平方函数的波形。从图 1 可看出, 自相关平方函数波形要比 $R(k)$ 的波形尖锐得多, 在峰值处尤为明显, 而且检测出这段语音为浊音, 其包括 3 个基音周期, 平均周期为 36 个采样点, 即 $T = 4.500$ ms。

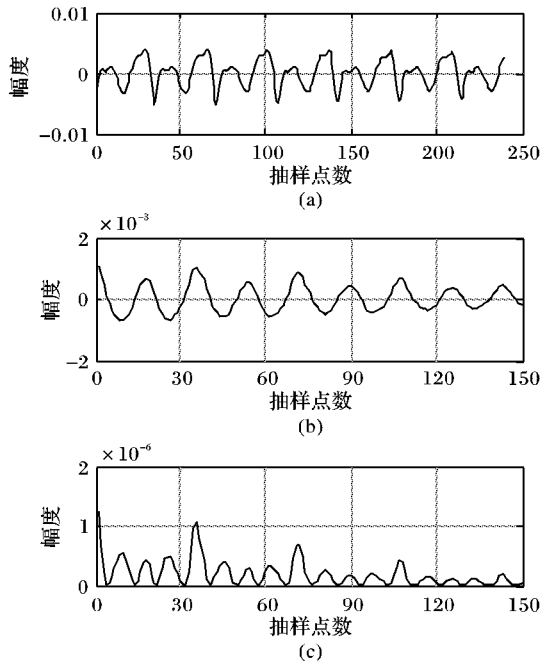


图 1 对一帧纯净语音信号自相关的基音检测

图 2 是用自相关方法对带噪语音信号进行基音检测的实验结果, 图 (a) 为加高斯白噪声的语音信号, 信噪比 (SNR) 为 -5 dB, 图 (b) 为其自相关函数的波形, 图 (c) 为其自相关平方函数的波形。从图 2 中可以看出, 由于噪声的干扰使得自相关函数法已无法求出语音的基音周期, 而是把它错误地判为清音, 即 $T = 0$ ms。

图 3 给出了本文算法对带噪语音信号进行基音检测的实验结果, 图 (a) 为加高斯白噪声的语音信号, 信噪比为 -5 dB, 图 (b) 为 $j = 3$ 时该段带噪语音信号二进小波变换的结果作为被处理信号而求出的自相关函数的波形, 图 (c) 为 $j = 3$ 时该段带噪语音信号二进小波变换的结果作为被处理信号而求

出自相关平方函数的波形。从图 3 可以看出, 本文方法能准确地提取带噪语音信号的基音周期, 平均周期为 36 个采样点, $T = 4.500$ ms, 检测结果和纯净语音信号检测的结果相同, 所以本文算法对噪声具有稳健性。

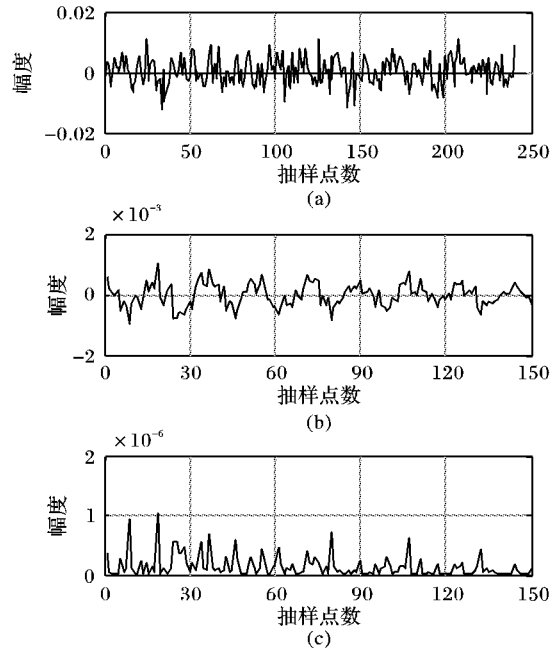


图 2 对一帧带噪语音信号自相关的基音检测

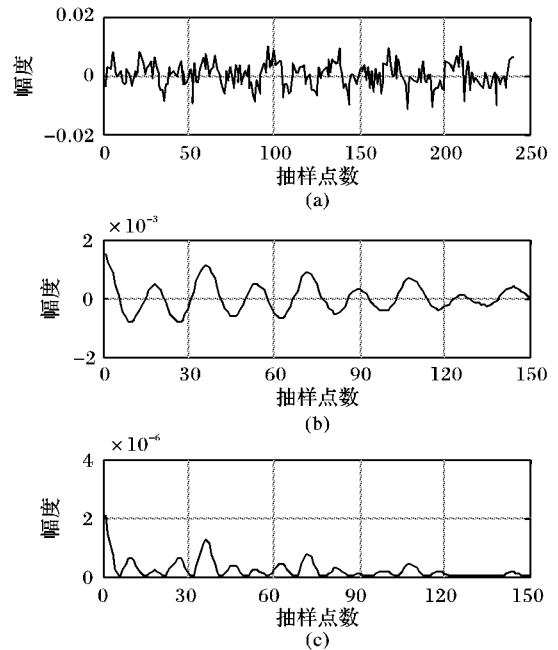


图 3 本文算法对一帧带噪语音信号的基音检测

对连续带噪语音信号的基音轮廓实验, 同时为了和女性基音周期进行比较, 语音信号分别采用一女性读“雅鲁藏布江”和一男性读“启动保密电话机”。按本文算法计算每一帧的基音周期, 最后拼合成整段语音信号的基音周期轮廓, 如图 4、5 所示。

图 4(a) 为加高斯白噪声的语音信号“雅鲁藏布江”, 图 (b) 为自相关方法对带噪语音信号进行基音检测的实验结果, 图 (c) 为本文算法对带噪语音信号进行基音检测的实验结果。图 5(a) 为加高斯白噪声的语音信号“启动保密电话机”, 图 (b) 为自相关方法对加噪语音信号进行基音检测的实验结果, 图 (c) 为本文算法对加噪语音信号进行基音检测的

实验结果。可以看出,本文算法不仅对噪声具有稳健性,而且对清浊音划分非常准确,基音周期轮廓很好地跟踪了基音周期的变化。从图中还可以看出女性的一般基音频率变化范围为 200~450 Hz,男性的一般基音频率变化范围为 60~200 Hz,多数情况下人的基音频率在 100~250 Hz。

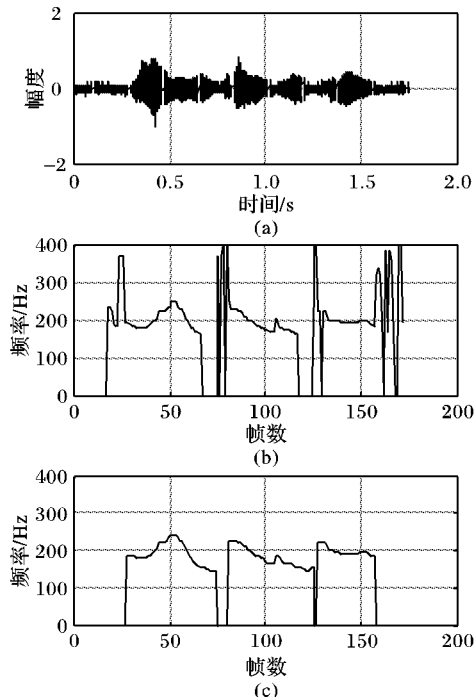


图 4 对女性连续带噪语音信号的基音检测

为了更直观地比较这两种方法,表 2 给出了两种方法对不同信噪比的语音信号的基音检测结果。以纯净语音信号的基音周期检测的结果为参考标准,语音信号为一女性读英文元音 [a:],与加入高斯白噪声后语音信号的基音检测结果进行比较。从表 2 可以看出,本文算法在低信噪比的情况下仍可以准确地提取基音周期。

表 2 两种方法对不同信噪比的语音信号的基音检测

检测方法	SNR/dB			
	∞	5	0	-5
自相关方法	4.5	4.5	2.1	0.0
本文算法	4.5	4.5	4.5	4.5

4 结语

本文将传统的基音检测的自相关方法和小波变换相结

(上接第 1432 页)

下一步将结合基于文本的自动构建技术,利用统计方法抽取图像情感领域中的概念词,利用本文提出的构建方法,构建更加完整的图像情感领域本体,为进一步实现基于本体的图像情感标注和检索奠定基础。

参考文献:

- [1] 杜小勇,李曼,王珊. 本体学习研究综述[J]. 软件学报, 2006, 17(9): 1837-1847.
- [2] ISO/IEC JTC1/SC29/WG11N6828, MPEG-7 Overview (version 10) [S]. 2004.
- [3] 陈贤明,王小铭. 基于本体与 MPEG-7 视频语义描述模型[J]. 华南师范大学学报: 自然科学版, 2007(2): 51-56.

合,提出了一种在噪声背景下的基音提取方法。实验结果表明,该方法比自相关方法具有更好的检测效果和噪声鲁棒性,并且计算复杂度低,有利于语音合成和编码的实时处理。

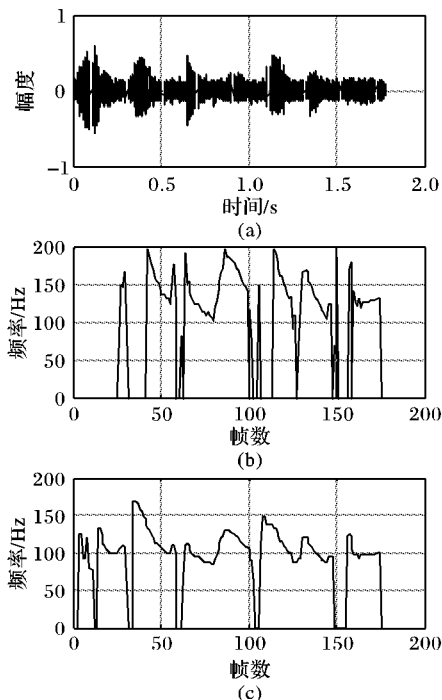


图 5 对男性连续带噪语音信号的基音检测

参考文献:

- [1] 楼红伟,赵建伟,胡光锐. 一种小波加权的基音检测方法[J]. 上海交通大学学报, 2003, 37(3): 447-449.
- [2] 俞卞民,李志均,金明录. 数字信号处理[M]. 西安: 西北工业大学出版社, 1998: 212-220.
- [3] CHUI C K. An introduction to wavelets[M]. New York: Academic Press, 1992.
- [4] 程俊,张璞,戴善荣,等. 小波变换用于信号突变的检测[J]. 通信学报, 1995, 16(3): 96-104.
- [5] 王艳芬,于洪珍,刘卫东,等. 基于时域自相关平方函数的基音周期估计[J]. 计算机工程与应用, 2004, 40(31): 83-84.
- [6] 李辉,戴蓓菁,陆伟. 基于前置滤波和小波变换的带噪语音基音周期检测方法[J]. 数据采集与处理, 2005, 20(1): 100-104.
- [7] 鲍长春,樊昌信. 基于归一化互相关函数的基音检测算法[J]. 通信学报, 1998, 19(10): 27-29.
- [8] KADAMBE S, BOUDREAUX-BARTELS G F. Application of the wavelet transform for pitch detection of speech signals[J]. IEEE Transactions on Information Theory, 1992, 38(2): 917-924.
- [4] MAREK O, VACLAV S, JAN S. Designing ontologies using formal concept analysis [C]// Proceedings of the International Conference on Communications in Computing. 2004: 302-308.
- [5] HWANG SUK-HYUNG, KIM HONG-GEE, YANG HAE-SOOL. A FCA-based Ontology construction for the design of class hierarchy [C]// Computational Science and Its Applications, LNCS 3482. Berlin: Springer-Verlag, 2005: 827-835.
- [6] 黄美丽,刘宗田. 基于形式概念分析的领域本体构建方法研究[J]. 计算机科学, 2006, 33(1): 210-212.
- [7] 王伟焱,余英林. 图像的情感语义研究进展[J]. 电路与系统学报, 2003, 8(5): 101-109.