

# Laguerre 滤波器在抗噪语音识别特征提取中的应用

黄丽霞, 张雪英

HUANG Li-xia, ZHANG Xue-ying

太原理工大学 信息工程学院, 太原 030024

College of Information Engineering, Taiyuan University of Technology, Taiyuan 030024, China

E-mail: tyzhangxy@163.com

HUANG Li-xia, ZHANG Xue-ying. Application of feature extraction in robust-noise speech recognition based on Laguerre filter. Computer Engineering and Applications, 2008, 44(18): 21-24.

**Abstract:** The FIR filter exists the defects which pass band or reject band character is bad and filter order is high. The defects make the speech recognition system performance become bad when FIR filter is used in front-end processing. Aiming at the problem, the paper uses Laguerre filter bank to replace FIR filter bank used in front-end processing of ZCPA (Zero-Crossings with Peak Amplitudes) feature extraction. After researching carefully FIR filter parameters determining method, the paper describes Laguerre filter's parameters determining method, and presents the computing results. The speech recognition experimental results of isolated words and nonspecific persons show that the speech recognition system using Laguerre filter has better robustness to noise than using FIR filter, and the order of Laguerre filter is much lower than one of FIR filter.

**Key words:** speech recognition; feature extraction; Laguerre filter

**摘要:** 为克服 FIR 滤波器存在的通阻带特性差、滤波器阶次高等缺点给语音识别系统带来的不利影响, 采用 Laguerre 滤波器组代替过零峰值幅度特征提取中使用的 FIR 滤波器组进行前端处理。在仔细研究 FIR 滤波器参数确定方法的基础上, 叙述了 Laguerre 滤波器原理及参数计算方法, 并给出了计算结果。孤立词、非特定人语音识别实验结果表明, 使用 Laguerre 滤波器不仅使识别系统抗噪性能优于使用 FIR 滤波器, 而且滤波器阶数也大为下降。

**关键词:** 语音识别; 特征提取; Laguerre 滤波器

DOI: 10.3778/j.issn.1002-8331.2008.18.007 文章编号: 1002-8331(2008)18-0021-04 文献标识码: A 中图分类号: TN912.3

## 1 引言

由于对人类听觉机理了解不够, 目前噪音环境下的语音识别系统尚不能达到完全实用化, 因此, 抗噪语音识别的研究是业界的焦点。语音识别系统通常包括预处理、特征提取及训练识别网络三大部分, 其中特征提取是一个重要部分。特征提取的前端信号处理常用的是滤波器组模型和 LPC 模型<sup>[1]</sup>。文献[2]提出一种基于人耳听觉特性的特征参数, 称之为过零峰值幅度 (Zero-Crossings with Peak Amplitudes, ZCPA) 特征, 它的前端信号处理采用的是 FIR 滤波器组模型。但由于 FIR 滤波器固有的缺点, 如通阻带特性差、滤波器阶次高, 会对系统性能产生不利影响。Laguerre 滤波器兼有 FIR 和 IIR 滤波器的优点<sup>[3]</sup>, 在信道均衡<sup>[4]</sup>、扩频通信<sup>[5]</sup>和频谱压缩等研究领域均获得好的应用效果。本文尝试将 Laguerre 滤波器用于语音识别特征提取的前端处理中。在深入研究 ZCPA 特征提取方法的基础上, 用 Laguerre 滤波器代替 FIR 滤波器进行特征提取, 并与后端的 RBF 神经网络结合进行孤立词非特定人识别, 实验证明改进后的特征较原特征具有更好的抗噪性能。

## 2 Laguerre 滤波器及参数计算

### 2.1 Laguerre 滤波器的结构

考虑 Laguerre 序列式

$$l_k(n, b) = \sqrt{1-b^2} \sum_{j=0}^k (-1)^{k+j} \frac{k!}{j! (k-j)!} \frac{(n+k-j)!}{k! (n-j)!} b^{n+k-2j} \quad n, k=0, 1, 2, \dots \quad (1)$$

式(1)是正交完备的, 式中  $|b| < 1$ , 称为 Laguerre 参数, 对式(1)进行  $z$  变换得到式(2)如下:

$$L_k(z, b) = \sqrt{1-b^2} \frac{(z^{-1}-b)^k}{(1-bz^{-1})^{k+1}} \quad |b| < 1 \quad (2)$$

由于 Laguerre 序列的完备性, 任何理想系统的传递函数  $\tilde{H}(z)$  都可以表示成如式(3)所示:

$$\tilde{H}(z) = \sum_{k=0}^{\infty} c_k(b) L_k(z, b) \quad (3)$$

其中

基金项目: 国家自然科学基金(the National Natural Science Foundation of China under Grant No.60472094); 山西省自然科学基金(the Natural Science Foundation of Shanxi Province of China under Grant No.20051039)。

作者简介: 黄丽霞(1981-), 女, 博士研究生, 主要研究方向为语音信号处理; 张雪英(1964-), 女, 博士生导师, 教授, 主要研究方向为语音信号处理。

收稿日期: 2008-01-04 修回日期: 2008-03-26

$$c_k(b) = \langle \tilde{H}(z), L_k(z, b) \rangle = \frac{1}{2\pi j} \oint_c \tilde{H}(z) L_k(z^{-1}, b) z^{-1} dz \quad (4)$$

$c_k(b)$ 称为Laguerre系数。在实际滤波器设计中,利用有限阶数(如取 $M+1$ 阶)Laguerre序列的 $z$ 变换 $H(z)$ 来近似理想系统的传递函数 $\tilde{H}(z)$ ,则有

$$H(z) = \frac{Y(z)}{X(z)} = \sum_{k=0}^M c_k(b) L_k(z, b) = \sum_{k=0}^M c_k(b) \sqrt{1-b^2} \frac{(z^{-1}-b)^k}{(1-bz^{-1})^{k+1}} = \sum_{k=0}^M c_k(b) \frac{\sqrt{1-b^2}}{1-bz^{-1}} \left( \frac{z^{-1}-b}{1-bz^{-1}} \right)^k = \sum_{k=0}^M c_k(b) A_0(z) (A_k(z))^k \quad (5)$$

式中的 $X(z)$ 和 $Y(z)$ 分别为系统的输入和输出,第一级 $A_0(z) = \sqrt{1-b^2} / (1-bz^{-1})$ 是一个简单的低通滤波器( $b>0$ ),其余的 $A_1(z) = (z^{-1}-b) / (1-bz^{-1})$ 都是统一的全通滤波器。与传统FIR滤波器比较是用全通传输函数 $A_1(z)$ 代替了单位延迟单元,该滤波器所有极点都在一点,且都位于单位圆内,具备稳定性。

### 2.2 Laguerre 参数的计算

实际中选取最优 $b$ 值的原则是使式(6)的均方误差 $\varepsilon(b)$ 最小。

$$\varepsilon(b) = \frac{1}{2\pi} \int_{2\pi} |\tilde{H}(z) - H(z)|^2 d\omega \quad (6)$$

令 $z=e^{j\omega}$ ,将式(5)代入等式(6)可得

$$\varepsilon(b) = \frac{1}{2\pi} \int_{2\pi} \left| \tilde{H}(e^{j\omega}) - \sum_{k=0}^M c_k(b) L_k(e^{j\omega}, b) \right|^2 d\omega \quad (7)$$

文献[3]证明了使式(7)取最小值的最优 $b$ 值为方程式 $c_{M+1}(b)=0$ 的根,利用牛顿-拉夫孙法<sup>[6]</sup>,可得到 $c_{M+1}(b)=0$ 的根,从而可求得 $c_k(b)$ 的值。

## 3 ZCPA 特征提取原理及前端滤波器参数的确定

### 3.1 ZCPA 特征介绍

ZCPA特征参数就是采用语音信号的过零率与峰值幅度来表征语音信号特征。图1是ZCPA系统的原理框图<sup>[2]</sup>,该系统由带通滤波器组、过零检测器、峰值检测器、非线性压缩和频率接收器组成。带通滤波器组由16个FIR滤波器组成,用来仿真耳蜗基底膜的作用;过零检测器、峰值检测器、非线性压缩部分则仿真听觉神经纤维的作用。从过零检测器获得频率信息,峰值检测器获得强度信息,经非线性压缩后,用频率接收器合成频率信息和强度信息,最后将16路所获得的信息合成为语音信号特征的完整输出。

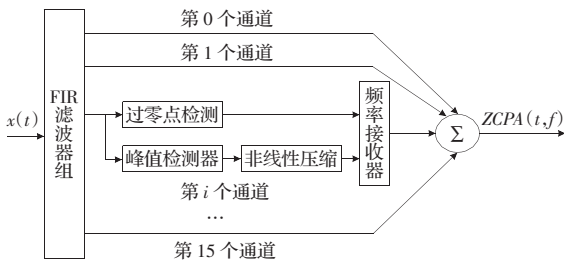


图1 ZCPA系统原理框图

### 3.2 滤波器中心频率、带宽和窗长的确定

文献[2]对FIR滤波器的参数确定进行了简单介绍,并未给出具体过程和计算结果,对它进行了细致分析和研究,计算出了结果,以便为用Laguerre滤波器代替FIR滤波器时进行相关参数值的确定提供参考。

#### 3.2.1 中心频率的确定

耳蜗在听觉系统中起着非常重要的作用,耳蜗中的基底膜对外来的声信号有频率选择和调谐的作用,其峰值出现在基底膜的不同位置。约800 Hz以上,声音频率沿基底膜按对数分布,其位移和频率的关系可用下式表示:

$$f = A(10^{ax} - 1) \quad (8)$$

其中 $f$ 是频率(Hz), $x$ 是基底膜的归一化距离, $x=0$ 表示基底膜的顶点, $0 \leq x < 1$ , $A$ 和 $a$ 分别是耳蜗的特征常量,可选为 $A=165.4, a=2.1$ 。

沿耳蜗基底膜选取16个点(也可选更多的点),对应一个由16个滤波器构成的带通滤波器组,其覆盖频率范围从200 Hz~4 000 Hz,中心频率根据式(8)确定。

#### 3.2.2 带宽的确定

滤波器的频带划分是基于文献[7]提出的等值矩形带宽(Equivalent Rectangular Bandwidth, ERB)来确定的,ERB频带划分对听觉滤波器中心频率的函数关系为:

$$ERB = 6.23f^2 + 93.39f + 28.52 \quad (9)$$

其中 $f$ 是每个通道的中心频率,单位是kHz,ERB是对应带宽。

#### 3.2.3 窗长的确定

用 $x_k(n)$ 表示第 $k$ 个带通滤波器的输出,用 $x_k(n, m)$ 表示 $x_k(n)$ 在 $m$ 时刻的输出帧信号,则

$$x_k(n, m) = x_k(n) w_k(m-n) \quad k=1, \dots, 16 \quad (10)$$

其中, $w_k(n)$ 是有限长度的Hamming窗函数。由于频带是非均匀划分的,所以窗函数长度也应是而非均匀的。如果用 $L_k$ 表示第 $k$ 个滤波器的窗长,则由式(11)确定窗长:

$$L_k \approx \frac{cf_s}{f_k} \quad (\text{样点数}) \quad (11)$$

这里取帧长 $c=10$  ms。 $f_s$ 是第 $k$ 个滤波器的中心频率, $f_s=11.025$  kHz是信号的采样频率。表1给出了依据式(8)~式(11)计算出的16个滤波器的中心频率、带宽以及窗长。

表1 16个滤波器的中心频率 $f_k$ 、带宽 $B$ 和窗长 $L_k$

通道数	$f_k/\text{Hz}$	$B/\text{Hz}$	$L_k/\text{样点数}$	通道数	$f_k/\text{Hz}$	$B/\text{Hz}$	$L_k/\text{样点数}$
0	200	46	549	8	1 172	151	94
1	264	53	417	9	1 408	176	78
2	340	61	324	10	1 685	206	65
3	429	71	256	11	2 011	241	54
4	534	82	206	12	2 395	283	46
5	657	95	167	13	2 845	331	38
6	802	111	137	14	3 376	389	32
7	972	125	109	15	4 000	456	27

## 4 Laguerre 滤波器相关参数值的确定

### 4.1 Laguerre 滤波器截止频率的确定

根据表1中的带宽和中心频率值,可利用公式(12)确定Laguerre滤波器的上、下载止频率 $\omega_2$ 和 $\omega_1$ :

$$\omega_1 = \frac{2\pi(f_k - \frac{B}{2})}{f_s}$$

$$\omega_2 = \frac{2\pi(f_k + \frac{B}{2})}{f_s}$$
(12)

### 4.2 时延以及滤波器阶数的确定

在通常的 FIR 滤波器设计中,为获得线性相位,要求时延  $\tau=(N-1)/2$ ,即  $N=2\tau+1$ , $N$  为窗内样点数。考虑这里的 Laguerre 滤波器组,这里希望在通带内也具有线性相位。由于频带划分的非均匀性,窗长(样点数)也是变化的,可将  $N=2\tau+1$  变为可变窗长  $N=2n\tau+1$  ( $n$  为大于零的整数)。设对应于第  $k$  个滤波器,窗长为  $L_k$ ,时延为  $\tau_k$ ,阶数为  $M_k$ ,则有

$$\tau_k = \frac{L_k - 1}{2n} \quad n=1,2,3,\dots$$
(13)

滤波器的阶数满足  $M_k > \tau_k$ 。本文选取海明窗(Hamming window)进行计算。

### 4.3 Laguerre 系数 $c_k(b)$ 的确定

有了最优  $b$  值后,可根据式(14)计算出相应的 Laguerre 滤波器系数式

$$c_k(b) = \frac{1}{2\pi j} \oint_c \tilde{H}(z) L_k(z^{-1}, b) z^{-1} dz$$

$$= \frac{\sqrt{1-b^2}}{2\pi j} \oint_c \tilde{H}(z) \frac{(z-b)^k}{(1-bz)^{k+1}} z^{-1} dz$$
(14)

对于理想带通滤波器,  $c_k(b)$  可变形为式(15)的形式<sup>[8]</sup>:

$$c_k(b) = \sqrt{1-b^2} \sum_{p=0}^k \sum_{n=p}^{2\tau} s(n) (-1)^{k-p} \frac{k!}{(k-p)! p!} \frac{(k+n-p)!}{(n-p)! k!} b^{k+n-2p} = \sqrt{1-b^2} P_k(b)$$
(15)

式中,  $P_k(b)$  是一个  $k+2\tau$  次多项式,  $s(n)$  (如式(16)):

$$s(n) = \begin{cases} \frac{\sin[\omega_2(n-\tau)] - \sin[\omega_1(n-\tau)]}{\pi(n-\tau)} & n \neq \tau \\ \frac{\omega_2 - \omega_1}{\pi} & n = \tau \end{cases}$$
(16)

至此,将每个滤波器对应的  $\omega_1, \omega_2, \tau$  代入式(15),就可得到 Laguerre 滤波器各级对应的  $c_k(b)$  系数。由于每个滤波器的阶数  $M$  不同,所以  $c_k(b)$  的个数也不尽相同。表 2 给出了 16 个

表 2 16 个 Laguerre 滤波器的上、下截止频率、时延、滤波器阶数以及最优  $b$  值

通道数	$\omega_1$	$\omega_2$	$\tau$	$M$	$b_{opt}$
0	0.100 9	0.127 1	54	58	0.077 4
1	0.135 4	0.165 6	52	56	0.054 2
2	0.176 4	0.211 2	54	58	0.149 2
3	0.224 3	0.264 8	44	48	0.137 9
4	0.281 0	0.327 7	51	54	0.190 0
5	0.347 4	0.401 5	41	44	0.100 1
6	0.425 4	0.488 7	34	38	0.174 1
7	0.539 4	0.612 9	54	58	0.036 6
8	0.624 9	0.711 0	47	50	0.257 8
9	0.752 3	0.852 6	39	43	0.304 0
10	0.901 6	1.019 0	32	36	0.014 5
11	1.077 4	1.214 7	27	30	0.244 2
12	1.284 3	1.445 6	23	26	0.130 4
13	1.527 1	1.715 7	19	22	0.138 5
14	1.813 1	2.034 8	16	18	0.204 0
15	2.149 7	2.409 6	13	16	0.104 6

Laguerre 滤波器的上、下截止频率、时延、滤波器阶数以及最优  $b$  值的计算结果,图 2 给出 Laguerre 滤波器和 FIR 滤波器的通阻带特性。从图中可以看出 Laguerre 滤波器的通带和阻带特性比 FIR 好。图 3 给出了所设计的 16 个 Laguerre 带通滤波器的频率响应。

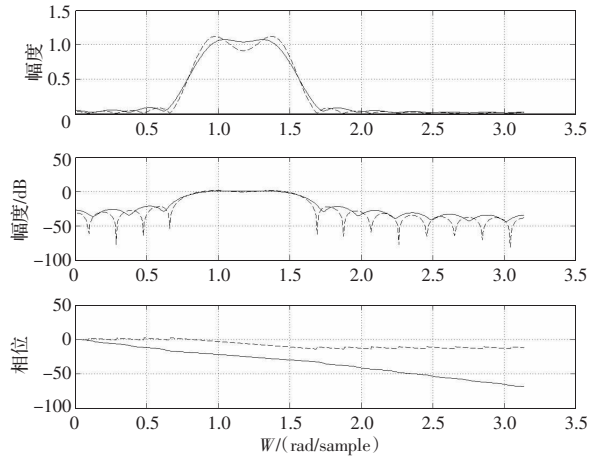


图 2 FIR 和 Laguerre 的带通滤波器幅频响应图 (实线是 Laguerre 滤波器的幅频曲线,虚线是 FIR 滤波器的幅频曲线)

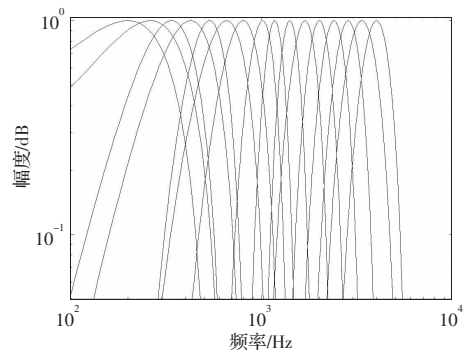


图 3 16 个 Laguerre 滤波器的频率响应

## 5 实验及结果

本文用 C++语言在 Windows 平台上实现了基于原 ZCPA 特征和基于 Laguerre 滤波器加 Hamming 窗的 HLZCPA 特征提取,后端的训练识别网络使用 RBF 神经网络,进行非特定人的孤立词语音识别。语音信号采样率为 11.025 kHz,帧长 10 ms,帧移 5 ms。实验共使用了 10 词、20 词、30 词、40 词、50 词,分别是 9 人在不同 SNR (15 dB、20 dB、25 dB、30 dB、无噪音)下的发音作为训练数据库,每人每个词发音 3 次。用另外 7 人在相应 SNR 下的发音进行识别,得到不同 SNR 下和不同词汇量下的基于 RBF 神经网络的语音识别结果。为处理方便,将提取的特征参数进行时间和幅度归一化处理,得到统一的 1 024 维的语音特征矢量序列,作为 RBF 网络的输入,即输入层节点数为 1 024。RBF 网络隐层节点数根据识别词汇量变化,10 词对应 10 个隐节点,50 词对应 50 个隐节点;输出层节点数与隐层节点数相同。关于 RBF 神经网络权值的训练和选取,详见文献[9]。

RBF 神经网络模型确定后,将测试集的单词输入网络进行识别。表 3 列出了在不同 SNR 和不同词汇量下的 ZCPA,HLZCPA 特征的识别结果。

从表 3 的结果可以看到:(1)基于 Laguerre 滤波器的特征 HLZCPA 的平均识别率高于 ZCPA,说明基于 Laguerre 滤波器

表3 基于 ZCPA、HLZCPA 特征的识别结果比较 %

词汇量	特征	SNR/dB				Clean
		15	20	25	30	
10 词	ZCPA	82.4	84.8	85.2	87.1	87.1
	HLZCPA	93.3	95.2	95.2	95.7	94.8
20 词	ZCPA	82.6	84.8	86.7	87.1	87.6
	HLZCPA	93.3	94.5	94.8	94.3	95.0
30 词	ZCPA	83.1	85.1	88.4	89.2	89.0
	HLZCPA	92.1	94.1	93.8	93.5	94.3
40 词	ZCPA	84.0	85.9	89.3	90.4	91.6
	HLZCPA	93.2	93.7	93.7	93.7	93.7
50 词	ZCPA	85.0	88.1	90.5	91.1	92.3
	HLZCPA	91.8	93.1	92.9	93.0	92.4

的特征总体上优于 ZCPA 特征;(2)当 SNR 在最大值和最小值间变化时,基于 HLZCPA 特征的识别率的变化范围较 ZCPA 识别率的变化范围更小,说明新特征对噪声不是很敏感,具有更好的鲁棒性。

表4是 FIR 滤波器组和 Laguerre 滤波器组中 16 个滤波器的阶数比较。可以看出,所有相对应的 Laguerre 滤波器的阶数都低于 FIR 滤波器的阶数,且随 FIR 滤波器阶数的增高,Laguerre 滤波器阶数的降低程度越高。

表4 FIR 滤波器组和 Laguerre 滤波器组中各级滤波器的阶数比较

滤波器序号	FIR	Laguerre	滤波器序号	FIR	Laguerre
0	100	58	8	64	50
1	100	56	9	54	43
2	100	58	10	46	36
3	100	48	11	40	30
4	100	54	12	34	26
5	100	44	13	28	22
6	86	38	14	24	18
7	74	58	15	20	16

(上接 8 页)

能找到更精确的描述信息,并选定更有代表性的索引词,则可能会对查询结果有进一步的帮助,这也是下一步要做的工作。

## 6 结束语

本文提出了一个基于 Web 信息提取的媒体文件描述符扩展算法,该算法可以有效地扩充文件描述内容和相关性信息,极大地改善了 P2P 系统中描述符不足,统计信息精度低的现象。同时,本文也提出了一个结构化 P2P 系统的描述符更新策略,将该描述符扩展算法跟结构化 P2P 系统进行了有效地结合,便于对系统和信息进行更新。实验表明,该算法可以有效地改善基于内容信息检索的结果,提高系统的查询精度。

## 参考文献:

- [1] Sharma S,Nguyen L T,Jia D.IR-wire:a research tool for information retrieval[R].SIGIR Workshop,2006:61-65.
- [2] Lu Y,Sebastien V.Peer-to-Peer media streaming application survey[C]//Proc UBIComm,2007:139-148.
- [3] Heinink V E P.Video-on-demand over the Internet [R].Delft University,2007.
- [4] Jin M,Cha B,Kang D.A design of a system for supporting keyword search in DHT-based P2P system[C]//International Conference on Internet Computing(IC),2005:247-254.
- [5] Li Y,Jagadish H V,Tan K.SPRITE:a learning-based text retrieval

## 6 结语

通过本文的研究和实验,证明了 Laguerre 滤波器用于 ZCPA 特征提取的前端处理时,不仅抗噪性能优于使用 FIR 滤波器,而且滤波器阶数也大为下降。代价是计算 Laguerre 滤波器系数的复杂度有所提高。

## 参考文献:

- [1] Rabiner L,Juang B H.Fundamentals of speech recognition[M].北京:清华大学出版社,1999:69-79.
- [2] Kim D S,Lee S-Y,Kil R M.Auditory processing of speech signals for robust speech recognition in real-world noisy environments[J].IEEE Transaction on Speech and Speech Audio Processing,1999,7(1):55-69.
- [3] Masnadi-Shirazi M,Aleshams M.Laguerre discrete-time filter design[J].Computers and Electrical Engineering,2003,29:173-192.
- [4] 贺双赤.用 Laguerre 滤波器实现多径衰落信道自适应均衡[J].电讯技术,2004,44(1):82-86.
- [5] 黄高勇,张家树.一种抑制直扩通信窄带干扰的新型非线性自适应预测滤波器[J].电子与信息学报,2007,29(6):1328-1331.
- [6] Silva Toe.On the determination of the optimal pole position of Laguerre filters[J].IEEE Signal Process,1995,4(9):2079-2087.
- [7] Oded Ghizta.Auditory models and human performance in tasks related to speech coding and speech recognition[J].IEEE Transactions on Speech and Audio Processing,1994,2(1):113-131.
- [8] 吉芳芳,张雪英.基于 Laguerre 网络的滤波器设计[J].太原理工大学学报,2007,38(4):283-286.
- [9] 白静,张雪英,侯雪梅.基于 RBF 神经网络的抗噪语音识别[J].计算机工程与应用,2007,43(22):28-30.
- [10] system in DHT networks[C]//ICDE 2007,2007:1106-1115.
- [6] Zhang Z,Zhou S,Qian W,et al.KEYNOTE:keyword search by node selection for text retrieval on DHT-based P2P networks [C]//Proceedings of DASFAA'06.[S.l.]:Springer,2006:797-806.
- [7] 肖建勇,张武生.Clair:一种基于 P2P 的 BitTorrent 关键词检索系统[J].计算机工程与应用,2006,42(18):139-142.
- [8] Garces-Erice L,Felber P A,Biersack E W,et al.Data indexing in Peer-to-Peer DHT networks [C]//24th IEEE International Conference on Distributed Computing Systems(ICDCS'04),2004:200-208.
- [9] Yee W G,Nguyen L T.Novel applications of information retrieval techniques to Peer-to-Peer file-sharing systems[C]//Proceedings of the International Workshop on Information Retrieval in Peer-to-Peer Networks,Arlington, Virginia, USA, 2006.
- [10] Jia D,Yee W G,Frieder O.Automatic tuning of file descriptors in P2P file sharing systems[C]//Proc Workshop on Web and Databases (WebDB),2006:50-55.
- [11] Kilgarriff A,Grefenstette G.Introduction to the special issue on the Web as corpus[J].Computational Linguistics (COLING),2003,29(3):333-348.
- [12] You Jia-li,Chen Yi-ning,Chu Min,et al.Identify language origin of personal names with normalized appearance number of Web pages[C]//INTERSPEECH-2006,2006:1353-1356.
- [13] The TREC-8 question answering track evaluation[C]//The Eighth Text Retrieval Conference(TREC-8).Washington DC:NIST,1999:77-82.