

个参数的分布规律和特点来进行汉语音节的声韵分割。对 115 个汉语孤立字进行了的声韵分割仿真实验, 仿真实验结果显示, 基于 GMP 分解的 Gabor 原子参数的汉语孤立字声韵分割的正确率达 80.87%。

2 匹配追踪算法原理

MP 算法的核心是将观察信号分解为一组最好匹配信号局部结构的时频原子的线性展开。该算法从冗余的函数字典 D 中选择时频原子, 然后将残差信号正交投影到 D 中与之最相似的原子上, 再将这部分投影从残差信号中减去, 如此循环, 直至残差信号的能量小于预先设定的阈值。从而, 信号第次的分解过程即^[6]

$$R^m f = R^{m-1} f - \langle R^m f, g_{\gamma_m} \rangle \quad (1)$$

式中, $R^m f$ 为第 m 次投影后的残差信号, g_{γ_m} 为由参数组 γ 定义的最佳时频原子, $\langle R^m f, g_{\gamma_m} \rangle$ 表示 $R^m f$ 与 g_{γ_m} 的内积。其中 g_{γ_m} 满足

$$\langle R^m f, g_{\gamma_m} \rangle = \sup_{\gamma \in \Gamma} |\langle R^m f, g_{\gamma} \rangle| \quad (2)$$

等式右边表示残差信号与原子库中所有原子内积的上确界。信号经过 m 步的分解式为

$$f = \sum_{n=0}^{m-1} \langle R^n f, g_{\gamma_n} \rangle g_{\gamma_n} + R^m f \quad (3)$$

原信号可近似表示为个原子的线性组合, 相应的 WVD 谱可表示为

$$WVD_f(t, \omega) = \sum_{n=0}^{m-1} WVD_{g_{\gamma_n}}(t, \omega) \quad (4)$$

时频原子是一组具有良好时频特性的信号, 由于它们之间的正交性不能被保证, 不能算作真正意义上的基, 因而改称原子。进行匹配追踪分解时, 为了用尽量少的原子表示待分解信号, 要求所选用的原子能够尽量适应待分解信号的局部结构。常用的原子主要有正弦原子、Gabor 原子和 Chirp 原子, 这里选用 Gabor 原子对语音信号进行分解。Gabor 原子由具有单位能量的高斯函数 $g(t) = 2^{1/4} e^{-\pi^2 t^2}$ 经过伸缩、平移、频率调制得到:

$$g_{\gamma} = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) e^{i\xi t} \quad (5)$$

其中, s 、 u 、 ξ 分别为尺度、平移和频率调制参数。上式所对应的实 Gabor 原子为:

$$g_{\gamma} = \frac{K}{\sqrt{s}} g\left(\frac{t-u}{s}\right) \cos(\xi t) \quad (6)$$

K 为归一化系数, 以便 $\|g_{\gamma}\| = 1$ 。上述 Gabor 原子在时域上以 u 为中心, 能量集中在 u 附近, 能量大小与尺度 s 成比例; 频域上以 $\omega = \xi$ 为对称中心, 能量集中在 ξ 附近, 大小与 $1/s$ 成比例。

MP 算法的计算复杂度虽然比现有的其它稀疏分解算法要低, 但是仍然庞大, 计算任务十分繁重, 因而研究的焦点集中在如何实现该算法的快速计算, 降低算法的复杂度, 以及选择何种类型的基构造合适的原子库两方面^[7]。

鉴于匹配追踪计算量过大的缺点, 本文采用遗传匹配追踪算法进行声韵分割。遗传算法是一种仿效大自然生物进化中物竞天择、适者生存自然选择过程的寻优算法。在一个有一定规模的种群里, 比较适应环境(适应度大)的个体, 可以获得更多的繁殖机会, 不适应环境的个体获得较少的繁殖机会。父代的繁殖通过交叉和变异等遗传操作传给下一代。通过引入遗传算法, 以模仿自然进化过程来寻找最佳匹配原子替代遍历式的耗尽搜索, 使得每个寻找最佳原子的过程由搜索整个原子库变为仅搜索原子库中一个随机产生的较小的子集, 从而大幅提高匹配追踪的运算速度。

3 算法实现

本文进行声韵分割的具体步骤如下:

1) 读入汉语音节的发音信号, 以帧长 256 点、帧移 128 点进行分帧, 将信号切分为具有一定重叠的连续片段。

2) 选取一帧语音信号, 对 Gabor 原子参数集 $\gamma = (s, u, \xi)$ 按照如下方式离散化: $\gamma = (a^j, pa^j \Delta u, ka^{-j} \Delta \xi)$, 其中, $a = 2$, $\Delta u = 1/2$, $\Delta \xi = \pi$, $0 < j \leq \log_2 N$, $0 \leq p \leq N2^{-j+1}$, $0 \leq k < 2^{j+1}$ 。

3) 利用遗传算法选择最佳原子。

① 编码并形成初始种群。对参数集进行编码, 并随机生成一个具有一定规模的初始种群(即参数集合)。这里设定种群规模 pop_num=80, 编码采用浮点数编码方式。

② 计算适应度。将初始种群解码并评价每个个体

(即每组参数)的适应度。这里的适应度就是由每组原子参数确定的时频原子分别与残差信号的内积。

③ 选择。依据每个个体的适应度,进行比例选择。

④ 交叉和变异。将选择后的个体分别按照交叉概率和变异概率进行交叉、变异。本文采用均匀算术交叉算子以及非均匀变异算子,取交叉概率 $pc=0.8$, 变异概率 $pm=0.01$ 。

⑤ 重复①~④到指定进化代数,选出与残差信号最相似(即与残差信号内积最大)的原子。本文设定进化代数 $T=10$ 。

4) 判断迭代终止条件。依公式(1)计算本次 MP 迭代后的残差信号,并根据事先设定的能量阈值判断是否终止迭代。若终止,则转入下一步;否则,重复 3)。

5) 提取原子特征。由于时域上,声母能量小于韵母致使其大尺度原子少于韵母;频域上,声母的高频成分多且分布分散致使其高频率原子多余韵母且原子的频率分布相对分散,分别将最佳原子参数中大尺度原子占原子总数的比例、高频率原子占原子总数的比例以及频移参数的标准差作为相应帧信号的特征输出。

6) 重复 2)~5), 取出各帧信号的原子特征。

7) 计算预分割帧位置。通过分别对原信号的各种时频原子特征参数分别求一阶导数,找出它们各自的突变帧位置 b_1 、 b_2 、 b_3 。求出 b_1 、 b_2 、 b_3 的均值,将其中偏离该均值最大的一个值略去,以另外两个的均值 b 作为预分割帧。

8) 确定分割帧。若信号的特征参数在预分割帧附近的一段时间内无明显变化,则视为零声母音节,重置 $b=1$; 否则,直接以第 7)步的预分割帧作为声韵母的过渡帧,完成分割。

4 仿真实验及结果分析

本文仿真实验所用的语音信号由 Cool Edit 软件在实验室环境下录制而成,录音水平为 16k/16bits,单声道。仿真实验过程中,静音部分被剪切掉。

表 1 中给出了汉语发音“雪”在其声韵母分割界限附近六帧的原子参数。由于数据量很大,这里取的六帧每相邻两帧之间间隔两帧,只列出每帧信号的前 10 个原子的尺度参数 s 和频移参数 ω 。

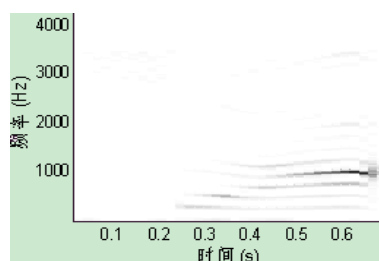
表 1 汉语发音“雪”的 Gabor 原子参数

帧数	24		26		28	
参数	s	w	s	w	s	w

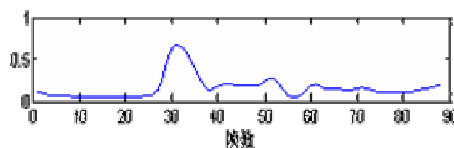
1	112.8	0.98	126.6	1.79	25.57	2.02
2	2.2	3.83	4.47	5.04	60.33	5.4
3	2.01	7.2	3.11	4.08	3.84	5.05
4	4.35	4.91	3.13	3.87	64.63	4.86
5	2.26	3.2	5.02	4.46	2.58	3.25
6	3.69	5.26	2.92	4.12	4.76	7.23
7	2.5	3.8	3.74	4.44	16.57	18.48
8	4.77	10.3	2.22	3.02	45.25	4.78
9	2.99	5.63	10.06	12.97	4.23	5.18
10	2.58	2.97	1.97	1.66	3.57	4.26
帧数	30		32		34	
参数	s	w	s	w	s	w
1	61.29	5.14	46.51	4.75	34.48	4.62
2	94.46	3.01	52.58	3.98	38.84	4.45
3	52.2	5.95	62.28	2.27	63.59	4.74
4	53.94	5.65	43.29	5.39	57.69	5.79
5	19.63	2.68	63.14	7.9	41.6	3.54
6	72.29	1.48	27.74	5.28	17.99	8.29
7	15.8	3.63	43.63	3.88	84.02	2.96
8	26.36	4.31	20.04	6.46	21.77	5.65
9	24.79	4.32	73.01	4.68	17.61	6.72
10	38.41	6.67	5.61	2.24	19.86	5.74

从上表中可以看出,对应于声母部分的 Gabor 原子,尺度参数普遍较小,频移参数分布零散;对应于韵母部分的 Gabor 原子,尺度参数相对较大,频移参数则相对集中。

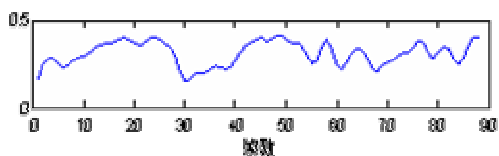
图 1 为汉语发音“雪”的声韵分割效果图。



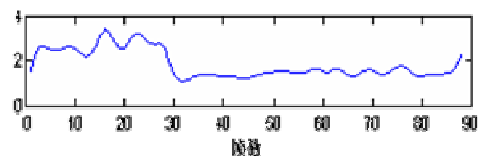
(a) 谱图



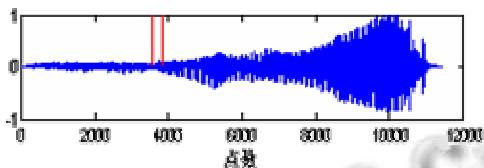
(b) 尺度参数大的样子比例



(c) 频移参数大的样子比例



(d) 频移参数的标准差



(e) 分割结果

图 1 孤立字“雪”的声韵分割效果图

从汉语发音“雪”的时频图上也可以看出，声母部分的频谱颜色很淡而韵母的明显明亮，与其在时域上声母的尺度明显小于韵母、频域上声母的各频率分量分布不集中相一致；韵母部分的条纹存在明显的横杠，且这些横杠主要集中于 200~2000Hz 之间，这与韵母含有明显的谐波成分且这些分量在频域上相对集中相一致。

从上面的分析可知，本文所提方法对具有明显声母部分的音节相当有效；但当发音较快或者声母本身时长过短时，分割效率则不甚理想，这主要是因为此时声韵母之间的过渡序列很短，而仿真实验中所设定窗长和窗移相对较长，不能准确地体现出原子参数的变化；对浊声母音节则不能正确分割，原因是浊声母的发音机制类似于韵母部分，其时频原子参数与韵母部分无明显差异，仿真实验中，容易误将第一帧作为分割帧。

通过统计，得 115 个汉语孤立字的仿真结果如表 2。

表 2 基于 MP 原子参数的孤立字声韵分割仿真结果

孤立字	字数	正确分割数	比例(%)
白	12	7	58.33
歌	17	14	82.35
画	11	7	63.64
全	10	9	90
山	15	13	86.67

天	15	13	86.67
雪	13	11	84.62
雨	9	7	77.78
最	13	12	92.31
共计	115	93	80.87

另外，从表 3 中容易看出，遗传算法的引入使运算速度大大提高。

表 3 遗传算法对运算速度的改进

孤立字	时长 (s)	传统 MP (s)	GMP (s)
白	0.346	588.9	64.98
歌	0.396	648.1	71.49
画	0.277	529.8	57.33
全	0.338	541.7	66.92
山	0.364	672.3	75.66
天	0.341	597.3	63.74
雪	0.465	961.7	101.9
雨	0.471	557.6	59.78
最	0.304	634.5	70.82

5 结论

汉语实际发音中，声母能量小，其频率成分多且分布分散；韵母能量相对较大，其频率成分较少且集中于中低频。本文通过运用遗传匹配算法将汉语语音分解至时频域，选用 Gabor 时频原子参数来刻画汉语孤立字语音的声韵母在时频域上的这种差异，依据分别对应于声母部分和韵母部分的 MP 原子参数的不同变化规律，实现了一种新的时频声韵分割方法。仿真实验结果表明，该方法是可行并且有效的，这也为下一步进行连续语流的声韵切分提供了基础；但本方法也有局限性，主要表现在对浊声母音节不能做出正确分割。

参考文献

- 1 Zhao XF, O'Shaughnessy D. A new hybrid approach for automatic speech signal segmentation using silence signal detection, energy convex hull, and spectral variation. Canadian Conference on Electrical and Computer Engineering. Niagara Falls, 2008.145-148.
- 2 吕军,马晓娜.汉语孤立词声韵分割算法的研究.安徽师范大学学报(自然科学版),2008,31(3):214-218.
- 3 Jurado RS, Gomez-Gil P, Garcia C. Speech text-independent

(下转第 203 页)

```

R_best(NC,:)=Tabu(pos(1,:));
L_ave(NC)=mean(L);
NC=NC+1
4) 更新蚂蚁信息及禁忌表清零
Delta_Tau=zeros(n,n);
for i=1:m
for j=1:(n-1)
Delta_Tau(Tabu(i,j),Tabu(i,j+1))=Delta_Tau(Tabu(i,
j),Tabu(i,j+1))+Q/L(i);
end
Delta_Tau(Tabu(i,n),Tabu(i,1))=Delta_Tau(Tabu(i,n)
,Tabu(i,1))+Q/L(i);
end
Tau=(1-Rho).*Tau+Delta_Tau;
Tabu=zeros(m,n); //禁忌表清零
End

```

5 结论

在基于上面基于蚁群算法实现的 VRP 路径问题解决方案中,我们采用 MATLAB 实现了该算法的性能测试。通过设定对应的配送成本和相关的服务水平值来测试蚁群算法解决 VRP 问题的收敛性,解决时间复

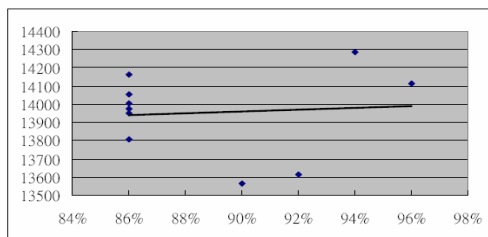


图 1 配送成本与服务水平关系图

杂度和空间复杂度问题,计算的配送成本与服务水平相关系数为 0.08527,显示出其配送总成本与服务水平,并无明显的相关性,如图 1 所示。

以我们执行 10000 次 VRP 配送问题进行测试,采用设定的物流配送服务水平为 90%以上,获得的 9797 次测试结果符合较低的配送成本,可以看到采用该算法实现的整体优化度为 98%。由于受到实际运算中路径规划其他相关因素的影响,可能部分路径的规划部能实现最优,整体上来说,基于该算法解决 VRP 问题是一种可行的方案。

参考文献

- 1 张云洲.物流企业信息化管理模式探析.中国物流与采购,2008.
- 2 杨丰,周广田.现代物流配送中心信息系统设计分析.信息技术,2008.
- 3 黄若男,向兆礼.浅论现代物流配送成本管理.邢台学院学报,2008.
- 4 沈金星,郑长江,徐鹏.结合交通特性的 VRP 新模型.交通与计算机,2008.
- 5 辛华.基于遗传算法的模糊信息 VRP 问题研究.中国水运,2007.
- 6 李怡,军涛.GIS 技术与配送车辆路线规划问题的一种结合方法.科技与管理,2006.
- 7 大鹏,徐瑞华.物流配送线路优化的改进遗传算法研究.交通运输系统工程与信息,2006.
- 8 胡小兵,黄席樾.蚁群算法在迷宫最优路径问题中的应用.计算机仿真,2005.
- 9 李祚泳,钟俊,彭荔红.基于蚁群算法的两地之间的最佳路径选择.系统工程,2004.
- 10 赵建有,闫旺,胡大伟.配送网络规划蚁群算法.交通运输工程学报,2004.

(上接第 236 页)

segmentation using an improvement method for identification of phoneme boundaries. International Conference on Electrical, Communications, and Computers. Cholula, Puebla: 2009. 20-24.

- 4 Wei T, Yan RQ, Gao RX. A hybrid diagnostic method based on wigner-ville distribution and wavelet packet transform. Wavelet Analysis and Pattern Recognition, International Conference, Qingdao, 2010. 386-391.

- 5 Stefanoiu D, Ionescu F. A genetic matching pursuit algorithm. Proc. of the Seventh International Symposium on Signal Processing and Its Applications. 2003: 577-580.
- 6 Mallat S, Zhang Z. Matching pursuit time-frequency distributions. IEEE Trans. on Signal Processing, 1993,41(12): 3397-3415.
- 7 邵君.基于 MP 的信号稀疏分解算法研究.成都:西南交通大学,2006.