

基于 EMD 的语音特征信息提取

王 辉 李生华

(贺州学院物理与电子信息工程系 贺州 542800)

摘 要 有效提取语音信号的特征信息是语音识别的关键。对语音信号采用经验模态分解法可得到语音的一系列本征模函数,提取本征模函数的过程是降低语音信号冗余度的过程。在语音识别的试验中以本征模函数为训练模型较传统的识别方法识别率更高。仿真结果表明:方法是有效的,用于提取语音的特征信息是可行的。

关键词 语音信号,经验模态分解,特征信息,语音识别,本征模函数

中图法分类号 TN912 **文献标识码** B

Feature Extraction of Speech Signal Based on Empirical Mode Decomposition

WANG Hui LI Sheng-hua

(School of Physical and Electronic Information Engineering, Hezhou University, Hezhou 542800, China)

Abstract Effectively extracting feature information of speech signal is the key of speech recognition. By empirical mode decomposition method, a series of intrinsic mode function can be got. Extraction process of intrinsic mode function is to reduce its redundancy. Experimental results show that the empirical mode decomposition method can decrease the recognition error rate. EMD is an effective method. It is feasible to extract the feature from speech signals with EMD.

Keywords Speech signal, EMD, Feature information, Speech recognition, Intrinsic mode function

1 引言

语音信号是一种非平稳的随机信号,其统计特征是语音信号随时间变化而变化,同时又是渐变的。基于语音信号在一小段时间范围里是平稳的,因此对语音信号的处理方法一般采用短时傅立叶变换、小波变换等方法^[1]。但这种短时处理方法在提升语音识别性能方面存在着很多的不足。

近年来 HHT(Hilbert-Huang Transform)法成为非线性信号处理的一种新的有效方法^[2,3]。经验模态分解法是一种能很好地处理非线性、非平稳信号的完全局部自适应时频分析方法,这种方法主要是利用信号的局部特征时间尺度,对信号进行逐级分解,这样得到的信号是具有不同特征时间尺度的数据序列,我们把通过逐级分解而得到的每个数据序列称为一个本征模函数(IMF)^[4]。得到的这一系列 IMF 分量能突出地体现数据的局部特征。

2 经验模态分解法(EMD)

通常情况下,一个信号都可以看成是由多个本征模函数(IMF)组成的。经验模态分解法就是从信号中提取出各个 IMF 分量。一个本征模函数必须满足下面 2 个条件^[5]:

(1)在整个函数中极值点和过零点的数目相差不能大于 1 个;

(2)在任何时间点上,局部极大值的包络和局部极小值的包络的均值必须为零。

同时经验模态分解是建立在以下 3 个假设下来进行的:

(1)信号至少有 2 个极值点,一个最小值点和一个最大值

点;

(2)特征时间尺度是通过 2 个相邻的极值点之间的时间间隔定义的;

(3)如果数据缺少极值点但有变形点,则可通过数据微分一次或几次获得极值点,然后再通过积分来获得分解结果。

经验模态分解方法在分解的过程中其基函数是基于数据本身的,分解的本质是由数据的特征时间尺度来获得信号本身征震荡的模式,并以此来分解数据,这一过程可以形象地称之为“筛”的过程。具体过程如下:

(1)设原始数据序列用 $S(t)$ 表示,找出 $S(t)$ 的极大值点和极小值点,通过 3 次抽样函数拟合得到 $S(t)$ 的上包络线和下包络线。

(2)求上包络线和下包络线的均值并用 m_1 来表示。用 h_1 表示 $S(t)$ 与 m_1 的差值即:

$$h_1 = S(t) - m_1$$

理想情况下 h_1 可以看成是一个 IMF,但实际情况中由于有欠调和过调现象的发生总会存在图形不对称的情况。因此,通常会把 h_1 看作是原始数据并多次重复以上操作直到最后得到的结果满足上文中提到的符合本征模函数的 2 个条件,并用 c_1 来表示 $c_1 = h_{1n}$ 。在实现的过程中依据以下 2 点来判断:

(I)过零点数和局部极值总的个数(包括局部极大值和局部极小值)相等;

(II)标准偏差 SD 的取值范围在 0.2 到 0.3 之间。

$$SD = \sum_{t=0}^T \left[\frac{(h_{1(n-1)}(t) - h_{1n}(t))^2}{h_{1(n-1)}^2(t)} \right]$$

(3)从原始数据中分离出分量 r_1 并记 $r_1 = S(t) - c_1$,再

对信号 r_1 重复(1)、(2)的过程得到 $c_2, c_3, \dots, c_n$, 这样就可以得到若干阶 IMF 分量和一个变化缓慢的趋势项 r_n . 进行以上分解后, 原始信号 $S(t)$ 可以表示为:

$$S(t) = \sum_{k=1}^n c_k(t) + r_n(t)$$

3 语音信号特征信息的提取

通过提取语音信号中的特征信息可以去除信息中的一些冗余量, 这样在数据处理过程中可以减少一定的计算量. 本文对语音信号通过经验模态分解的方法来获得语音的特征信息, 采用经验模态分解法提取语音信号特征信息的示意图如图 1 所示.

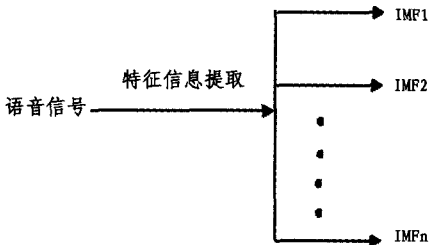


图 1 特征信息提取示意图

图 2 示出语音信号数字 3(在普通工作环境下采用单声道, 8kHz, 8Bit 量化得到的)的时域波形.

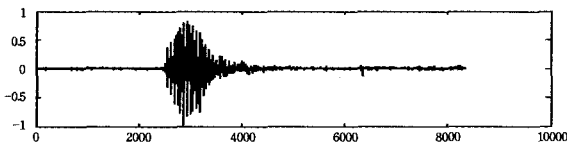


图 2 语音信号的时域波形图

对图 2 所示语音信号进行 EMD 分解后得到的一组本征模函数(IMF1—IMF6), 如图 3 所示, 对应的频谱图如图 4 所示.

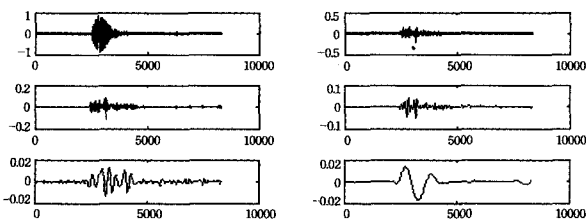


图 3 语音信号采用 EMD 分解得到的本征模函数波形图(IMF1—IMF6)

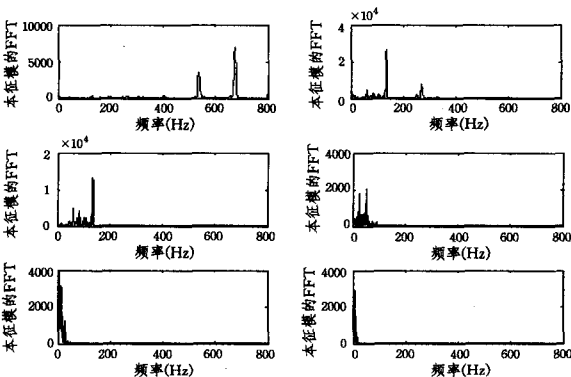


图 4 语音信号本征模函数(IMF1—IMF6)的频谱图

对图 2 所示输入的语音信号采用 EMD 分解后将得到一系列的本征模函数, 我们取前 6 个分量, 如图 3 所示. 将得到的 IMF1—IMF6 这些函数恢复成语音信息后发现, 第一个基信号(IMF1)所得到的声音比较高, 第二个基信号(IMF2)得到的声音比较低沉, 第三个比第二个更加低沉, 以后的基信号还原成语音信号后将听不到清楚的发音. 通过图 4 中所示频谱图也能清楚地看到随着分解的进行得到的基信号的频率会越来越低并逐渐趋向于零. 根据这些性质, 可以推测前面的几个基信号中包含了比较多的原始语音信息, 后面几个基信号可能是发音者的特性.

4 实验过程与分析

下面把采用经验模态分解法得到的语音本征模函数用到语音识别中去, 与传统的语音识别进行比较得到的实验结果如表 1 所列.

表 1 语音识别的结果

不同的训练模型	传统方法	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6
识别率	82.0%	96.0%	84.0%	82.0%	68.0%	54.0%	32.0%

上述语音识别过程所采用的方法都是基于隐马尔可夫(HMM)模型的, 语音录制是在普通工作环境下采用单声道, 8kHz, 8Bit 量化得到的. 语料部分是由 5 个人每人 10 个数字(1—10)总共 50 个词构成的. 在识别的过程中采用的方法是将语料通过经验模态分解后得到一系列的本征模函数 $IMF_1, IMF_2, \dots, IMF_n$, 再将各本征模函数训练成模型分别记为 $M_1, M_2, \dots, M_n$, 识别时将待识别的语料也经过经验模态分解得到对应的本征模函数来进行识别. 分析表 1 中的数据结果可以看出前 3 个本征模函数中含有比较多的语音特征信息, 识别率也比较高. 后面的基函数中包含的则可能是说话人的特性.

另外对于发音比较相似的语音信号(比如数字 4、10), 在识别的实验中发现以 IMF2 作为训练模型其识别率是最高的. 为了很好地区分相似音, 我们构建以下的模型来进行训练:

$$0.2 * IMF_1 + 0.8 * IMF_2$$

识别结果如表 2 所列.

表 2 相似音识别的结果

不同的训练模型	传统方法	IMF1	IMF2	$0.2 * IMF_1 + 0.8 * IMF_2$
识别率	80.0%	86.0%	94.0%	92.0%

结束语 对语音信号采用基于 EMD 的方法提取出来的本征模函数中含有原始语音的特征. 提取本征模函数的过程是一个降低语音信号冗余度的过程, 同时, 在处理发音相似的语音信号时也得到了比较好的识别结果. 可见, 基于 EMD 的语音本征模信息提取是一种有效、可行的方法.

参考文献

[1] 韩继庆. 语音信号处理[M]. 北京: 清华大学出版社, 2004
[2] Huang N E, Shen Z, Long S R. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proc R Soc Lond A, 1998, 454: 903-995
[3] Hawley S D, Atlas L E, Chizeck H J. Some properties of an empirical mode type signal decomposition algorithm[J]. Proceed-

dings of IEEE International Conference on Acoustics, Speech and Signal Processing, 2008, 8(9):3625-3628

- [4] 张德祥, 吴小培. 基于经验模态分解和 Teager 峭度的语音端点检测[J]. 仪器仪表学报, 2010, 3(005):101-104
- [5] 李晋, 王景芳, 高金定. 基于经验模态分解和递归图的语音端点检测算法[J]. 计算机工程与应用, 2010, 46(34):132-135

- [6] 宋倩倩, 于凤芹. 基于 EMD 和改进双门限法的语音端点检测[J]. 电声技术, 2009, 8(33):60-63
- [7] 谈雪丹, 顾济华, 赵鹤鸣. 基于 HHT 瞬时能频值的耳语音端点检测[J]. 计算机工程与软件, 2010, 46(29):147-150
- [8] 李凌, 曾以成, 雷雄国. EMD 在说话人辨识中的应用[J]. 湘潭大学自然科学学报, 2006, 3(28):108-111

(上接第 422 页)

```
uchar i, *p;
EA=1;
ES=1;
REN=1;
/*****设置短信模式*****/
while(1)
{
    clearSBUF(tempSBUF,100);
    countSBUF=0;
    sendAT(AT_CMGF1);//将模式改为中文模式
    for(i=0;i<20;i++)
        delay(60000);//时延等待
    p=strstr(tempSBUF,AT_OK);//判断是否收到 OK
    if(p!=NULL)
        break;
    else
    {
        for(i=0;i<20;i++)
            delay(60000);//延时等待
    }
}
temp3[3]=*num;
temp3[7]=*(num+1);
temp3[15]=*(num+3);
temp3[19]=*(num+4);
number_change(phone_NUM);
clearSBUF(tempSBUF,100);
countSBUF=0;
sendAT(AT_CMGS);
sendAT("52");
send byte('\r');
for(i=0;i<20;i++)
    delay(60000);
sendAT(PDU0);
sendAT(pdu_number);
sendAT(temp1);
sendAT(temp3);
sendAT(temp2);
send_bye(0x1a);
for(i=0;i<100;i++)
    delay(60000);
/*****设置短信模式*****/
while(1)
```

```
{
    clearSBUF(tempSBUF,100);
    countSBUF=0;
    sendAT(AT_CMGF);
    for(i=0;i<20;i++)
        delay(60000);
    p=strstr(tempSBUF,AT_OK);
    if(p!=NULL)
        break;
    else
    {
        for(i=0;i<20;i++)
            delay(60000);//延时等待
    }
}
```

结束语 GSM 网络覆盖地域广,目前移动通信网在全国的覆盖率在 95%以上,非边远地区的覆盖率几乎 100%;通信距离远,网络可靠性高稳定性好。短信远程监控系统利用成熟的 GSM 网络,实现对家居设备电源通断、远程监视与控制。利用工作在 ISM(2.4GHz)的无线模块既能简单、快捷地实现系统与家居设备的链接,还能构成更大的无线家居设备控制网络。最后,短信费用低廉和手机的灵活方便使得利用手机短信远程控制家居设备具有广阔的市场前景与极高的应用价值。

参考文献

- [1] 王景景, 宫崎. 基于 SMS 的无线家电控制系统[J]. 微计算机信息, 2006(10-2):252-253
- [2] 王丽, 黎英, 廖育梅. 利用手机短信控制多种家电及接收报警[J]. 昆明理工大学学报, 2005, 30:135-137
- [3] 池涛, 范靖, 梁福元. GSM/SMS 短信息技术在石油生产办公中应用[J]. 信息技术与信息化, 2007, 5:127-128
- [4] 王晓娟, 黄忠全, 张根保. 短消息系统设计与实现[J]. 重庆大学学报, 2004, 27(5):96-98
- [5] Siang B K, Bin Ramli A R, Prakash V, et al. SMS gateway interface remote monitoring and controlling via GSM SMS[C]// Telecommunication Technology, 2003. NCTT 2003 Proceedings. 4th National Conference on. 2003:84-87
- [6] ETSI TS 100 585 v7.0 [S]. European Telecommunications Standard Institute, 1999
- [7] SIEMENS. AT Command Set TC35i Siemens Cellular Engine. Version:00.01 DocID:TC35i_ATC_V00.01