空 军 工 程 大 学 学 报(自然科学版)
第1卷第2期 JOURNAL OF AIR FORCE ENGINEERING UNIVERSITY Vol.1 No.2
2000年6月 (NATURAL SCIENCE EDITION) Jun. 2000

语音时变频谱分析的FFT技术

段艳丽， 郑 荣
(空军工程大学 电讯工程学院，陕西 西安 710077)

摘 要：针对语音信号频谱分析实际上是时变频谱分析的特性，详细地讨论了用FFT技术对语音进行频谱分析过程中的方法问题。

关键词：语音信号；时变频谱；FFT
中图分类号：TN912.3 **文献标识码**：A **文章编号**：1009-3516(2000)02-0026-03

虽然使用FFT技术对瞬态过程和平稳过程的频谱分析已有很长的历史，但是，语音过程与一个稳定的元音或擦音不同。当激励和声道特性改变时，所得到的语音信号特性随时间发生变化。因此，适用于平稳随机信号的标准傅里叶变换不能直接用于语音信号。对语音频谱分析应能得到时变频谱参数。但在相对短的时隙内，语音信号可看作准周期序列或随机噪声/激励一个线性时不变系统产生。将短时分析思想应用于语音频谱分析，可得到语音时变频谱。本文针对语音时变频谱特性，对各种分析频率点数的频谱计算进行了详细讨论，给出了计算流程框图。

1 语音时变频谱的定义及抽样率

语音时变傅立叶变换定义为：

$$X_n(e^{j\omega}) = \sum_{m=-\infty}^{\infty} w(n-m)x(m)e^{-j\omega m} \tag{1}$$

这里，$x(n)$为语音取样序列，取样频率为F_S。$w(n)$为一窗序列，窗长为N。

可见，时变傅立叶变换$X_n(e^{j\omega})$是n和ω的二维函数。当n固定不变时，$X_n(e^{j\omega})$是$w(n-m)x(m)(-\infty<m<\infty)$序列的标准傅立叶变换，这时$X_n(e^{j\omega})$是$\omega$的以$2\pi$为周期的连续函数。当$\omega$为固定时，$X_n(e^{j\omega})$是$n$的离散序列，这时，$X_n(e^{j\omega})$可看作$w(n)$与$x(n)e^{j\omega n}$的卷积。

要得到非混迭的$X_n(e^{j\omega})$表示，以便使$x(n)$能由之准确恢复，就应对$X_n(e^{j\omega})$在时域和频域正确抽样。由时域抽样定理，$X_n(e^{j\omega})$时域抽样速率$F'_S \geqslant 2B$，B为窗的单边带宽。对汉明窗$B=2F_S/N$(Hz)，所以，

$$F'_s \geqslant 4F_s/N \tag{2}$$

$X_n(e^{j\omega_k})$为$X_n(e^{j\omega})$在$[0,2\pi]$间的L点等间隔采样，

$$\omega_k = \frac{2\pi}{L}k \qquad (0 \leqslant k \leqslant L-1) \tag{3}$$

由频域抽样定理，

$$L \geqslant N \tag{4}$$

因此，对语音信号频谱分析，就是使窗$w(n-m)$沿着$x(m)$序列滑动，窗的滑动速率为F'_s，选择加窗段$w(n-m)x(m)$，对其进行L点离散傅氏变换(DFT)。分析表明[1]，当$L<N$(即频域欠速率采样)时，若窗w

收稿日期：1999-12-17
作者简介：段艳丽(1963-)，女，讲师，硕士.

(n)的特性满足

$$w(n)=1/L, \qquad n=r_0L$$
$$=0, \qquad n=rL(r\neq r_0, r=0, \pm 1, \pm 2, \cdots) \tag{5}$$

仍能从 $X_n(e^{j\omega_k})$中恢复出原始语音 $x(n)$。

2 直接计算 $X_n(e^{j\omega_k})$的运算量

由(1)式：

$$X_n(e^{j\omega_k})=X_n(e^{j\omega})\big|_{\omega=\frac{2\pi}{L}k} \qquad k=0,1,\cdots,L-1$$
$$=\sum_{m=-\infty}^{\infty}x(m)w(n-m)e^{-j\frac{2\pi}{L}km} \tag{6}$$

n 固定时，计算某一频率点 ω_k 的 $X_n(e^{j\omega_k})$，需 $4N$ 次实数乘法，$2N$ 次实数加法，那么计算 M 个频率点的 $X_n(e^{j\omega_k})$，需 $4MN$ 次实数乘法，$2MN$ 次实数加法。当 $M=N=256$ 时，需 262,144 次实数乘法，131,816 次实数加法，可见运算量是相当大的。当 $F_s=10\text{kHz}$ 时，每秒钟实数乘法次数为：$(4F_s/L)\times 262144=4,096,000$ 次。

3 用 FFT 技术进行短时谱分析

(a)当 $L=N$ 时，由(1)式得：

$$X_n(k)=X_n(e^{j\omega_k})=e^{-j\frac{2\pi}{L}kn}\sum_{m=-\infty}^{\infty}x(n+m)w(-m)e^{-j\frac{2\pi}{L}km}$$
$$=e^{-j\frac{2\pi}{L}kn}\sum_{m=-\infty}^{\infty}x_n(m)e^{-j\frac{2\pi}{L}km}$$

这里，$x_n(m)=x(n+m)w(-m)$，是将序列 $x(n+m)w(-m)$的原点定义为 n 抽样而得到的有限长序列，若 $w(m)$为非因果窗

$$\text{上式}=e^{-j\frac{2\pi}{L}kn}\sum_{m=0}^{L-1}x_n(m)e^{-j\frac{2\pi}{L}km}$$
$$=e^{-j\frac{2\pi}{L}kn}DFT[x_n(m)] \tag{7}$$

所以，
$$X_n(k)=DFT[x_n((m-n))_L] \tag{8}$$

(b)当 $L>N$ 时，

$$X_n(k)=e^{-j\frac{2\pi}{L}kn}\sum_{m=0}^{N-1}x_n(m)e^{-j\frac{2\pi}{L}km}$$

给 $x_n(m)$补 $L-N$ 个零，使其长为 L，则

$$X_n(k)=e^{-j\frac{2\pi}{L}kn}\sum_{m=0}^{L-1}x_n(m)e^{-j\frac{2\pi}{L}km} \tag{9}$$

由此可见，$X_n(k)$是对补零后 $x_n(m)$循环移位 n 个抽样的离散傅立叶变换。

(c)当 $L<N$ 时，令 $m=Lr+q$，$(-\infty<r<\infty, 0\leqslant q\leqslant L-1)$，则(9)式变为：

$$X_n(k)=e^{-j\frac{2\pi}{L}kn}\sum_{r=-\infty}^{\infty}\Big[\sum_{q=0}^{L-1}x_n(Lr+q)e^{-j\frac{2\pi}{L}k(Lr+q)}\Big]$$
$$=e^{-j\frac{2\pi}{L}kn}\sum_{q=0}^{L-1}\Big[\sum_{r=-\infty}^{\infty}x_n(Lr+q)\Big]e^{-j\frac{2\pi}{L}kq}$$

定义：$u_n(q)=\sum_{r=-\infty}^{\infty}x_n(Lr+q)$，则

$$X_n(k)=e^{-j\frac{2\pi}{L}kn}\sum_{q=0}^{L-1}u_n(q)e^{-j\frac{2\pi}{L}kq}$$
$$=DFT[u_n((q-n))_L] \tag{10}$$

短时频谱分析的流程如图 1：

说明：

(1)窗的移动步长 l 由窗的移动速率 F'_s 和 $x(m)$ 的采样率 F_s 决定。对于汉明窗，

$$l = \frac{F_s}{F'_s} \leqslant \frac{F_s}{4F_s/L} = \frac{L}{4} \tag{11}$$

(2)由 N 长序列 $x_n(m)$ 构成 L 长序列 $u_n(q)$ 时，若 N/L 不为整数，可给 $x_n(m)$ 补零，使 N/L 为整数。

(3)根据要求调整 L 大小，可结合清/浊音判断进行。若为清音帧，L 可取的小一些，若为浊音帧，L 可取的大一些。这样，我们就得到了随着时间变化、不同频率点的时变频谱。

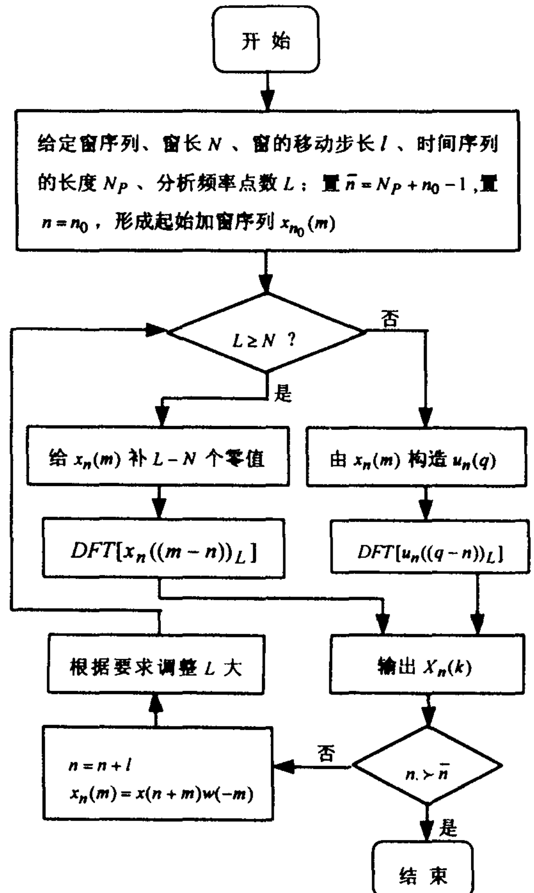

图 1 短时频谱分析流程图

4 讨论

(1)FFT 算法的计算量。若 L 为 2 的幂次，则 L 点 FFT 需 $2L\log_2 L$ 次实数乘法和实数加法，形成窗选序列 $x_n(m)$ 需 N 次乘法，构成 $u_n(q)$ 序列需 N 次加法。所以，计算 L 个点的离散傅氏变换需 $N+2L\log_2 L$ 次实数乘法和实数加法。当 $L=N=256$ 时，$N+2\log_2 L=4352$ 次。可见，运算量远小于直接计算的 262,144 次。

(2)由式(7)、(9)、(10)可知，$X_n(k)$ 是窗选序列或组合序列 $u_n(q)$ 的 DFT 与指数序列 $e^{-j\frac{2\pi}{L}kn}$ 的乘积。若仅考虑频谱的幅度特性，可使计算更为简单，无须对序列循环移位，直接进行 DFT 即可。但是，当需要由 $X_n(k)$ 恢复 $x(n)$ 时，计算中应考虑相位 $e^{-j\frac{2\pi}{L}kn}$。

参 考 文 献

[1] L. R. 拉宾纳，R. W. 谢弗著. 朱雪龙译. 语音信号处理[M]. 北京：科学出版社，1978.

[2] A. V. 奥本海姆，R. W. 谢弗著. 董士嘉等译. 数字信号处理[M]. 北京：科学出版社，1983.

FFT Technique on Speech Time-Varied Spectrum Analysis

DUAN Yan-li, ZHENG Rong

(The Telecommunication Engineering Institute, AFEU., Xi'an 710077, China)

Abstract: Speech signal spectrum analysis is the time-varied spectrum analysis. FFT technique on speech signal spectrum analysis are discussed in detail in this paper.

Key words: speech signal; time-varied spectrum; FFT