

LPC 方法提取语音信号共振峰的分析

郁伯康, 郁梅

(杭州电子工业学院, 浙江 杭州 310037)

【摘要】通过对 LPC(线性预测编码)方法提取语音信号共振峰进行的研究表明,采用相-频特性与对数幅-频特性同样能提取语音信号共振峰。与对数幅-频特性的二次导数相比,相-频特性的三次导数有更高的频率分辨率,能更有效地解决共振峰合并的问题,提取更精确的共振峰参数。

【关键词】共振峰; 语音信号; 线性预测编码(LPC)

【Abstract】The study on speech signal formant abstracted by LPC shows that it can also be abstracted by adopting phase-frequency characteristics and logarithmic amplitude-frequency characteristics. Compared with the second derivative of logarithmic amplitude-frequency characteristics, the third derivative of phase-frequency characteristics has better frequency resolution, so it can solve the combination of formant more effectively and abstract formant parameters more precisely.

【Key words】formant; speech signal; linear prediction coding(LPC)

1 引言

肺部气流冲击声带,通过声道的响应,形成语音。不同的声道形状产生不同的声道响应,导致不同的语音。就声道的数学模型,主要有两种观点,(1)把声道看作由多个不同截面积的声管串联而成,即声管模型;(2)把声道看作谐振腔,共振峰就是该腔体的谐振频率,即共振峰模型^{[1],[2]},因人耳听觉的柯替氏器官就是按频率感受而排列其位置的,因而,实践证明共振峰模型方法是非常有效的^[3]。共振峰是描述语音信号特征的重要参数,所以,准确有效的共振峰提取算法对语音信号的分析、合成、编码有重要意义。

线性预测编码(LPC)是进行语音信号分析、语音信号编码最有效的技术之一。其重要性在于提供了一组简洁的语音信号模型参数,比较精确地表征了语音信号的幅度谱,而分析它们所需的计算量相对而言并不大^[3]。语音信号共振峰的 LPC 分析方法的一个主要特点在于能够由预测系数构成的多项式中精确地估计共振峰频率和带宽^[2]。在过去的研究中,已有许多利用 LPC 分析提取共振峰方法被提出^{[5]-[10]}。McCandless 提出了由线性预测谱提取语音信号前三个共振峰的算法^[9],Christensen 等对三种由线性预测谱提取语音信号共振峰的算法进行了比较分析,即:寻找传递的函数的极点法,传递函数

对数幅度谱的峰值法,和传递函数对数幅度谱的二次导数的峰值法。其比较结果表明,传递函数对数幅度谱的二次导数的峰值法显示了较好的性能^[10]。在共振峰提取方法中,如何分离线性预测谱中的合并共振峰是一个重要问题。

2 LPC 方法语音信号提取共振峰

在语音信号的 LPC 模型中,语音信号样本 $s(n)$ 可由如下差分方程表示:

$$s(n) = \sum_{k=1}^p a_k s(n-k) + Gu(n) \quad (1)$$

式中, $u(n)$ 为激励函数, G 是增益, $\{a_k; k=1, 2, \dots, p\}$ 是 LPC 系数。相应的数字滤波器传递函数 $H(z)$ 为

$$H(z) = \frac{S(z)}{Gu(z)} = \frac{1}{1 - \sum_{k=1}^p a_k z^{-k}} \quad (2a)$$

上式还可表示为 p 个极点的级联形式:

$$H(z) = \frac{1}{\prod (1 - z_k z^{-1})} \quad (2b)$$

式中, $z_k = r_k \exp(j\theta_k)$ 是 $H(z)$ 在 z -平面上第 k 个极点。若 $H(z)$ 是稳定的,其所有极点都在 z -平面的单位圆内。则第 k 个共振峰的频率 F_k 和带宽 B_k 分别为 $F_k = \theta_k / 2\pi T$, $B_k = -\ln(r_k) / \pi T$, T 为语音信号采样周期。

语音信号的共振峰能由数字激光器传递函数 $H(z)$ 进行估计,最直接的方式是对 $H(z)$ 进行多项式

【基金项目】浙江省自然科学基金资助项目(698064)

求根,由所求的根来判断共振峰或谱形状极点^[5]。然而,该方法难以快速而有效地找到根值。另一有效的获取共振峰的途径是语音信号 LPC 谱的谱峰检测的方法进行估计共振峰,该方法需要解决共振峰合并的问题。McCandless 试图采用在 z -平面单位圆内重复计算 $H(z)$ 的方式来分离合并的共振峰^[9],另一个有效的方法是采用对数 LPC 谱的二阶导数进行估计共振峰^[10],取得较好效果。

下文就 LPC 对数幅-频特性的二次导数、相-频特性的一次导数和三次导数进行分析。这里, LPC 对数幅-频特性的二次导数^[10]的表述被进一步改进。首先考虑单个谐振腔, $H(z)$ 有两个共轭极点($p=2$):

$$H(z) = \frac{1}{1 - a_1 z^{-1} - a_2 z^{-2}} = \frac{1}{(1 - z_1 z^{-1})(1 - z_1^* z^{-1})} \quad (3)$$

式中, $z_1 = r_1 \exp(j\theta_1)$, z_1^* 是 z_1 的共轭复数。 $H(z)$ 的频率响应为

$$H(e^{j\omega}) = \frac{1}{(1 - z_1 e^{-j\omega})(1 - z_1^* e^{-j\omega})} \quad (4a)$$

$$|H(e^{j\omega})| = \frac{1}{\sqrt{1 + r_1^2 - 2r_1 \cos(\omega - \theta_1)} \sqrt{1 + r_1^2 - 2r_1 \cos(\omega + \theta_1)}} \quad (4b)$$

$$L(\omega) = -10 \left\{ \lg [1 + r_1^2 - 2r_1 \cos(\omega - \theta_1)] + \lg [1 + r_1^2 - 2r_1 \cos(\omega + \theta_1)] \right\} \quad (4c)$$

$$\varphi(\omega) = \arctan \frac{-r_1 \sin(\omega - \theta_1)}{1 - r_1 \cos(\omega - \theta_1)} + \arctan \frac{-r_1 \sin(\omega + \theta_1)}{1 - r_1 \cos(\omega + \theta_1)} \quad (4d)$$

ω 表示角频率, $\omega = 2\pi FT$, F 是频率。 $|H(e^{j\omega})|$, $L(\omega)$ 和 $\varphi(\omega)$ 分别为幅度谱, 对数幅度谱和相位谱。令 $c = -20 \lg e$, $L(\omega)$ 的二次导数、 $\varphi(\omega)$ 的一次导数和三次导数分别为:

$$L^{(2)}(\omega) = c \left\{ \frac{r_1(1+r_1^2)\cos(\omega-\theta_1)-2r_1^2}{[(1+r_1^2-2r_1\cos(\omega-\theta_1))^2] +} + \frac{r_1(1+r_1^2)\cos(\omega+\theta_1)-2r_1^2}{[(1+r_1^2-2r_1\cos(\omega+\theta_1))^2] +} \right\} \quad (5a)$$

$$\varphi^{(1)}(\omega) = \frac{r_1^2 - r_1 \cos(\omega - \theta_1)}{1 + r_1^2 - 2r_1 \cos(\omega - \theta_1)} + \frac{r_1^2 - r_1 \cos(\omega + \theta_1)}{1 + r_1^2 - 2r_1 \cos(\omega + \theta_1)} \quad (5b)$$

$$\varphi^{(3)}(\omega) = r_1(1-r_1^2) \times$$

$$\left\{ \frac{(1+r_1^2)\cos(\omega-\theta_1)-2r_1[(1+\sin^2(\omega-\theta_1))] +}{[(1+r_1^2-2r_1\cos(\omega-\theta_1))^3]} + \frac{(1+r_1^2)\cos(\omega+\theta_1)-2r_1[(1+\sin^2(\omega+\theta_1))] +}{[(1+r_1^2-2r_1\cos(\omega+\theta_1))^3]} \right\} \quad (5c)$$

性质-1 $|H(e^{j\omega})|$, $L(\omega)$ 为偶对称, 而 $\varphi(\omega)$ 为奇对称。 $L^{(2)}(\omega)$, $\varphi^{(1)}(\omega)$ 和 $\varphi^{(3)}(\omega)$ 都为偶对称。

性质-2 对于 $|H(e^{j\omega})|$, $L(\omega)$, $L^{(2)}(\omega)$, $\varphi^{(1)}(\omega)$ 和 $\varphi^{(3)}(\omega)$ 显然其正峰或负峰出现在 θ_1 附近, 这是峰值法估计共振峰的依据。

性质-3 考虑到语音信号的共振峰频率及带宽, 若角频率 ω 从 0 到 π 变化, 则 $L^{(2)}(\omega)$, $\varphi^{(1)}(\omega)$ 和 $\varphi^{(3)}(\omega)$ 式中的第二项可以被忽略。如, 当 $F_1=200$ Hz, $B_1=90$ Hz, $L^{(2)}(\omega)$, $\varphi^{(1)}(\omega)$ 和 $\varphi^{(3)}(\omega)$ 式中的第一项与第二项的比值分别为 81.61, 627.08 和 2 170.15。

性质-4 若 $\omega = \theta_1$, 则 $L^{(2)}(\theta_1)$, $\varphi^{(1)}(\theta_1)$ 和 $\varphi^{(3)}(\theta_1)$ 近似表示为

$$\begin{aligned} L^{(2)}(\theta_1) &\approx \frac{cr_1}{(1-r_1)^2}, \\ \varphi^{(1)}(\theta_1) &\approx \frac{-r_1}{1-r_1}, \\ \varphi^{(3)}(\theta_1) &\approx \frac{r_1(1+r_1)}{(1-r_1)^3} \end{aligned} \quad (6)$$

据以上关系, 通过检测 $L^{(2)}(\omega)$, $\varphi^{(1)}(\omega)$ 或 $\varphi^{(3)}(\omega)$ 的有效峰值, 估计出 r_1 , 故得到相应的共振峰带宽。图 1 示出单个谐振腔(共振峰频率 F_1 及带宽 B_1 分别为 2 000 Hz, 80 Hz, $T=0.0001$ s)的 $|H(e^{j\omega})|$, $L^{(2)}(\omega)$, $\varphi(\omega)$, $\varphi^{(1)}(\omega)$ 和 $\varphi^{(3)}(\omega)$ 。这里, $L^{(2)}(\omega)$, $\varphi^{(1)}(\omega)$ 和 $\varphi^{(3)}(\omega)$ 式中的第一项与第二项的比值分别为 5 728.75, 79.69, 27 555 640.69。由 $L^{(2)}(\omega)$, $\varphi^{(1)}(\omega)$ 和 $\varphi^{(3)}(\omega)$ 估计得到共振峰频率都为 2 000 Hz, 估计的共振峰带宽分别为 80.01 Hz, 81.00 Hz 和 80.00 Hz。

令 b_{1L} , b_{1L2} , $b_{1\varphi1}$, $b_{1\varphi3}$ 分别为 $L^{(2)}(\omega)$, $\varphi^{(1)}(\omega)$ 和 $\varphi^{(3)}(\omega)$ 波形中主波峰的带宽, 图 1 中 b_{1L} , b_{1L2} , $b_{1\varphi1}$, $b_{1\varphi3}$, 分别为 478 Hz, 28 Hz, 52 Hz 和 20 Hz, $b_{1L} > b_{1\varphi1} > b_{1L2} > b_{1\varphi3}$ 。这意味着 $\varphi^{(3)}(\omega)$ 随 $|\omega - \theta_1|$ 增加而衰减得最快, $\varphi^{(3)}(\omega)$ 具有最高的频率分辨率。定义 $\beta_{1L} = b_{1L}/B_1$, $\beta_{1L2} = b_{1L2}/B_1$, $\beta_{1\varphi1} = b_{1\varphi1}/B_1$, $\beta_{1\varphi3} = b_{1\varphi3}/B_1$, 图 2 给出了单个谐振腔(共振峰频率 F_1 为 2 000 Hz, 带宽 B_1 变化从 20 到 500 Hz)的 β_{1L} , β_{1L2} , $\beta_{1\varphi1}$ 和 $\beta_{1\varphi3}$ 的变化曲线。图 2 中, β_{1L2} , $\beta_{1\varphi1}$ 和 $\beta_{1\varphi3}$ 小于 1, 这表明采用 $L^{(2)}(\omega)$, $\varphi^{(1)}(\omega)$ 和 $\varphi^{(3)}(\omega)$ 进行波峰法估计共振峰

比直接用 $|H(e^{j\omega})|$ 有更高的频率分辨率,同时, β_{1q3} 小于 β_{1L} , β_{2L} 和 β_{1q1} ,因而, $\varphi^{(3)}(\omega)$ 的频率分辨率最高。这对更好地解决合并的共振峰是至关重要的。

考虑 $H(z)$ 为谐振腔的传递函数($p=4$), $H(z)=\frac{1}{(1-z_1z^{-1})(1-z_1^*z^{-1})(1-z_2z^{-1})(1-z_2^*z^{-1})}$ 。则相应的 $L^{(2)}(\omega)$, $\varphi^{(1)}(\omega)$ 和 $\varphi^{(3)}(\omega)$ 分别为

$$L^{(2)}(\omega) = \sum_{n=1}^2 c \left[\frac{r_n(1+r_n^2)\cos(\omega-\theta_n)-2r_n^2}{[(1+r_n^2-2r_n\cos(\omega-\theta_n))]^2} + \frac{r_n(1+r_n^2)\cos(\omega+\theta_n)-2r_n^2}{[(1+r_n^2-2r_n\cos(\omega+\theta_n))]^2} \right] \quad (7a)$$

$$\varphi^{(1)}(\omega) = \sum_{n=1}^2 \left[\frac{\frac{r_n^2-r_n\cos(\omega-\theta_n)}{1+r_n-2r_n\cos(\omega-\theta_n)} + \frac{r_n^2-r_n\cos(\omega+\theta_n)}{1+r_n-2r_n\cos(\omega+\theta_n)} \right] \quad (7b)$$

$$\varphi^{(3)}(\omega) = \sum_{n=1}^2 r_n(1-r_n^2) \times$$

$$\left[\frac{(1+r_n^2)\cos(\omega-\theta_n)-2r_n[(1+\sin^2(\omega-\theta_n))]^3}{[(1+r_n^2-2r_n\cos(\omega-\theta_n))]^3} + \frac{(1+r_n^2)\cos(\omega+\theta_n)-2r_n[(1+\sin^2(\omega+\theta_n))]^3}{[(1+r_n^2-2r_n\cos(\omega+\theta_n))]^3} \right] \quad (7c)$$

式中, $z_1=r_1\exp(j\theta_1)$, $z_2=r_2\exp(j\theta_2)$ 。

性质-5 若共振峰 θ_1 (或 F_1)足以远离共振峰 θ_2 (或 F_2),则 $L^{(2)}(\theta_n)$, $\varphi^{(1)}(\theta_n)$ 和 $\varphi^{(3)}(\theta_n)$, $n=1$ 或 2 ,可由式(6)进行求解。否则,计算如下:

$$L^{(2)}(\theta_1) \approx c \left[\frac{r_1}{(1-r_1)^2} + \frac{r_2(1+r_2^2)\cos(\theta_1-\theta_2)-2r_2^2}{[(1+r_2^2-2r_2\cos(\theta_1-\theta_2))]^2} \right] \quad (8a)$$

$$\varphi^{(1)}(\theta_1) \approx \frac{-r_1}{1-r_1} + \frac{r_2-r_2\cos(\theta_1-\theta_2)}{1+r_2-2r_2\cos(\theta_1-\theta_2)} \quad (8b)$$

$$\varphi^{(3)}(\theta_1) \approx \frac{r_1(1+r_1)}{(1-r_1)^3} +$$

$$r_2(1-r_2^2) \frac{(1+r_2^2)\cos(\theta_1-\theta_2)-2r_2[(1+\sin^2(\theta_1-\theta_2))]^3}{[(1+r_2^2-2r_2\cos(\theta_1-\theta_2))]^3} \quad (8c)$$

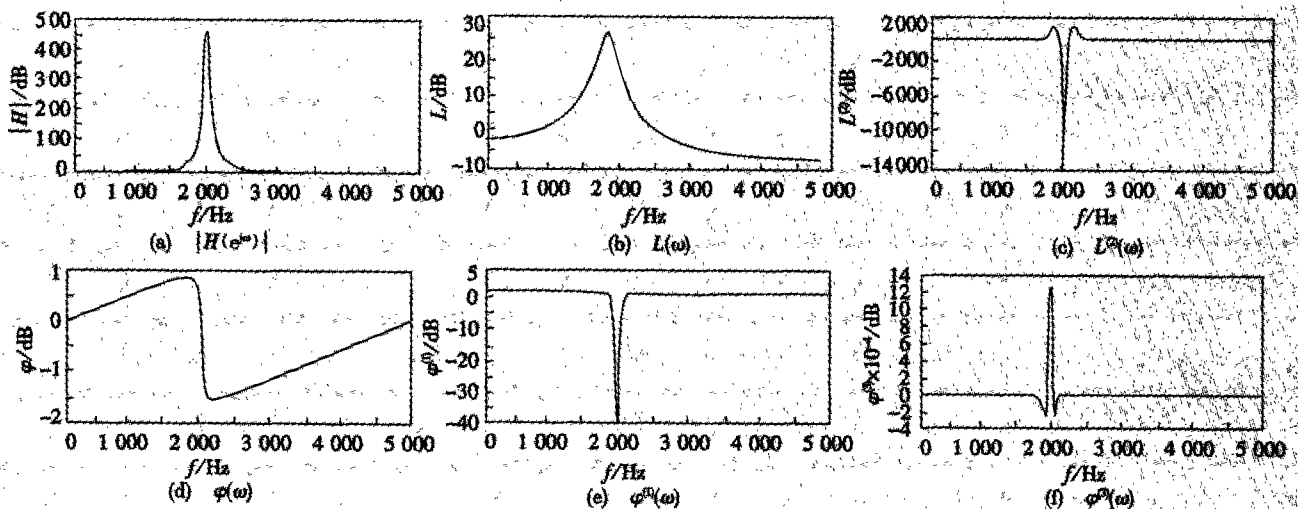


图1 单个谐振器的频率响应($F_1=2\ 000\text{ Hz}$, $B_1=80\text{ Hz}$)

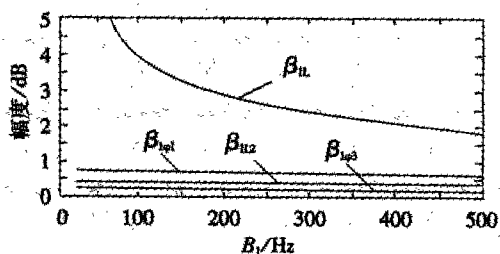


图2 单个谐振腔($F_1=2\ 000\text{ Hz}$)的 β_{1L} , β_{2L} , β_{1q1} 和 β_{1q3} 的变化曲线

图3,4和5显示了三种不同带宽情况下两个合并共振峰检测的结果。图中两个共振峰频率 F_1 , F_2

分别为1900 Hz,2000 Hz。图3(a),4(a),和5(a)分别给出了三种不同带宽情况下的对数幅度谱 $L(\omega)$,这里,相应的(B_1 , B_2)分别为(90 Hz,110 Hz),(140 Hz,160 Hz),和(190 Hz,210 Hz)。图3(b),4(b),和5(b)为 $L^{(2)}(\omega)$,图3(c),4(c),和5(c)为相应的 $\varphi^{(1)}(\omega)$,而图3(d),4(d),和5(d)对应于 $\varphi^{(3)}(\omega)$ 。很明显,合并的共振峰难以直接由 $L(\omega)$ 或 $|H(e^{j\omega})|$ 分离并提取,但由于 $\varphi^{(1)}(\omega)$, $L^{(2)}(\omega)$ 和 $\varphi^{(3)}(\omega)$ 有较高的频率分辨率,所以仍能以不同的程度分离合并的共振

峰并提取相应的参数。当两个共振峰带宽较小时,如图 3 所示,上述三种方法均能够分离合并的共振峰;随着带宽增大,如图 4 所示,虽然 $\varphi^{(1)}(\omega)$ 不能识别合并的共振峰,但 $L^{(2)}(\omega)$ 和 $\varphi^{(3)}(\omega)$ 有效地分离了合并的共振峰;当带宽进一步增大如图 5, $\varphi^{(1)}(\omega)$ 和 $L^{(2)}(\omega)$ 都不能分离了合并的共振峰,但 $\varphi^{(3)}(\omega)$ 仍旧检测出合并的共振峰。三种不同带宽情况下三种方法提取的共振峰参数如表 1 所示。表 1 中,对于图 3 情况,真实的 (B_1, B_2) 为 (90 Hz, 110

Hz), 因 $\varphi^{(1)}(\omega)$ 提取的共振峰频率分别为 1 903 Hz 和 1 994 Hz, 若直接由式 (6) 对带宽估计, 其结果为 (78.07, 94.42), 采用修正后的式 (7), 则检测结果为 (93.29, 117.42)。若所用方法能够分离合并的共振峰, 分别给出由式 (6) 和修正后的式 (7) 检测带宽的结果, 为方便辨认, 在表 1 中对由式 (7) 检测的结果加上阴影。从表 1 中可知, 采用 $\varphi^{(3)}(\omega)$ 的方法得到共振峰参数检测结果与实际的共振峰参数最接近。

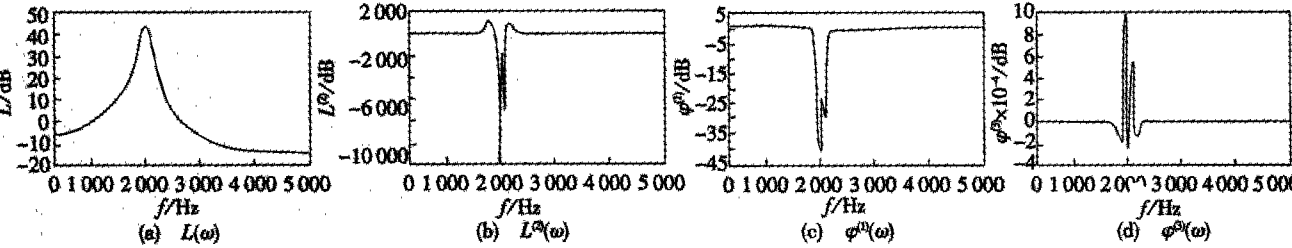


图 3 两个合并共振峰(F₁, F₂)=(1 900, 2 000)的检测:(B₁, B₂)=(90, 110)

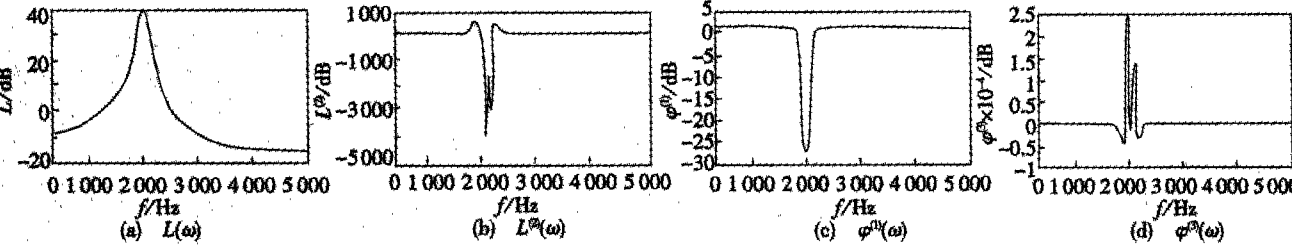


图 4 两个合并共振峰(F₁, F₂)=(1 900, 2 000)的检测:(B₁, B₂)=(140, 160)

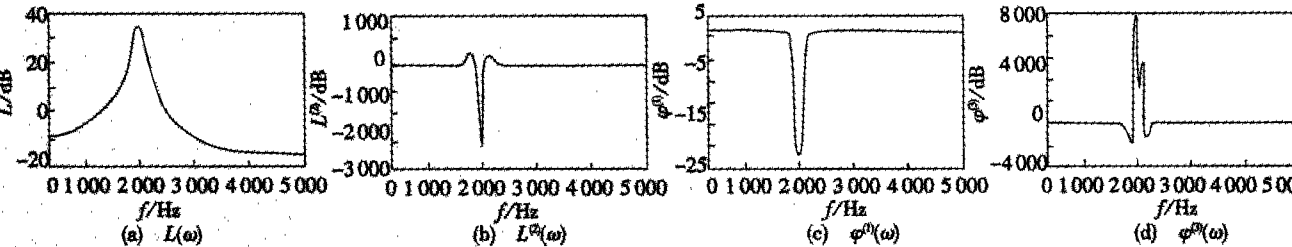


图 5 两个合并共振峰(F₁, F₂)=(1 900, 2 000)的检测:(B₁, B₂)=(190, 210)

表 1 三种方法提取不同带宽合并共振峰的比较 (F₁=1 900Hz, F₂=2 000Hz) (单位: Hz)

	$\varphi^{(1)}(\omega)$		$L^{(2)}(\omega)$		$\varphi^{(3)}(\omega)$	
	(F ₁ , F ₂)	(B ₁ , B ₂)	(F ₁ , F ₂)	(B ₁ , B ₂)	(F ₁ , F ₂)	(B ₁ , B ₂)
(B ₁ , B ₂)=(90, 110)	(1903, 1994)	(78.07, 94.42)	(1 900, 2 001)	(94.02, 120.55)	(1 900, 2 001)	(91.91, 114.75)
		(93.29, 117.42)		(90.65, 110.17)		(90.09, 110.07)
(B ₁ , B ₂)=(140, 160)	F ₁ =1 913	B ₁ =107.44	(1 902, 1 998)	(144.79, 173.20)	(1 899, 20 001)	(147.58, 177.45)
				(142.50, 161.35)		(141.46, 160.62)
(B ₁ , B ₂)=(190, 210)	F ₁ =1 930	B ₁ =131.15	F ₁ =1 909	B ₁ =186.52	(1 900, 2 002)	(203.31, 240.55)
						(194.85, 211.61)

对于式 (2b) 中的 $H(z)$, $L^{(2)}(\omega)$, $\varphi^{(1)}(\omega)$ 和 $\varphi^{(3)}(\omega)$ 分别表示如下:

$$L^{(2)}(\omega)=\sum_{n=1}^p c\left(\frac{r_n(1+r_n^2)\cos(\omega-\theta_n)-2r_n^2}{\left[(1+r_n^2-2r_n\cos(\omega-\theta_n))\right]^2}\right)\tag{9a}$$

$$\varphi^{(1)}(\omega) = \sum_{n=1}^p \left(\frac{r_n^2 - r_n \cos(\omega - \theta_n)}{1 + r_n^2 - 2r_n \cos(\omega - \theta_n)} \right) \quad (9b)$$

$$\varphi^{(3)}(\omega) = \sum_{n=1}^p r_n (1 - r_n^2) \times \frac{(1 + r_n^2) \cos(\omega - \theta_n) - 2r_n \left[(1 + \sin^2(\omega - \theta_n)) \right]}{\left[1 + r_n^2 - 2r_n \cos(\omega - \theta_n) \right]^3} \quad (9c)$$

当共振峰 θ_k 与其他共振峰不相近时, $L^{(2)}(\theta_k)$, $\varphi^{(1)}(\theta_k)$ 和 $\varphi^{(3)}(\theta_k)$ 的计算相似于式(6)。否则, 计算如下:

$$L^{(2)}(\theta_k) \approx 2c \left(\frac{r_k}{(1 - r_k^2)^2} + \frac{r_j(1 + r_j^2) \cos(\theta_k - \theta_j) - 4r_j^2}{\left[(1 + r_j^2 - 2r_j \cos(\theta_k - \theta_j)) \right]^2} \right) \quad (10a)$$

$$\varphi^{(1)}(\theta_k) \approx \frac{-r_k}{1 - r_k^2} - \frac{r_j^2 - r_j \cos(\theta_k - \theta_j)}{1 + r_j^2 - 2r_j \cos(\theta_k - \theta_j)} \quad (10b)$$

$$\varphi^{(3)}(\theta_k) \approx \frac{r_k(1 + r_k)}{(1 - r_k)^3} + r_j(1 - r_j^2) \times \frac{(1 + r_j^2) \cos(\theta_k - \theta_j) - 2r_j \left[1 + \sin^2(\theta_k - \theta_j) \right]}{\left[(1 + r_j^2 - 2r_j \cos(\theta_k - \theta_j)) \right]^3} \quad (10c)$$

这里, 假定 θ_k 与 θ_j 相近。

3 实验分析

由上述讨论和分析, 对三种提取语音信号共振峰方法进行了计算机模拟实验。首先, 对已知共振峰的提取进行三种方法比较。图 6(a) 为已知的 5 对共振峰参数构成的传递函数 $H(z)$ 的对数幅度谱 $L(\omega)$, 图 6(b), 6(c) 和 6(d) 分别为相应的对数幅度谱二次导数 $L^{(2)}(\omega)$ 、相位谱一次导数 $\varphi^{(1)}(\omega)$ 和相位谱三次导数 $\varphi^{(3)}(\omega)$ 。表 2 显示了采用三种方法提取得到的共振峰参数值, 其中第一栏为实际的 5 对共振峰参数值。图 6 和表 2 显示 $L^{(2)}(\omega)$ 和 $\varphi^{(1)}(\omega)$ 难以分离合并的第一、第二共振峰, 而 $\varphi^{(3)}(\omega)$ 则较好地分离了两共振峰并且得到相应的共振峰参数值。根据表 2, 三种方法提取的共振峰带宽都大于实际值。采用 $\varphi^{(1)}(\omega)$ 和 $L^{(2)}(\omega)$ 提取的共振峰带宽偏离实际值较远; 与 $L^{(2)}(\omega)$ 和 $\varphi^{(1)}(\omega)$ 相比, 由 $\varphi^{(3)}(\omega)$ 提取的共振峰参数值更接近于实际的 5 对共振峰参数值, 因而, 其估计的精度最高。

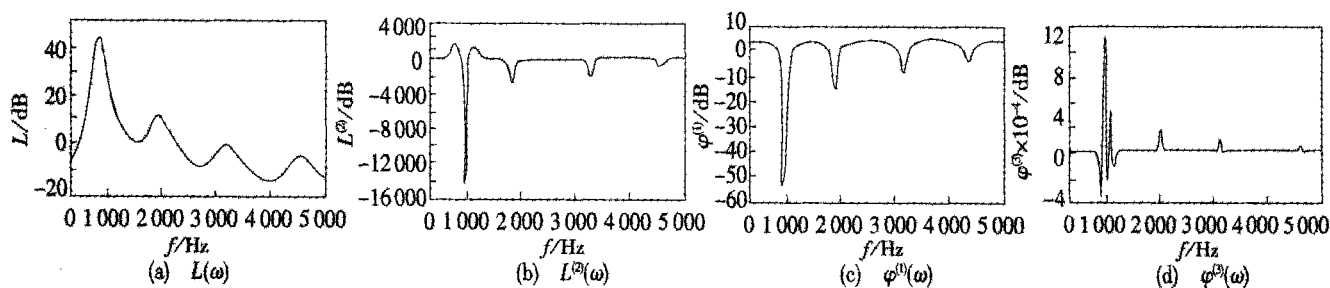


图 6 三种方法提取 5 对共振峰的比较

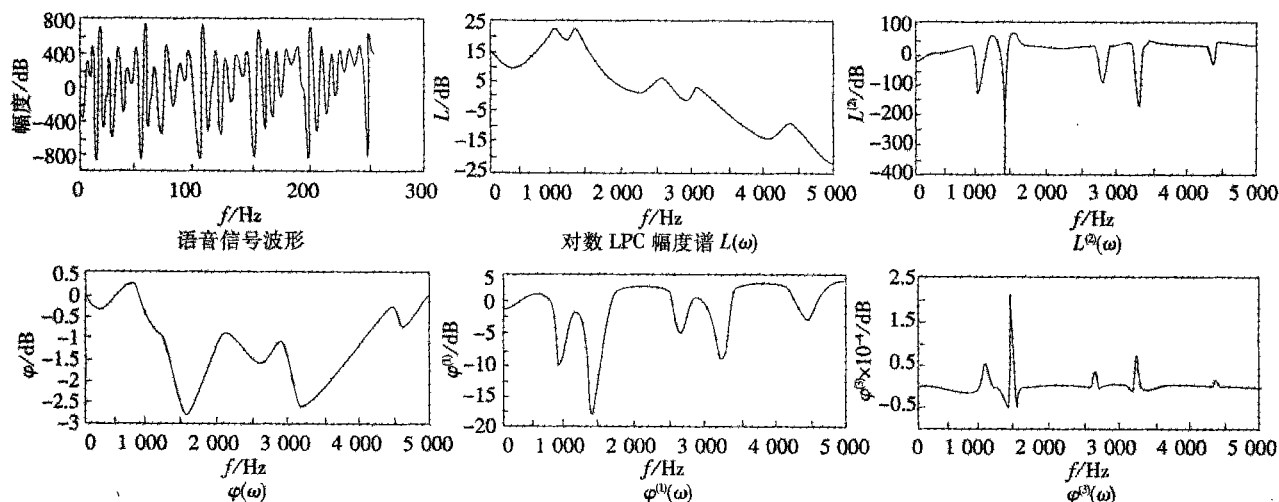


图 7 实际语音信号共振峰的提取

其次,对实际语音信号共振峰参数的提取进行实验。语音信号先进行低通滤波器滤波,再进行 10 kHz 频率采样。每个语音帧长为 25.6 ms,以 Hanning 窗进行加权,由自相关法求取 LPC 系数 $\{a_k; k=1, 2, \dots, p\}$ 。采用语音信号共振峰 LPC 分析方法的一个主要特点在于能够根据预测系数构成的多项式估计出共振峰频率和带宽。额外的极点通常由于其带宽与语音信号共振峰带宽相比要大很多而被孤立。在传递函数谱分析上,与幅度谱相比,采用相位谱的优点在于它与各个极点存在着一定的线性关系。式(2a)的 LPC 谱函数可以表示为:

$$|H(e^{j\omega})| = \frac{1}{\sqrt{\left(1 - \sum_{k=1}^p a_k \cos k\omega\right)^2 + \left(\sum_{k=1}^p a_k \sin k\omega\right)^2}} \quad (11a)$$

$$\varphi(\omega) = \arctan \frac{\sum_{k=1}^p a_k \sin k\omega}{1 - \sum_{k=1}^p a_k \cos k\omega} \quad (11b)$$

相应的 $L^{(2)}(\omega)$, $\varphi^{(1)}(\omega)$ 和 $\varphi^{(3)}(\omega)$ 能被计算。图 7 一帧语音信号时域波形、对数 LPC 幅度谱 $L(\omega)$, $L^{(2)}(\omega)$, LPC 相位谱 $\varphi(\omega)$, $\varphi^{(1)}(\omega)$ 和 $\varphi^{(3)}(\omega)$ 。表 3 给出了三种方法提取的共振峰参数值。

表 2 三种方法提取 5 对共振峰的比较(单位:Hz)

	$\varphi^{(1)}(\omega)$	$L^{(2)}(\omega)$	$\varphi^{(3)}(\omega)$
$(F_1, B_1) = (750, 80)$	(756, 60.25)	(752, 78.15)	(750, 81.66)
$(F_2, B_2) = (800, 120)$			(805, 123.27)
$(F_3, B_3) = (1\ 800, 150)$	(1\ 800, 186.34)	(1\ 800, 151.45)	(1\ 800, 150.00)
$(F_4, B_4) = (3\ 200, 200)$	(3\ 200, 271.40)	(3\ 200, 202.26)	(3\ 200, 200.01)
$(F_5, B_5) = (4\ 500, 250)$	(4\ 500, 368.58)	(4\ 500, 255.15)	(4\ 500, 250.08)

表 3 实际语音信号共振峰的提取 (单位:Hz)

	$\varphi^{(1)}(\omega)$	$L^{(2)}(\omega)$	$\varphi^{(3)}(\omega)$
(F_1, B_1)	(991, 294.43)	(989, 251.29)	(990, 239.45)
(F_2, B_2)	(1\ 412, 168.79)	(1\ 413, 148.78)	(1\ 413, 146.19)
(F_3, B_3)	(2\ 567, 425.12)	(2\ 566, 276.31)	(2\ 566, 264.52)
(F_4, B_4)	(3\ 196, 289.87)	(3\ 197, 209.64)	(3\ 197, 204.75)
(F_5, B_5)	(4\ 340, 862.88)	(4\ 340, 377.56)	(4\ 340, 360.10)

4 结论

文中就 LPC 方法提取语音信号共振峰进行研究。研究讨论了数字滤波器 $H(z)$ 在 z -平面上的幅-频、相-频特性,分析比较了对数幅-频特性的二次导数、相-频特性的一次导数和三次导数的频率分辨率。理论分析和实验结果表明,采用相-频特性与对数幅-频特性同样能够提取语音信号共振峰。并且,与 $L^{(2)}(\omega)$, $\varphi^{(1)}(\omega)$ 相比, $\varphi^{(3)}(\omega)$ 随 $|\omega - \theta_k|$ 增加而衰减得最快,因而, $\varphi^{(3)}(\omega)$ 具有最高的频率分辨率,能更有效地解决共振峰合并的问题,提取更精确的语音信号共振峰参数。

[参考文献]

- [1] L.R.Rabiner, V.B.H.Juang. Fundamentals of Speech Recognition. PTR Prentice Hall, 1993.
- [2] L.R.Rabiner, R.W.Schafer. Digital Processing of Speech Signals. Prentice Hall, 1978.
- [3] 陈永彬, 王仁华. 语音信号处理. 合肥: 中国科学技术大学出版社, 1990.
- [4] J.D.Markel, A.H.Gray, Jr. Linear Prediction of Speech. Springer-verlag, New York, 1976.
- [5] B.S.Atal, S.L.hananer. Speech Analysis and Synthesis by Linear Prediction of the Speech Wave. J.Am. Acoust.Soc., 1971, 50:637-655.
- [6] J.D.Markel. Digital Inverse Filtering A New Tool for Formant Trajectory Estimation. IEEE Trans.on Audio and Electroacoustics, 1972, 20:129-137.
- [7] J.D.Markel. Application of a Digital Inverse Filter for Automatic Formant and F0 Analysis, IEEE Trans.on Audio and Electroacoustics, 1973, 21:149-153.
- [8] J.Markel. Spectral Analysis of Speech by Linear Prediction. IEEE Trans.on Audio and Electroacoustics, 1973, 21:140-148.
- [9] S.S.McCandless. An Algorithm for Automatic Formant Extraction Using Linear Prediction Spectra. IEEE Trans.on ASSP, April 1974, 22:135-141.
- [10] R.L.Christensen, W.J.Sreong, E.P.Palmer. A Comparison of Three Methods of Extracting Resonance Information from Predictor-coefficient Coded Speech. IEEE Trans. on ASSP 1976, 24:8-14.

[收稿日期] 1999-12-28

欢迎广大作者、读者、编委踊跃投稿