

自适应格型滤波器新算法及其在语音 线性预测合成中的应用¹

江太辉

(五邑大学信息科学研究室 广东江门 529020)

摘 要 本文从自适应格型滤波器导出自适应线谱对 (LSP) 滤波, 提出了使用最小均方 (LMS) 型自适应算法逐级更新计算线谱对系数的方法。实验表明, 该算法与其它算法比较, 具有更高的收敛率和较低的失调。用该算法计算得到的 LSP 系数进行语音线性预测合成, 获得比使用 PARCOR 系数更好的结果。

关键词 线谱对滤波器, 自适应格型滤波器, 线性预测, LMS 算法

中图号 TN713, TN912.3

1 引 言

自适应滤波器由于能自动调节其参数, 使滤波器的输出误差最小, 从而达到最佳滤波的效果, 所以它已广泛用于系统模型识别、通信信道自适应均衡、雷达和声纳波束形成、噪音抵消、谱线增强和线性预测等。自适应滤波器的传统形式为抽头延时线的横向结构数字滤波器, 然而横向结构的自适应滤波器已被证明其收敛速度慢和权系数对有限字长舍入误差灵敏等缺点^[1,2]。为了克服这些缺点, 一种格型自适应滤波器被提出来了^[1,3-6]。格型滤波器是一种正交结构滤波器, 它将原始观察数据转换为前向和后向预测误差的两个正交序列, 对整个系统最小化, 可以化成一系列独立的对每一级局部最小化。这种格型结构滤波器可以被解释为 Durbin 算法的结构^[3,4]。它具有许多优点, 在语音处理等方面有重要应用。格型自适应数字滤波器采用 Durbin 算法得到的系数称为偏自相关 (PARCOR) 系数^[3,4]。PARCOR 系数是目前进行语音合成的常用系数。

本文在格型自适应滤波器基础上导出线谱对 (Line Spectral Pair, LSP) 自适应滤波器, 使用一种最小均方 (LMS) 型的逐级修正的算法来计算 LSP 系数, 该算法可避免计算自相关参数或矩阵操作。实验表明, 与其它算法相比, 它提供了更高的收敛率和较低的失调。用该算法计算得到的 LSP 系数, 在 LSP 结构自适应滤波器上进行线性预测语音合成, 效果好于用 PARCOR 系数合成的结果。

2 从格型滤波器到线谱对分析

在语音的线性预测中, 语音的产生可以用如下的传递函数的全极点谱为模型^[3]:

$$H(z) = G/A(z). \quad (1)$$

¹ 1995-03-09 收到, 1995-07-18 定稿
广东高教厅重点科研项目资助课题

这里

$$A(z) = 1 + \sum_{k=1}^p a_k z^{-k}, \quad (2)$$

$A(z)$ 是 p 阶逆滤波器, G 是增益因子, a_k 是预测系数, p 为模型的极点数 (阶数) 或预测系数的数目。假如 $H(z)$ 稳定且是最小相位的, 则 $A(z)$ 能用一个格型滤波器来实现^[3]。图 1 虚线框中表示一个 p 阶格型滤波器, 用 $e^{(i)}(n)$ 和 $b^{(i)}(n)$ 表示时间为 n 时 i 阶格点输出的前向预测误差 (或称残差) 和后向预测误差。从图 1 我们有

$$e^{(0)}(n) = b^{(0)}(n) = S(n), \quad (3)$$

$$e^{(i)}(n) = e^{(i-1)}(n) - k_i b^{(i-1)}(n-1), \quad (4)$$

$$b^{(i)}(n) = b^{(i-1)}(n-1) - k_i e^{(i-1)}(n), \quad (5)$$

这里 $S(n)$ 为输入信号, k_i 是第 i 级的 PARCOR 系数。PARCOR 系数与 (2) 式中的线性预测系数 a_k 有如下的关系^[3,4]:

$$a_i^{(i)} = k_i, \quad 1 \leq i \leq p. \quad (6)$$

若滤波器是稳定的, 则要求 $-1 \leq k_i \leq 1$ 。 k_i 的计算, Itakura 建议为^[3]

$$k_i = \frac{\sum_{n=0}^{N-1} e^{(i-1)}(n)b^{(i-1)}(n)}{\left\{ \sum_{n=0}^{N-1} (e^{(i-1)}(n))^2 \cdot \sum_{n=0}^{N-1} (b^{(i-1)}(n))^2 \right\}^{1/2}}. \quad (7)$$

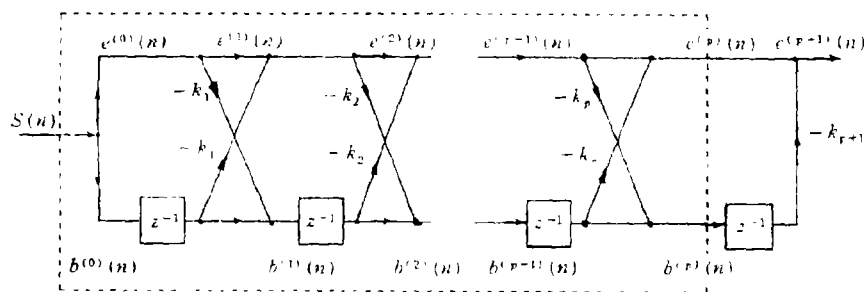


图 1 扩展了的格型滤波器

在图 1 虚线框内的 p 阶格型滤波器中, 如果输入信号 $S(n)$ 与前向预测误差信号 $e^{(p)}(n)$ 之间的传递函数表示为

$$A(z) = E^{(p)}(z)/S(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_p z^{-p}, \quad (8)$$

则输入信号 $S(n)$ 与后向预测误差信号 $b^{(p)}(n)$ 之间的传递函数为

$$\frac{B^{(p)}(z)}{S(z)} = z^{-p} A(z^{-1}) = a_p + a_{p-1} z^{-1} + a_{p-2} z^{-2} + \cdots + a_1 z^{-(p-1)} + z^{-p}. \quad (9)$$

注意, 这里 $a_k (k=1, 2, \dots, p)$ 与 (2) 式中的 a_k 一样, 都为线性预测系数, a_k 与 PARCOR 系数 k_i 的关系由 (6) 式确定。

当一个格型滤波器的 PARCOR 系数 k_i 的绝对值小于 1 或等于 1 时, $A(z)$ 的所有零点一定在 z 平面的单位圆内或单位圆上。现在我们将格型滤波器扩展额外的一级, 如图 1 所示, 且使 $k_{p+1} = \pm 1$, 这样扩展了的格型滤波器的所有零点都在单位圆上^[7]。若用 $F(z)$ 和 $G(z)$ 分别表示 $k_{p+1} = +1$ 和 -1 的扩展了的格型滤波器的传递函数, 则有

$$F(z) = A(z) - z^{-(p+1)} \cdot A(z^{-1}), \quad (10)$$

$$G(z) = A(z) + z^{-(p+1)} \cdot A(z^{-1}). \quad (11)$$

当 p 为偶数时, $F(1) = 0, G(-1) = 0$ 。设 $A(z)$ 在 $z = +1$ 和 -1 处没有根, 则 $F(z)$ 和 $G(z)$ 的根除了可以等于 1 和 -1 外, 其余的根仅以复共轭出现。因此 $F(z)$ 和 $G(z)$ 可以表示为^[7]

$$F(z) = (1 - z^{-1}) \prod_{i=1}^{p/2} (1 + c_i z^{-1} + z^{-2}), \quad (12)$$

$$G(z) = (1 + z^{-1}) \prod_{i=1}^{p/2} (1 + d_i z^{-1} + z^{-2}), \quad (13)$$

式中

$$c_i = -2 \cos \phi_i, \quad i = 1, 2, \dots, p/2; \quad (14)$$

$$d_i = -2 \cos \theta_i, \quad i = 1, 2, \dots, p/2; \quad (15)$$

而 $\phi_i = \cos^{-1}(-c_i/2)$, $\theta_i = \cos^{-1}(-d_i/2)$ 。 c_i 和 d_i (或 ϕ_i 和 θ_i) 称为 LSP 系数^[7]。由 (10) 式和 (11) 式可得

$$A(z) = [F(z) + G(z)]/2. \quad (16)$$

这样 LSP 系数 c_i 和 d_i (或 ϕ_i 和 θ_i) 可以用来代替 $a_i (i=1, 2, \dots, p)$ 表示 (1) 式的预测滤波器的传递函数 $A(z)$ 。当 $A(z)$ 为最小相位时, $F(z)$ 和 $G(z)$ 的零点在单位圆上交替出现, 且有

$$0 < \theta_1 < \phi_1 < \theta_2 < \dots < \theta_{p/2} < \phi_{p/2} < \pi. \quad (17)$$

从 (12), (13) 和 (16) 式可以得到一个线谱对自适应滤波器的结构, 如图 2 所示。图中 p 为偶数。

3 线谱对的更新算法

图 2 所示的 LSP 滤波器结构是用非递归的二阶节级联来实现的。这个滤波器有一个输出误差信号 $e(n)$ 。每个二阶节有两个误差信号 $f^{(i)}(n)$ 和 $g^{(i)}(n)$, ($i=1, 2, \dots, p/2$)。输出误差信号表示为

$$e(n) = [f^{(p/2)}(n) + g^{(p/2)}(n)]/2. \quad (18)$$

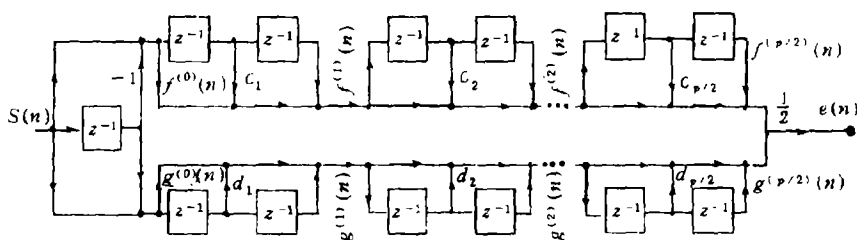


图 2 LSP 自适应滤波器结构

$f^{(i)}(n)$ 和 $g^{(i)}(n)$ 为

$$f^{(i)}(n) = f^{(i-1)}(n) + c_i f^{(i-1)}(n-1) + f^{(i-1)}(n-2), \quad (19)$$

$$g^{(i)}(n) = g^{(i-1)}(n) + d_i g^{(i-1)}(n-1) + g^{(i-1)}(n-2), \quad (20)$$

且有

$$f^{(0)}(n) = S(n) - S(n-1), \quad (21)$$

$$g^{(0)}(n) = S(n) + S(n-1). \quad (22)$$

现在我们的目标是调节图 2 滤波器的系数 c_i 和 d_i ($i = 1, 2, \dots, p/2$), 使输出误差达到最小均方值。这个最小均方值可以通过对每个二阶节进行局部最小化来达到, 对此, 我们使用 LMS 型算法来更新滤波器的系数, 使每一级的输出达到最小化。这就是使 $[f^{(i)}(n)]^2$ 和 $[g^{(i)}(n)]^2$ 分别对系数 c_i 和 d_i ($i = 1, 2, \dots, p/2$) 在负梯度方向上进行迭代计算, 直到 $e(n)$ 获得最小均方值为止。

为了方便, 我们只对图 2 中的一节滤波器进行研究, 表示在图 3 中。若对这一节滤波器使用 LMS 的修正算法, 则 c_i 和 d_i 可以按下式更新。

$$c_i(n+1) = c_i(n) - \mu_i \partial [f^{(i)}(n)]^2 / \partial c_i(n), \quad (23)$$

$$d_i(n+1) = d_i(n) - \mu_i \partial [g^{(i)}(n)]^2 / \partial d_i(n), \quad (24)$$

其中 μ_i 是一个控制收敛速度和稳定性的常数^[1]。若使用 (19) 和 (20) 式, 则可得

$$\partial [f^{(i)}(n)]^2 / \partial c_i(n) = 2f^{(i)}(n)f^{(i-1)}(n-1), \quad (25)$$

$$\partial [g^{(i)}(n)]^2 / \partial d_i(n) = 2g^{(i)}(n)g^{(i-1)}(n-1). \quad (26)$$

这样我们可得到 LSP 自适应滤波器逐级逐个样本的修正算法为

$$c_i(n+1) = c_i(n) - 2\mu_i f^{(i)}(n)f^{(i-1)}(n-1), \quad (27)$$

$$d_i(n+1) = d_i(n) - 2\mu_i g^{(i)}(n)g^{(i-1)}(n-1). \quad (28)$$

于是, 从 (19)-(22) 式, (27) 和 (28) 式, 通过逐级迭代更新计算出 c_i 和 d_i , 使每一级达到最小化, 最后使输出 $e(n)$ 达到最小均方值。

表 1 LSP 的终点和逐级算法的学习特性 ($\mu = 0.002$)

算法	收敛周期 (迭代次数)	失调 (%)
LSP 终点	65	84.6
LSP 逐级	31	53.4

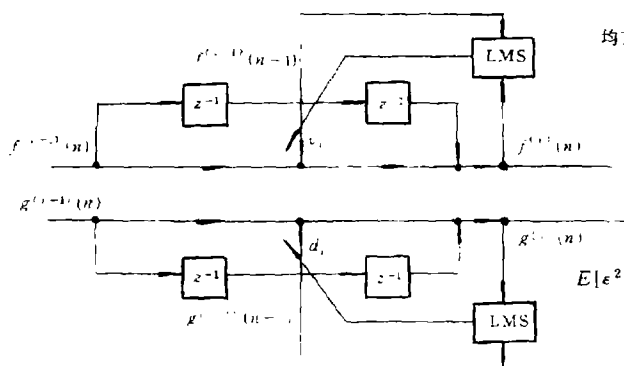


图 3 一节 LSP 滤波器的 LMS 修正算法

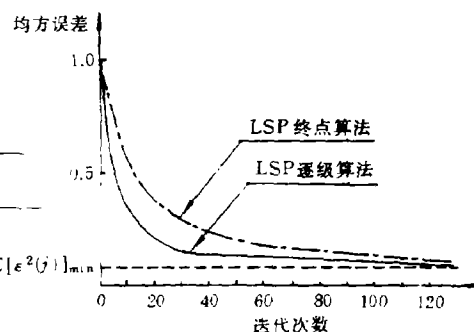


图 4 两种自适应算法的学习特性比较

4 实验和应用

我们在 486 微型机上进行一系列的实验, 研究本文提出的逐级更新的 LSP 滤波器算法的收敛性。计算出格型滤波器的 PARCOR 系数和 LSP 滤波器的线谱对系数和相应的误差信号, 并用这些参数在 486 微型机上进行语音合成试验, 比较它们的特性。

我们采用文献 [4] 附录 8 中提供的日语元音“ウ”以频率 10kHz 抽样, 并予以 10bit 量化后得到的 256 个数据做为实验的语音信号输入数据。对文献 [7] 介绍的 LSP 终点误差 (end-point error) 算法和本文使用的逐级更新算法进行试验, 每 256 个数据作为一个周期, 实验进行 40 次, 并从这 40 次的单个学习曲线计算出其综合平均。图 4 和表 1 总结了 LSP 终点误差算法和 LSP 逐级更新算法的实验结果。这里 μ 取相同的值, 即 $\mu = 0.002$ 。

从表 1 和图 4 可看出, LSP 逐级更新算法的收敛周期 (31 次迭代) 大大小于 LSP 终点误差算法的收敛周期 (65 次迭代)。其次, LSP 逐级更新算法的失调系数也小于 LSP 终点误差算法的失调系数。失调系数按下式计算 [1]:

$$M = (\xi_{ss} - \xi_{\min}) / \xi_{\min}$$

式中 $\xi_{\min} = E[\epsilon^2(j)]_{\min}$ 为最小均方误差值, $\xi_{ss} = \lim_{j \rightarrow \infty} E[\epsilon^2(j)]$ 为稳态均方误差值。

另外, 我们仍用上述的 256 个数据作为语音信号的输入数据, 使用 Durbin 算法 [3,4], 按 (7) 式计算出格型线性预测的 PARCOR 系数 k_i ($i = 1, 2, \dots, p$), 同时按图 2 所示的 LSP 结构, 使用 LSP 逐级更新算法, 求得 LSP 系数 c_i 和 d_i ($i = 1, 2, \dots, p/2$), 如表 2 所示, 这里级数 $p = 14$ 。

表 2 PARCOR 系数和 LSP 系数

PARCOR 系数				LSP 系数		
i	$k(i)$	i	$k(i)$	i	$c(i)$	$d(i)$
1	0.9609	8	-0.9075	1	-1.5102	-1.5471
2	-0.6136	9	-0.2598	2	-1.2054	-1.3390
3	-0.2789	10	0.0920	3	-1.0986	-1.1875
4	-0.3503	11	0.4665	4	0.3083	0.1103
5	0.2204	12	0.0644	5	0.6704	0.5083
6	-0.2953	13	-0.1250	6	1.3455	0.9346
7	-0.1091	14	0.0330	7	1.7818	1.5685

表 3 由 PARCOR 滤波器求出的误差值

n	$e(n)$	n	$e(n)$
1	1.8134	11	1.2813
2	-1.5237	12	2.3144
3	1.2561	13	2.4237
4	0.6719	14	2.1781
5	-0.5517	15	1.9643
6	0.6801	16	1.6555
7	2.3205	17	-3.4462
8	2.7458	18	-12.6301
9	1.5441	19	2.7716
10	1.4026	20	2.2430

表 4 由 LSP 滤波器求出的误差值

n	$e(n)$	n	$e(n)$
1	2.1363	11	1.3567
2	1.1585	12	4.3561
3	-0.6345	13	-2.4890
4	1.4211	14	0.4347
5	0.7734	15	0.3121
6	-0.4713	16	0.2817
7	0.8014	17	0.1932
8	2.3040	18	0.2675
9	1.7316	19	0.2283
10	1.5242	20	0.1907

表 5 PARCOR 滤波器合成语音值

n	$S(n)$	n	$S(n)$
1	-1.4337	11	0.5667
2	0.6901	12	0.4659
3	0.1358	13	0.0392
4	0.3921	14	-0.2351
5	0.3095	15	-0.0392
6	0.4656	16	-0.1758
7	0.5661	17	0.1560
8	0.2898	18	0.6818
9	0.1634	19	0.8966
10	0.7804	20	1.3110

表 6 LSP 滤波器合成语音值

n	$S(n)$	n	$S(n)$
1	-1.4430	11	0.5659
2	-0.6825	12	0.4688
3	0.1358	13	-0.0390
4	0.3903	14	-0.2344
5	0.3127	15	-0.0390
6	0.4688	16	-0.1753
7	0.5662	17	0.1566
8	0.2925	18	0.6819
9	0.1624	19	0.8986
10	0.7811	20	1.3074

表 7 元音“ウ”原来的语音值

n	$S(n)$	n	$S(n)$
1	-1.4454	11	0.5663
2	-0.6822	12	0.4687
3	0.1361	13	-0.0391
4	0.3905	14	-0.2344
5	0.3125	15	-0.0391
6	0.4687	16	-0.1755
7	0.5663	17	0.1563
8	0.2928	18	0.6822
9	0.1625	19	0.8984
10	0.7813	20	1.3090

为了语音合成, 我们再以原来 256 个数据作为输入, 使用表 2 的 PARCOR 系数和 LSP 系数, 分别计算出自适应格型滤波器和自适应 LSP 滤波器的误差(残差)信号 $e(n)$, 如表 3 和表 4 所示。表中只列出前 20 个值。

为了实现语音合成, 还应该求出所需要的音源参数, 如音调周期, 音源脉冲强度, 有声/无声判别参数等^[3,4]。音调周期和有声/无声判别参数都可以采用误差信号的自相关函数(或称变形自相关函数)求得, 具体方法见文献[4]。求得音源参数后, 利用表 3 和表 4 的误差信号以及表 2 的 PARCOR 系数和 LSP 系数, 分别按格型 PARCOR 滤波器的结构和本文给出的逐级更新算法的 LSP 结构(图 2), 根据文献[4]提供的方法, 在 486 微型机上将元音“ウ”的数值合成出来, 结果如表 5 和表 6 所示(仅给出前 20 个值)。为了比较

方便,我们从文献[4]附录8中元音“ウ”抽样量化后的数据计算出其幅值,并列于表7中(仅给出前20个值)。从表5,6和7看出,由LSP系数合成出来的数值更接近原来的语音信号值,这说明由LSP系数通过线性预测进行语音合成的结果好于用PARCOR系数合成的结果。

参 考 文 献

- [1] 王宏禹. 随机数字信号处理. 北京: 科学出版社, 1988, 368-408.
- [2] Makhoul J. IEEE Trans. on ASSP, 1977, ASSP-25(5): 423-427.
- [3] L. R. 拉宾纳, R. W. 谢弗著, 朱雪龙, 等译. 语音信号数字处理. 北京: 科学出版社, 1983, 303-347.
- [4] 安居院 猛 中嶋正之著, 周迪伟, 等译. 计算机语音处理. 北京: 国防工业出版社, 1987, 36-79.
- [5] Furui S. Digital Speech Processing, Synthesis and Recognition. New York: Marcel Dekker, 1989, 218-356.
- [6] Gibson C J, Haykin S. IEEE Trans. on ASSP, 1980, ASSP-28(6): 681-691.
- [7] Cheetham B M G. Electron. Lett., 1987, 23(2): 89-90.

A NEW ALGORITHM FOR ADAPTIVE LATTICE FILTER AND ITS APPLICATION IN THE SPEECH LINEAR PREDICTIVE SYNTHESIS

Jiang Taihui

(Wuyi University, Jiangmen, Guangdong 529020)

Abstract An adaptive line spectral pair filter is derived from an adaptive lattice filter. A least-mean-square(LMS) type adaptive algorithm used to calculate directly the line spectral pair coefficients on a stage-by-stage is proposed. Experimental results show that the algorithm has higher convergence rate and lower misadjustment as compared with other algorithms. Using LSP coefficients calculated by the algorithm, speech linear predictive synthesis work has been carried out, and better results have been obtained as compared with that using PARCOR coefficients.

Key words Line spectral pair filter, Adaptive lattice filter, Linear prediction, LMS algorithm

江太辉: 男, 1945年生, 副教授, 现从事数字信号处理和语音数字处理的研究和教学工作。