Corrections to

Digital Speech Transmission
Enhancement, Coding and Error Concealment

by

Peter Vary & Rainer Martin

Version 1.12

2.3.2 Digital All-Pole Model of the Vocal Tract

Page 22, first line: The *lattice structure* also requires four multiplications and two additions ...

Equation (2.27-d) should read

$$= \begin{pmatrix} +1 & -r_1 \\ -r_1 \cdot z^{-1} & +z^{-1} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \frac{V_L}{g_r} \tag{2.27-d}$$

Page 31: (thanks to Arne Leijon)

$$\frac{b}{\mathrm{Bark}} = 13 \arctan\left(0.76 \frac{f}{\mathrm{kHz}}\right) + 3.5 \arctan\left(\frac{f}{7.5\,\mathrm{kHz}}\right)^2 \tag{2.37}$$

$$\frac{\Delta f_c}{\mathrm{Hz}} = 25 + 75 \cdot \left[1 + 1.4\left(\frac{f}{\mathrm{kHz}}\right)^2\right]^{0.69} \tag{2.38}$$

Table 2.1:

b/Bark	f_c/Hz	Δf_c/Hz
0.5	50	100
1.5	150	100
2.5	250	100
3.5	350	100
4.5	450	110
5.5	570	120
6.5	700	140
7.5	840	150
8.5	1000	160
9.5	1170	190
10.5	1370	210
11.5	1600	240

b/Bark	f_c/Hz	Δf_c/Hz
12.5	1850	280
13.5	2150	320
14.5	2500	380
15.5	2900	450
16.5	3400	550
17.5	4000	700
18.5	4800	900
19.5	5800	1100
20.5	7000	1300
21.5	8500	1800
22.5	10500	2500
23.5	13500	3500

3.2 Fourier Transform of Discrete Signals

Page 37, after (3.4): and the sifting property of the Dirac impulse . . . (thanks to Arne Leijon)

3.4.4 z-transform Analysis of LSI-Systems

Page 46: where we have set $\widetilde{a}_0 = 1$ and $\widetilde{a}_\mu = -a_\mu$.

4.1 Spectral Analysis Using Narrowband Filters

Page 76, line over equation (4.7): decimated sequence at sampling rate f_s' *without* ...

4.1.4 Short-Term Spectral Analysis and Synthesis

Page 87, Figure 4.10-b), first block: $h_\mu^{\mathrm{BP}}(k)$ instead of $h_\mu(k)$

$$\overline{Y}_\mu\left(e^{j\Omega}\right) = \tilde{X}_\mu\left(e^{j\Omega}\right) \cdot G\left(e^{j\Omega}\right) \tag{4.20-a}$$

Bibliography

Page 115: Boite, R.; Leich, H. (1981). A New Procedure for the Design of High-order Minimum-phase FIR Digital or CCD Filters, ~~IEEE Transactions on~~ Signal Processing, vol. 3, pp. 101–108.

6.1 Vocal Tract Models and Short-Term Prediction

Page 166, Equation (6.7-b) should be

$$C(z) = \sum_{i=1}^{m} c_i \cdot z^{-i} \tag{6.7-b}$$

6.3.1 Block-Oriented Adaptation

Page 178, Figure 6.8-a): Arrow is missing on the top left line ($x(k)$)

6.3.1.3 Levinson–Durbin Algorithm

Page 185, fourth bullet point: ... the reflection coefficients k_p can be computed ...

7.2 Uniform Quantization

Page 205, Figure 7.3-a), $i = 1, 2, \ldots$, the same also holds for the caption, where: $\hat{x}_i = (2i-1)\frac{\Delta x}{2}; i = \pm 1, \pm 2, \pm 3, \ldots$

7.4 Optimal Quantization

Page 222, Description of the LMQ, second line: levels $\hat{x}_i$ ($i = 1, 2, \ldots, 2^w$).

7.6.5 Gain-Shape Vector Quantization

Page 237, below Equation (7.95): ..., the ~~minimum~~ mean square error can be calculated ...

8.3.1 First-Order DPCM

Page 246, Figure 8.4-c). Vertical axis label should be $|G(e^{j\Omega})|$

8.3.2 Open-Loop and Closed-Loop Prediction

Page 249, Figure 8.6: output signal must read $y(k)$ instead of $d(k)$

8.5.3 Analysis by Synthesis: CELP

Page 294: Table 8.3, Item a): $K(L+1)$ operations (as $H \cdot c_i$ has already been calculated in the denominator)

a)	Numerator (8.85)	$K(L+1)$
b)	Demoninator (8.85)	$K \cdot (L(L+1)/2 + L)$
c)	Division (8.85)	$K \cdot 16$
d)	Division (8.83)	$1 \cdot 16$
	Sum:	$K(L(L+5)/2 + 17) + 16$

Page 294:

$$\mathrm{CE} = \frac{K(L(L+5)/2 + 17) + 16}{L} f_s \tag{8.86}$$

8.6 Adaptive Postfiltering

$$z_{\infty i} = \rho e^{j\Omega_{\infty i}} \quad \text{with} \quad \rho = \sqrt[N_0]{\eta}\,; \quad \Omega_{\infty i} = \frac{2\pi}{N_0} i \tag{8.107}$$

$$z_{0i} = \zeta e^{j\Omega_{0i}} \quad \text{with} \quad \zeta = \sqrt[N_0]{\eta}\,; \quad \Omega_{0i} = \frac{\pi}{N_0}(2i+1) \tag{8.108}$$

Bibliography

Page 311:
Itakura, F. (1975). Line Spectral Representation of Linear Prediction Coefficients of Speech Signals, *Journal of the Acoustical Society of America*, vol. 57, no. S1, pp. S35.

9.6 Further Improvements

Page 353: “Finally it should be mentioned that the concept of soft decision source decoding opens up possibilities for iterative source-channel decoding, e.g., [Görtz 2000], [Hindelang 2000], [Adrat et al. 2002], [Perkert et al. 2001].”

Figure 9.18, “Soft decision source decoder” instead of “Softbit source decoder”

10.3.2 Spectral Envelope Estimation

Equation (10.10) should read

$$\hat{\mathbf{a}} = \arg\max_{\mathbf{a}} p(\mathbf{a}|\mathbf{b}), \tag{10.10}$$

10.3.3 Extension of the Excitation Signal

Figure 10.7: output of the block $1 - \hat{A}(z)$ should be $\hat{d}_{nb}$ instead of $\hat{u}_{nb}$

11.1 Introduction

Page 390: ... the short noise bursts around $k = 40000$ and $k = 100000$ are not removed.

11.4.4 MMSE Estimation

Page 415, Equation (11.63) should read:

$$\mathrm{E}\left\{\mathrm{Re}\{S_\mu\} \mid \mathrm{Re}\{Y_\mu\}\right\} = \frac{1}{\sigma_{N,\mu}\sigma_{S,\mu}\pi p\left(\mathrm{Re}\{Y_\mu\}\right)} \exp\left(-\frac{\mathrm{Re}\{Y_\mu\}^2}{\sigma_{N,\mu}^2}\right) \cdot \int_{-\infty}^{\infty} u \exp\left(-\frac{\sigma_{S,\mu}^2 + \sigma_{N,\mu}^2}{\sigma_{S,\mu}^2\sigma_{N,\mu}^2}u^2 + \frac{2\mathrm{Re}\{Y_\mu\}}{\sigma_{N,\mu}^2}u\right) \mathrm{d}u$$

Page 415, the first line of the equation after “the result”:

$$\mathrm{E}\left\{\mathrm{Im}\{S_\mu\} \mid \mathrm{Im}\{Y_\mu\}\right\} = \frac{1}{\sigma_{N,\mu}\sigma_{S,\mu}\pi p\left(\mathrm{Im}\{Y_\mu\}\right)} \exp\left(-\frac{\mathrm{Im}\{Y_\mu\}^2}{\sigma_{N,\mu}^2}\right) \cdot \int_{-\infty}^{\infty} u \exp\left(-\frac{\sigma_{S,\mu}^2 + \sigma_{N,\mu}^2}{\sigma_{S,\mu}^2\sigma_{N,\mu}^2} u^2 + \frac{2\mathrm{Im}\{Y_\mu\}}{\sigma_{N,\mu}^2} u\right) \mathrm{d}u$$

11.7 Computation of Likelihood Ratios

First line of equation (11.91) should read:

$$\Lambda_\mu = \frac{1-q_\mu}{q_\mu} \frac{\sigma_{N,\mu}^2}{\sigma_{N,\mu}^2 + \mathrm{E}\left\{A_\mu^2 \mid H_\mu^{(1)}\right\}} \cdot \exp\left(-\frac{R_\mu^2}{\sigma_{N,\mu}^2 + \mathrm{E}\left\{A_\mu^2 \mid H_\mu^{(1)}\right\}} + \frac{R_\mu^2}{\sigma_{N,\mu}^2}\right) \tag{11.91}$$

11.9 Dual Channel Systems

Left hand side of (11.151) should be $|\gamma_{y_1 y_2}(\mathrm{e}^{j\Omega})|$

Figure 11.21 x-axis: $\frac{f}{\mathrm{Hz}}$ instead of kHz

11.9.1.1 Implementation of the Adaptive Noise Canceller

The right hand side of (11.159) should be $\mathbf{h}(k) + \beta(k)e(k)\mathbf{y}_1(k)$

12.4.2 Filter-and-Sum Beamforming

Page 479: to make the notation consistent with later sections, we must introduce the Hermitian (conjugate) transpose already here (thanks to Arne Leijon)

$$H(e^{j\Omega}, \mathbf{u}_s) = \sum_{\ell=1}^{N_M} H_\ell(e^{j\Omega}) \exp(j\Omega f_s \Delta\tau_\ell) = \mathbf{H}^H(e^{j\Omega})\mathbf{a} \tag{12.40}$$

Page 480: to make the notation consistent with later sections, the Hermitian (conjugate) transpose also has to be used here

$$\mathbf{H}(e^{j\Omega}) = (H_1(e^{j\Omega}), H_2(e^{j\Omega}), \ldots, H_{N_M}(e^{j\Omega}))^H \tag{12.41}$$

12.5.2 Directivity Pattern

Page 482: The directivity pattern depicts the power gain (~~attenuation~~) of the sound energy from a given direction. (thanks to Arne Leijon)

13.7 The Affine Projection Algorithm

Figure 13.16, $p = 1$ instead of $p = 0$ in the figure.

13.8 Least-Squares and Recursive Least-Squares Algorithms

Page 532, second paragraph: However, a recursive approximation, the *recursive least-squares (RLS) algorithm*, avoids...