Experimental Analysis of Different Implementations for a 3-tap Comb Filter for the Prefiltering and Postfiltering Operations of the CELT codec

Daniele Giacobello

March 31, 2011

1 Introduction

Consider the form of the current 3-tap comb filter:

$$P(z) = 1 - g_p(\alpha z^{-T_p+1} + \beta z^{-T_p} + \alpha z^{-T_p-1}),$$

where T_p is the pitch lag and g_p is the pitch gain. The current constraints give the filter coefficients the current values of $\alpha = 0.26795$ and $\beta = 0.46410$ (see Appendix). In this document we analyzed and provided an objective (through PEAQ score) and subjective (through careful listening) evaluation of other possibilities for the fixed values of the filter taps.

2 Experimental Analysis

In order to obtain usable data, we analyzed a diverse database composed of 16 files (low-pitched, mid-pitched, high-pitched audio, and male and female speech) sampled at $f_s = 48$ kHz. CELT was run at 42 bytes/packet and 240 samples per packet at 48 kHz, producing a rate of 67.2 kbit/s. We modified the CELT source code by changing only the values of α and β. We then ran several experiments to evaluate several different values of these two coefficients, corresponding to different frequency behaviors. The values used are shown in Table 2. The first approach (I_1 to I_4, Figure 2) was to modify the cut-off frequency (0 dB gain) in the filter design. Since we are considering only the 48 kHz case, the cut-off frequency used were, respectively, $f_{c_1} = 20$

kHz, $f_{c_2} = 21.33$ kHz, $f_{c_3} = 22.67$ kHz, $f_{c_4} = 24$ kHz. The second approach (I_4 to I_7, Figure 2) was to define more pronounced frequency behaviors for the pitch filter. We started with a relative gain $G_r = G(f = f_s/2) - G(f = 0)$ of -6 dB (I_4) to a relative gain of -0.45 dB, with a quasi-flat frequency behavior.

The PEAQ results, shown in Table 2, suggest that the changing the values to I_4 implementation seems to be beneficial and increase the objective quality. The subjective test, through careful listening, did not reveal any significant difference.

Table 1: Different configurations for the 3-tap pitch prefilter values α and β.

METHOD	α	β	f_c	G_r [dB]
I_1	0.26795	0.464100	0.8333	-6.67
I_2	0.25800	0.484000	0.8889	-6.30
I_3	0.25190	0.496200	0.9444	-6.09
I_4	0.25000	0.50000	1.0000	-6.00
I_5	0.17500	0.60000	-	-3.74
I_6	0.10000	0.80000	-	-1.93
I_7	0.05000	0.90000	-	-0.45

Table 2: Results from all the analyzed files. The Δ_{wc} and Δ_{bc} are the worts-case and best-case difference in PEAQ score.

METHOD	PEAQ score	Δ	Δ_{wc}	Δ_{bc}
I_1	-2.6950	-	-	-
I_2	-2.6837	+0.0113	-0.0103	+0.1254
I_3	-2.6773	+0.0484	-0.0200	+0.0773
I_4	-2.6618	+0.0332	-0.0359	+0.1979
I_5	-2.6958	-0.0008	-0.0201	+0.0351
I_6	-2.7101	-0.0150	-0.1129	+0.0987
I_7	-2.7560	-0.0610	-0.1272	+0.0834

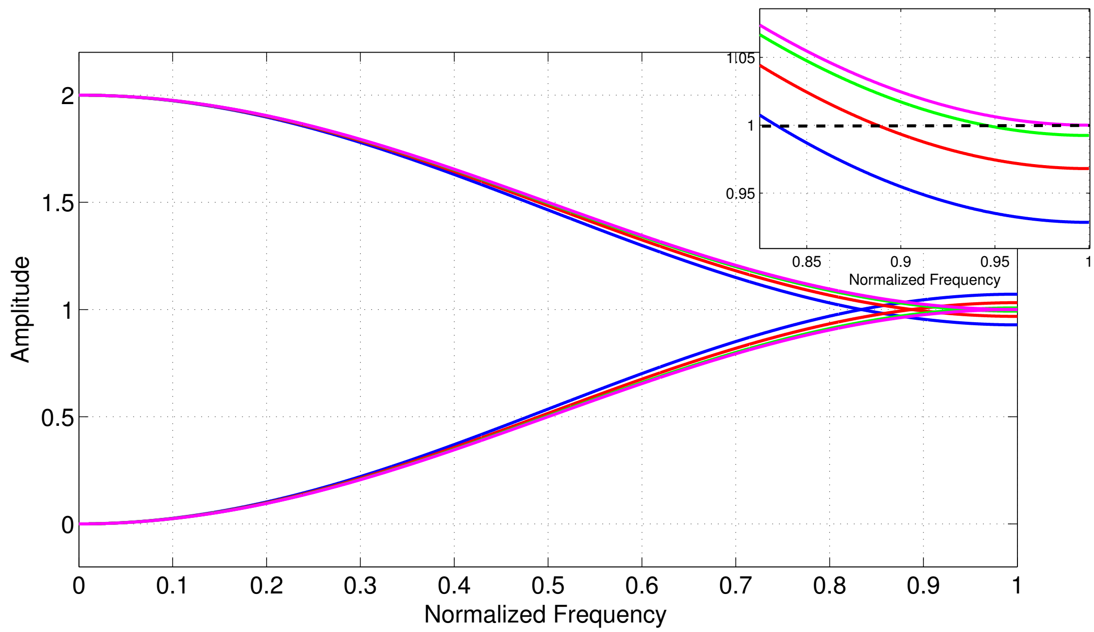

Figure 1: Spectral envelopes (linear scale) of the different filter implementation, presented in Table 2, based on changing the cut-off frequency. I_1 in blue, I_2 in red, I_3 in green, and I_4 in magenta. A detail that shows the different cutoff frequencies is shown in the smaller box.

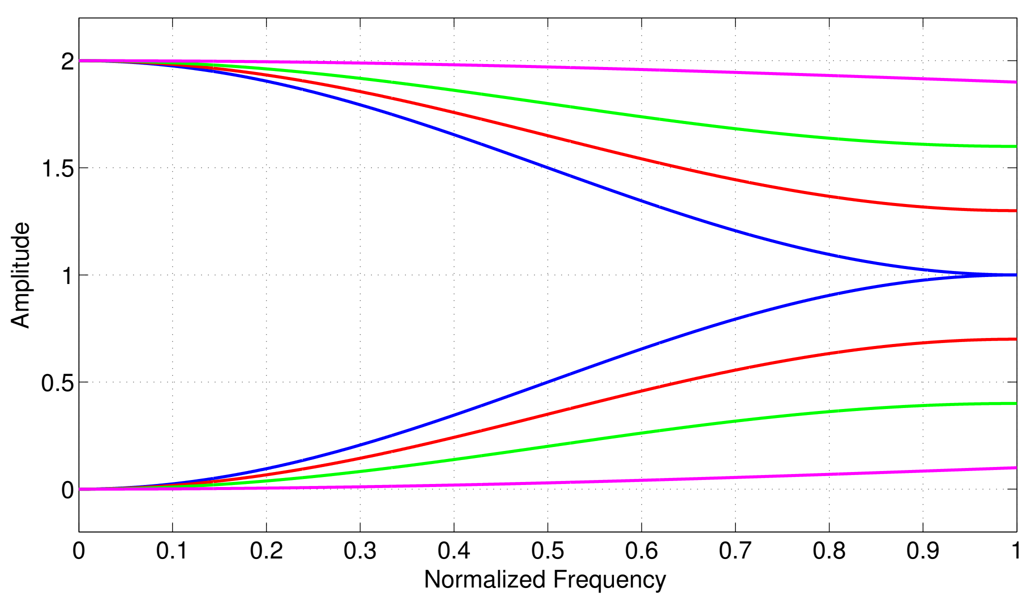

Figure 2: Spectral envelopes (linear scale) of the different filter implementation, presented in Table 2, based on changing the gain at the Nyquist frequency. I_4 in blue, I_5 in red, I_6 in green, and I_7 in magenta.

Appendix: Current Configuration of the Comb Filter

The coefficients of perceptual comb filter used for prefiltering and postfiltering are currently subject to two constraints:

- a 0 dB gain at $f_c = 20$ kHz (at $f_s = 48$ kHz),
- 1-norm of the filter coefficient vector less than one (sufficient condition for stability).

Let us consider the comb filter center around T_p:

$$P(z) = 1 + g_p(\alpha z^{-T_p+1} + \beta z^{-T_p} + \alpha z^{-T_p-1}). \tag{1}$$

The first constraint is to have a 0 dB gain at f_c ($f_c \leq f_s/2$). Let us analyze (1) as a function of ω_c:

$$\log|1 + g_p(\alpha e^{-j\omega_c(T_p+1)} + \beta e^{-j\omega_c T_p} + \alpha e^{-j\omega_c(T_p+1)})| = 0, \tag{2}$$

this can be rewritten as:

$$|1 + g_p(\alpha e^{-j\omega_c(T_p-1)} + \beta e^{-j\omega_c T_p} + \alpha e^{-j\omega_c(T_p+1)})| = 1, \tag{3}$$

factorizing:

$$|1 + g_p e^{-j\omega_c T_p}(\alpha e^{-j\omega_c} + \beta + \alpha e^{j\omega_c})| = 1. \tag{4}$$

Since we are interested only in the envelope, there is no need to compute the actual absolute value, we can simply remove the modulation term $e^{-j\omega_c T_p}$. Using the hyperbolic trigonometry identity:

$$e^{-j\omega_c} + e^{j\omega_c} = 2cos(\omega_c), \tag{5}$$

since $\omega_c \leq \pi/2$ and $\alpha, \beta \geq 0$, we can rewrite (4) as:

$$g_p\,(2\alpha\cos(\omega_c) + \beta) = 0. \tag{6}$$

The second constraint is $\|\mathbf{p}\|_1 = 1$, a sufficient condition for stability of an all-pole filter $1/P(z)$ with coefficient vector $\mathbf{p}$, In our case:

$$g_p\,(2\alpha + \beta) = 1. \tag{7}$$

We can ignore g_p in both (6) and (7) as this is calculated adaptively and does not affect the stability ($g_p < 1$).

Putting together (6) and (7) in the same system of equations $\mathbf{Ax} = \mathbf{b}$:

$$\begin{cases} 2\alpha\cos(\omega_c) = -\beta, \\ 2\alpha + \beta = 1, \end{cases} \tag{8}$$

where

$$\mathbf{x} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \mathbf{A} = \begin{bmatrix} 2\cos(\omega_c) & -1 \\ 2 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

then:

$$\begin{cases} \alpha = \dfrac{1}{2\cos(\omega_c) + 2} \\ \beta = \dfrac{2\cos(\omega_c)}{2\cos(\omega_c) + 2} \end{cases} \tag{9}$$

$$\tag{10}$$

Currently, $1 - \frac{f_c}{f_s/2} = 1/6$, which defines $\alpha = 0.2679$ and $\beta = 0.4641$.