FILTER TRANSITIONS IN ADAPTIVE IIR APPROXIMATE FILTERING

Rajiv Adhikary and S. Hamid Nawab

ECE Department, Boston University
Boston, MA 02215
email: hamid@bu.edu

ABSTRACT

The effects associated with the switching of filter orders in an incrementally adaptive IIR approximate filtering technique are evaluated in the context of speech signals. Through listening experiments, we have found that the perceptual quality of the speech at the filter output is sensitive to the initial conditions used in initiating a transition to a higher filter order. If zero initial conditions are used there is a significant crackling sound due to error bursts at points where switches to higher order filters take place. If output samples from the pre-transition filter are used as initial conditions for the post-transition filter, the amplitude of the error bursts are found to decrease significantly. An analysis is presented to account for these observations.

1. BACKGROUND

Reduction of power consumption in digital integrated circuits is a prime consideration in various digital signal processing implementations. Approximate processing as a means to achieve these goals is presented in [1] and [2]. An approximate filter structure $\mathcal{H}$ is defined as a collection of frequency selective digital filters each having a filter order N in a given range $N_{min} \leq N \leq N_{max}$. The filters in $\mathcal{H}$ possess similar frequency response characteristics and higher order filters in the filter structure have higher average stopband attenuation. A related issue is that of finding the optimal constituent element in the filter structure $\mathcal{H}$ which would guarantee a minimum tolerable SNR at the filter's output. A low-cost adaptation strategy [1] for updating the filter order is shown in Fig. 1. It involves calculation of input and output signal power estimates from the current block of L samples, which are given to the decision module D, that chooses the best constituent element to be used. Under certain assumptions, it is shown in [1] that an estimate

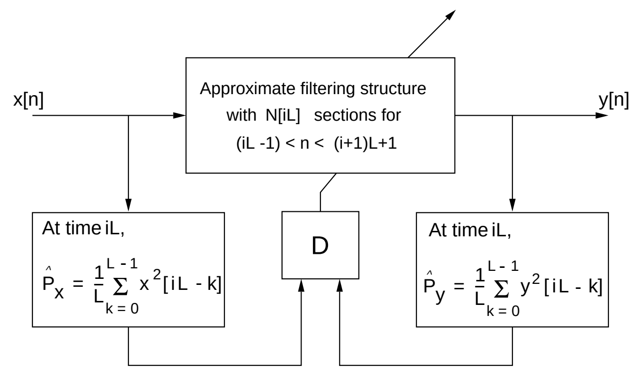

Fig. 1. Concept of approximate filtering.

of output SNR for a order N filter is given by:

$$\mathrm{O\widehat{SN}R}[N] = \left[\frac{\mathbf{y}^T\mathbf{y}}{\mathbf{x}^T\mathbf{x} - \mathbf{y}^T\mathbf{y}}\right]\left(\frac{\int_{SB}\left[1 - |H_{N_0}(\omega)|^2\right]d\omega}{\int_{SB}|H_N(\omega)|^2 d\omega}\right) - \left(\frac{\int_{SB}|H_{N_0}(\omega)|^2 d\omega}{\int_{SB}|H_N(\omega)|^2 d\omega}\right),$$

where N_0 is the filter order currently being used, $\mathbf{x}$ and $\mathbf{y}$ represent the length L vectors of the current block of input and output samples respectively. The decision module compares this to a minimum tolerable output SNR ($OSNR_{tol}$) to determine the optimal constituent filter to be used for the next block.

The strategy just described requires an expensive search over the stored values in order to obtain an optimal minimum filter order. An *incremental* approach has been suggested in [2] to get around this issue. Suppose the current block of L input samples $\mathbf{x}$ is being processed by an order N_0 filter. Based on the output block $\mathbf{y}$ of this filter and assuming the signal quality does not change very much, we need to estimate the filter order to be used to process the next block of input samples. If an N_0 order filter is again used to process the next block, the output SNR estimate is given by,

$$\widehat{\text{OSNR}}_{\text{est}} = \left[\frac{\mathbf{y}^T\mathbf{y}}{\mathbf{x}^T\mathbf{x} - \mathbf{y}^T\mathbf{y}}\right]\left(\frac{\int_{SB}\left[1 - |H_{N_0}(\omega)|^2\right]d\omega}{\int_{SB}|H_{N_0}(\omega)|^2 d\omega}\right) - 1.$$

The decision rule for updating the filter order for next block of samples is then as follows,

$$N = \begin{cases} N_0 - 1 & \text{if } \widehat{\text{OSNR}}_{\text{est}} > OSNR_{tol} + \delta \\ N_0 + 1 & \text{if } \widehat{\text{OSNR}}_{\text{est}} < OSNR_{tol} - \delta \\ N_0 & \text{else} \end{cases}$$

with the constraint that the order determined must be in the range $N_{min} \leq N \leq N_{max}$. It is to be noted that the incremental approach involves an extra application specific parameter δ.

Incremental adaptation of filter order becomes particularly suitable for truncation filter structures [2]. A truncation filter structure is defined as an approximate filter structure where the set of pole/zero pairs for each filter is constrained to be a subset of the pole/zero pairs defining any higher order filter. Truncation structures can be implemented as cascade interconnections of second-order sections. The later second-order sections can be *powered down* in order to affect a decrease in stopband attenuation and consequent power savings.

2. SPEECH EXPERIMENTS

Experiments were performed to measure the influence of initial state conditions on the error encountered when the filter order is switched.. Two speech signals sampled at $44100\ Hz$ and bandlimited to $0 \leq \omega \leq \frac{3\pi}{8}$ in the digital frequency domain were considered. The two speech samples were "*I know that I have*" and "*long as we got some time*". The second one was single sideband amplitude modulated to $\frac{5\pi}{8} \leq \omega \leq \pi$, and they were mixed to construct the input signal. These frequency bands correspond to the passband and the stopband for the Butterworth IIR truncation structure being used. Truncations of a 20^{th} order Butterworth IIR filter with half power frequency at $\frac{\pi}{2}$ were used to construct the filter structure. The ordered sequence of second-order sections used is as defined in [3]. The incremental adaptation strategy for approximate filtering discussed in section 1 was used to determine the number of second-order sections to be used for processing each block of $L = 60$ samples. The output SNR tolerance was set to $OSNR_{tol} = 2000$ and $\delta = 100$. In one case the initial conditions of added filter sections were set to zero at transition instances and in the other case the initial conditions were set equal to past outputs of the pre-transition filter. The speech quality of output signals under the two different sets of initial conditions was compared aurally. The output obtained while using zeroed out initial state had significantly more *crackling* sound present than compared to the case when past outputs were used for initial state.

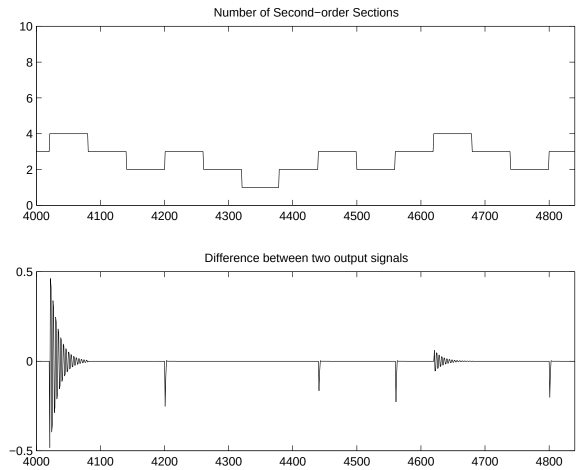

Fig. 2. Top plot shows time evolution of number of second-order sections used in the adaptive system. The second plot shows the difference between the output signals for the zero initial conditions case and the case where the outputs of the pre-transition filter are used as initial conditions. The horizontal axis in each plot represents the time index for signal samples.

The top-half of Fig. 2 shows a portion of the time evolution of the number of second-order sections used from the approximate filter structure. The lower-half shows the output of adaptive system when transitions are initiated with zeroed out states and when they are with past outputs. Since the same filters are being used to process the input signal, the outputs are expected to be the same except for the transient errors at each transition. From Fig. 2 we observe that large difference signal values occur only at transitions involving a switch to a higher number of second-order sections. In light of our listening experiments and the analysis presented in the next section, the fluctuating and rapidly decaying difference signal at such transitions may be attributed to the larger error present in the output for zero initial conditions. Also, the largest transition errors were observed when a second-order section comprising the pole pair closest to the unit circle was added, such as transition $3 \rightarrow 4$ in Fig. 2.

3. ANALYSIS

The corruptive error at transition can be modeled as an additive noise term in the adaptive filter output signal by relating the output of the adaptive filter to the output of a fixed digital filter. The additive corruption is termed *State Transition Error* (STE). Suppose N_1 order filter is being used to process the current block of L samples, and at time $n = 0$, the decision module D determines N_2 to be the order used to process the next block. The output of this filter can be viewed as,

$$y_{N_1N_2}[n] = \underbrace{y_{N_2N_2}[n]}_{\text{Desired output}} + \underbrace{y_{tr}[n]}_{\text{STE}} \quad \text{for } n \geq 0$$

where $y_{N_2N_2}[n]$ is the output of the order N_2 fixed digital filter. Quantifying the corruption due to state transition error is an important step in characterizing the performance of the filter structure being used.

3.1. STE Bounds

Added post-transition second-order sections (SOS) can be started with different sets of initial state conditions leading to differing state transition errors. It is to be noted that when a single or multiple sections are removed from the filter structure at transition instances, the STE remains at zero provided the values at state nodes of the remaining SOS are not altered.

We briefly define the notations used in our analysis below,

1. a_{is_j}— i^{th} coefficient of polynomial corresponding to the poles of the S_j^{th} SOS.
2. b_{is_j}— i^{th} coefficient of polynomial corresponding to the zeros of the S_j^{th} SOS.
3. $y_{s_is_j}[n]$ — Output of S_j^{th} SOS when pre-transition filter used S_i SOSs.
4. $w_{s_is_j}[n]$ — State information in the S_j^{th} SOS when pre-transition filter used S_i SOSs.

Assuming a transition $S_1 \rightarrow S_2 = S_1 + 1$ at time $n = 0$, two different sets of initial state conditions will be considered,

- A. Using Zero initial conditions

$$w_{S_1S_2}[-2], w_{S_1S_2}[-1] = 0, 0$$

- B. Using the past output values

$$w_{S_1S_2}[-2], w_{S_1S_2}[-1] = y_{S_1S_1}[-2], y_{S_1S_1}[-1]$$

The difference in the state of the S_2^{th} SOS on a transition $S_1 \rightarrow S_2$ as compared to the case when S_2 sections are used for all time is given by,

$$\begin{aligned} w_{tr}[n] &= w_{S_1S_2}[n] - w_{S_2S_2}[n] \\ &= \sum_{k=1}^{2} a_{kS_2}\left[w_{S_1S_2}[n-k] - w_{S_2S_2}[n-k]\right] \\ &= \sum_{k=1}^{2} a_{kS_2} w_{tr}[n-k] \quad \text{for } n \geq 0 \end{aligned}$$

with initial conditions for $w_{S_1S_2}[n]$ as defined before. Thus for $n \geq 0$, $w_{tr}[n]$ is a *Zero Input Response* running on the initial difference in states and it decays over time. The STE can be expressed as a linear combination of the values in $w_{tr}[n]$,

$$y_{tr}[n] = \sum_{k=0}^{2} b_{kS_2} w_{tr}[n-k] \quad \text{for } n \geq 0$$

Using the triangular inequality the magnitude of STE can be shown to have a bound as follows,

$$|y_{tr}[n]| \leq B_I \sum_{k=0}^{2} |b_{ks_2}| B_w[n-k] \quad \text{for } n \geq 0$$

where

1. $B_w[n]$ is a positive number dependent entirely on the filter coefficients used,

$$B_w[n] = \left[\left| \frac{a_{1s_2} + a_{2s_2} z_1^{-1}}{1 - z_2 z_1^{-1}} z_1^n + \frac{a_{1s_2} + a_{2s_2} z_2^{-1}}{1 - z_1 z_2^{-1}} z_2^n \right| + \left| \frac{a_{2s_2}}{1 - z_2 z_1^{-1}} z_1^n + \frac{a_{2s_2}}{1 - z_1 z_2^{-1}} z_2^n \right| \right]$$

for $n \geq 0$ and equal to 1 for $n = -2, -1$. The z_i are the pole locations of S_2^{th} SOS.

2. B_I represents the effect of initial conditions at transition and includes the bound B_X on the input signal $x[n]$.

   - Using Zero initial conditions.

$$B_I = B_X B_{S_1} B_Z$$

   - Using past output values.

$$B_I = B_X B_{S_1}(B_Z - 1)$$

where

$$\begin{aligned} B_{S_1} &= \sum_{m=0}^{\infty} |h_{s_1}[m]| \\ B_Z &= \sum_{m=0}^{\infty} |h_Z[m]| \end{aligned}$$

$h_{s_1}[n]$ is the impulse response of S_1 SOSs, and

$$h_Z[n] \leftrightarrow \frac{1}{1 - \sum_{k=1}^{2} a_{kS_2} z^{-k}}$$

are z-transform pairs.

Table 1. Maximum of STE bounds for Incremental Butterworth Truncation Filter Structure

Transition	*Zero initial condition*		*Using Past values*	
$S_1 \to S_2$	Theoretical	Simulation	Theoretical	Simulation
$1 \to 2$	0.9726	0.7535	0.0385	0.0305
$2 \to 3$	0.8884	0.7216	0.0124	0.0098
$3 \to 4$	12.8262	0.9647	10.9600	0.7716
$4 \to 5$	1.8564	0.8409	0.0029	0.0014
$5 \to 6$	6.6246	1.0010	4.1170	0.6077
$6 \to 7$	2.9846	0.9943	0.4062	0.1376
$7 \to 8$	5.2274	1.0725	2.3338	0.4663
$8 \to 9$	4.7908	1.1319	1.5025	0.3858
$9 \to 10$	4.0430	1.1851	0.8592	0.2559

The bound on STE can be numerically evaluated based on the filter coefficients used and the input signal bound for different n, the maximum of these values gives an indication of the maximum value STE can assume. Table 1 shows the theoretical bounds for Butterworth truncation filter structure under the two *initial conditions*. An input signal bound of $B_X = 1$ was used. The bounds are better if pre-transition output values are used as starting state for the SOS added and we expect the filter structure to give better overall performance for this case. Moreover the factor $B_Z(\sum_{k=0}^{2} |b_{k S_2}| B_W[n-k]|)$ in the bound for $|y_{tr}[n]|$ depends only on the SOS added and increases as the pole pair of the SOS added moves closer to the unit circle. We also expect the error to grow if the SOS added has poles closer to the unit circle.

4. SIMULATIONS

In order to gain more insight into the nature of the bound, Matlab based simulations were performed. A random input signal bounded by $B_X = 1$ with independent and identically distributed samples according to a uniform probability distribution is fed to a IIR Butterworth truncation filter structure with pre-transition order S_1. At time $n = 0$ a state transition $S_1 \to S_2$ occurs. The maximum value of the STE is measured for all pairs of S_1,S_2 and for the two different initial state conditions. The procedure was repeated for $10,000$ such random input signals. Table 1 lists the simulation results. Even though the simulation bounds are lower than the theoretical bounds derived using the triangular inequality, they are in proportion with the theoretical results.

5. CONCLUSION

In this paper we examined the effects of incrementally switching the filter order during adaptive IIR approximate filtering. We found that every switch to a higher filter order causes a transient error burst as the new filter is launched. The magnitude of the error burst was found to be highly dependent upon the initial condition used to launch the new second-order section of the higher order filter. The error bursts were perceptually significant in speech signals when zero initial conditions were used. The error burst was found to diminish significantly if past outputs from the lower order filter were used as initial conditions. Accompanying analysis was used to explain these experimantal observations.

6. REFERENCES

[1] J. T. Ludwig, “Low Power Digital Filtering Using Adaptive Approximate Processing,” *Doctoral dissertation*, Massachusetts Institute of Technology, Cambridge, Sept 1997.

[2] J. T. Ludwig, S. Hamid Nawab and A. P. Chandrakasan, “Low-Power Digital Filtering Using Approximate Processing,” *IEEE Journal of Solid State Circuits*, vol 31, no. 3, March 1996

[3] J. T. Ludwig, S. H. Nawab, A. Chandrakasan, “Convergence Results on Adaptive Approximate Filtering,” *Proc. SPIE Int. Sym. on Advanced Signal Processing Algorithms, Architectures, and Implementations,* vol. 2846, Denver, CO, Aug 1996