Dynamic non-uniform filter bank constructing algorithms for reconfigurable speech processing system based on the FPGA-device and TMS320C31

PhD-student Krzysztof Bielawski ¹ PhD-student Alexey Petrovsky ²

¹ Department of Computer Science, Bialystok University of Technology
Wiejska 45A, 15-351 Bialystok, Poland, e-mail: biel@ii.pb.bialystok.pl
² Department of Computer Engineering, Belarusian State University of Informatics and Radioelectronics
6, P. Brovky Str., 220013, Minsk, Republic of Belarus, e-mail: palex@it.org.by

Abstract
Programmable logic, in the form of Field Programmable Gate Arrays (FPGA’s), offers an attractive solution for the computational-intensive functions found in digital signal processing applications. By computing the flexibility of a general-purpose, programmable digital signal processor with the speed and density of a custom hardware implementation, FPGAs can provide increase DSP system performance and reduce system cost [1].

Traditional digital signal processing task have been implemented using programmable processor or custom circuit. In case of well known and standard design filter bank the DSP processor deliver the good and stable solution. But need for new class of reconfigurable and nonuniform frequency spaced filter bank combined with additional highly complex processing in bands increase the need for expensive, very fast DSP processors with a lot of memory space. In this paper we present the model, design technique, examples and application of the reconfigurable nonuniformly frequency spaced filter bank build based on polyphase decomposition and all-pass transformation technique. Analysis of the polyphase structure reveals that we have achieved a significant computational reduction in exchange for an increase in warping algorithm complexity and control.

Introduction
The m-band modulated filter and wavelet bank are becoming popular in many signal and image processing applications [2]. Consider the block transform audio coder. The input blocks are N-component vector \( \mathbf{r}(n) \) which is the output at the \( k \) th branch of the serial to parallel converter (S/P): \( x(nM-k) \), \( k = 1, N \), so the input block is selected-clockwise: \( 0, M-1, \ldots, 1, 0, M-1, \ldots \). The output rate of an S/P converters is \( 1/M \) of the input rate. Each block is transformed to length \( M \) by \( y(n) = P^T \mathbf{r}(n) \). Its equivalent to uniform analysis filter bank \( H(z) \) with \( M \) channels. The subband signal \( y(n) \) is quantized, entropy coded, and stored.

In synthesis, the stored signal is decoded (to produce \( \tilde{y}(n) \)) and then transformed by matrix \( P \). This operation is equivalent to uniform synthesis filter bank \( \tilde{F}(z) \) with \( M \) channels. The synthesis filters are time-reversed versions of the analysis filters. The reconstructed signal \( \tilde{x}(n) \) is a linear combination of the basis vectors. Thus, the audio coder contains three parts: analysis and synthesis filter banks with \( M \) channels, and processing module.

The analysis and synthesis \( H_k(z) \) and \( F_k(z) \) filters in the DFT filter bank version[1] are uniformly shifted versions \( (z \rightarrow zW^k) \) of the lowpass filters. When the polyphase components of \( F_0(z) \) and \( H_0(z) \) are related by \( R_k(z) = \prod \{E_i(z) \} \), the overall distortion is \( T_0(z) = z^{-(M-1)} \prod \{E_i(z^M) \} \).

The output should be equalized.

A popular method to design \( M \) channel filter bank is to cascade smaller systems. A two-level wavelet packets use two-channel systems. An example, the six-channel filter bank is obtained from cascading a two-channel system with two three-channel systems. Typical relations between filter bank \( H_k(z) \) and wavelet packet \( H_k(z) \) and \( G_k(z) \) are

\[
H_0'(z) = H_0(z) \cdot G_0(z^2) \\
H_0''(z) = H_0(z) \cdot G_0(z^4)
\]

The six-channel filter bank is perfect reconstruction (PR) if and only if the two-channel and three-channel filter banks are PR[1]. One obtains nonuniform filter banks by cascading systems with different decimation factors \( D_k \). A nonuniform bank with decimation factors \( (6, 6, 6, 4, 4) \) uses a three-channel and a two-channel system at the second tree level. Thus, the nonuniform filter bank is the filter bank where the input signal is channelized by a set of filters with unequal bandwidths, and downsampled.

The non-uniformly frequency spaced reconfigurable filter bank is an important function in speech processing. In this article, we have discussed
the design technique and application of filter bank
with non-linear frequency warped, in particular, bark
scaled frequency bands based on the polyphase filter
with all-pass transform, and wavelet packet
decomposition. For large filter length, subband
processing is superior to full-band processing in terms
of computational complexity. In designing the filter
bank, we must find a compromise between the
prototype filter length, subsampling rate, stop-band
attenuation and aliasing.

FPGA-based digital signal processing provides a
well-balanced compromise between the flexibility of a
programmable processor and the performance of a
custom DSP system. However, effective FPGA-based
DSP implementations must be tailored to the
architecture of the FPGA device, using techniques
such as distributed arithmetic algorithms.

The proposed technique optimize system
performance, and, in addition, provides a convenient
framework within which ongoing research in the areas
of non-uniform polyphase filter bank applied to
speech enhancement and speech/audio coding
algorithms and reconfigurable architectures can be
synergistically combined to enable the design of
reconfigurable high-performance DSP systems.

Non-uniform frequency spaced filter
bank
It is obvious that filter bank with uniform
frequency resolution can be implemented very
efficiently by using polyphase network for modulated
bandpass filters, which have been derived from a
common prototype lowpass and the Discrete Fourier
Transformation (DFT). If the magnitude response
of the prototype lowpass FIR filter is denoted as

\[ H(e^{j\omega}) \]

then the frequency response of this
uniform bank are given by equation:

\[ H_m(e^{j\omega}) = H(e^{j(\omega - \frac{2\pi m}{N})}) , \]

where \( H_m(e^{j\omega}) \) is \( m \)-th filter of the DFT filter bank
has centre frequencies

\[ \omega_m = \frac{2\pi}{M} m , \quad m = 0,\ldots,M-1 , \]

where \( M \) is a number of bands.

Proposed modification of this uniform filter bank,
leads to the non-uniformly spaced frequency filter
bank using allpass filter instead of delay elements at
signal input. In proposed reconfigurable filter bank [3]
the allpass transformation of first order is used but to
exploit elasticity of the solution the higher order all-
pass is shown. The delay elements are substituted by
the casual and stable allpass,

\[ A(z) = \frac{z^{-1} - a}{1 - \frac{a}{z}} \]

where \( a = ae^{i\alpha} \) is a complex parameter with
\( |a| < 1 \), and phase of this allpass

\[ F(\omega) = -\omega + 2\arctan \frac{a\sin(\omega - \alpha)}{a\cos(\omega - \alpha) - 1} . \]

This leads to the following bandpass filters

\[ H_m(e^{j\omega}) = H(e^{j(-F(\omega) - \frac{2\pi m}{N})}) . \]

Solving this equation adequately to the smooth and
monotonic warping (mapping) function \( F(\omega) \) we
get desired frequency warping transformation. So we
can map the uniform frequency resolution to the new
representation of the uniform frequency resolution on
the new scale (e.g. Log, Bark, ERB-scale) which has a
non-uniform resolution when observed from old scale
(Hz). Figure 1 presented some warping characteristic
where used first order all-pass with real coefficient.
The prototype filter is pictured in figure 2.

![Fig. 1. Frequency warping characteristic of the first order all-pass section for different values of real value warping parameter](image)

![Fig. 2. Low pass filter prototype characteristic](image)
pass filter prototype (128 coefficients, Hamming window) shown in fig. 2.

For making allpass transformation useful to create the cochlea spaced filter bank we decide to use the real parameter $a = a$, to make a compromise between the complexity introduce by the allpasses and perfect approximation of the human inner ear model. Concerning cochlear model and idea of critical bands we created the frequency resolution of the human ear in band of telephone speech (300-3400Hz) for 8kHz sampling frequency. The 3dB bandwidth characteristics of cochlear filter banks is illustrated in plot of fig. 6.

Knowing limitation of the assumed approach (real value of the parameter $\alpha$), we must decide which frequencies (lower or upper) of the model will be approximated better by the function of the all-pass transformation. We decide that lower frequencies resolution must be more exact, so parameter $\alpha = 0.4034$ is chosen (note that for different sampling frequencies results different value of $\alpha$ approximated by the following analytic expression

$$\alpha = 1.0211 \left( \frac{2}{\pi} \arctan(0.076 f_s) \right)^{1/2} - 0.19877.$$ 

The solid line in fig. 7 represents the approximation of the cochlea resolution frequency in this case.

The advantage of that assumption is based on perception property of the human ear, which is able to separate low frequencies more precisely than high frequencies. The interest in low frequencies are also motivated by the acoustical environment properties in moving vehicle which will be shown in application example, where the noise field is highly energetic in low frequencies and most noise power is concentrated below 500Hz. It is related to the road nature, so in case of moving car can not be predicted by statistic tool. Please notice that in low frequencies human auditory system has high resolution (100 Hz band width) in terms of the critical bands. So evaluating noise reduction technique in close manner to the ear can be worth the effort.

Fig. 3. Warped filter bank characteristic for $a=0.5$

Fig. 4. Warped filter bank characteristic for $a=-0.9$

Fig. 5 shows the logarithmic Fourier transformations of the impulse responses of the two filter banks (analysis plus synthesis). The frequency response of the allpass-transformed polyphase filter bank shows a lifting of nearly 0.4 dB.

Fig. 6. Logarithmic Fourier transformation of the impulse response of the analysis-synthesis structure for bank from fig. 2.

**Cochlear spaced filter bank**

Since the human perception of audio signals works in logarithmic scale [4], it is appropriate to analyze the signal by a non-uniformly (cochlear) spaced filter bank. This way of signal processing seems to be a perfect solution for speech enhancement system based on psychoacoustical properties of the human ear.
For an example let's consider a 256 taps FIR filter as prototype to projected cochlear spaced filter band with DFT length of 32, the we get the filter bank presented in Fig. 8 for selected even numbered bands (16 barks). In order to exploit the enhanced flexibility of the all-pass transformation the complex and higher order all-pass can be used [5] and the design are not more complicated. However as it is visible in example designer must specify the usefulness of the achieve filter bank.

**Dynamic wavelet packet decomposition**

Optimum time-frequency decomposition are very useful in audio coding applications, because the signal energy can be maximally concentrated even for the wide variety of audio signal characteristics. Moreover, this signal representation is particularly well suited for a perceptual weighting of the quantization noise. The well known tree structure of cascaded 2-channel filter banks allows a very flexible optimization, leading to a signal adaptive, dynamic wavelet packet decomposition. As initial basis we use the wavelet packet transform from [2] with a frequency resolution approximating the critical bandwidths of the human ear for a sampling frequency 44.1 kHz. This basis has the highest possible time resolution within the given frequency bands.

Assume that for one frame there exist $2^M$ time slots in a particular band. If we allow any finer frequency resolution up to the possible maximum, which depends on the actual frame length, i.e. if we build any possible branching of the subband coding tree, we achieve $C(M)$ possible bases for this band $C(n)$ id defined by the nonlinear, recursive equation [6]

$$C(n) = C^2(n - 1) + 1, C(0) = 0.$$ 

Since $C(n)$ rapidly increases with n, an extremely high bit rate would be necessary to transmit the chosen basis. The ensemble of all possible transform bases builds our wavelet packet library. So, if we permit a variable bit rate, our aim is to find the optimum basis for a decomposition of the input signal vector that minimizes the bit rate under the constraint of an imperceptible quantization noise.

In the fig. 9 and fig. 10 are shown the tree of the wavelet packet transform and its magnitude response approximating the critical bandwidths of the human ear respectively.
Computational load

The computational complexity of the $M$-subband system will be measured by the number of real multiplications per input sample. In DSP applications such as FIR filtering and FFTs, additions are usually performed in conjunction with multiplications using MAC instruction (number of multiply-accumulate instructions), which is vital to many DSP applications [1].

Polyphase filter. Computational complexity for the $M$-channel polyphase filter system is the complexity for analysis and synthesis filtering. An example, the polyphase uniform DFT filter bank: the prototype lowpass filter is length $L$, downsampled factor is $D$, and $M/D$ is an integer. Then the complexity for analysis and synthesis filtering $C_{asf}$ is computed as follows. There are a total of $M$ polyphase filters, each of length $L/D$ operating at a downsampled rate of $1/D$ in the filter bank, thus requiring $(LM/D^2)$ real multiplications per input sample. This operation is performed two times: for the analysis filtering and for the synthesis filtering. The $M$-point DFT and IDFT are implemented (assuming $M$ is a power of 2) with radix-2 FFT. For real data, the $M$-point IDFT can be realized with an $M/2$-point FFT and $M/2$ complex multiplications. This relates to $M \log_2(M/2)$ real multiplications for the analysis filtering. A similar realization holds for the synthesis filtering. Thus the total number of real multiplications (real MAC operations) for subband filtering per input sample is

$$C_{asf} = 2M(L/D^2 + \log_2(M/2)),$$

which shows that the polyphase form of the filter bank is more effective with point of view computational complexity and the computational complexity for filter-decimator with small rectangular properties is defined as the computational complexity of DFT or FFT.

Wavelet packet decomposition. For nonuniform filter bank, an example the $n$-level DWT tree, the first level is computed in each input sample period, the second level is calculated one time per two input sample period, and $n$th level – one time per $2^{n-1}$ input sample period. So, the following schedule algorithm can be offered:

if the second level of the $n$-level DWT tree is computed on each $2k$th input sample period, the third level – in each $(2^2k+2)$th input sample period, and $n$th – in each $(2^{n-1}k-2^{n-2})$th input sample period then only two levels will be computed during one input sample period: the first level and a some $i$th level. Here $k$ is 1,2,...

Thus the computational complexity of this schedule algorithm is constant per input sample period and independent from number of the DWT tree levels, and also is determined by the sum of the first level computational complexity and of a some $i$th level maximum computational complexity per input samples.  

Polyphase filter bank with the all-pass transformed. The computation load of the all-pass transformed polyphase filter bank corresponds to the usual polyphase filter bank plus calculation in the all-passes. The all-pass operation must be carried out in the original sampling rate which highly increase the computation load. Assuming $M$-as a band number, $L$-prototype filter length, $R$-down-sampling factor we get the 

$$2(L-1) + \frac{L}{R} + 16\log_2 M$$

real operation per sample in analysis part and the same number operation in synthesis part, which comparing to the uniform polyphase filter bank load is about 16 time larger. But its acceptable for non-uniform filter bank with the flexible design.

Reconfigurable signal processing system

Often a complete system solution have been proposed as well DSP-specific systems. The techniques specific to DSP system include computational complexity reduction (at the algorithmic level) and pipelining and parallel processing (at the architecture level). All of these techniques are applied during the VLSI design phase and their implementation is time variant. Thus, this class methods are referred to as static techniques. Recently, dynamic techniques both at the circuit level and algorithmic level have been proposed. These techniques are based on principle that the input is usually nonstationary, and hence, it is better to adapt the algorithm and architecture to the input. Such systems are referred to as reconfigurable signal processing systems [1]. Hybrid architecture based on FPGA’s and general-purpose DSP’s is the topic of our research.
We have proposed dynamic algorithm transform (DAT) for the design of a joint echo cancellation/noise reduction systems and audio signals coding based on the non-uniform filter bank. From an implementation perspective, a DAT-based reconfigurable DSP system has the filter bank algorithm (analysis + synthesis) implemented in a reconfigurable hardware (see fig. 11) (such as FPGA plus distributed arithmetic [7]), whereas the input state and state transitions are monitored and in subband processed by control signals monitoring (in our case it’s warping parameter $a$) and in subband processing (different algorithms based on the adaptive filtering or psychoacoustic spectral subtraction etc.) algorithms respectively are mapped in a general DSP processor.

![Diagram](image)

Fig. 11. DAT-based reconfigurable signal processing system

The digital part’s of the reconfigurable signal processing system basic structure is presented in fig. 12. The host part of the given system includes DSP processor (TMS320C31), data RAM (32K), ROM of loadable programs for DSP inside memory and FPGA1, and also clock-time generator is included. A digital filters bank is realized on the FPGA1 (FPGA is SRAM based, and distributed arithmetic is used [7]) two buffers memory (for real-time processing without missing of input data). The PLD serves as an interface between the embedded DSP system and the outside world.

The acoustic echo and noise control device based on the idea [3] can be realised using the given system. Concerning single microphone system gives echo attenuation at 30 dB. A two microphone system gives better results that pervious one. It has echo attenuation at about 40 dB. The perceptual audio coding based on the flexible tree-structured time-varying wavelet packet decomposition with 4 times and 16 times compression were realised on this system, too.

The combination of DSP and FPGA reconfigurable systems offers several advantages over competing alternatives:
- faster and smaller than general purpose hardware solutions;
- lower development cost than dedicated hardware solutions;
- dynamic reconfigurable supports multiple algorithms within a single application;
- multi-purpose architecture generates volume demand for a single hardware design.

**Conclusion**

The combination of DSP and FPGA reconfigurable systems offers several advantages over competing alternatives: faster and smaller than general purpose hardware solutions; lower development cost than dedicated hardware solutions; dynamic reconfigurable supports multiple algorithms within a single application; multi-purpose architecture generates volume demand for a single hardware design.

The proposed techniques optimize system performance, and, in addition, provides a convenient framework within which ongoing research in the areas of non-uniform filter bank applied to speech enhancement and speech/audio coding algorithms and reconfigurable architectures can be synergistically combined to enable the design of reconfigurable high-performance DSP systems.

**References**


This work was supported by Bialystok Univ. Of Tech., the grant W/II/2/00
Fig. 12. The block diagram of the reconfigurable signal processing system.