Filter Banks I

Prof. Dr. Gerald Schuller

Fraunhofer IDMT & Ilmenau Technical University
Ilmenau, Germany
Structure of perceptual Audio Coders

Encoder

Audio Input → Analysis-Filterbank → Quantisation & Coding → Bitstream-Multiplexer → Bitstream Output

Perceptual Model

Decoder

Bitstream Input → Bitstream-Demultiplexer → Inverse Quantisation → Synthesis-Filterbank → Audio Output
Filter Banks

- essential element of most audio coders
- transform from time to frequency domain and vice versa

- Goal:
  - Good filter bank
  - Compress audio signals

- Approach:
  - Redundancy Reduction
  - Irrelevance Reduction
Critically sampled Analysis and Synthesis Filter Bank

Example: 44,1 kHz sampling

Lowest frequency

44,1 Hz sampling

44,1 kHz places audio components at right frequencies

Example: 10,000 samples

Convolution

N=1000

10,000 samples

no increase in data/sample rate!

10,000,000 samples

use Nyquist

44,1 kHz

Prof. Dr.-Ing. K. Brandenburg, bdg@idmt.fraunhofer.de Prof. Dr.-Ing. G. Schuller, shl@idmt.fraunhofer.de

Page 4
Down-Sampling

• The operation of “down-sampling” by factor N describes the process of keeping every Nth sample discarding the rest.

\[ x[m] \rightarrow \downarrow N \rightarrow y[n] \]
Up-Sampling

- The operation of “up-sampling” by factor N describes the insertion of N-1 zeros between every sample of the input.

\[ x[m] \xrightarrow{\uparrow N} y[n] \]
Filter Bank Structure - The Analysis Filter Bank

Example:
- N=1024 filters
- $f_s=44100\text{Hz}$ sampling frequency
- $f_g=22050\text{Hz}$ Nyquist frequency

\[
\frac{f_g}{N} = \frac{f_s / 2}{N} = 21.5\text{Hz}
\]

If we have N filters, and no down-samplers, then we would have $N \times f_s$ samples per second after filtering – more than input!
- hence down-samplers.
- with down-samplers: number of samples stays constant.
- “critical sampling”
Filter Bank Structure - The Synthesis Filter Bank

- Up-sample each subband by N to restore original sampling rate
- Apply passband filter to each subband signal
- Add each subband signal to generate output signal
Definition: Perfect Reconstruction

• The property of the output signal out of cascaded analysis and synthesis filter bank being identical to the input signal (except for a time shift $n_d$) is called “Perfect Reconstruction”:

$$output = x(n - n_d)$$

• A filter bank having this property is called a “Perfect Reconstruction Filter Bank”.

Prof. Dr.-Ing. K. Brandenburg, bdg@idmt.fraunhofer.de  Prof. Dr.-Ing. G. Schuller, shl@idmt.fraunhofer.de
Filter Bank Structure – Perfect Reconstruction

Thought experiment: ideal `brick wall filters` 

Brick wall: magnitude in passband is one, otherwise zero

Nyquist Theorem: we can down-sample the subband signals by factor N without loss of information

With suitable brick wall synthesis filters, perfect reconstruction (input = output) could be achieved
Bandpass Nyquist

Goal:

• Keep critical downsampling (downsampling rate is \( N \), equal to number of subbands).

• \( \rightarrow \) No increase in number of samples.

• Still want to obtain perfect reconstruction!

• \( \rightarrow \) Ideally aliasing cancels!
Example Bandpass Signal

Bandpass Nyquist: sampling with twice the Bandwidth ($f_b$)!
After Downsampling and Upsampling:

Reconstruction: apply ideal bandpass filter for original frequency range ("fish out" original).

Problem: ideal bandpass filters are not realizable!
Basic Principle: z-Transformation (1)

• Goal: Realizable FB with critical sampling and perfect reconstruction (PR)

• Problems with ideal filter banks:
  – Brick wall filters not realizable (infinite delay!)

• Approach:
  – Find a suitable mathematical description for realizable Perfect Reconstruction Filter banks
Basic Principle: z-Transformation (2)

• For a sampled causal function

\[ u(t) = \sum_{n=0}^{\infty} c(n) \delta(t - nt_0) \]

• The Laplace-transformation is

\[ C(p) = \sum_{n=0}^{\infty} c(n) e^{-np_t_0} \]

• By substitution \( z = e^{p_t_0} \), it is called z-transformation

\[ C(z) = \sum_{n=0}^{\infty} c(n) z^{-n} \]
Basic Principle: z-Transformation (3)

- Turns a convolution $a(n) * b(n)$ into a multiplication in the z-domain:

$$A(z) \cdot B(z)$$

- Important: multiplication is invertible, unlike convolution.
Relationship between Time Domain ↔ z-Domain

Convolution:
\[ a(n) * b(n) = \sum_{m} a(n - m)b(m) \]

Transformation into z-Domain:
\[ A(z) = \sum_{n} a(n) \cdot z^{-n} \]
\[ B(z) = \sum_{n} b(n) \cdot z^{-n} \]
\[ A(z) \cdot B(z) = \sum_{n} z^{-n} \sum_{m} a(n - m)b(m) \]

→ Convolution corresponds to polynomial multiplication.
Polyphase Description (1)

Output of k-th filter of length L, no down-sampling:

\[ y_k(n) = x(n) \ast h_k(n) = \sum_{i=0}^{L-1} x(n-i) \cdot h_k(i) \]

Symbol * means convolution

- \( y_k(n) \) is the \( k^{th} \) sub-band signal
- \( x(n) \) is our input signal (in time domain)
- \( h_k(n) \) is the \( k^{th} \) filter
- \( L \) is the filter length
Polyphase Description (2)

Signal after down-sampling \(( n \Rightarrow mN)\)

\[
y_k(mN) = \sum_{i=0}^{L-1} x(mN - i) \cdot h_k(i)
\]

Simplify the indexing for \(y\):

\[
y_k(m) = \sum_{i=0}^{L-1} x(mN - i) \cdot h_k(i) \quad \text{(subband signal)}
\]

Problem: Because of the down-sampling (term \(mN\)) sum is no longer a simple convolution.

Approach: treat \(m\) as simple index for larger units, the “blocks”. Also rewrite index \(i\) in terms of blocks.
Approach: View Signals in „Blocks”, „Blocking”

Rewrite index \( i \) in terms of a block index \( l \) and an index inside the block, \( n' \), to obtain:

\[
i = lN + n'
\]

Using this notation we obtain a double sum.

Index inside a block: \( n' \), \( 0 \leq n' \leq N - 1 \)
Polyphase Description (3)

Further write the filter length as: $L = KN$

With the block index $l$ and the inside block index $n'$ we obtain the double sum:

$$y_k(m) = \sum_{l=0}^{K-1} \sum_{n'=0}^{N-1} x(mN - lN - n') \cdot h_k(lN + n')$$

Observe: the inner sum can be reformulated as vector multiplication with the following vectors:

$$x(m) = [x(mN), x(mN - 1), ..., x(mN - N + 1)]$$

$$h_k(m) = [h_k(mN), h_k(mN + 1), ..., h_k(mN + N - 1)]^T$$

(In z-Domain: $H_k(z) = \sum_{m} h_k(m)z^{-m}$)
Polyphase Description (4)

The inner sum simplifies to a vector product:

\[
\sum_{n'=0}^{N-1} x(mN - lN - n') h_k (lN + n') = x(m - l) \cdot h_k (l)
\]

The convolution and down sampling then becomes a simple convolution again:

\[
y_k (m) = \sum_{l=0}^{K-1} x(m - l) \cdot h_k (l)
\]
Polyphase Description (5)

Or, using the convolution notation:

\[ y_k(m) = x(m) * h_k(m) \]

Now use the z-transform:

\[ Y_k(z) = X(z) \cdot H_k(z) \]

Arrange the N impulse response vectors \( H_k(z) \) of length N into a \( N \times N \) square matrix (can be invertible!):

\[
H(z) = [H_0(z), H_1(z), \cdots, H_{N-1}(z)]^{\text{subbands}}
\]

\[
Y(z) = [Y_0(z), Y_1(z), \cdots, Y_{N-1}(z)]
\]
Polyphase Description, Analysis (1)

• Hence the form of the polyphase matrix for analysis is (Type 1 polyphase):

\[
H(z) = \begin{bmatrix}
H_{0,0}(z) & H_{0,1}(z) & \cdots \\
H_{1,0}(z) & \ddots & \\
\vdots & & \ddots \\
H_{N-1,0}(z) & & & H_{N-1,N-1}(z)
\end{bmatrix}
\]

• and each subband filter can hence be written as

\[
H_k(z) = \sum_{n=0}^{N-1} z^{-n} H_{n,k}(z^N)
\]
Polyphase Description, Analysis (2)

Block diagram:

Mathematically very simple operation for entire filter bank including down sampling

Final equation of analysis filter bank:

\[ Y(z) = X(z) \times H(z) \]

- \( X(z) \): Vector of polynomials, contains input samples
- \( H(z) \): Analysis Polyphase Matrix, \( NxN \)

Prof. Dr.-Ing. K. Brandenburg, bdg@idmt.fraunhofer.de Prof. Dr.-Ing. G. Schuller, shl@idmt.fraunhofer.de
Polyphase Description, Synthesis (1)

• The polyphase matrix for synthesis is (Type 2 polyphase):

\[
G(z) = \begin{bmatrix}
G_{0,0}(z) & G_{0,1}(z) & \cdots \\
G_{1,0}(z) & \ddots & \\
\vdots & & \ddots \\
G_{N-1,0}(z) & & & G_{N-1,N-1}(z)
\end{bmatrix}
\]

• Now each filter has its polyphase components along the rows, and each subband filter can be written as

\[
G_k(z) = \sum_{n=0}^{N-1} z^{-n} G_{k,n}(z^N)
\]
Polyphase Description, Synthesis (2)

Perfect reconstruction (PR) results if

Since by substitution we get

Observe: PR requires ‘only’ a matrix inversion

Problem: How to invert a matrix of polynomials?
Example: Construction of $H(z)$

- How to obtain a $H(z)$, which has only real numbers?
  - design $h_k(n)$ such that $H_k(z)$ is real valued

\[
h_k(mN+n) \quad H_k(z)
\]

\[
1, 2, x \cdots \rightarrow 1 + 2 \cdot z^{-1} + x \cdot z^{-2} + \cdots
\]

\[
1, 0, 0, \cdots \rightarrow 1 + 0 \cdot z^{-1} + 0 \cdots = 1
\]

\[
a, 0, 0, \cdots \rightarrow a
\]

\[
m=0, 1, 2, \ldots
\]
Examples: The DFT as a filter bank

The Discrete Fourier Transform is defined as

\[
Y_k(m) = \sum_{i=0}^{N-1} x(mN - i) \cdot e^{-j\frac{2\pi}{N} ki}
\]

This is a critically sampled filter bank with the impulse response (design trick: \(h_k(n)\) is only as long as one block)

\[
h_k(n) = e^{-j\frac{2\pi}{N} kn}
\]

\(n, k = 0 \ldots N - 1\)

The analysis polyphase matrix of the DFT is identical to the DFT transform matrix:

\[
H(z) = F = DFT - Matrix
\]

\(\rightarrow\) Perfect reconstruction, but filters not good enough!
Examples: DFT

The fourier matrix is defined as

\[ F_{n,k} = W^{nk} \]

\[ F = \begin{bmatrix}
    W^0 & W^0 & \cdots & W^0 \\
    W^0 & W^1 & \cdots & W^{(N-1)} \\
    \vdots & \vdots & \ddots & \vdots \\
    W^0 & W^{(N-1)} & \cdots & W^{(N-1)(N-1)}
\end{bmatrix} \]

is identical to the polyphase matrix of the DFT viewed as a filter bank

with \[ W = e^{-j \frac{2 \pi}{N}} \]
Example: The DCT_{IV}

The Discrete Cosine Transform type 4 is defined as:

\[ Y_k(m) = \sum_{i=0}^{N-1} x(mN - i) \cdot \cos \left( \frac{\pi}{N} \left( k + \frac{1}{2} \right) \left( i + \frac{1}{2} \right) \right) \]

with impulse response (N: block length):

\[ h_k(n) = \cos \left( \frac{\pi}{N} \left( k + \frac{1}{2} \right) \left( n + \frac{1}{2} \right) \right) \quad n, k = 0 \ldots N - 1 \]

Special property: filter bank is orthogonal

\[ \frac{2}{N} H^T (z^{-1}) = z^{-d} \cdot H^{-1}(z) \]

for Perfect Reconstruction, hence the synthesis is the transposed time reversed matrix
DCT Type 4, with 8 Subbands (N=8)

Filter for subband 0:

- Impulse response
- Magnitude response

Bad filter

Nyquist frequency of input sampling rate
DCT Type 4, with 8 Subbands (N=8)

Filter for subband 1:

Impulse response

Magnitude response

passband

ideal

stopband attenuation

stopband

stopband

Sample n

Omega/π
Problem

- Except for the zeros, the stopband attenuation is not very high (still PR!)

- Problem especially for audio, since the zeros are not sufficient for good selectivity.

- Approach: design filter banks with longer filters, with better ability for higher stopband attenuation.

  → Really use z-domain for longer filters