Filterbanks and transforms


Filterbanks
- Introduction
- Critical sampling, half-band filter
- QMF-filterbank, prototype filter
- Avoiding aliasing errors, modified bank
- Perfect reconstruction, complementary filters
- MCF-filterbank

Transforms
- Introduction
- Discrete Fourier transform
- Discrete cosine transform
- Modified DCT
- Filterbanks vs. transforms

1 Introduction

- Filterbanks are used for example in
  - Perceptual audio coding
  - Multi-band equalizers
  - Bandwise dynamic range control
  - Machine hearing and audio content analysis
- The human auditory system performs frequency analysis
  - Critical bands in hearing, structure of the inner ear
  - One reason why we encounter filterbanks in many audio processing applications

Introduction

Filterbanks in audio coding

- In perceptual audio coding, the input signal is processed at subbands
  - That allows us to utilize the auditory masking phenomenon
    → A filterbank is required, in other words, a set of filters that select neighbouring narrow subbands that cover the entire frequency range
- The filterbanks used in audio coding consist of
  - Analysis filterbank that decomposes a signal into subbands
  - Synthesis filterbank that reconstructs a wideband signal to the output
- In audio coding, typically critically sampled, perfect reconstruction filterbanks are used
  - **Critical sampling:** if the filterbank subdivides the frequency range into \( K \) bands, the signal at each band is downsampled by factor \( 1/K \)
    → Amount of data does not increase
  - **Perfect reconstruction:** if no processing takes place at subbands, the signal can be reconstructed without errors using a synthesis filterbank

2 Critical sampling at two subbands

- Figure: block diagram of a two-band critically sampled analysis-synthesis filterbank

- Figure: magnitude responses of the filters applied at the two bands
Critical sampling at two subbands

2.1 Decimation in the analysis bank

- What happens in the analysis filterbank?
  - LP ↓2: Lowpass filter and downsample by factor 2
  - HP ↓2: Highpass filter and downsample by factor 2

- When the upper half-band \([f_s/4, f_s/2]\) is decimated, it is aliased (mirrored) to the lower frequencies \([0, f_s/4]\)
  - The aliasing does not corrupt spectral information since the lower frequency components were filtered out using a highpass filter
  - Figure: (a) original signal spectrum, (b) highpass filtered (HP) spectrum, (c) highpassed and decimated (HP + ↓2), aliased spectrum
  - Note the lower sampling rate in panel c) \(\text{Nyquist frequency is } 4\text{ kHz}\)

Interpolation in the synthesis bank

- What happens in the synthesis filterbank?
  - ↑2 + LP: Upsample by factor 2 and lowpass filter
  - ↑2 + HP: Upsample by factor 2 and highpass filter

- ↑2 operation in practice:
  - Add zeros between the sample values in the signal (vector of numbers)
  - Multiply the signal by 2 in order to keep its level unchanged

- Let's look again at the upper half-band

- Figure: (c) spectrum of the signal that was highpassed and decimated in analysis bank, (d) spectrum obtained by interpolating (↑2) signal in c, (e) after interpolating and highpass filtering (↑2 + HP) the signal in c

3 Several subbands uniformly distributed

- Also the higher half-band was (almost exactly) reconstructed despite the decimation at the subband
- The entire original signal can be reconstructed by summing the upper and lower half-bands at the output of the synthesis bank

Also the higher half-band was (almost exactly) reconstructed despite the decimation at the subband
- The entire original signal can be reconstructed by summing the upper and lower half-bands at the output of the synthesis bank

- Two subbands are not of much use yet
- The principle scales easily to \(n\) uniformly distributed subbands

- Figure: \(n\) subbands, each decimated by factor \(k\)
  - Bandwidths must be equal in order that all of them can be decimated by factor \(k\) \(\Rightarrow\) uniformly distributed subbands
  - Critical sampling if \(n=k\)
Several subbands uniformly distributed

- What happens in the analysis bank at \( n \) subbands?
  - Spectrum in range \([0, f_s/2]\) is divided into \( n \) bands, each of width \((f_s/2)/n\)
  - Bandpass filter \( H_m(f) \) in the analysis bank selects band \( m \)
  - Band \( m \) covers the frequencies
    \[
    \left[ \frac{mf_s}{2n} - \frac{m+1}{2n} \frac{f_s}{2}, \frac{m+1}{2n} \frac{f_s}{2} \right], \quad m = 0, 1, \ldots, n - 1
    \]
  - In downsampling, the band is aliased to frequencies \([0, f_s/(2n)]\)
    - No problem, since those frequencies were filtered out by \( H_m(f) \)

- In the synthesis bank
  - Interpolation by factor \( k \) (\( k=n \)) replicates the subband \([0, f_s/(2n)]\) at all subbands
  - Each subband is selected at its correct frequency range using synthesis bandpass filter \( G_m(f) \) (same passband as \( H_m(f) \))

4 Half-band FIR filter

- Let’s return to the two-band filterbank
- Half-band filters have the nice property that they can be implemented very efficiently computationally
- Figure: example of an 80-tap FIR filter that passes the upper half-band \([f_s/4, f_s/2]\)
  - Even-numbered filter coefficients are zero, except for the middle value
  - Impulse response is symmetric
  - Convoluting with \( n \)-length FIR requires only \( n/4 + 1 \) multiplications for each input signal sample
  - Can design relatively steep transition bands using FIR filters

Half-band filter

- Figure: magnitude and phase response of the highpass half-band filter presented on the previous slide
  - Advantage of FIR: phase response in passband is exactly linear
  - Filtered signal is a delayed copy of the original signal in passband

Aliasing error

- It is clear that in a critically-sampled filterbank some unwanted aliasing happens at the subbands
  - Filters are not ideal (transition band, not step function)
  - For example when downsampling by factor 2, the part that exceeds the new Nyquist frequency \( \pi/2 = f_s/4 \) is aliased
- Figure below shows the low- and highpass filters applied in the two-band critically sampled filterbank
The filterbanks used in audio coding are usually designed so that the synthesis bank eliminates the aliasing that occurs at the subbands. Achieves perfect or near-perfect reconstruction despite the unwanted aliasing at subbands. Reconstruction errors in the filterbank are insignificant compared to the errors introduced in the encoding and decoding. In the following we study the QMF filterbank that eliminates the aliasing in the synthesis bank.

A shortcoming of the above-described filterbank is that perfect reconstruction is not possible due to unwanted aliasing.

Figure: Two-band analysis-synthesis Quadrature Mirror Filter (QMF) bank
- \( H_0(z) \) and \( F_0(z) \) are lowpass filters.
- \( H_1(z) \) and \( F_1(z) \) are highpass filters.

Idea of QMF: filters \( H_0(z), F_0(z), H_1(z), F_1(z) \) and be designed so that the aliasing in analysis part is eliminated in the synthesis part. Perfect reconstruction is achieved (if not processing at subbands).

Figure: QMF filter bank subdividing a signal into octave bands using critical sampling (octave=frequency range \([f, 2f]\))
- successive lowpass/highpass subdivisions into half bands
- decimation of the half bands by factor 2 after each subdivision
- lower band is recursively subdivided

Acronyms:
- QMF bank= Quadrature Mirror Filter bank
- SP = signal processing
- LP = lowpass filtering
- HP = highpass filtering
- \( \downarrow 2 \)= downsampling
- \( \uparrow 2 \)= upsampling

As a result, we get subbands \( Y_1 \ldots Y_N \)

The frequency boundaries of the bands are given by

\[ \Omega_{ck} = 2^{-k} \pi \quad \Rightarrow \quad f_{ck} = 2^{-k} \left( f_s / 2 \right) \]

where \( k = 1, 2, \ldots, N - 1 \)

The above frequency range subdivision results from:
- The downsampled upper half-band is ready as it is
- The lower half-band is recursively split into two half-bands (see the block diagram on the previous slide)
Subbands in the QMF-bank

- **QMF-bank has non-uniform frequency resolution**
- Bandwidths and sampling rates are different at each subband (halving at each step towards lower frequencies)
  - Lower subbands are subjected to multiple downsampling steps (and multiple upsampling steps in the synthesis bank)
- **Note that the amount of data remains the same** as a result of the subdivision into subbands and downsampling
  - Amount of data in the original signal is $S \cdot f_s$
  - Where $S$ is signal duration in seconds and $f_s$ is the sampling rate
  - Amount of data after the division into subbands is:
    - $S \cdot (0.5f_s + 0.25f_s + 0.125f_s + ... ) \approx S \cdot f_s$
  - Where the sampling rates have been given starting from the highest subband
  - Critical sampling regardless of the number of subbands

5.2 Prototype filter

- The same lowpass and highpass filter can be used in all successive lowpass/highpass subdivisions
  - The term “prototype filter” is used
- Response of a filter is always given proportional to the sampling rate
  - For example the cutoff frequency defined as $0.3f_s$
- When the sampling rate of the lower half-band is decreased by factor 2, we get a new sampling rate $\tilde{f}_s = f_s / 2$
  - In the new sampling rate, the lower sub-band $[0, \tilde{f}_s / 4]$ fills the entire frequency range $[0, \tilde{f}_s / 2]$
  - Now the cutoff frequency of the same filters is $\tilde{f}_s / 4 = f_s / 8$
  - By applying the same prototype filter as it is in the new sampling rate, it subdivides the lower half-band again
  - From the viewpoint of the original sampling rate, the passband gets twice narrower, but also the transition band gets twice steeper
    - So-called multirate signal processing: steep filters with few coefficients

6. Avoiding aliasing errors

- Aliasing errors in the QMF-bank output part were avoided by appropriate synthesis filter design
- However in many applications it is not enough that the synthesis part compensates for the aliasing that happens in the analysis part
  - Multi-band equalizers, dynamic range control, audio content analysis
  - We have to design an analysis filterbank, where aliasing does not happen at subbands

Avoiding aliasing errors

- The analysis filterbank can be slightly modified in order to avoid aliasing at subbands
  - Harmful aliasing occurs at the boundary between the low- and highpass bands, since downsampling by factor 2 is applied even though the filters are not perfect
- Figure: response of the modified prototype filter (multi-complementary filter bank, MCF) compared to the previous bank (QMF)
  - In the modified filterbank, the half-band boundary has been moved from $\pi / 2$ to a lower frequency
    - When downsampling the lowpass band by factor 2, no aliasing occurs
      - The upper band cannot be downsampled at all
## Avoiding aliasing errors

### 6.1 Modified filterbank

- **Figure:** modified filterbank where the new subband division is repeated
  - Difference to QMF-bank is that the upper band is not downsampled

- **Subband boundaries:**
  \[
  \Omega_{k+1} = \frac{\pi}{3} 2^{k+1}
  \]

## Perfect reconstruction

### 7.1 Complementary filters

- **Figure:** implementation of complementary filters and the interdependency of their frequency responses
  - The complementary filter of a certain FIR filter \( H_c(z) \) is obtained by (i) filtering the input signal with \( H_c(z) \) (ii) subtracting the filtered signal from the original input signal which has been delayed by the amount of (group) delay caused by \( H_c(z) \)
Perfect reconstruction

7.2 MCF-bank

- Multi-complementary filter bank is a filterbank where
  - The half-band boundary is moved from \( \pi/2 \) to a lower frequency to avoid harmful aliasing, and the upper half is not downsampled (amount of data doubles in the subband division)
  - The basic lowpass / highpass divisions blocks are implemented using complementary filters to enable perfect reconstruction

8 Transforms

- In frame-based processing, the signal is windowed, after which a (frequency) transform can be applied
- The transform consists of calculating an inner product between the windowed signal and basis functions of the transform
- In audio signal processing, the basis functions are typically sines and cosines with different frequencies, or complex exponents
- Output represents the spectrum of the signal in the frame
- Efficient algorithms exist for computing the transform at all frequencies simultaneously (e.g. fast Fourier transform)

8.1 Discrete Fourier transform (DFT)

- Basis functions are complex exponents
  \[
  X(k) = \frac{1}{\sqrt{2M}} \sum_{n=0}^{2M-1} x(n) e^{-jn\pi/M}, \quad 0 \leq k \leq 2M - 1
  \]
- Efficient algorithm: FFT
- Frequency responses:

Discrete Fourier transform (DFT)

- First few basis functions (blue: real part, red: imag. part)
8.2 Discrete cosine transform (DCT)

- Basis functions are cosines:
  \[ X(k) = c(k) \sqrt{\frac{2}{M}} \sum_{n=0}^{M-1} x(n) \cos \left( \frac{\pi}{M} \left( n + \frac{1}{2} \right) k \right), \quad 0 \leq k \leq M - 1 \]
- Where \( c(0) = 1/\sqrt{2} \)
  ja \( c(k) = 1, \quad 1 \leq k \leq M - 1 \)
- Efficient algorithm:
  Fast cosine transform
- Frequency responses:

8.3 Modified discrete cosine transform (MDCT)

- Cosine basis functions:
  \[ X(k) = \sqrt{\frac{2}{M}} \sum_{n=0}^{M-1} x(n) \cos \left( \frac{\pi}{M} \left( n + \frac{1}{2} \right) \frac{M+1}{2} \left( k + \frac{1}{2} \right) \right), \quad 0 \leq k \leq M - 1 \]
- Efficient algorithm
- Differences to DCT
  - Input has twice more samples than output \( \rightarrow \) transform causes time-domain aliasing
  - Adjacent frames cancel out the aliasing in the synthesis bank when a proper window function is used
  - Enables perfect reconstruction, critical sampling, and overlapping frame \( \rightarrow \) very useful in audio coding!