An audio watermarking method robust to time and frequency fluctuation

R. Tachibana, S. Shimizu, S. Kobayashi, T. Nakamura
Tokyo Research Laboratory
IBM Japan

Robustness Requirements for Audio Watermarking

- ✓ MPEG compression
- ✓ Additive noise
- ✓ Lowpass filtering
- ✓ D/A and A/D conversion
- ✓ Echoing

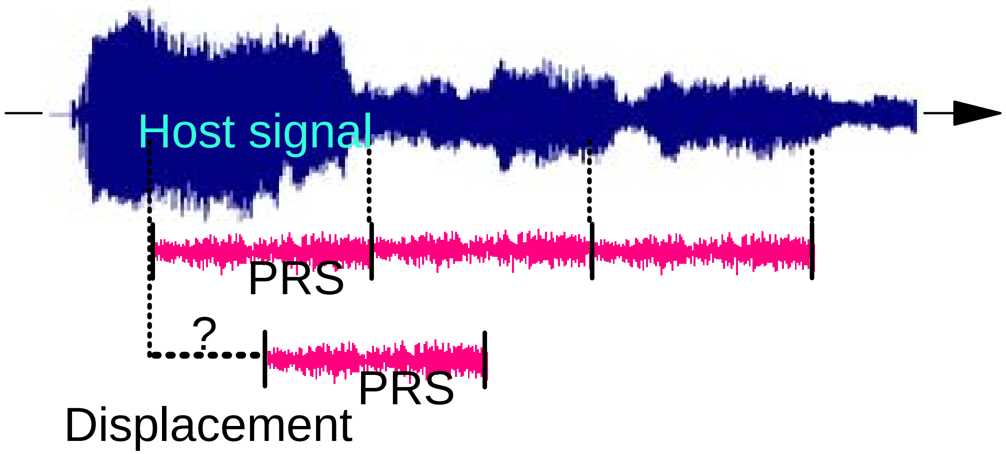
- × Random sample cropping
- × Pitch shifting
- × Wow-and-flutter

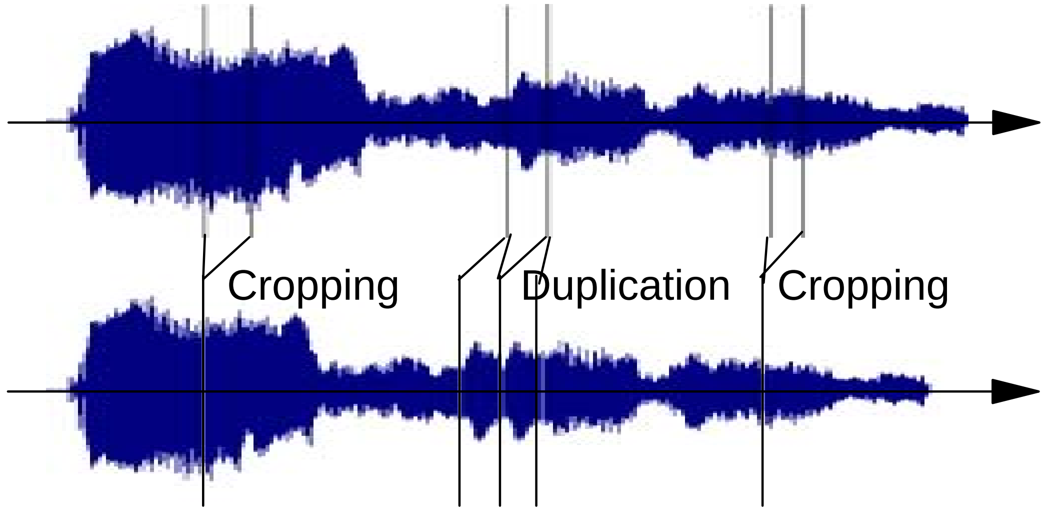
↓

Time and frequency fluctuation

What's difficult with time fluctuation?

- ► Pseudo-random sequence (PRS) mismatch along time
  - Detection requires matching of embedded PRS and detection PRS



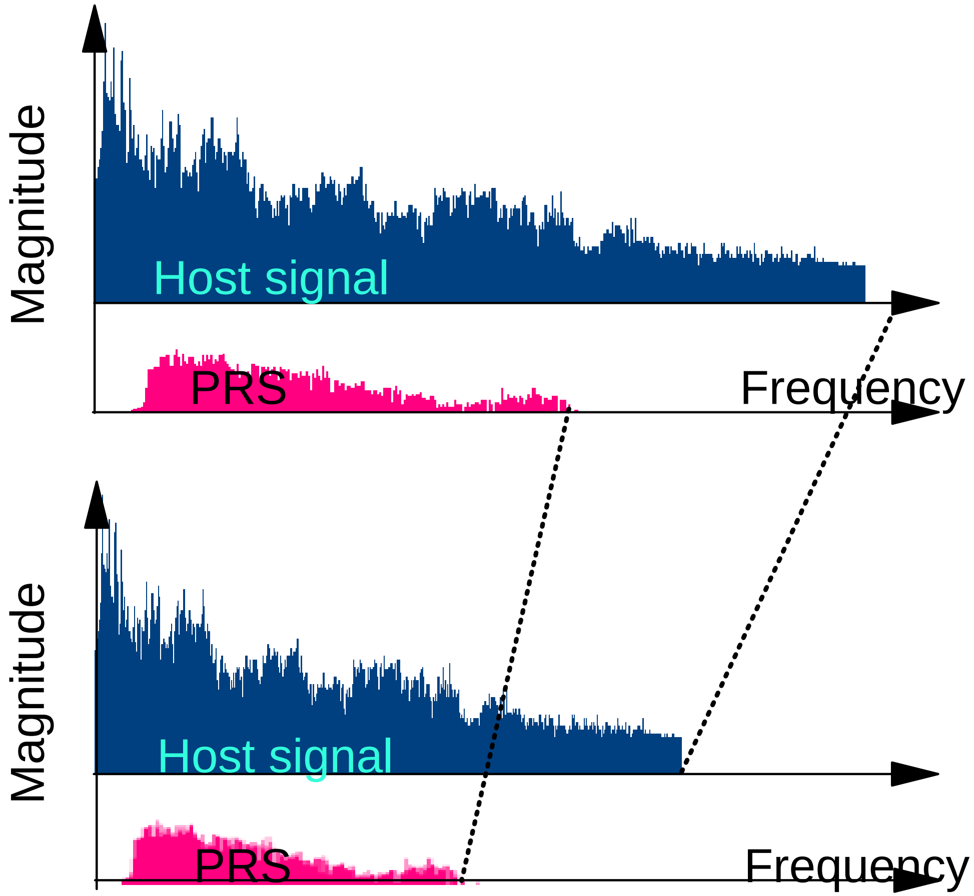


Random stretching

- × Sample-by-sample exhaustive search costs a lot

What's difficult with frequency fluctuation?

- PRS mismatch along frequency



Basic Concepts

1. Magnitude modification

2. Geometric wide subbands

3. Two-dimensional pseudo-random array

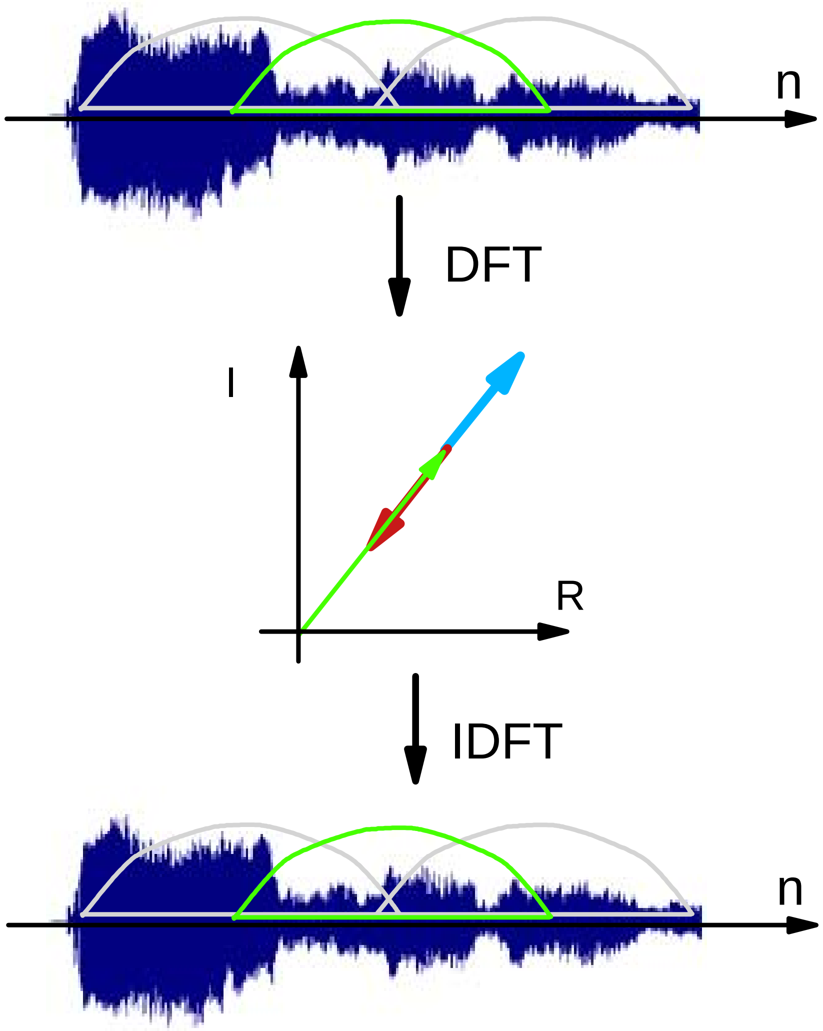
Basic concept 1 :

Magnitude Modification

- Watermark observable at displaced DFT windows
- Embedding
  - Modifies magnitude of the content in the frequency domain
- Detection
  - Correlates the magnitudes and the PRS

Basic concept 1 :

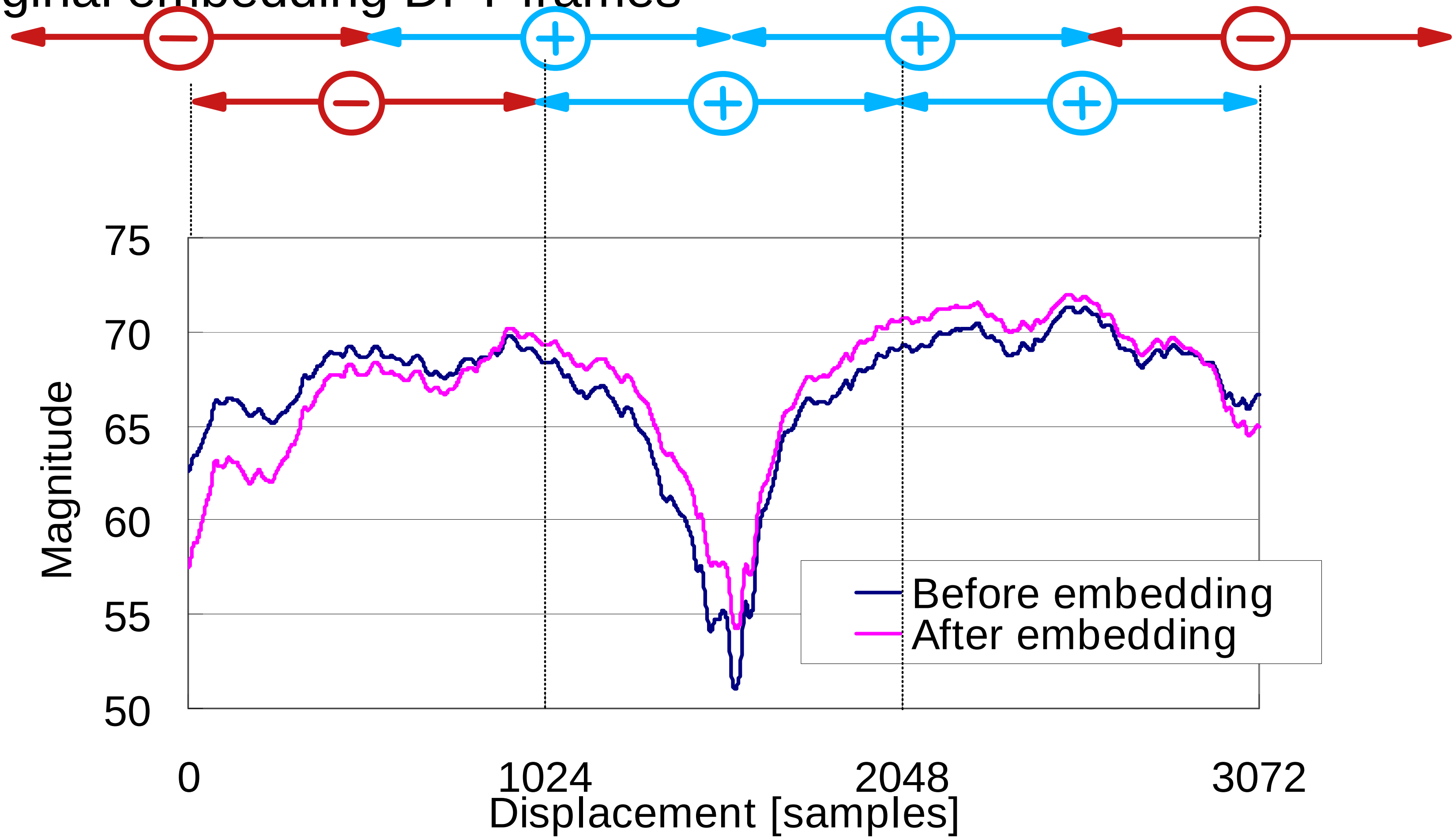
Magnitude Modification



- Embedding
  - Modifies amplitudes of frequency bins
  - Multiplies windowing function
  - Overlaps adjacent DFT frames

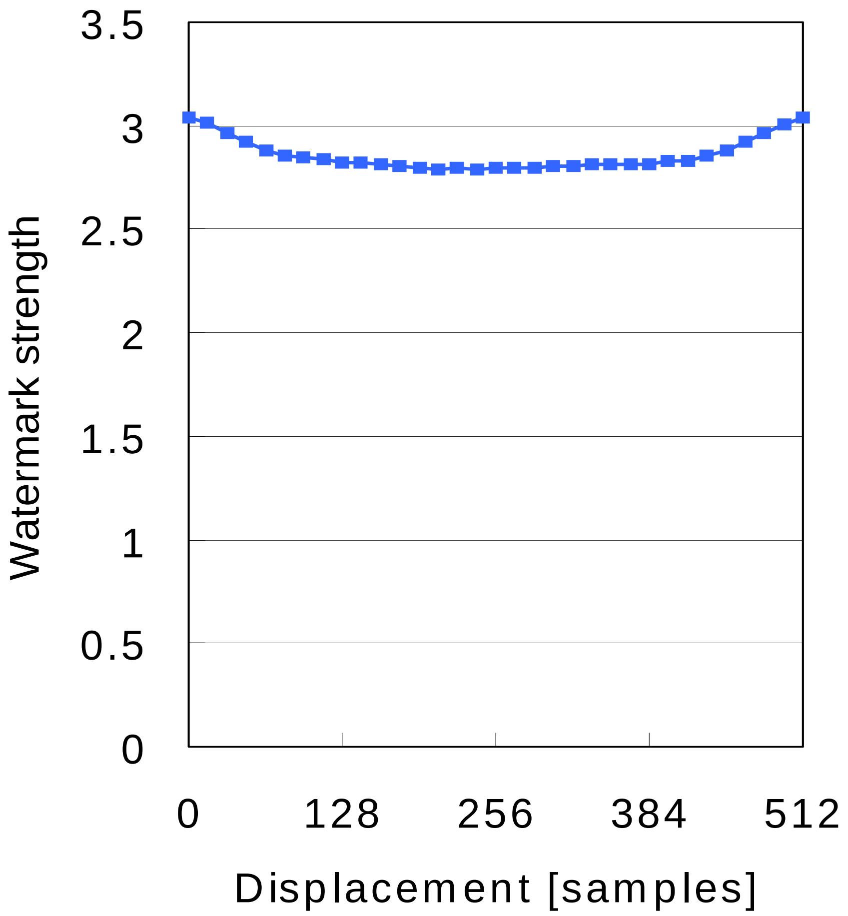
Modified magnitude observable at displaced DFT windows

Original embedding DFT frames



Observed magnitude of a frequency bin at displaced DFT windows

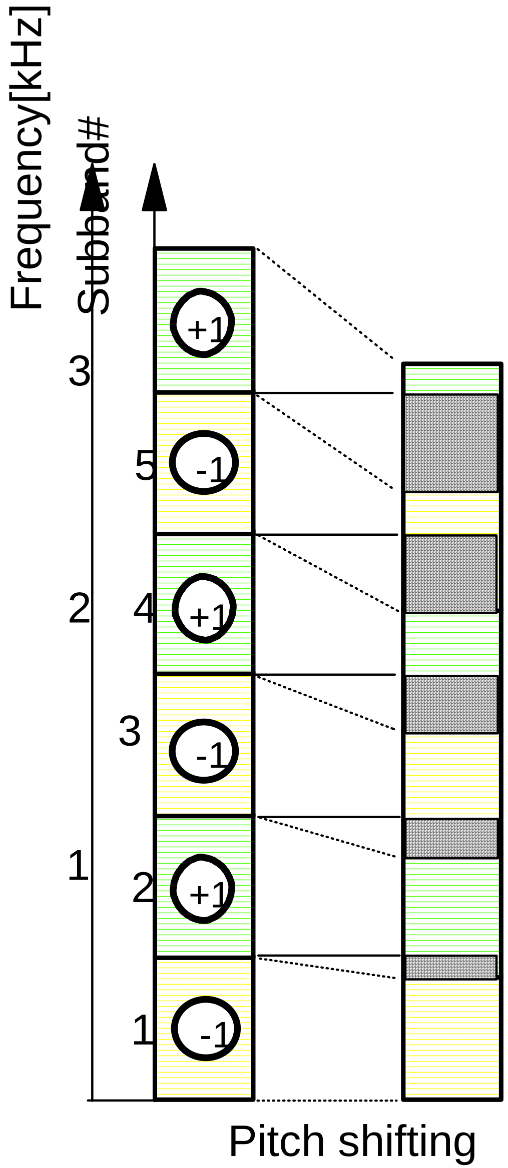
Watermark strength remains high with displaced DFT windows



- Detection
  - Worst displacement is 256-sample

Basic concept 2 :

Geometric wide subbands



► Watermark insensitive to pitch shifting

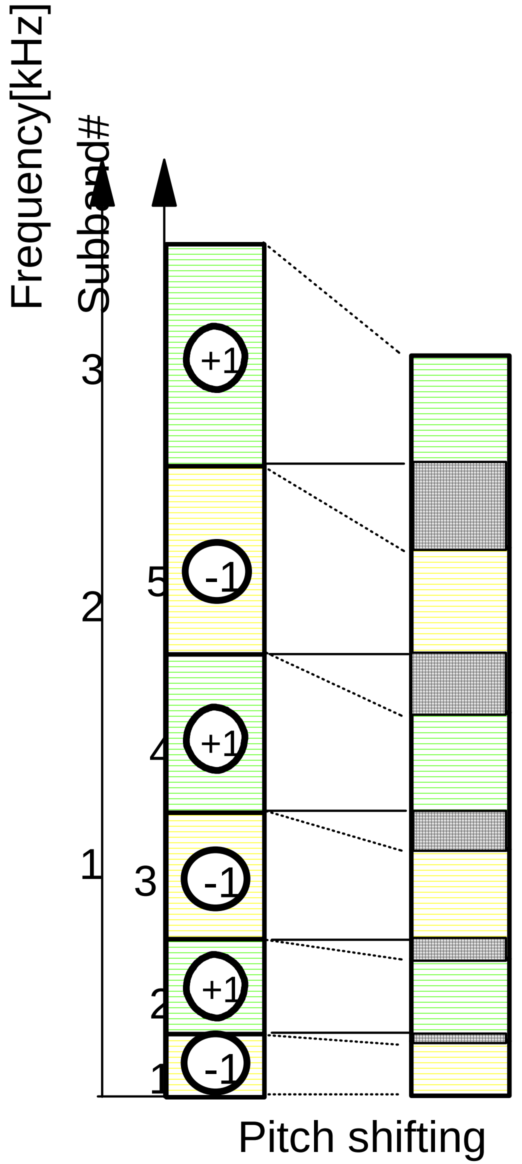
- Freq. bins remaining in a same subbands can contribute to watermark detection
- Wider subbands, more robust against pitch shifting

× Arithmetic Subband design (ASD)

- Subbands with the same bandwidth
- High frequency subbands weak at pitch shifting

Basic concept 2 :

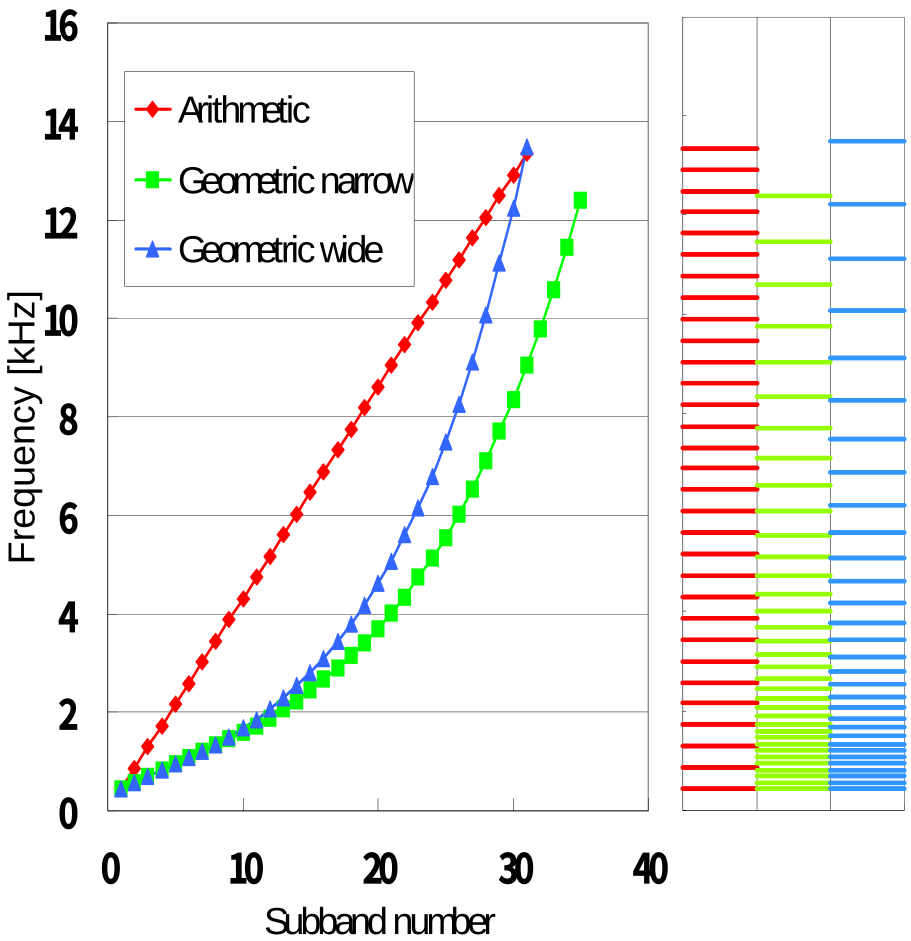
Geometric wide subbands

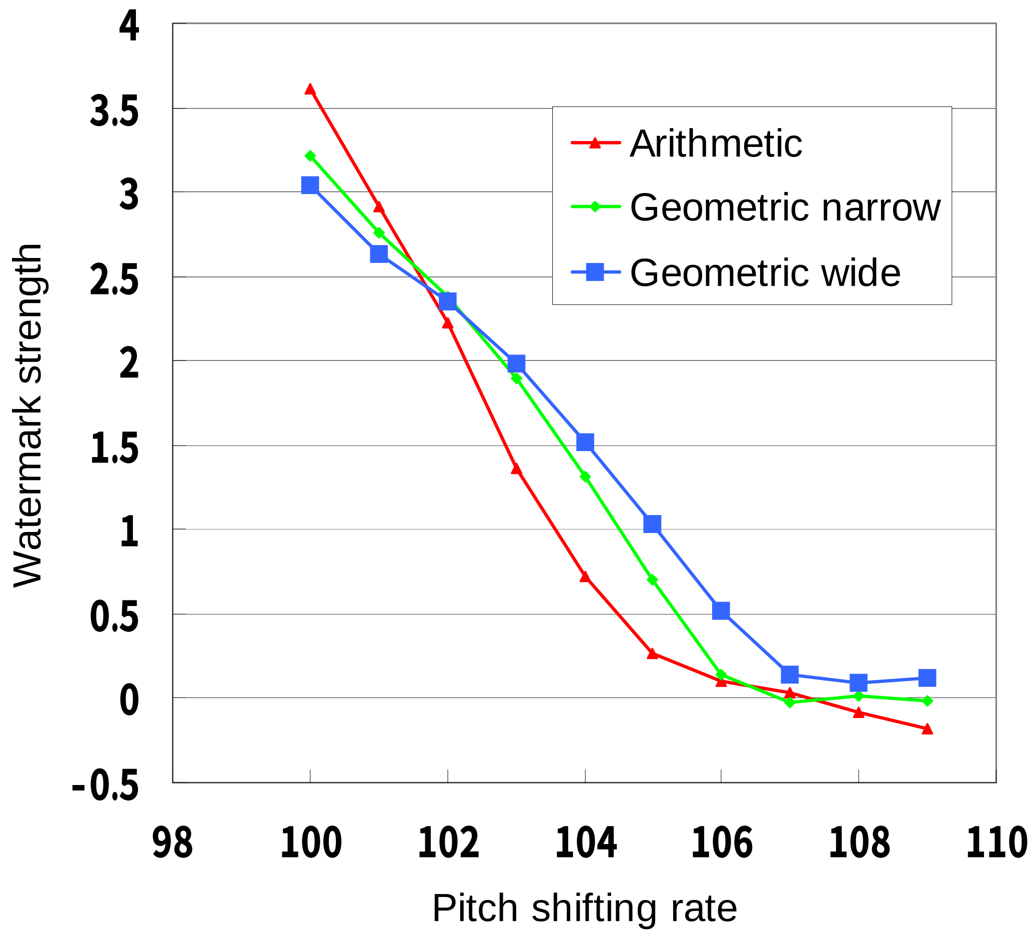


- ►Watermark insensitive to pitch shifting
  - Freq. bins remaining in a same subbands can contribute to watermark detection
  - Wider subbands, more robust against pitch shifting

- Geometric subband design (GSD)
  - Subbands with logarithmically increasing bandwidth
  - All subbands degrade at the same rate

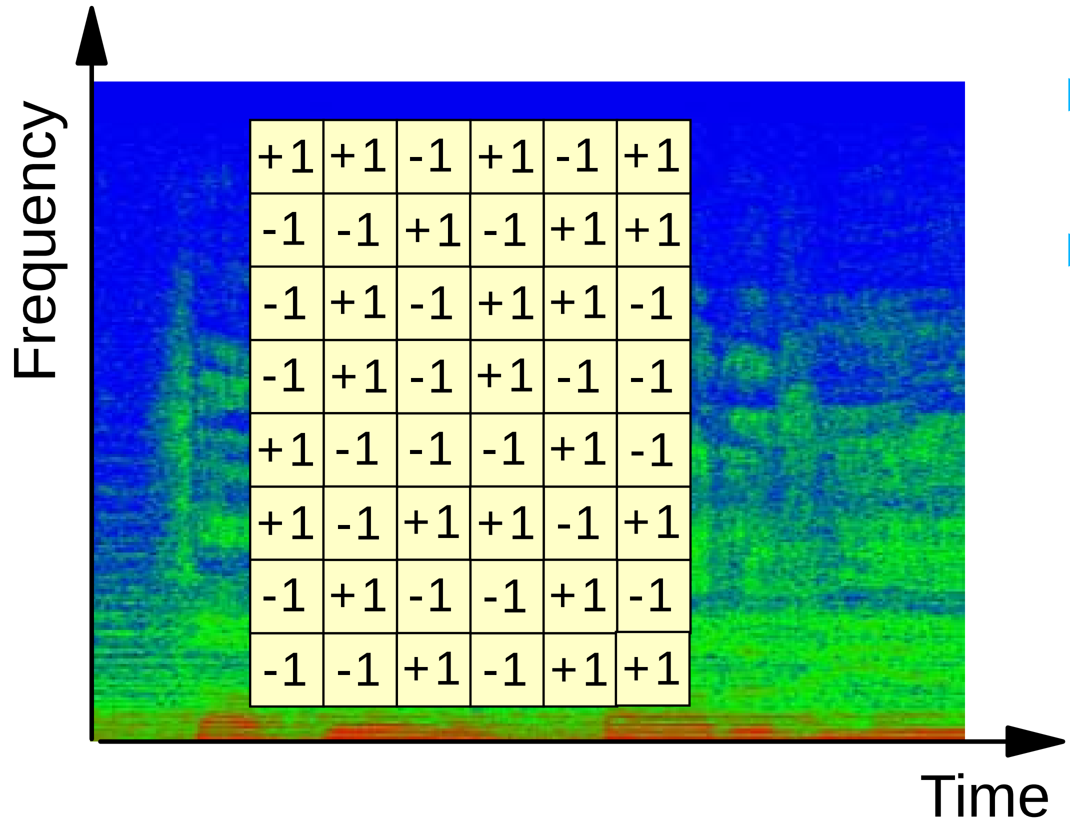
GSD survives better





Basic concept 3 :

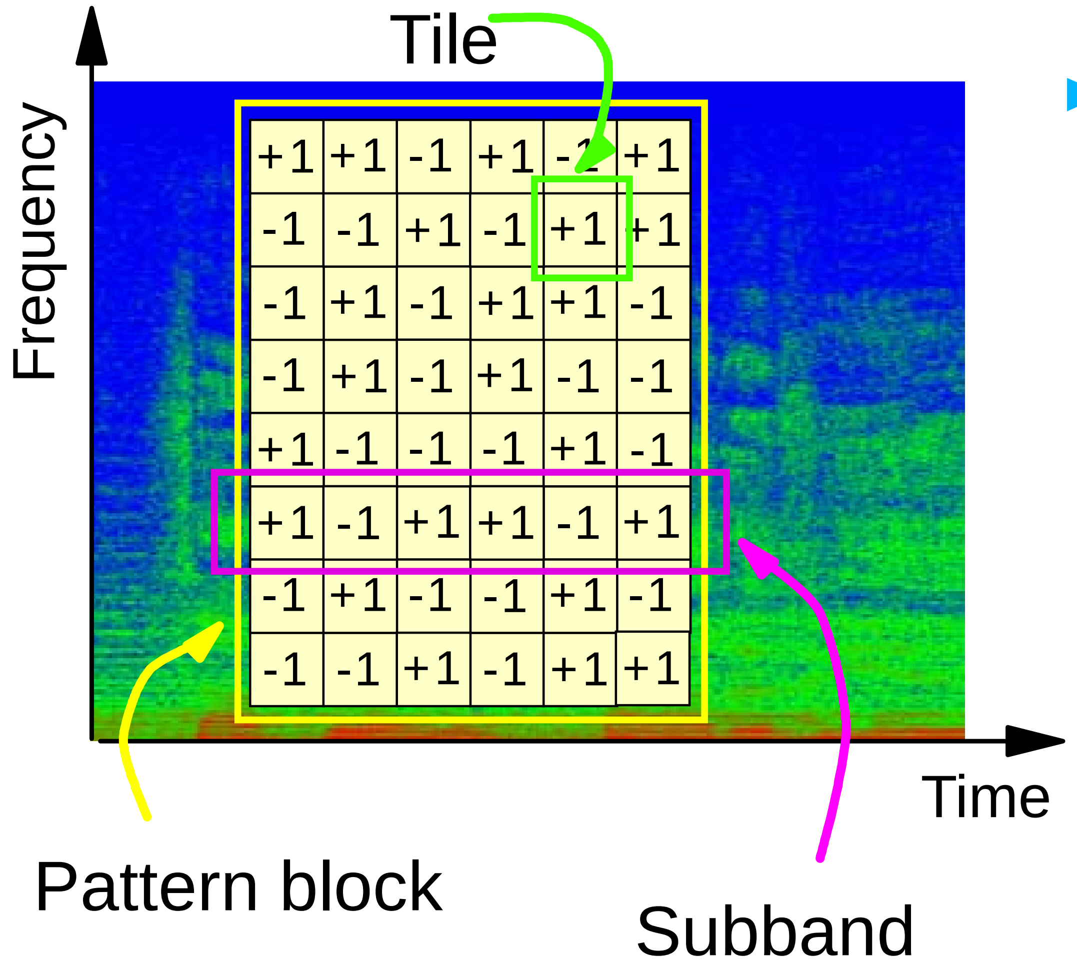
Two-dimensional Pseudo-Random Array (PRA)



- More pseudo-random numbers
- Less possibility to miss block

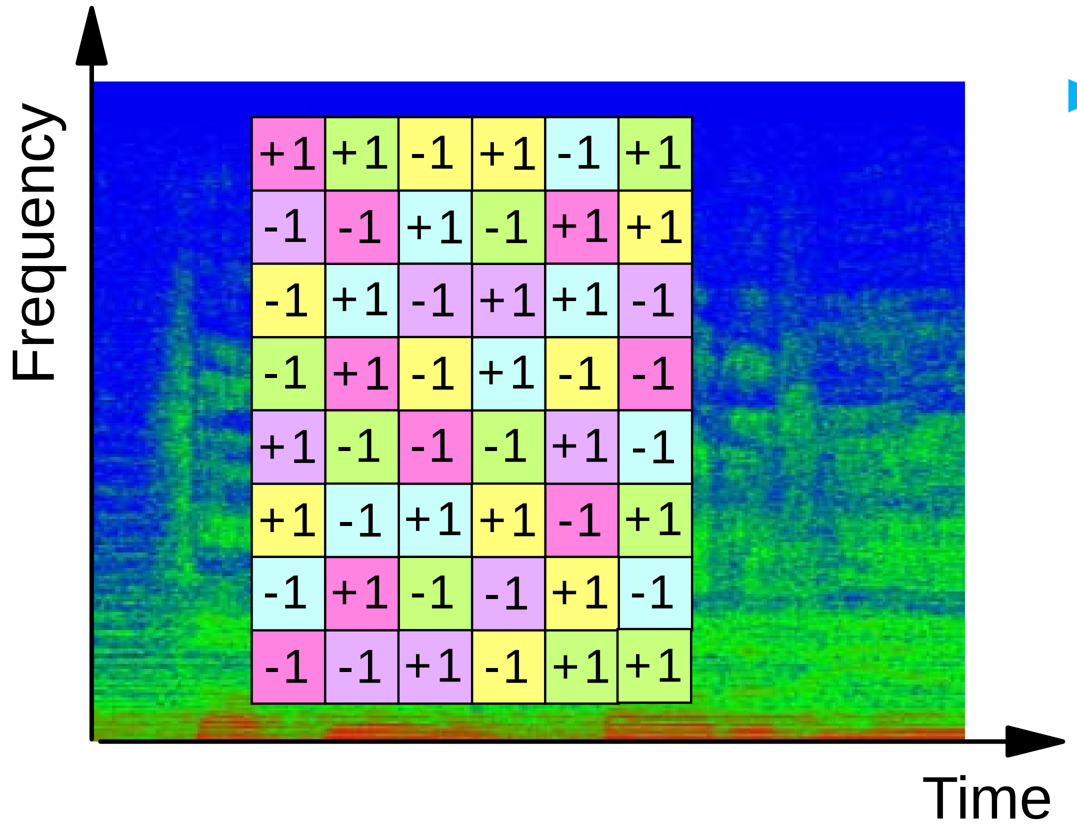
Basic concept 3 :

Two-dimensional Pseudo-Random Array (PRA)



- 30 (H) x 9 (W) = 270 tiles

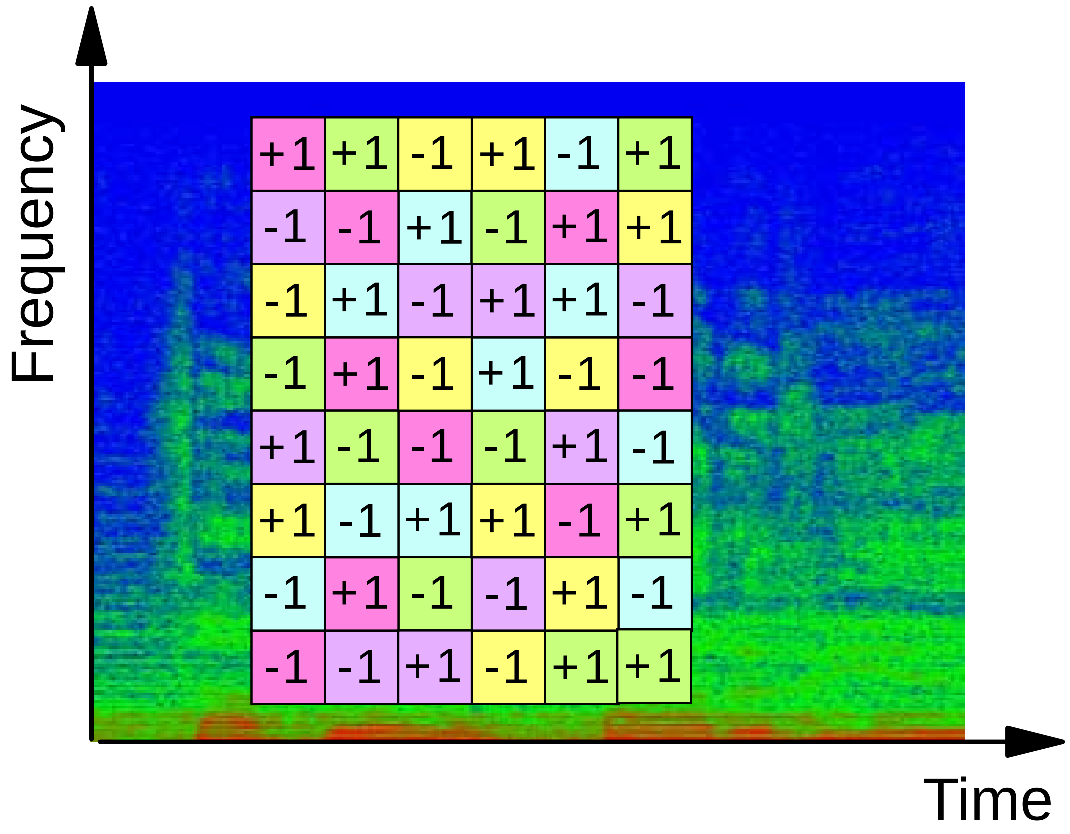
Basic concept 3 :
Two-dimensional Pseudo-Random Array (PRA)



- 4 bits and sync signal
  - 270 tiles = 30 x 4 + 150

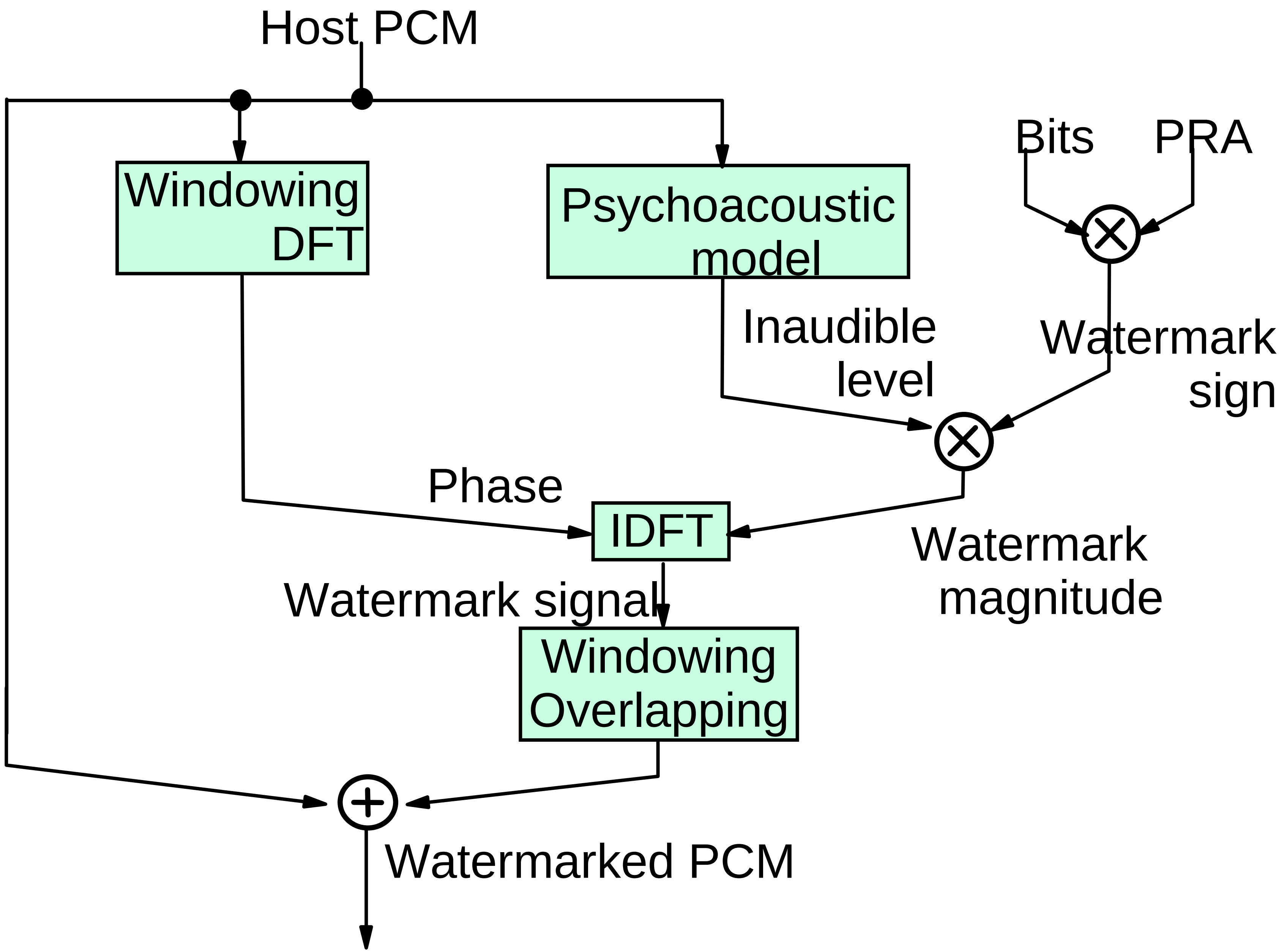
Basic concept 3 :

Two-dimensional Pseudo-Random Array (PRA)



- Block Synchronization
  - Search for the start of the pattern block
  - Maximizing sync signal strength
  - Frame-by-frame search
  - No additional DFT required

Embedding Algorithm



Multiple-bit message

64-bit message

Cyclic Redundancy Check (CRC) parity bits

71

Bose-Chaudhuri-Hocquenghem (BCH) encoding

127

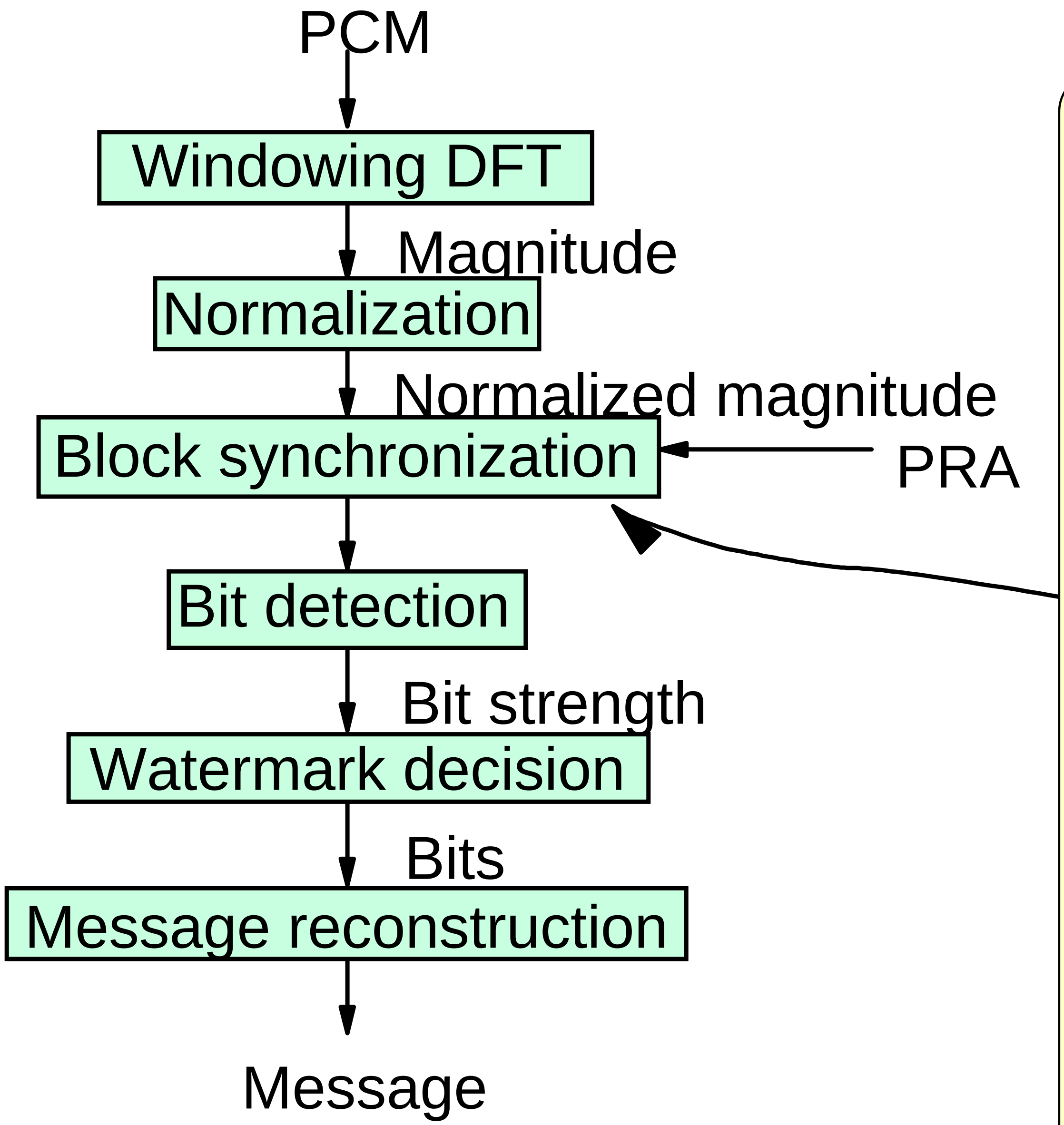
System bits for indicating the beginning of the message

140

4 bits per pattern block

35 pattern blocks

Detection Algorithm



Correlate the magnitudes with PRA

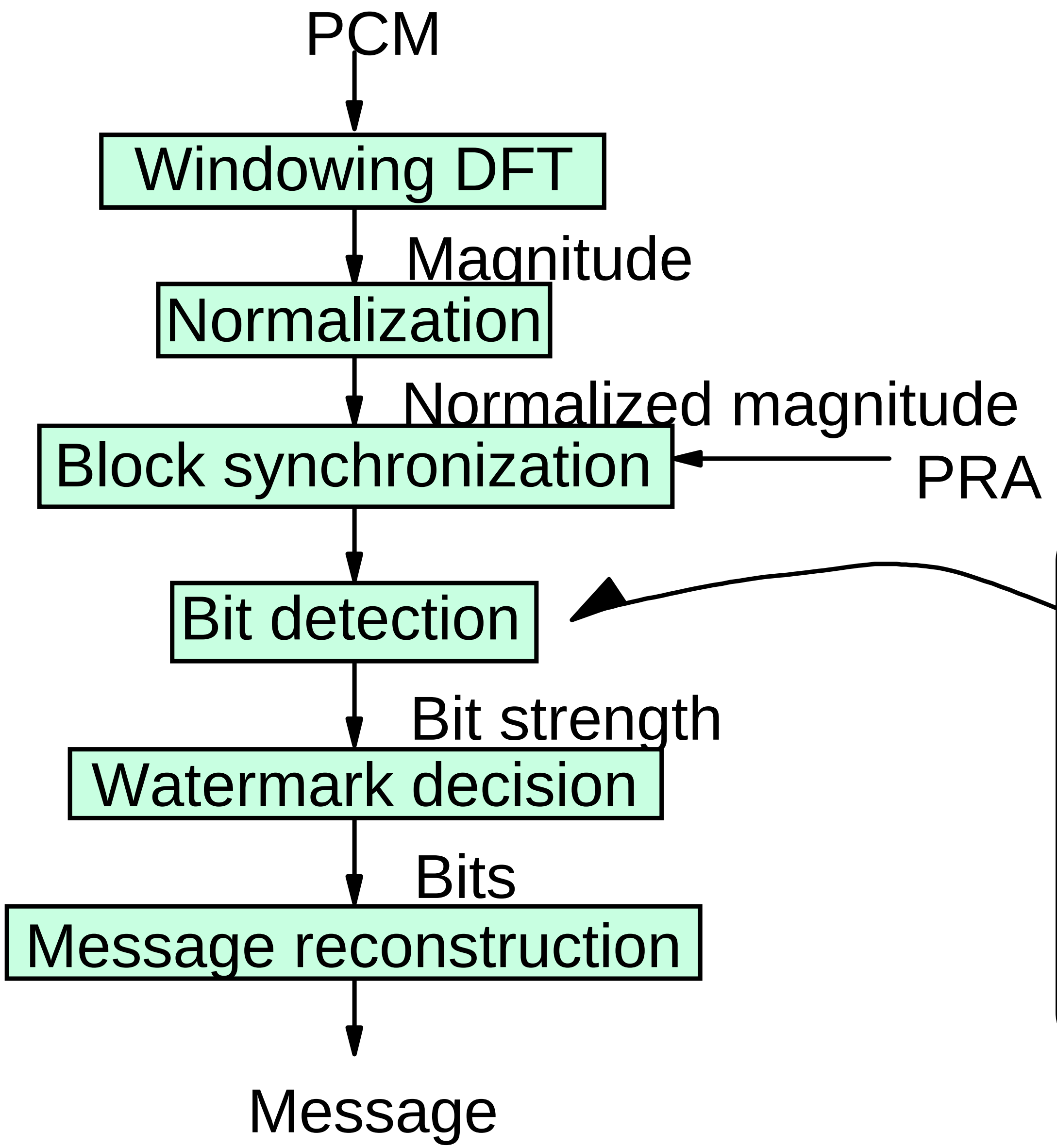
$$\hat{Q}_{t,f} = Q_{t,f} - \overline{Q}$$

$$\overline{Q} = \frac{\sum_{k=1}^{D_S} Q_{t,k}}{D_S}$$

$$S_s = \frac{\sum_{k=1}^{D_S} \omega_{S,k} \hat{Q}_{t,f}}{\sqrt{\sum_{k=1}^{D_S} \left\{ \omega_{S,k} \hat{Q}_{t,f} \right\}^2}}$$

Find the position maximizing
the synchronization strength

Detection Algorithm



Correlate the magnitudes with PRA

$$X_j = \frac{\sum_{k=1}^{D_j} \omega_{j,k}\left(Q_{s+t,k} - \overline{Q}\right)}{\sqrt{\sum_{k=1}^{D_j} \left\{\omega_{j,k}\left(Q_{s+t,k} - \overline{Q}\right)\right\}^2}}$$

Robustness of 64-Bit Message

- Ten 100-second music samples
  - CD-quality (44.1kHz and 16bit sampling)
- One detection in every 30 seconds
- Performance
  - On Pentium III 600 MHz and Windows NT 4.0 with 256MB RAM
  - Embedding in 36% of playback speed
  - Detection in 8% of playback speed

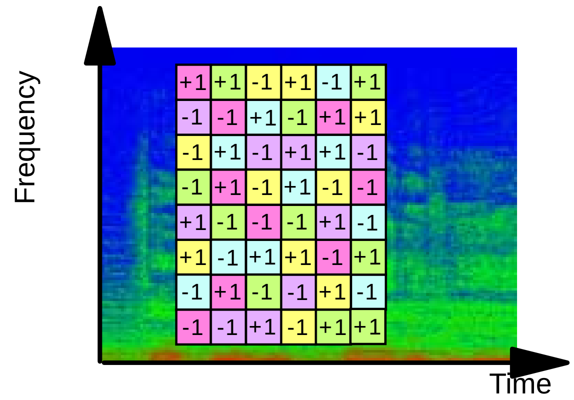
Robustness of 64-Bit Message

Original watermark	100%	Random stretching 104%	100%
Wow-and-flutter	100%	Random stretching 102%	100%
Echo 50 msec 0.3	100%	Random stretching 98%	100%
Echo 100 msec 0.5	97%	Random stretching 96%	100%
Noise -40dB	97%	Pitch shifting 104%	90%
Noise -30dB	87%	Pitch shifting 102%	100%
MP3 96kbps	100%	Pitch shifting 98%	100%
		Pitch shifting 96%	83%

Conclusion

Robust audio watermarking

- ► Modifies the magnitudes of the content
- ► Uses a two-dimensional pseudo-random array
- ► Robust against;
  - Random stretching (Time fluctuation)
  - Pitch shifting (Frequency fluctuation)
  - Wow-and-flutter (Time and frequency fluctuation)



- ► Fast
  - No need for sample-by-sample search