
Audio Watermarking for Live Performance

Ryuki Tachibana

(ryuki@jp.ibm.com)

Tokyo Research Laboratory

IBM Japan

Background

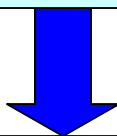
- Applications of Audio Watermarking (AWM)
 - Copy control of digital music
 - Ownership identification
 - Searching for illegally copied music

- How can we make AWM easier to use?
- What other applications of AWM?

Target of embedding

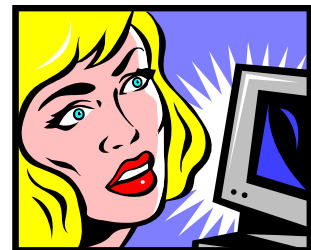
File Embedding:

Digitally stored music *file*



Real-Time Embedding:

(Analogue or Digital) Audio *stream*



Purpose

- By *Real-Time Embedding*, make AWM easier to use, and find its new applications.

Real-Time Embedding Composition Methods

4 Groups of Composition Methods

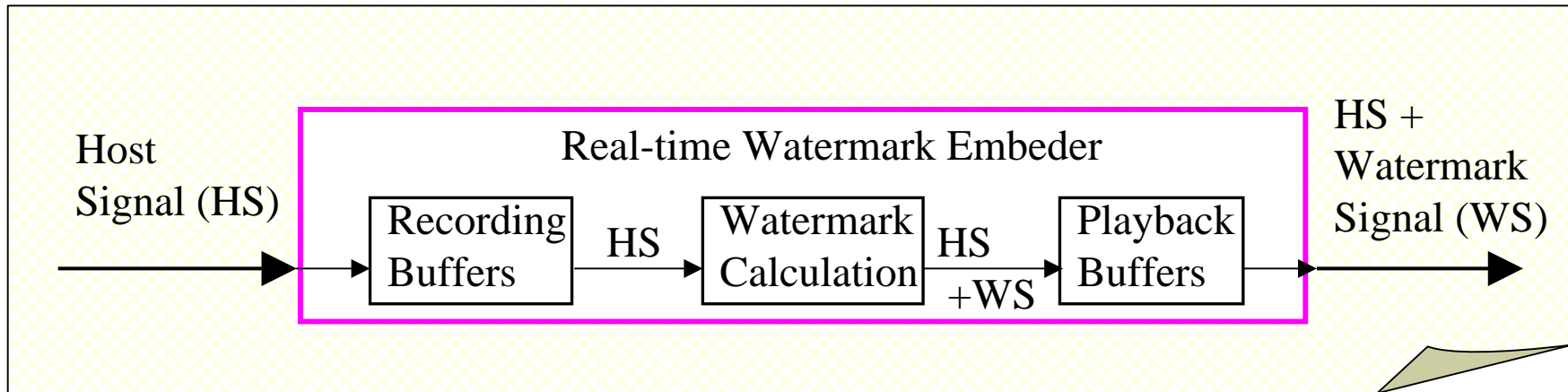
(C1) Naïve Composition

(C2) Analogue Watermarking

(C3) Sonic Watermarking

(C4) Sonic Watermarking with Sampling

C1 : Naïve Composition



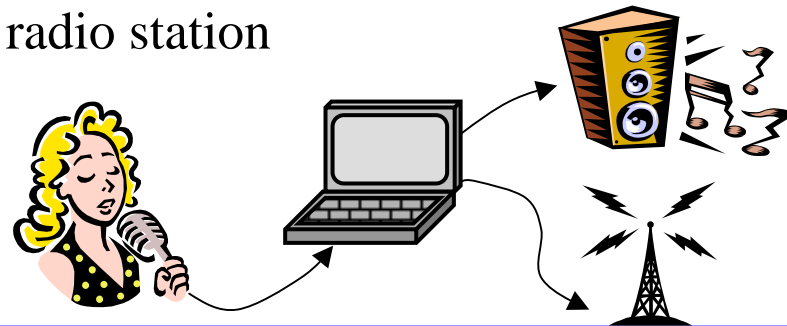
Embedder receives HS and produces HS+WS.

Merits

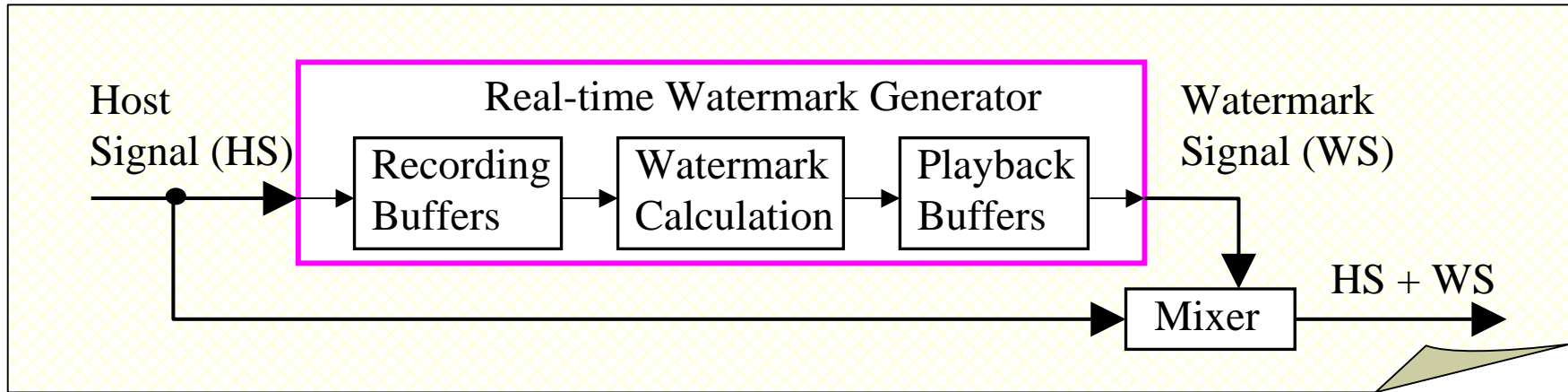
- ☺ Applicable to any AWM algorithm
- ☺ Instant listening of watermarked sound in a music authoring process
- ☺ Live broadcast of watermarked sound in a radio station

Drawbacks

- ☹ Risk of interrupting of the playback
- ☹ Inevitable delay of the host signal (HS)



C2 : Analogue Watermarking

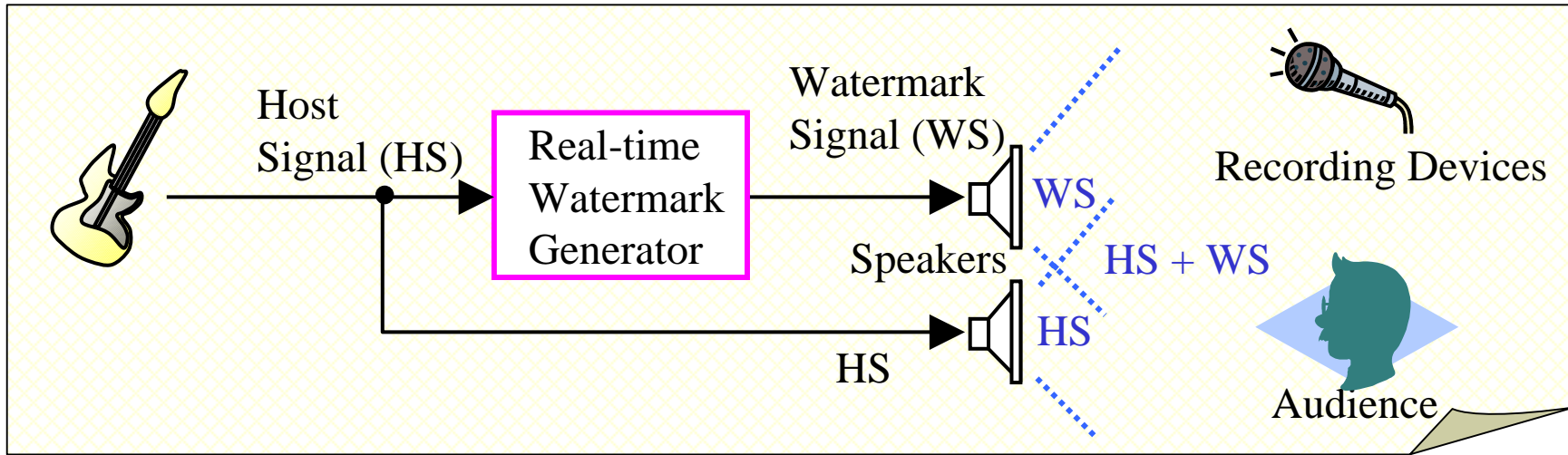


Watermark generator receives HS and produces WS.

Conventional analogue mixer mixes HS and WS on the outside.

- ☺ No delay for the HS
- ☺ No risk of interrupting of the playback of the HS
- ☺ Easier control of the watermark volume
- ☹ Delay for the WS to the HS

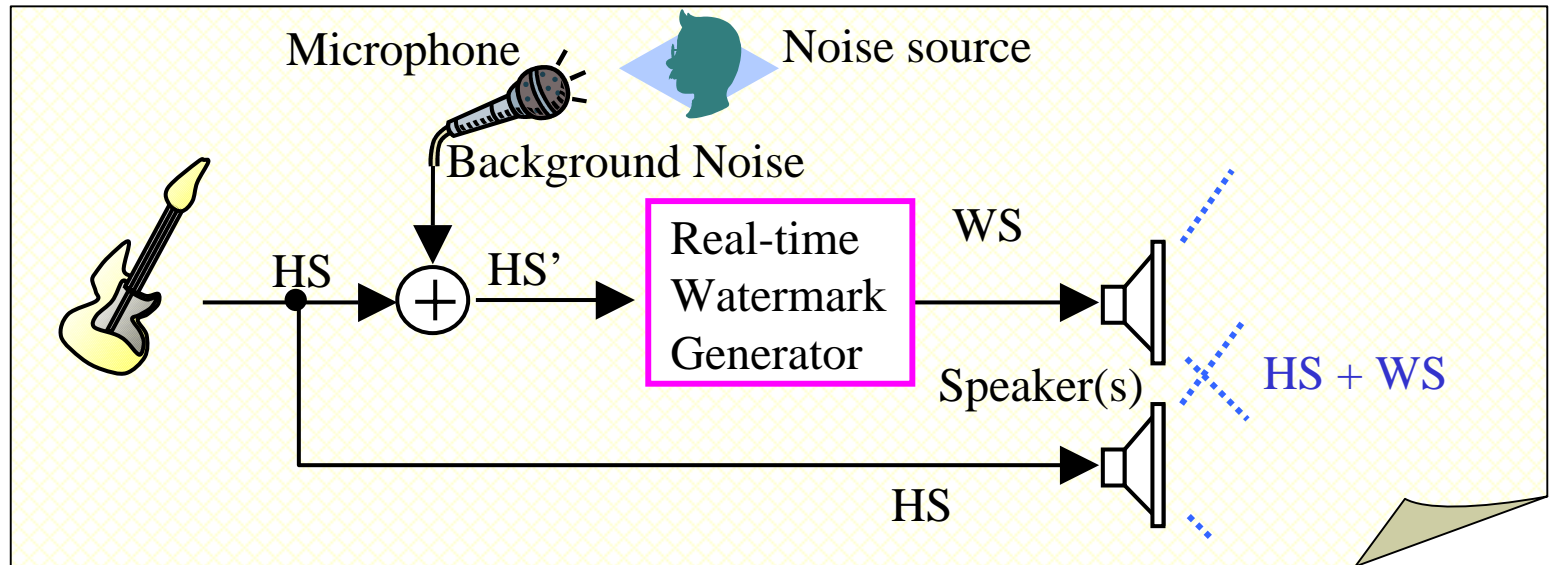
C3 : Sonic Watermarking



The sounds of WS and HS mix in the air.

- ☺ Applicability to live concerts
- ☺ Identifying *bootleg recordings* (illegal recordings by audience members)
- ☺ Mixing unnecessary
- ☹ Detection disturbed by background noises

C4a : Sonic Watermarking with Sampling (1)

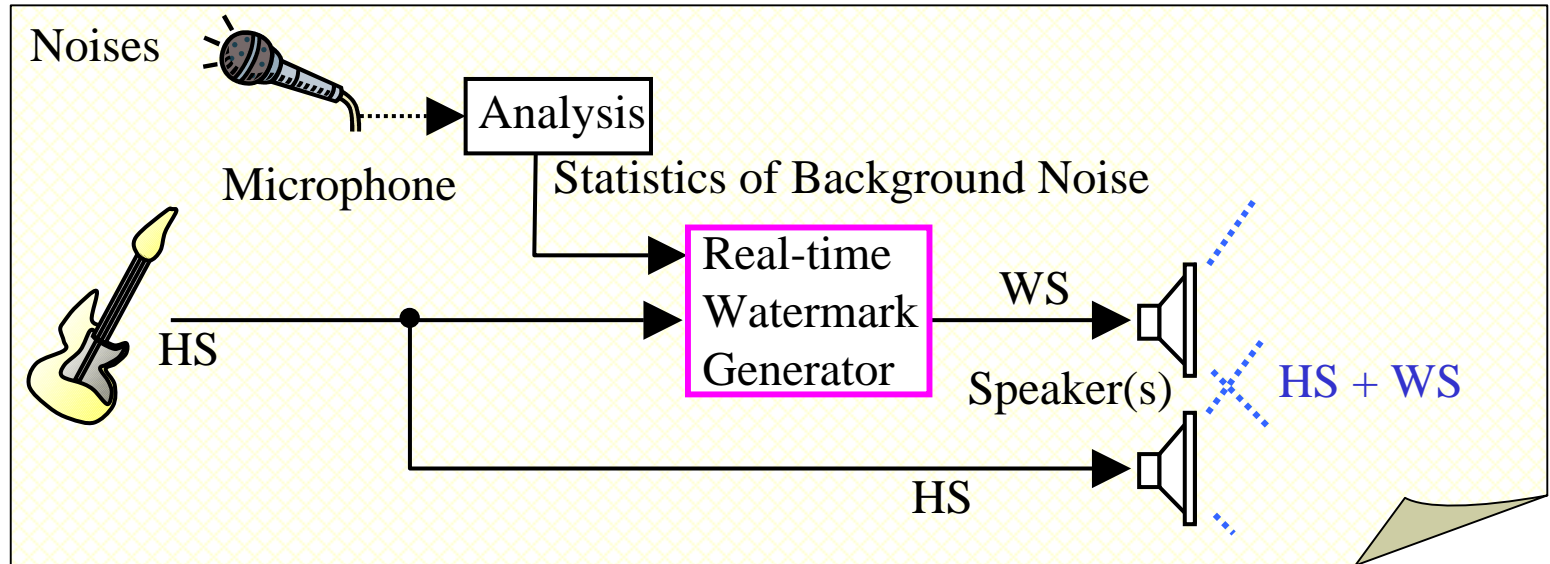


Frequency masking effect *of the noises* is also considered.

Background noises are recorded by microphone and added to HS.

- ☺ More chance of successful detection
- ☺ Dynamic response to changing the background noise
- ☹ Detection disturbed by noises made near recording devices
- ☹ Difficulty to separate the noise from the HS sound

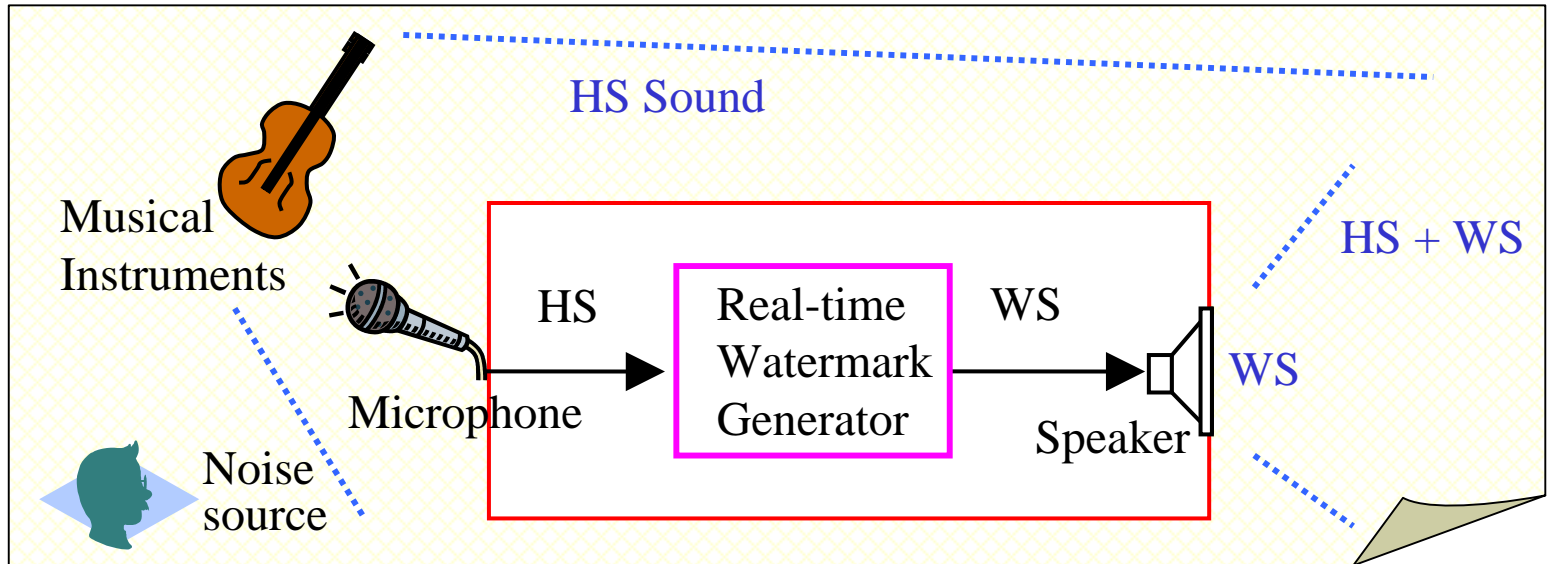
C4b : Sonic Watermarking with Sampling (2)



Frequency masking of the noises is predicted by their statistics analyzed prior to the performance.

- 😊 Separation of the HS and the background noise
- 😞 Stationary background noise assumed

C4c : Sonic Watermarking with Sampling (3)



HS sound is also recorded by the microphone.

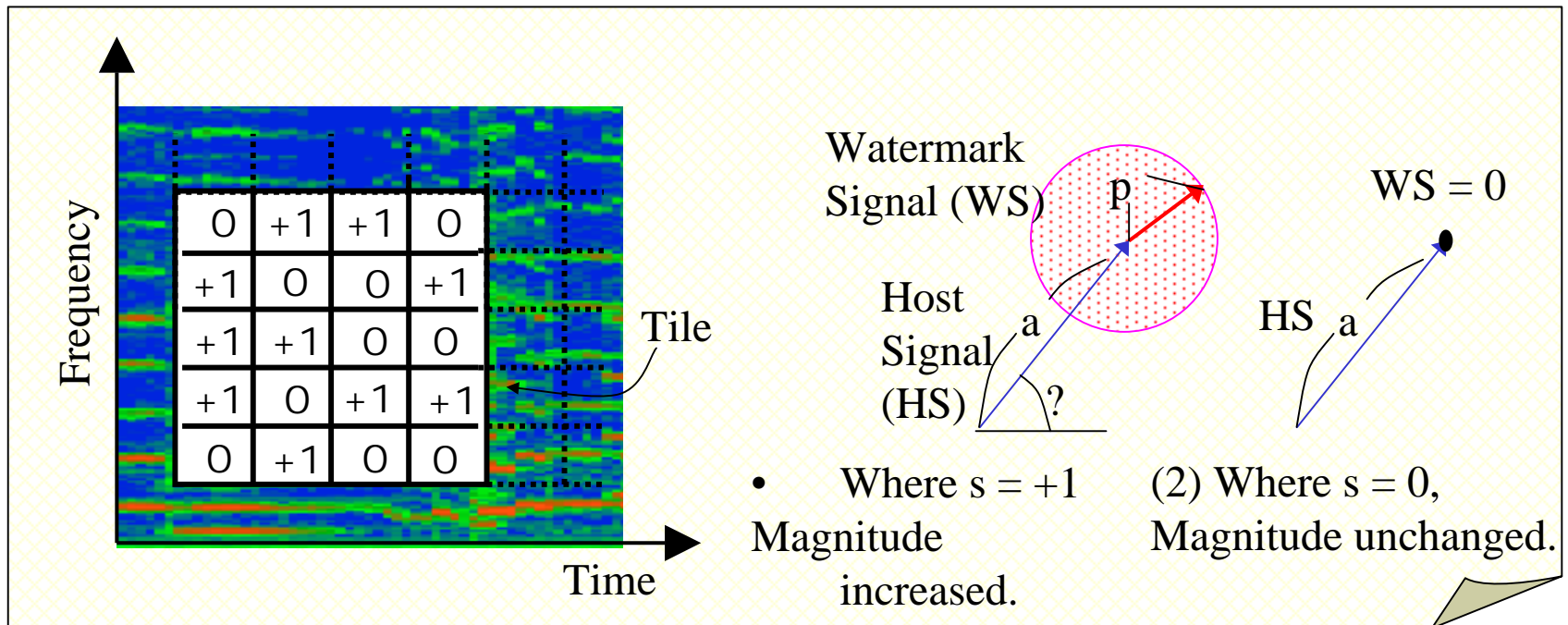
- ☺ Applicable to musical instruments without electric amplification
- ☹ Different source positions for HS sound and WS sound

Experiments

- A Spread Spectrum algorithm used
- 64-bit message embedded in 30-second music pieces
- Message encoded in 448 signal bits
 - Cyclic Redundancy Check (CRC) parity bits
 - Turbo Coding
 - Repetition
- 4 conditions compared
- “CDR” and “Mean” observed
 - CDR : the correct detection rate of the message
 - Mean : the average of detected watermark strengths for 448 signal bits
- *Stronger WS* required than file embedding
 - Root mean square power of WS is -20.4dB relative to that of HS

Watermark Generation

- Produce non-uniformity of the magnitude distribution in the frequency domain using short-term DFT
 - Increase the magnitudes in the “+1” tiles
 - Leave the magnitudes in the “0” tiles unchanged



Detection

- Calculate the logarithmic magnitude for all tiles by DFT
- Correlates them with the pseudo-random array

$$X = \frac{\sum_{\text{assigned tile}} \varpi_{\text{tile}}(Q_{\text{tile}} - \bar{Q})}{\sqrt{\sum_{\text{assigned tile}} \{\varpi_{\text{tile}}(Q_{\text{tile}} - \bar{Q})\}^2}}$$

X : Detected Watermark Strength

ϖ_{tile} : Pseudo-Random Number

Q_{tile} : Magnitude of a tile

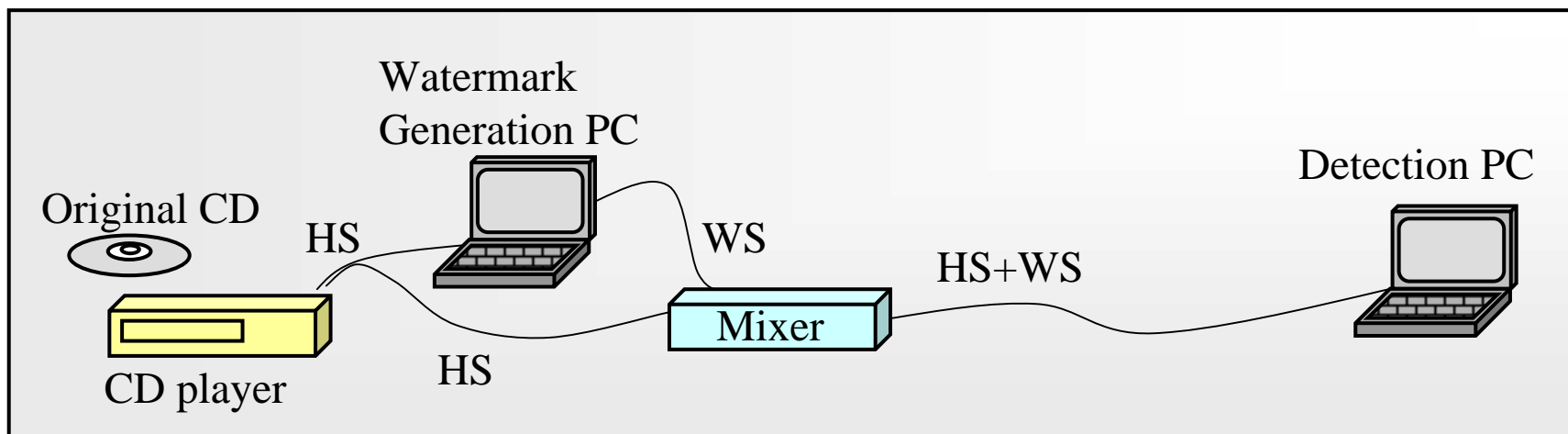
\bar{Q} : Average of Q_{tile}

Condition1 : File Embedding

- Non-real-time embedding
- HS and WS digitally mixed on a computer
- WS delayed to HS by 512 samples

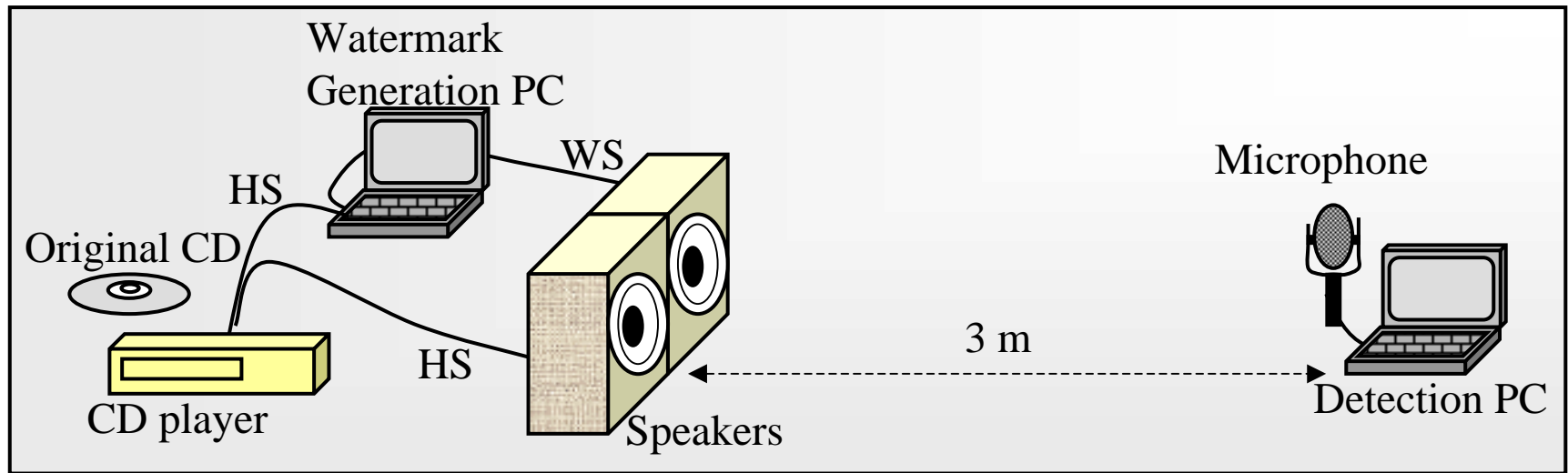
Condition2 : Analogue Watermarking

- Real-time embedding
- HS and WS mixed by an analogue mixer



Condition3 : Sonic Watermarking

- HS and WS mixed in the air
- A soundproof room with a 30 dB(A) background noise
 - dB(A) is a unit for A-weighted sound level.

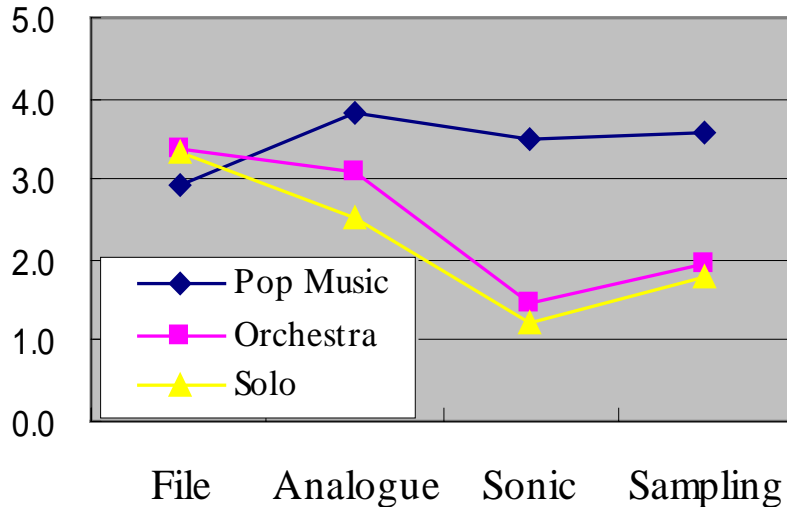


Condition4 : Sonic Watermarking with Sampling(2

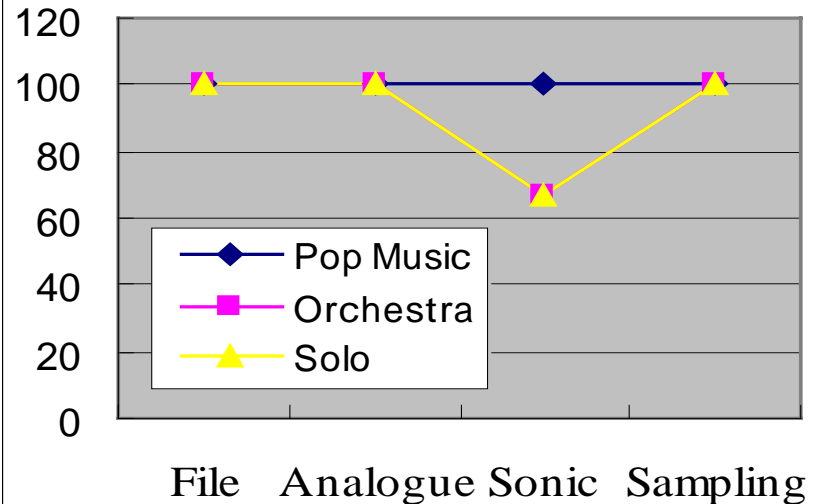
- Stationary background noise analyzed prior to the HS playback
- WS increased assuming the existence of the background noise

Result

Detected Watermark Strength

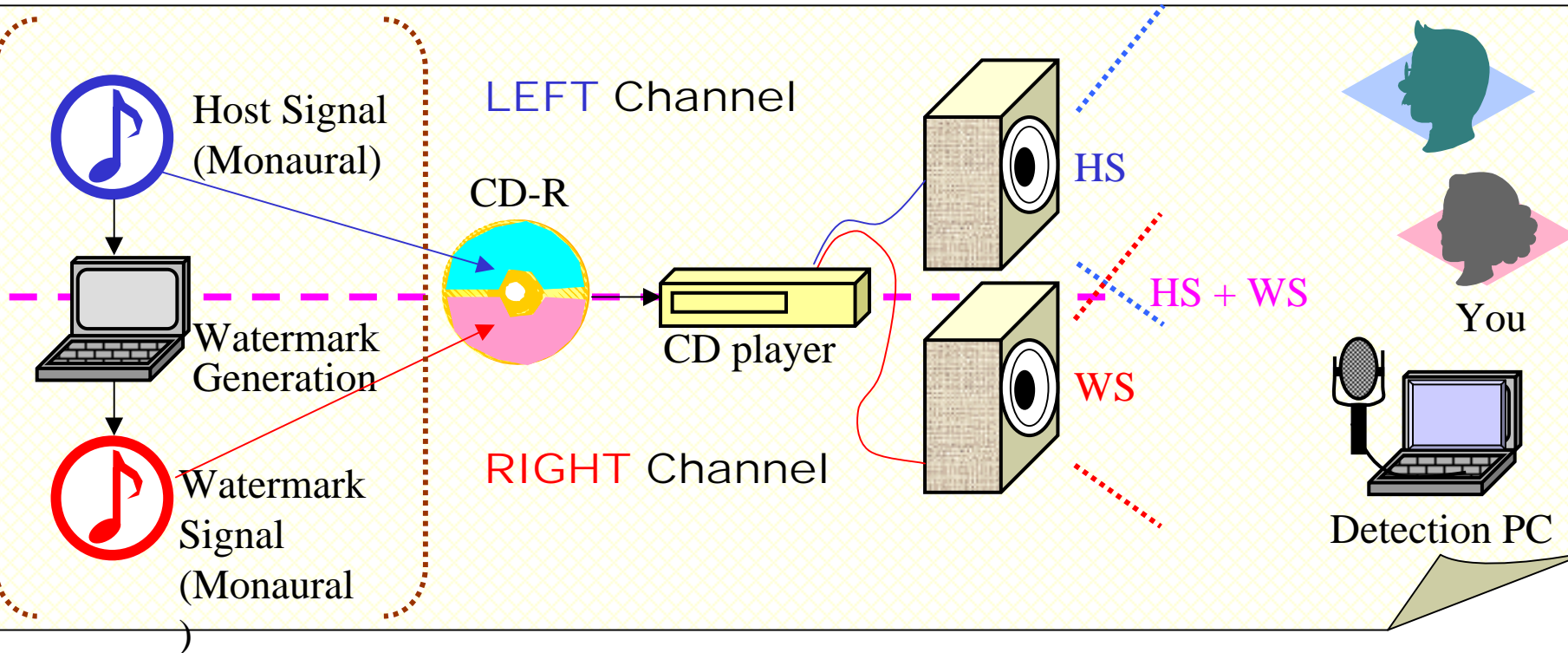


Correct Detection Rate



- Weak strength for Orchestras and Solo instruments
 - Because of the existence of 30dB(A) background noise
- Strength recovered by Sampling
- A little hoarse sound when HS played

Demonstration of Non-Real-Time Sonic Watermarking

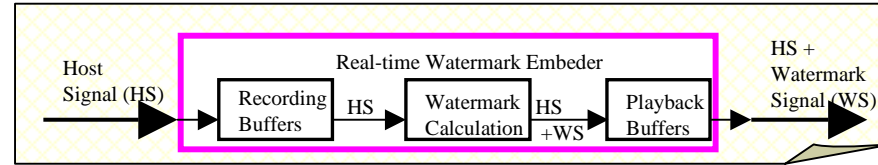


- Burned HS in the left, WS in the right of a CD
- Delayed WS to HS by 819 samples (18.5ms)
- Mix sounds of both channels in the air

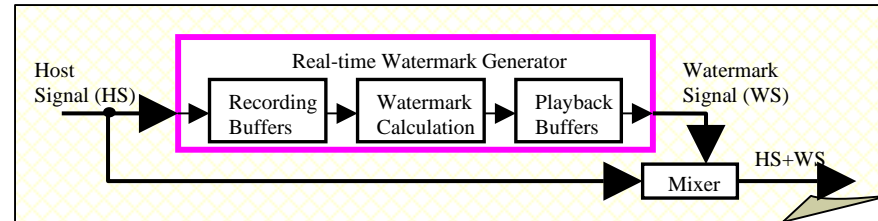
Conclusion

- Easier usages and new application scenarios of Audio Watermarking using Real-Time Embedding
- 4 groups of composition methods
- Experimental results supporting feasibility of the composition methods
- More robustness and transparency necessary for practical use

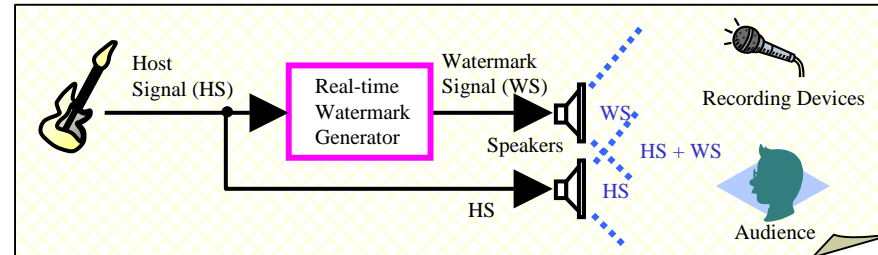
C1 : Naï ve Composition



C2 : Analogue Watermarking



C3 : Sonic Watermarking



C4c : Sonic Watermarking with Sampling (3)

