

# Improving Audio Watermarking Robustness Using Stretched Patterns Against Geometric Distortion

Ryuki Tachibana

Tokyo Research Lab., IBM Japan

# Digital Watermarking

A means to embed a message to digital multimedia content

Applications;

- Device control
- Searching of illegal content
- Fingerprinting
- Broadcast Monitoring

Requirements;

- Robustness
- Security
- Reliability
- Capacity
- Transparency

Robustness against *Geometric Distortion* is a big issue.

# Geometric Distortion

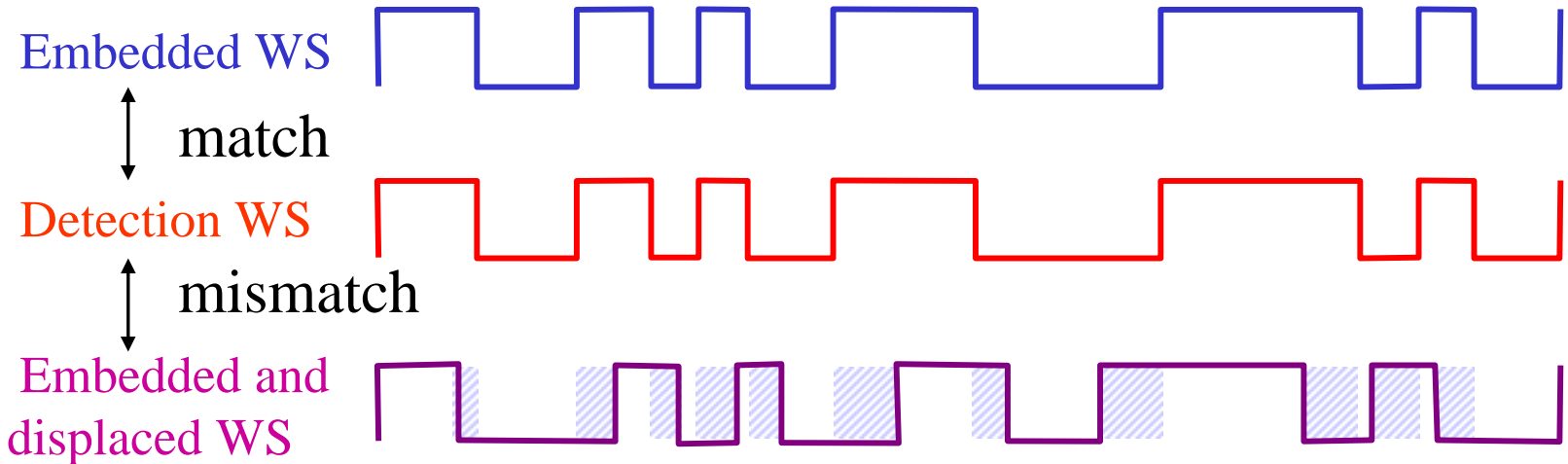
- Examples

Rotation, scaling, translation,  
Stirmark random bending, etc



- Why is it **difficult**?

Spread spectrum watermarking requires matching of embedded watermark signal (WS) and detection WS.



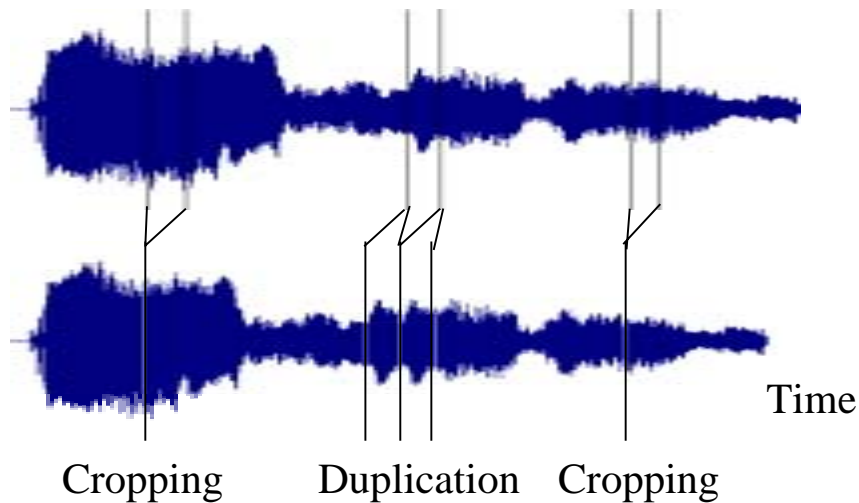
- What is geometric distortion of audio?

Time scaling, pitch shifting, and wow-and-flutter

# Geometric Distortions of Audio

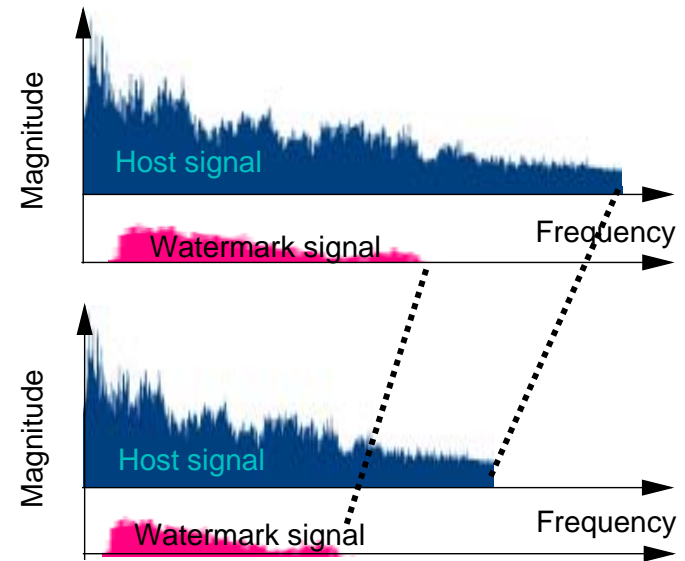
## □ Time scaling

- Changes the length of content by duplication and cropping
- Causes mismatch along time



## □ Pitch shifting

- Changes the pitch of content by linear interpolation
- Causes mismatch along frequency

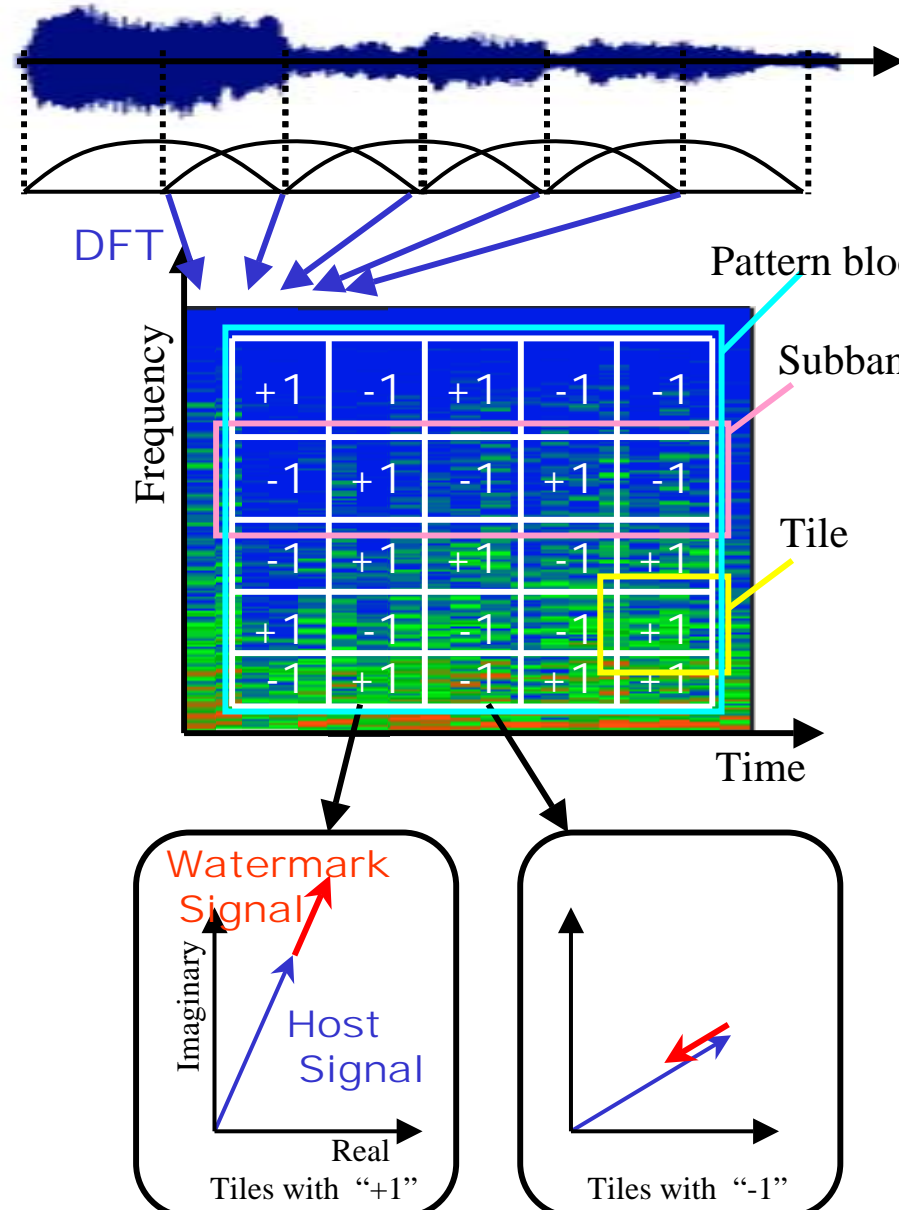


✗ Sample-by-sample exhaustive search costs a lot

# Previous Method (Embedding)

- Perform time-frequency analysis by short-term DFTs
- Assign two-dimensional pseudo-random array
- Increase or decrease magnitudes
- Convert to the time domain by inverse DFTs

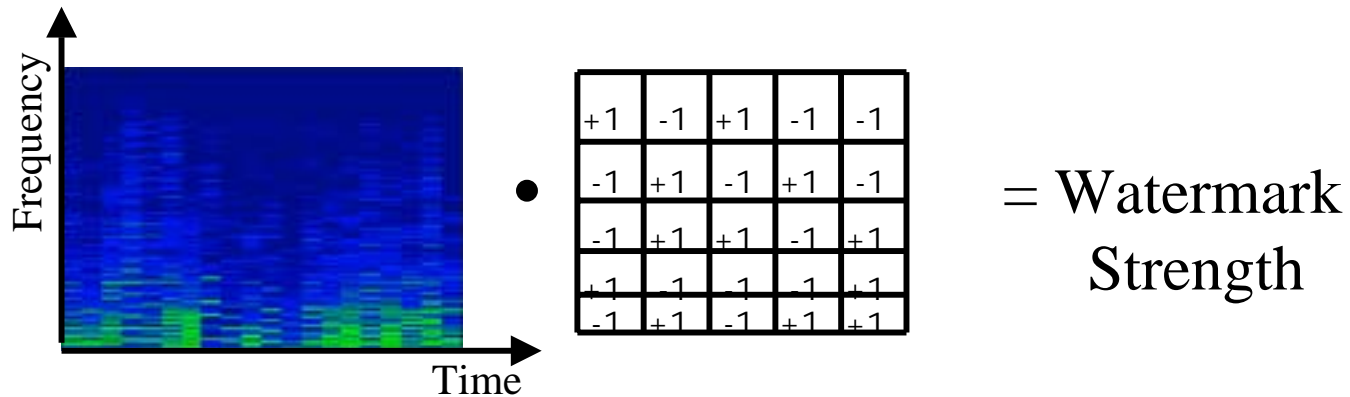
Non-uniformity in magnitude distribution of content



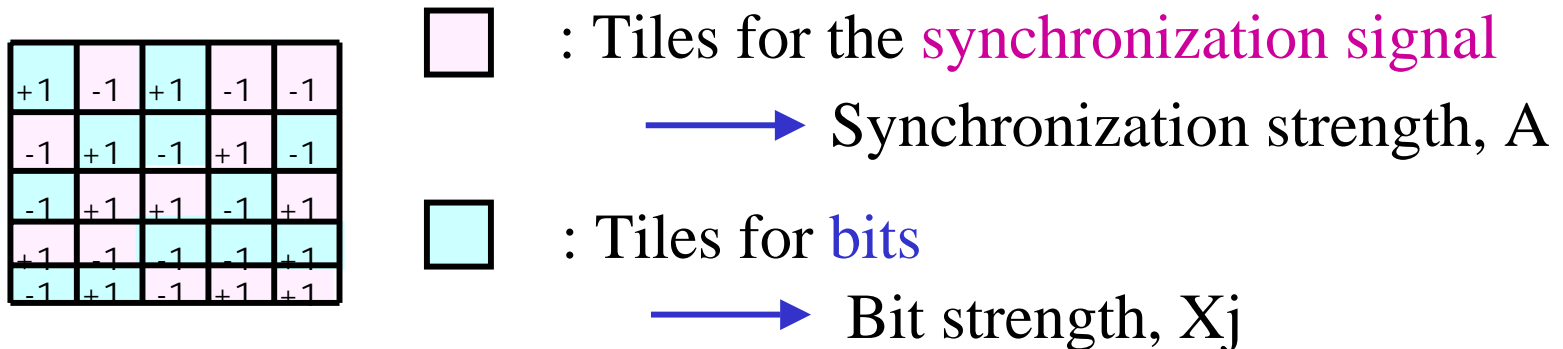
# Previous Method (Detection)

Detects the non-uniformity of the magnitude distribution of content.

- Correlation with the pseudo-random array

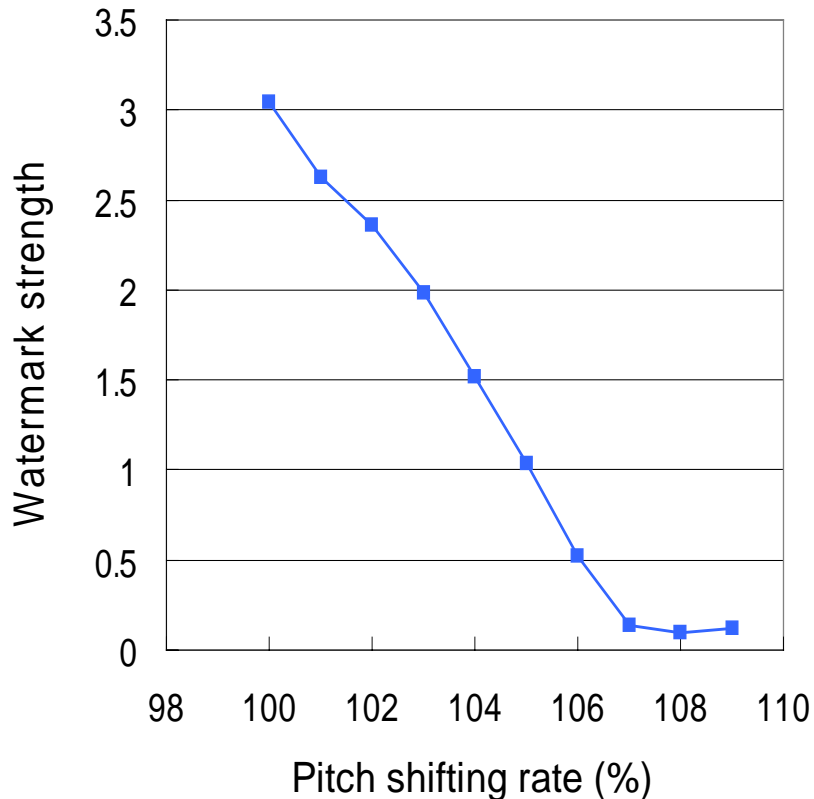


- Synchronization necessary to find the first frame of a block



# Results of the Previous Method

- Robustness against pitch shifting and random stretching up to  $\pm 4\%$  was achieved



Time scaling 104%	100%
Time scaling 102%	100%
Time scaling 98%	100%
Time scaling 96%	100%
Pitch shifting 104%	90%
Pitch shifting 102%	100%
Pitch shifting 98%	100%
Pitch shifting 96%	83%

Correct detection rate of 64-bit message

- Detection from more heavily distorted content was impossible

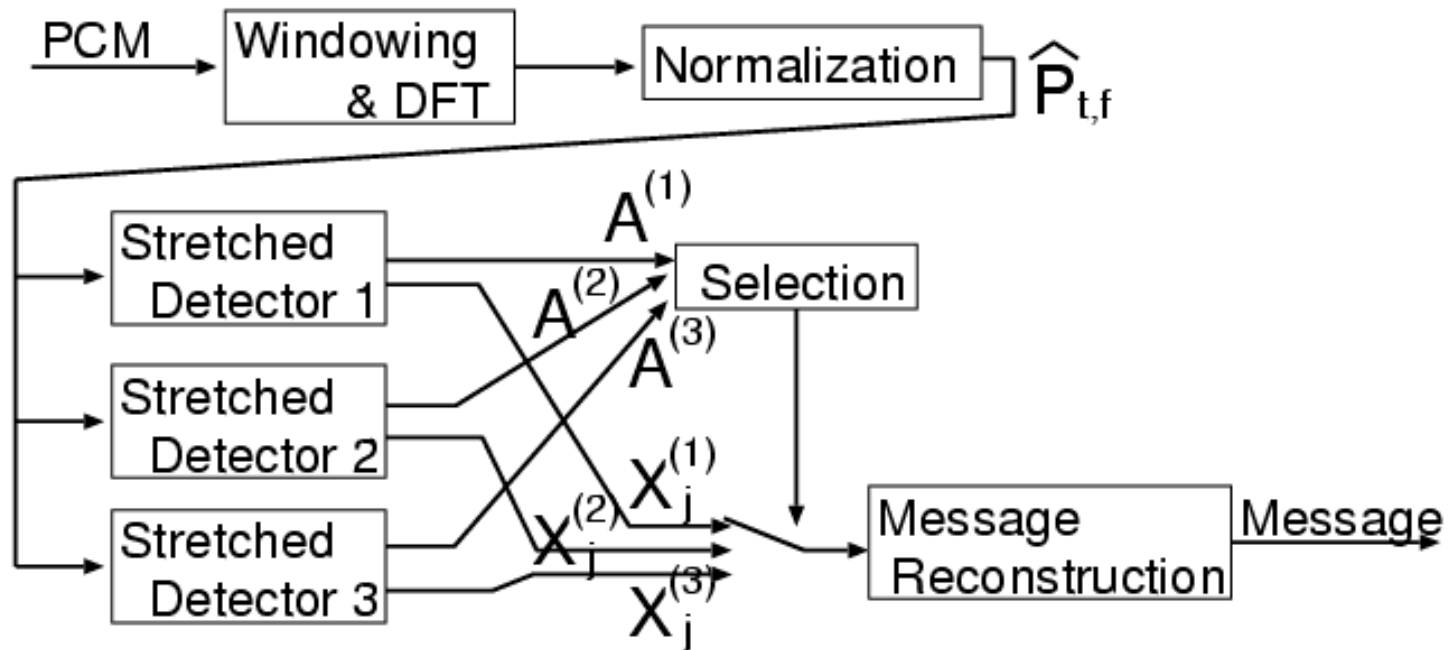
# Basic Concept

To detect watermark from heavily distorted content, use multiple *Stretched Detectors*,  $D(r_t, r_f)$ , in parallel.

A *Stretched Detector* uses a pattern stretched at;

- Time expansion rate  $r_t$
- Frequency shifting rate  $r_f$

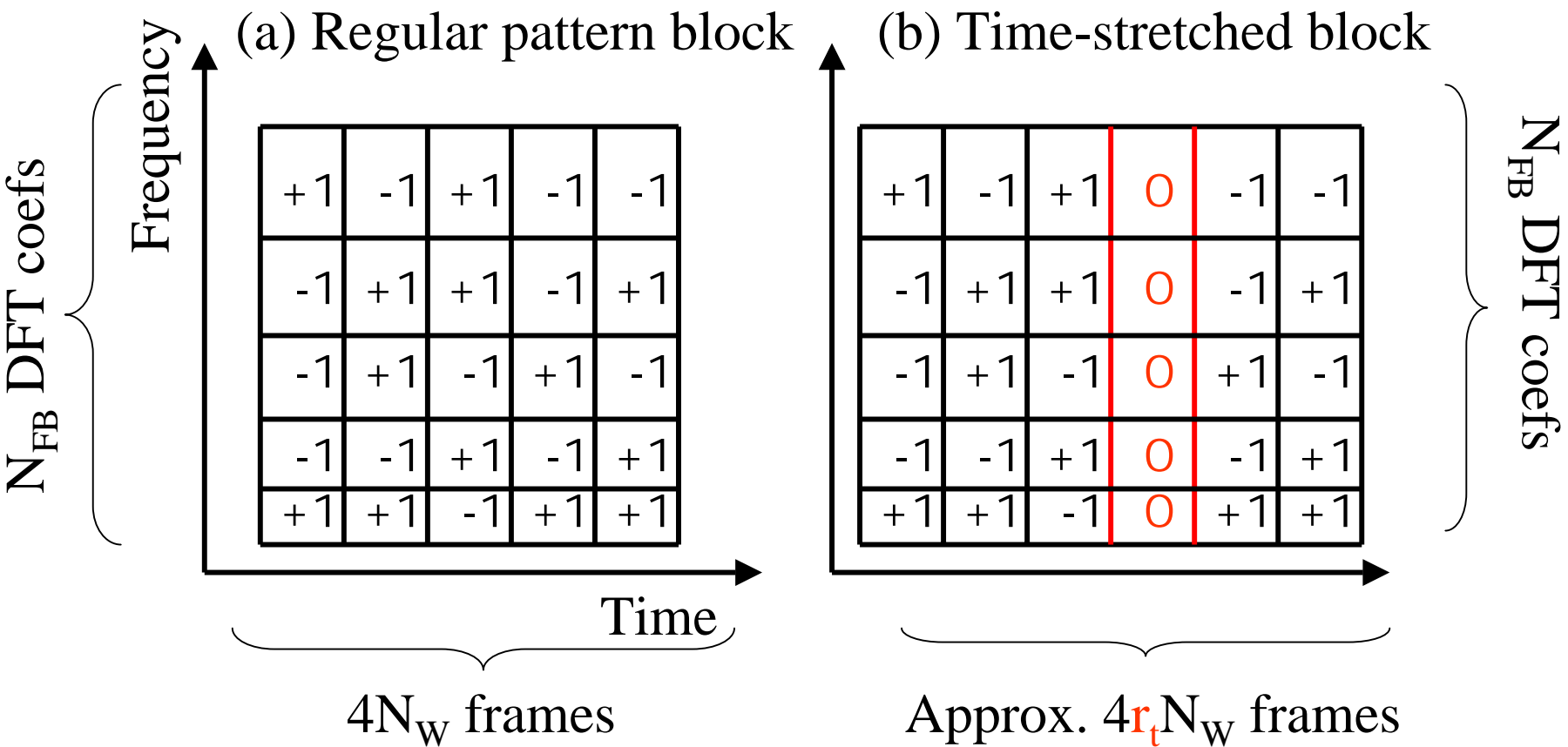
# Detection Flow Using Multiple Stretched Detectors



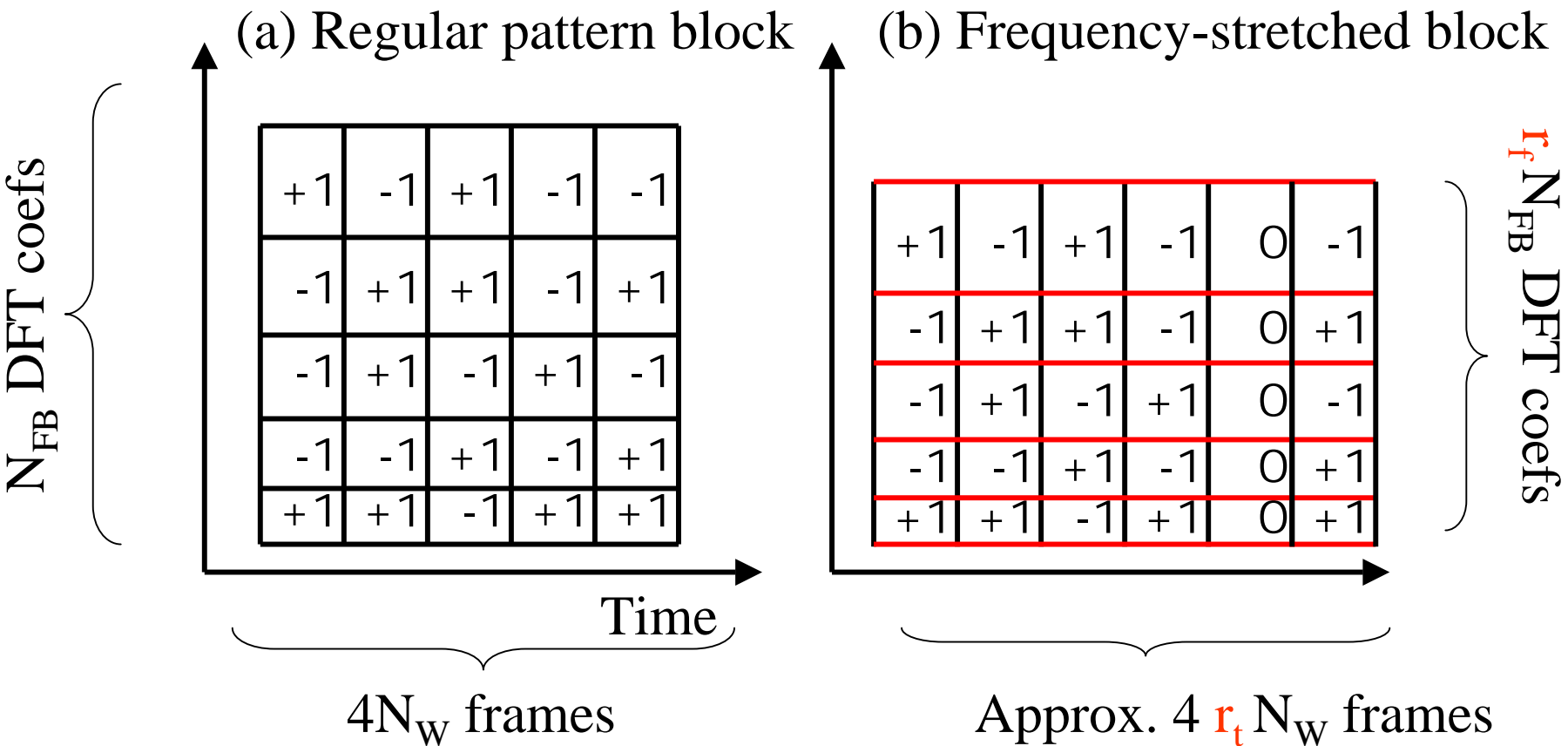
$A^{(i)}$  : Sync. Strength for i-th stretched detector

$X_j^{(i)}$  : Bit Strength of j-th bit for i-th stretched detector

# Stretched Pattern Block for Time Scaling



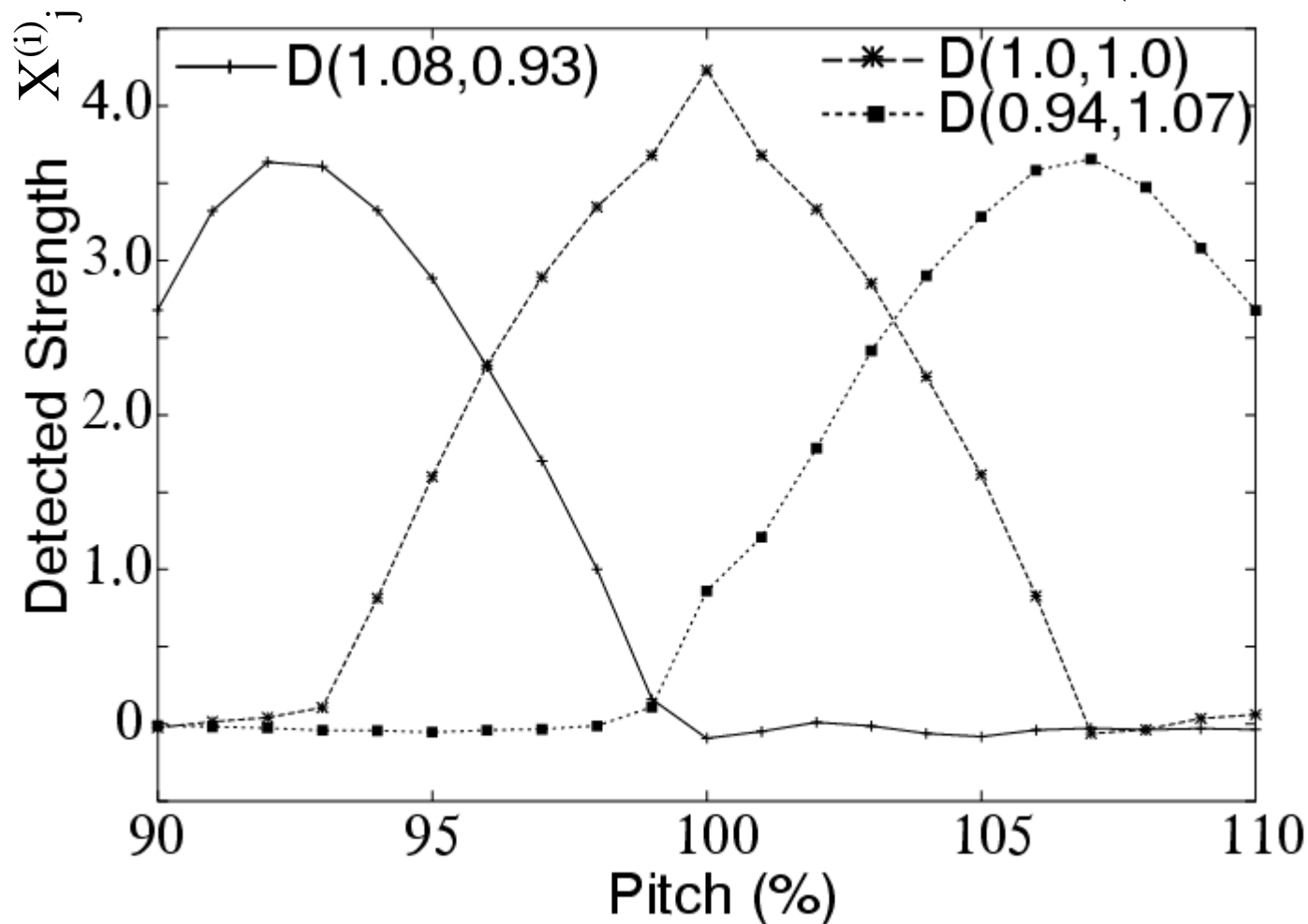
# Stretched Pattern Block for Pitch Shifting



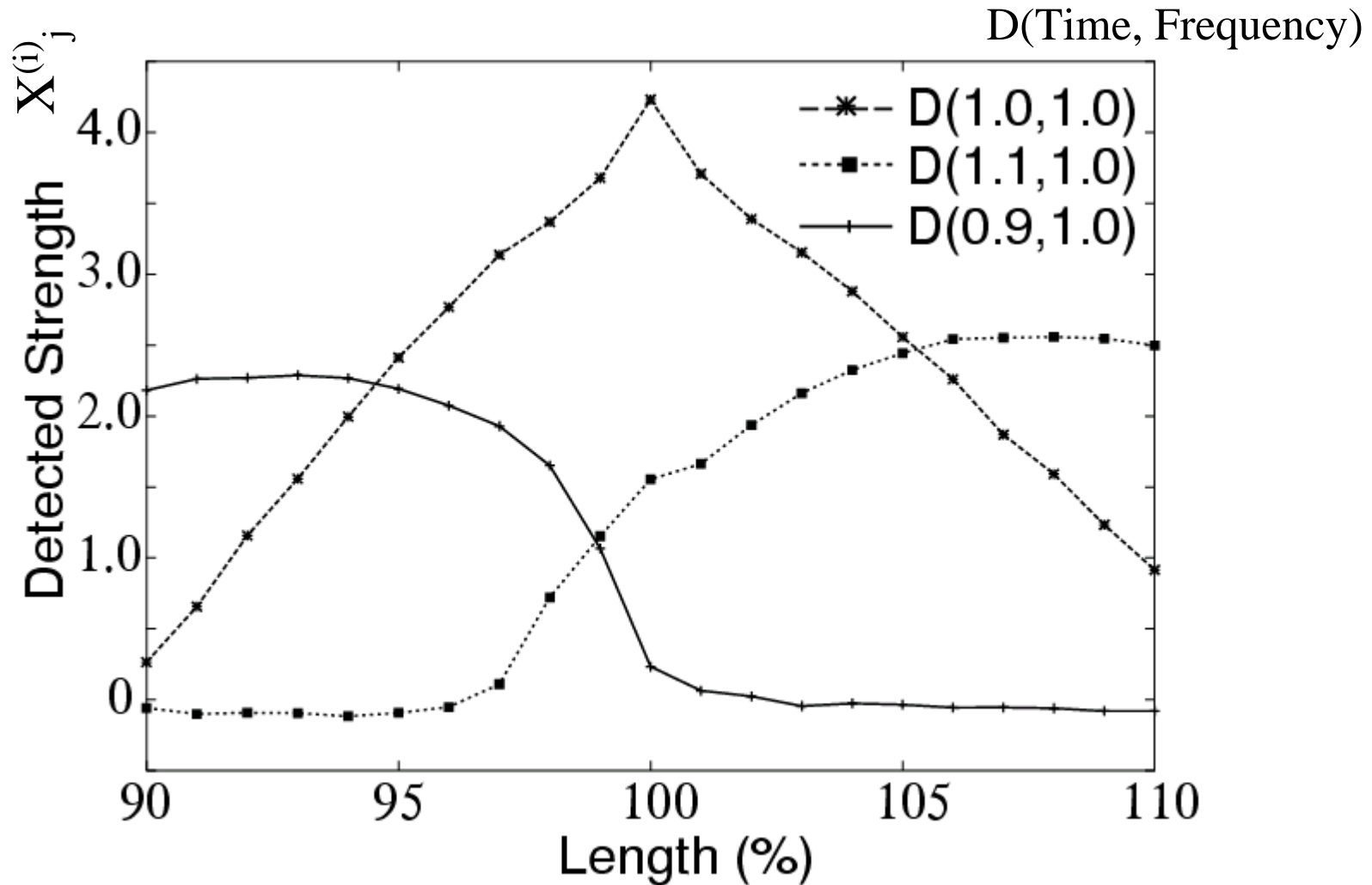
Where  $r_t = 1/r_f$

# Detected Bit Strength After Pitch-Shifting

$D(\text{Time}, \text{Frequency})$

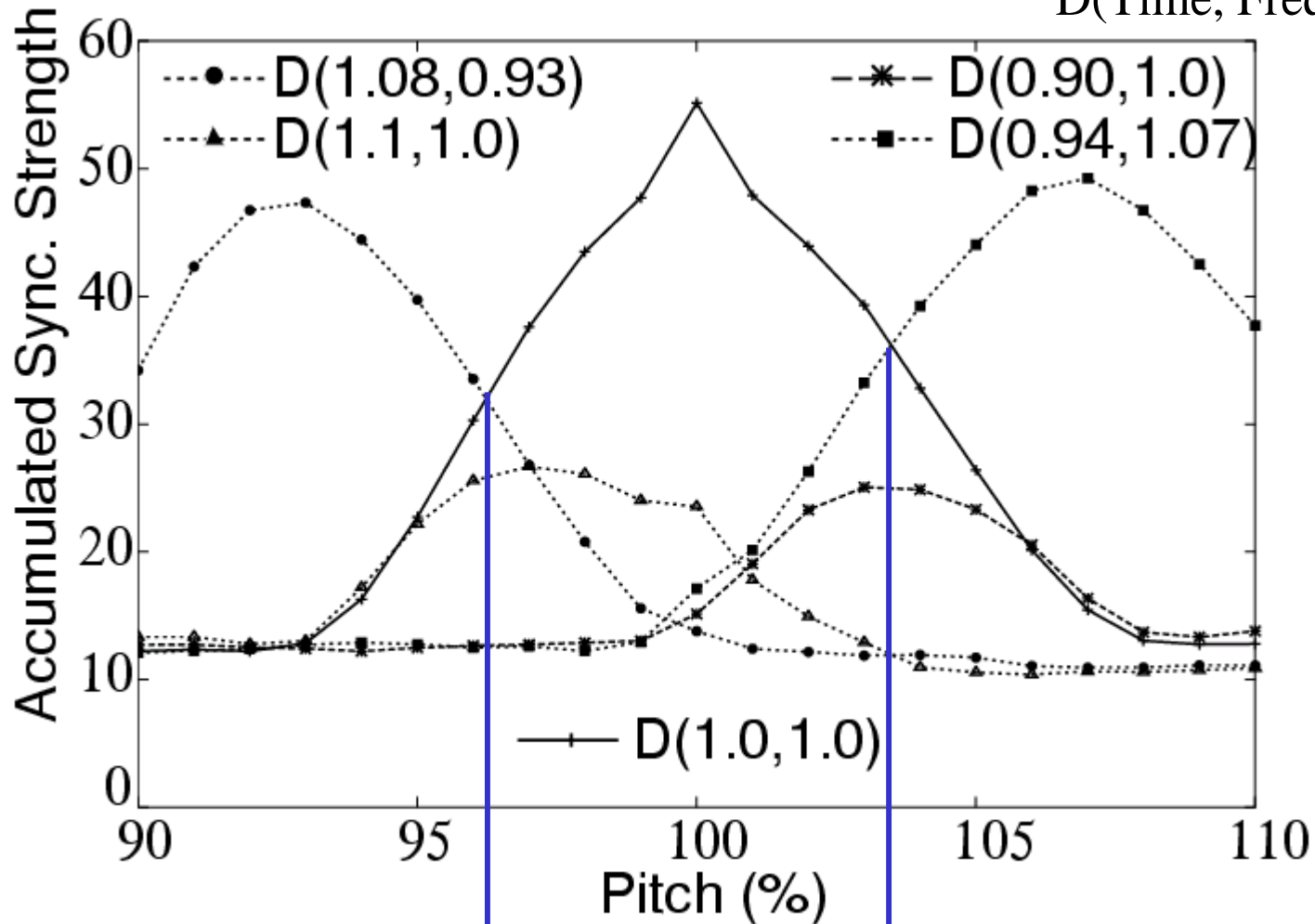


# Detected Bit Strength After Time Scaling



# Synchronization Strength After Pitch-Shifting

$A^{(i)}$  D(Time, Frequency)



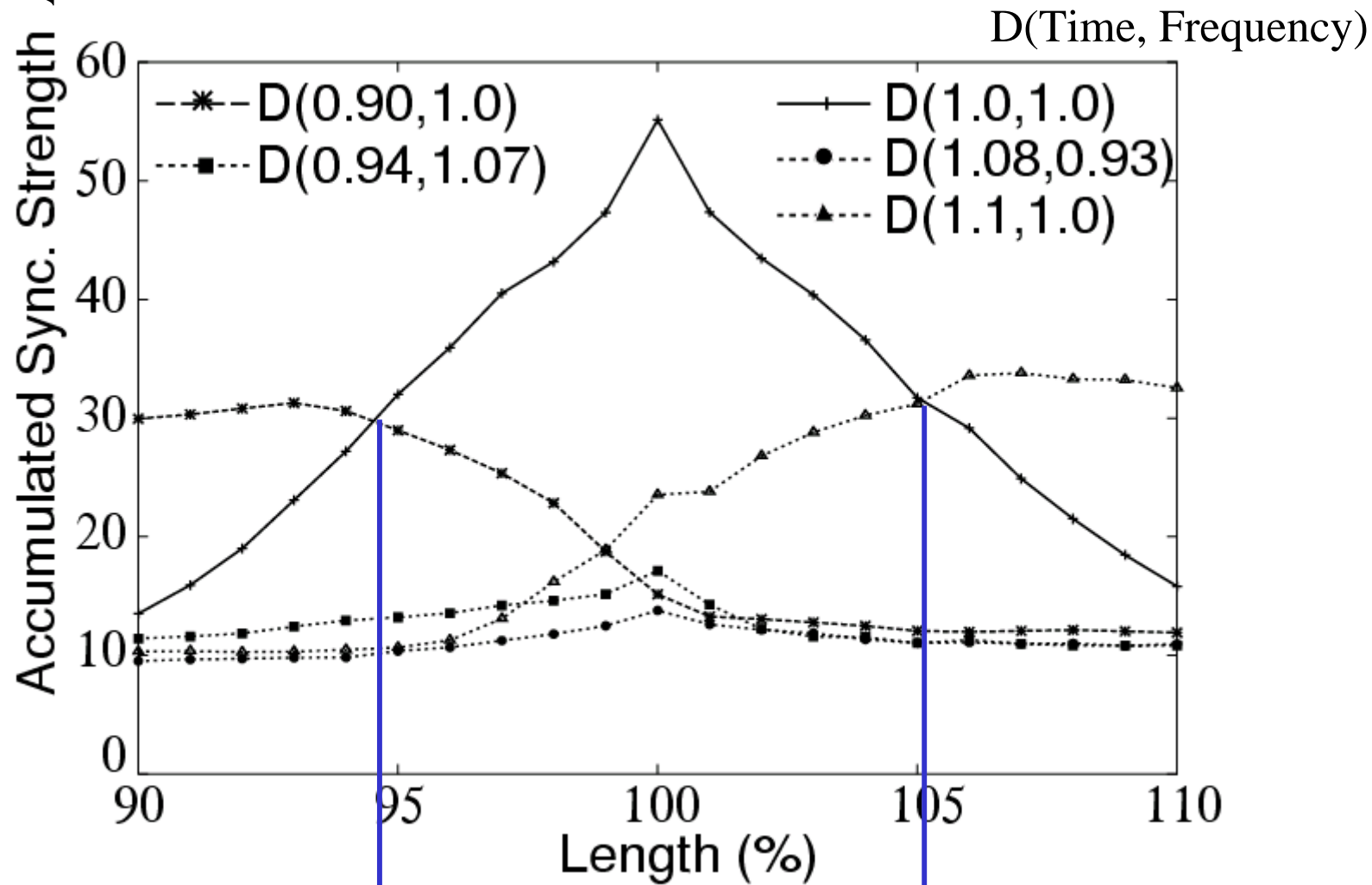
Selected  
Stretched  
Detector :

$D(1.08, 0.93)$

$D(1.0, 1.0)$

$D(0.94, 1.07)$

# Synchronization Strength After Time Scaling



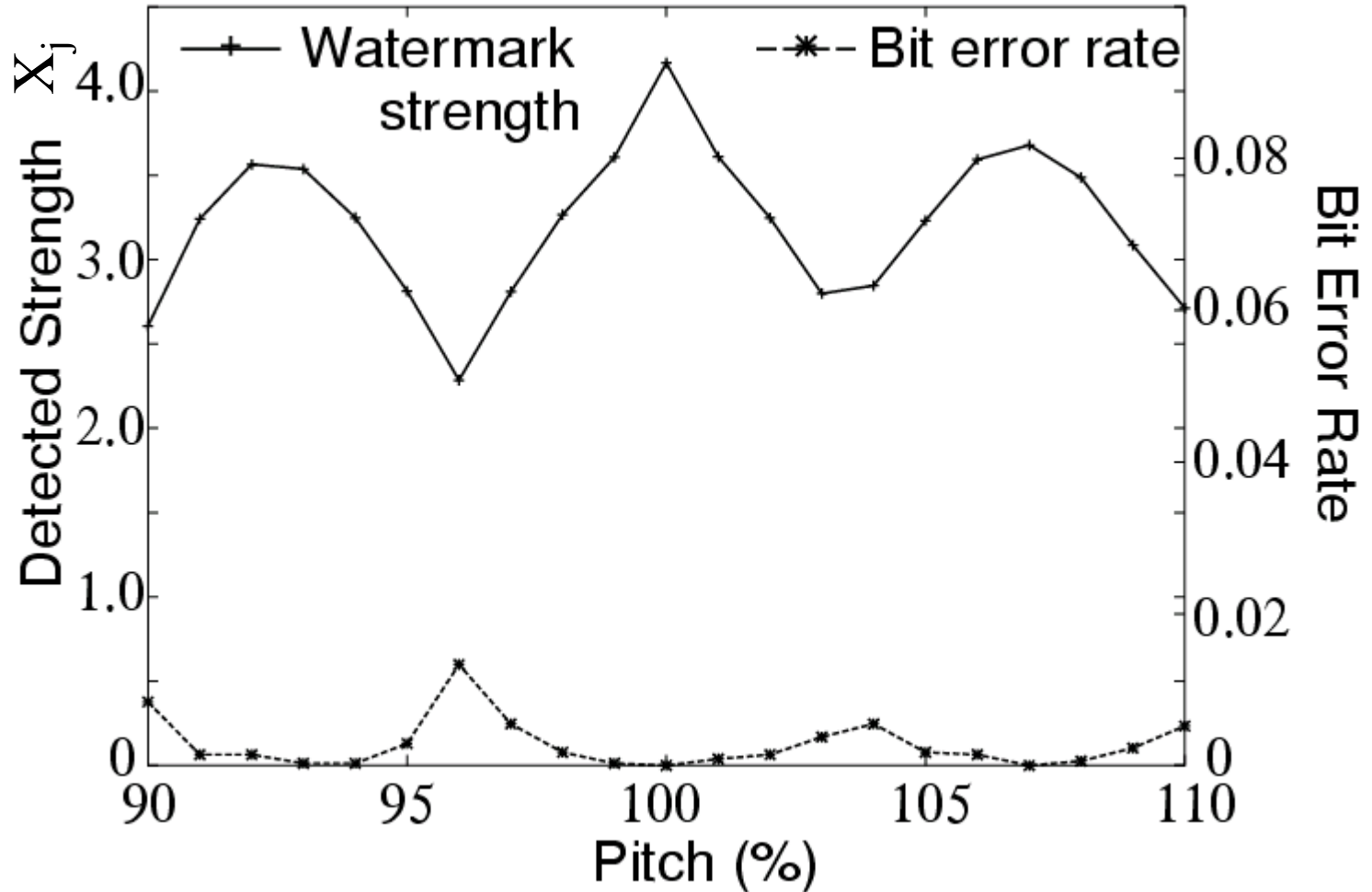
Selected  
Stretched  
Detector :

$D(0.90, 1.0)$

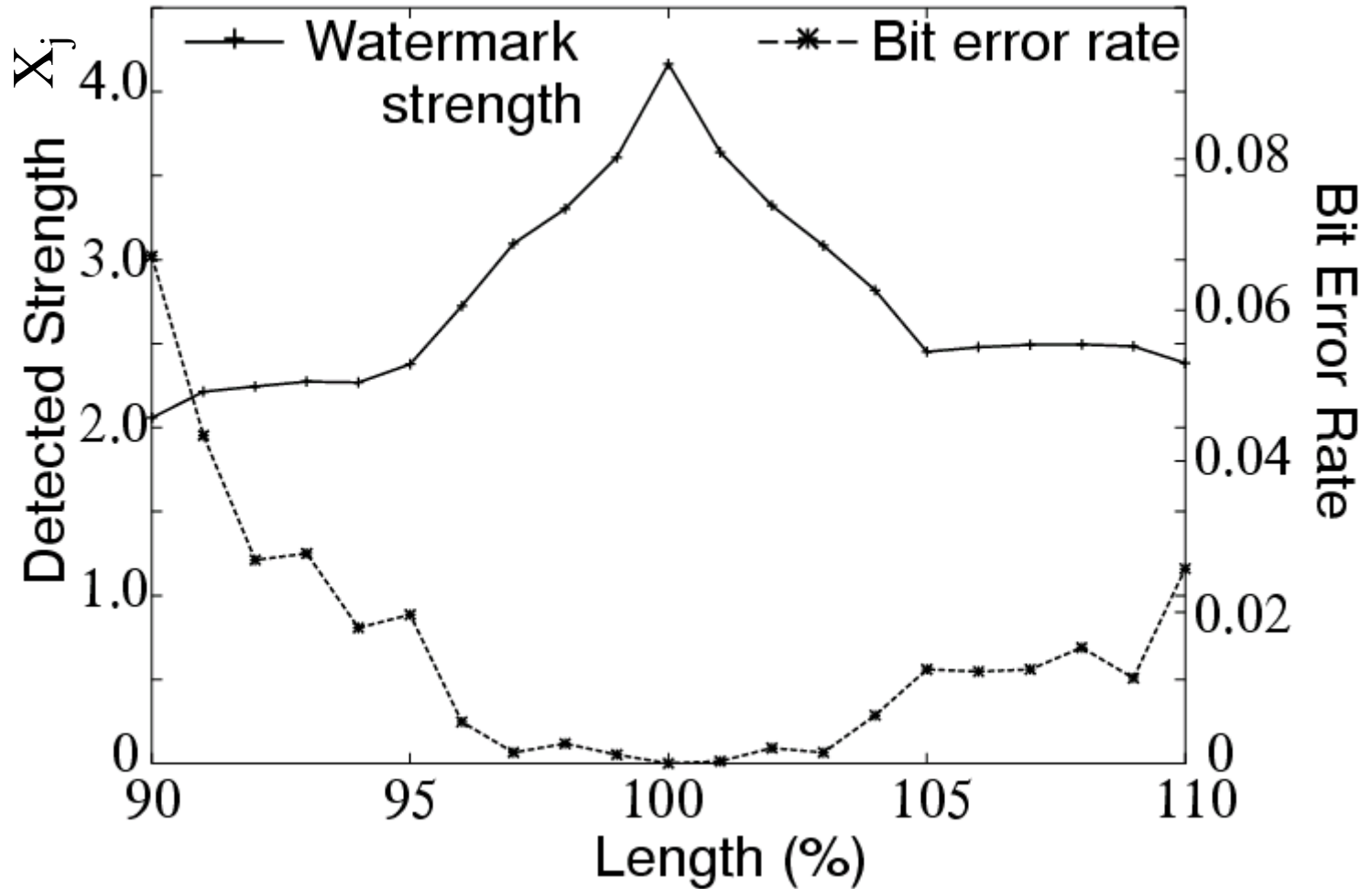
$D(1.0, 1.0)$

$D(1.1, 1.0)$

# Detected Bit Strength of *Selected* Stretched Detector After Pitch-Shifting



# Detected Bit Strength of *Selected* Stretched Detector After Time Scaling



# Robustness Against Pitch Shifting

Processing		Mean	BER	CDR
Pitch shifting	-10%	2.61	0.008	96%
Pitch shifting	-8%	3.56	0.001	100%
Pitch shifting	-6%	3.25	0	100%
Pitch shifting	-4%	2.29	0.013	90%
Pitch shifting	-2%	3.27	0.002	100%
Original watermark	0%	4.17	0	100%
Pitch shifting	2%	3.25	0.001	100%
Pitch shifting	4%	2.85	0.005	100%
Pitch shifting	6%	3.59	0.001	100%
Pitch shifting	8%	3.49	0.001	100%
Pitch shifting	10%	2.71	0.005	100%

CDR : Correct Detection Rate of 64-bit message

BER : Bit Error Rate for each bit

Mean : Detected Strength

# Robustness Against Time Scaling

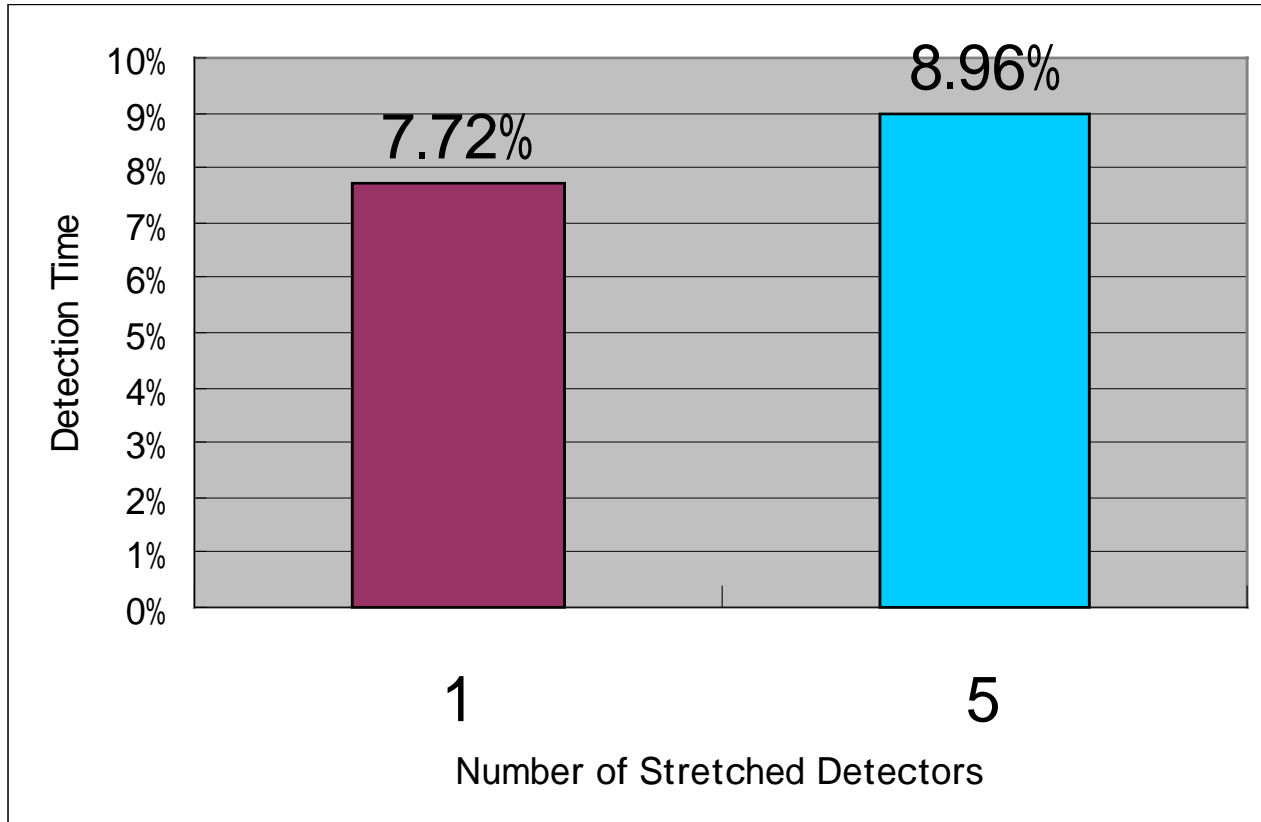
Processing		Mean	BER	CDR
Time scaling	-10%	2.06	0.067	83%
Time scaling	-8%	2.25	0.027	87%
Time scaling	-6%	2.27	0.018	87%
Time scaling	-4%	2.72	0.005	100%
Time scaling	-2%	3.3	0.003	100%
Time scaling		4.17	0	100%
Time scaling	2%	3.32	0.002	100%
Time scaling	4%	2.82	0.006	100%
Time scaling	6%	2.48	0.012	93%
Time scaling	8%	2.49	0.015	93%
Time scaling	10%	2.38	0.026	87%

CDR : Correct Detection Rate of 64-bit message

BER : Bit Error Rate for each bit

Mean : Detected Strength

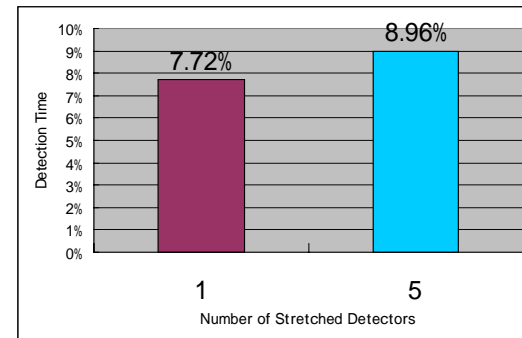
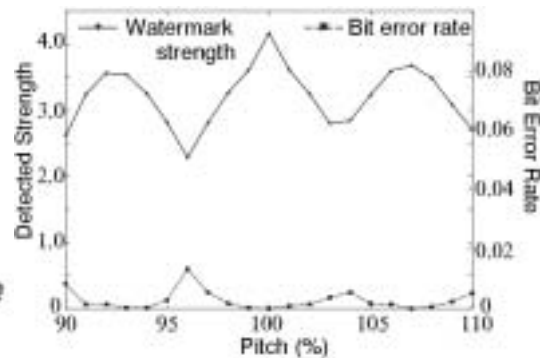
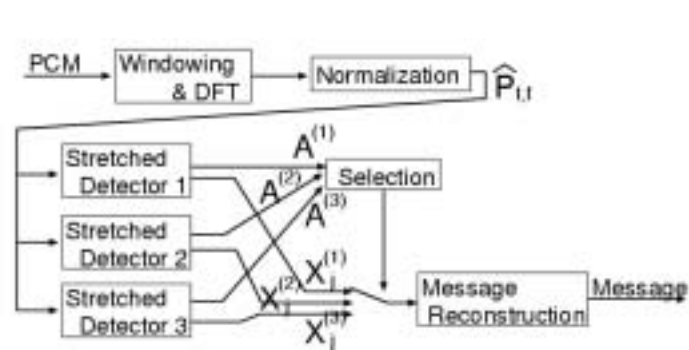
# Computational Cost for Detection



On 600 MHz Pentium III running Windows NT.  
Increase of detection time is limited,  
because DFT is shared for stretched detectors.

# Conclusion

- The previous method based on magnitude modification was robust to geometric distortion to some extent.
- Use of multiple *Stretched Detectors* make it possible to detect watermark from more heavily distorted content.
- Additional computational cost by multiplicity of stretched detectors is limited.



*Thank you!*