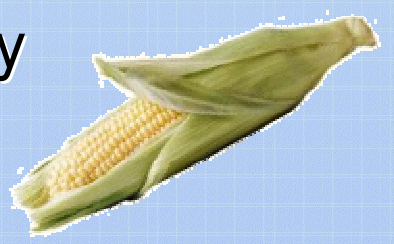
Audio Watermarking

Adam Brickman
abrick@mail.utexas.edu
Department of Electrical & Computer Engineering
The University of Texas at Austin
March 5, 2003

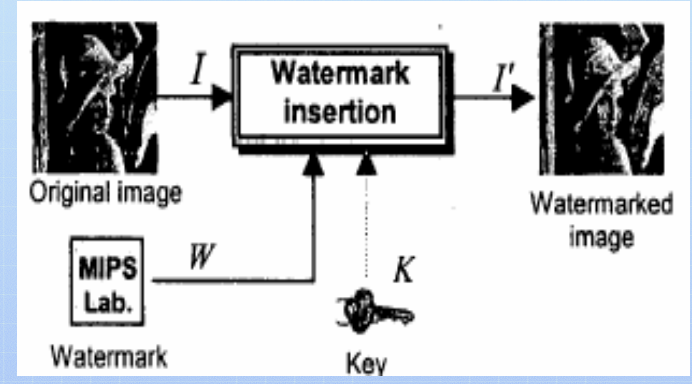
What is Audio Watermarking?

- What
  - Insertion of information into an audio file
  - May be done *audibly* or *inaudibly*
  - May be done in time or frequency
- Why
  - Copyright protection
  - Control or hidden description
  - Covert communications
- Challenge: Ear sensitivity



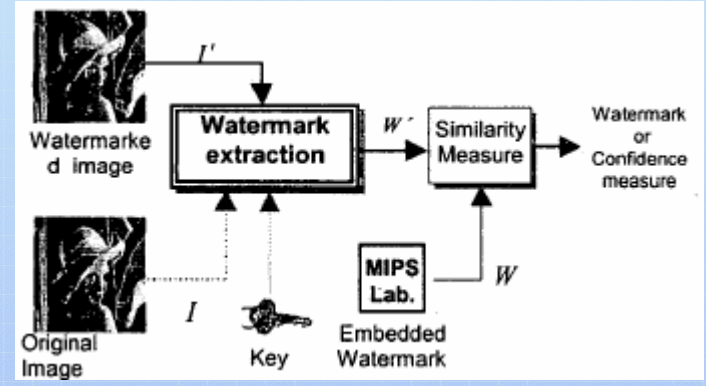
Basic Method

Insertion



$$A' = A + f(A, W)$$

Extraction



$$sim(w, w^*) = \frac{w' \cdot w}{\sqrt{w' \cdot w'}}$$

Compare to threshold T

Previous Method #1

- Independent Multiband Phase Modulation [Kuo *et al*, 2002]
  - Alter phase in limited amount
  - Covert (non-blind)
  - Intellectual property protection

Previous Method #1

- Insertion in *time*,
- Modulation in *frequency*
- Time Window

$$win(n) = \sin\left(\frac{\pi(n+0.5)}{N}\right), 0 \le n \le N-1$$

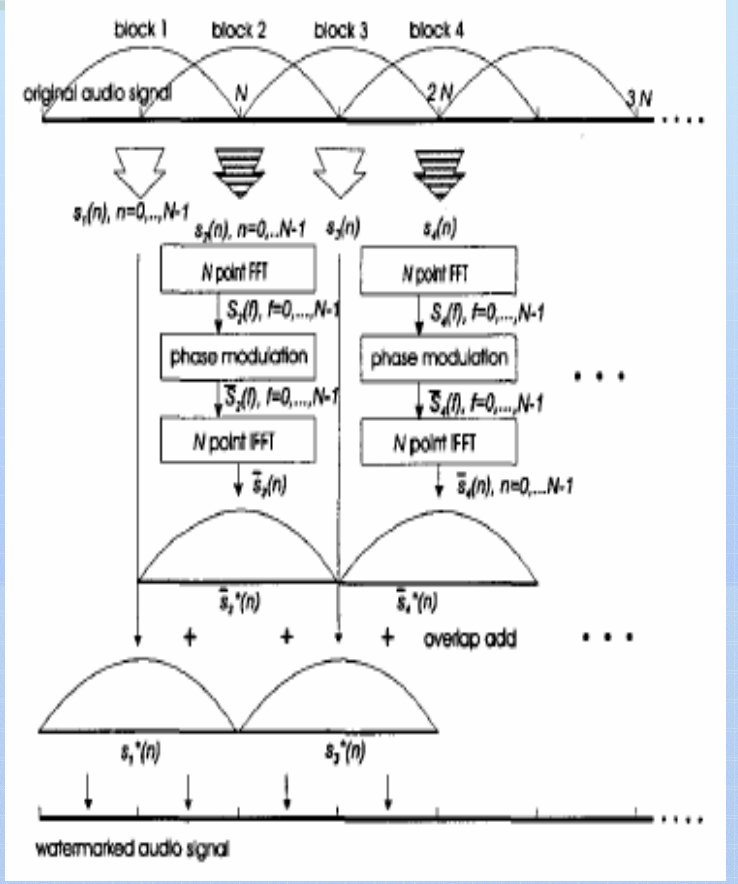
- Phase Window

$$\phi(b) = \sin^2\left(\frac{\pi(b+1)}{2}\right), -1 \le b \le 1$$

$$\left|\frac{d\phi}{db}\right| < 30°$$

- Phase modulation

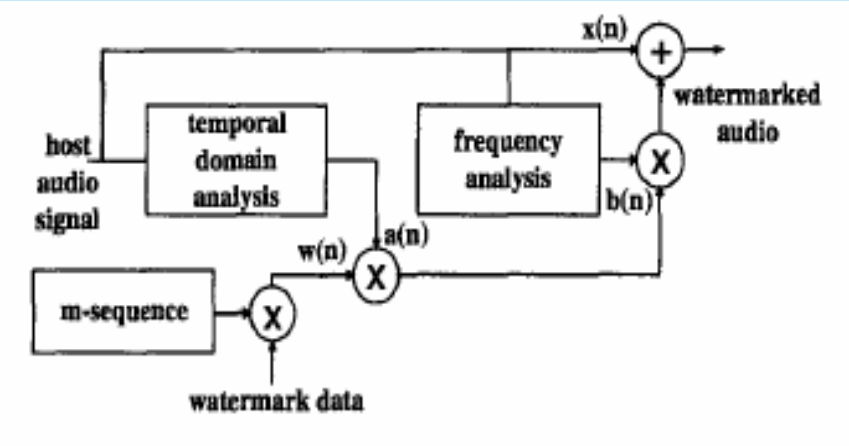
$$\Phi_k(b) = \sum_{i=1}^{I} a_i \phi(b-i), 0 \le b \le I$$



Previous Method #2

- m-Sequences and Temporal Masking [Cvejic *et al*, 2001]
  - Can work in real-time (fast algorithm)
  - Attack resistant
  - Employs temporal masking
- Watermark bits spread in time
  - Modulation through pseudo-noise sequence
  - M-sequence

Previous Method #2



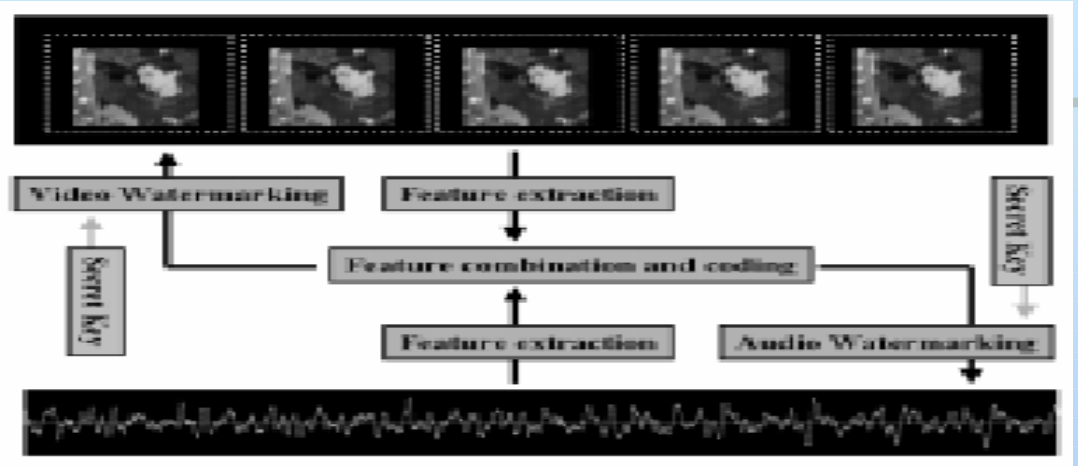
$$y(n) = x(n) + a(n) \cdot b(n) \cdot w(n)$$

- Temporal Analysis
  - Helps determine level of power for watermark sequence
  - Audio frames ≈ 7.6 ms in length
- Frequency Analysis
  - Via zero-crossings
  - Computes information about audio spectrum
  - Higher b(n) values where host audio has higher frequency

Previous Method #3

- Video Watermarking Cross-Correlation [Dittmann, Mukherjee, and Steinebach, 2000]
  - “Content-fragile” watermarking
  - MPEG files – image + audio
  - File integrity
- Watermark images via audio information
- Watermark audio via image information

Previous Method #3



- Video Domain
  - Extract an image
  - Pattern from audio features
  - Apply pattern to luminance
- Audio Domain
  - Extract a frame
  - Pattern from image edge characteristics
  - Apply pattern to audio scale factors

Conclusion

- Inaudibility possible by exploiting characteristics of Human Auditory System
- Type of watermark to use dependent on application
  - Robustness vs. Capacity
  - Security
  - Transparency (audibility)
- Future work will build upon cross-correlation method applied toward stereo (multi-channel) audio files
- Goal is to produce an inaudible watermark