Spread Spectrum-based Watermarking for Tardos Code-based Fingerprinting for H.264/AVC Video

Zafar SHAHID, Marc CHAUMONT and William PUECH
LIRMM, UMR CNRS 5506, Université Montpellier II
zafar.shahid@lirmm.fr, marc.chaumont@lirmm.fr, william.puech@lirmm.fr

Problem Statement

OBJECTIVE: Protect H.264/AVC video with fingerprinting against illegal distribution.
PROPOSED SOLUTION: Embed Tardos fingerprinting code in H.264/AVC using spread spectrum (SS) embedding while taking into account the reconstruction loop.
OUR APPROACH: The proposed solution consists of the following steps:

- Tardos fingerprinting code is generated offline for each user.
- Fingerprinting code is embedded into DC coefficients of all 4×4 blocks.
- Spread spectrum watermarking is used for embedding fingerprinting codes.
- Reconstruction loop is taken into account during the embedding process.

H.264/AVC

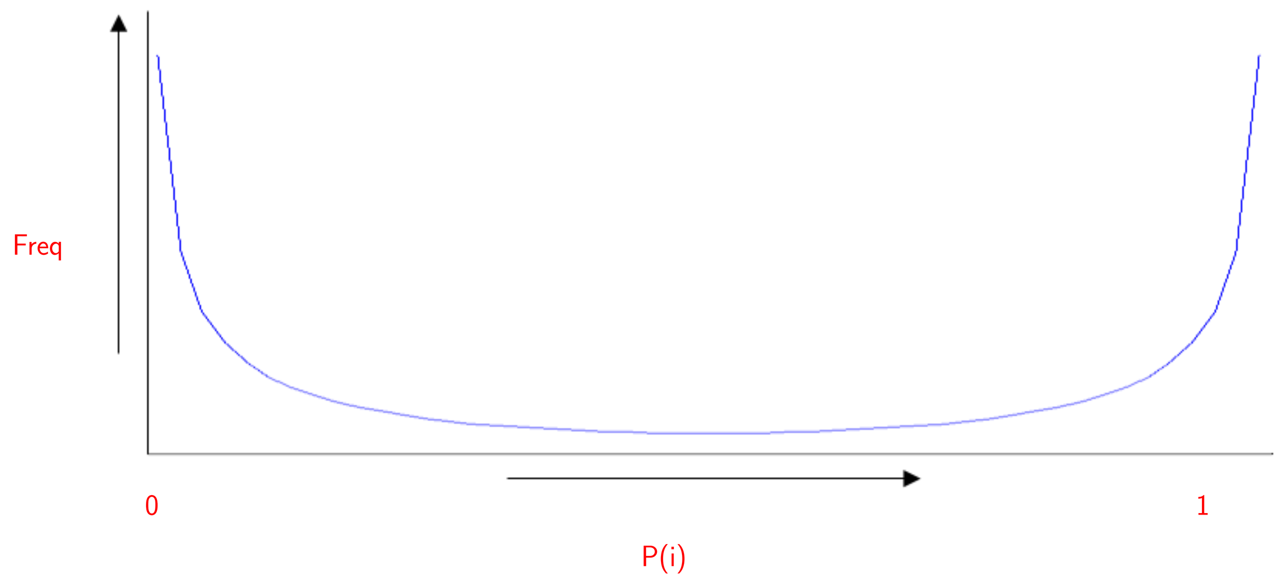
H.264/AVC [1] is the state of the art video codec and performs better than previous standards owing to many new tools including:

- 4X4 integer transform.
- Better entropy coding techniques:
  - CAVLC
  - CABAC
- Quarter pixel motion estimation.
- Multiple block size motion estimation
- Multiple reference frames.

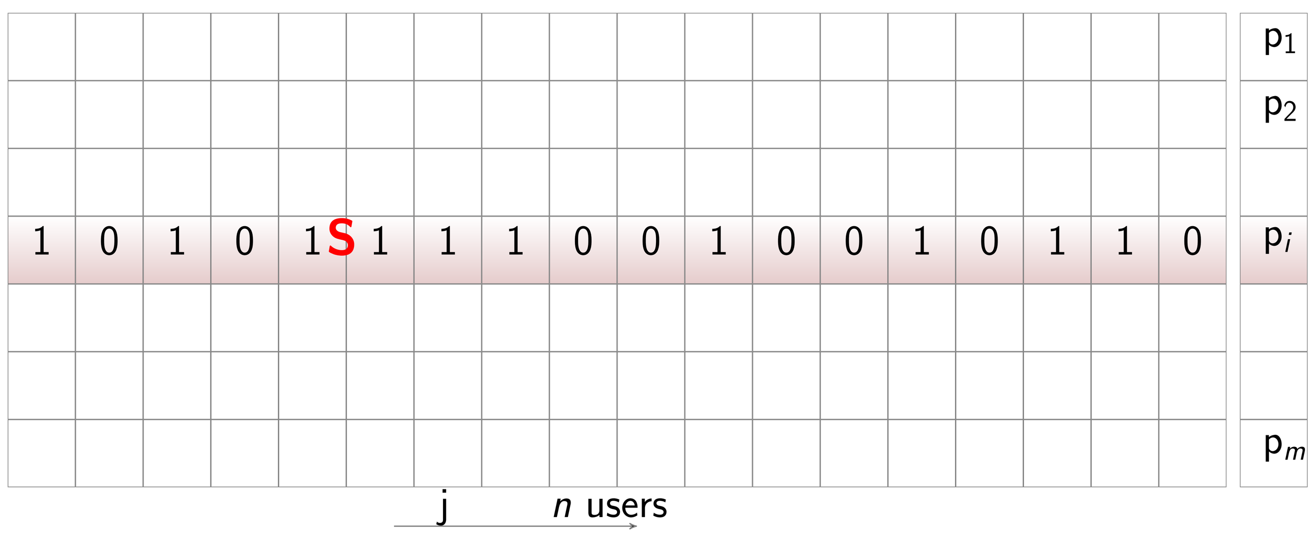
Tardos fingerprinting code generation

A fingerprinting code is analyzed based on its code-length m, maximum number of colluders c_0, false-positive (ϵ_1) and false-negative (ϵ_2) values. For binary asymmetric Tardos code, the length of code is given as $m = 100\, c_0^2 \ln\left(\frac{1}{\epsilon_1}\right)$ and ϵ_1 and ϵ_2 is given as $\epsilon_2 = \epsilon_1^{\frac{c_0}{4}}$, as proposed by Tardos [2].
For code generation, we have the following three steps:

- For a code of length m, we generate random and independent probabilities $\{p(i)\}_{1 \le i \le m}$ with the distribution $f(p) = \frac{1}{\pi\sqrt{p(1-p)}}$ for $p \in [0,1]$. Practically p is between t and $1-t$ with $t = 10^{-3}$ and has high frequency on the edges:



- The next step is to generate Tardos code. For the case of binary Tardos code, each line of S is filled with 0 or 1 with $Prob[S(i,j) = 1] = p(i)$. Each column is a fingerprinting code for separate user. For n users with m code-length, it is a matrix of size $m \times n$:



- For accusation process, a sequence Z is extracted from the pirated copy and an accusation score A_j is associated with user j given as:

$$A_j = \sum_{i=1}^{m} U\left(Z(i), S(i,j), p(i)\right),$$

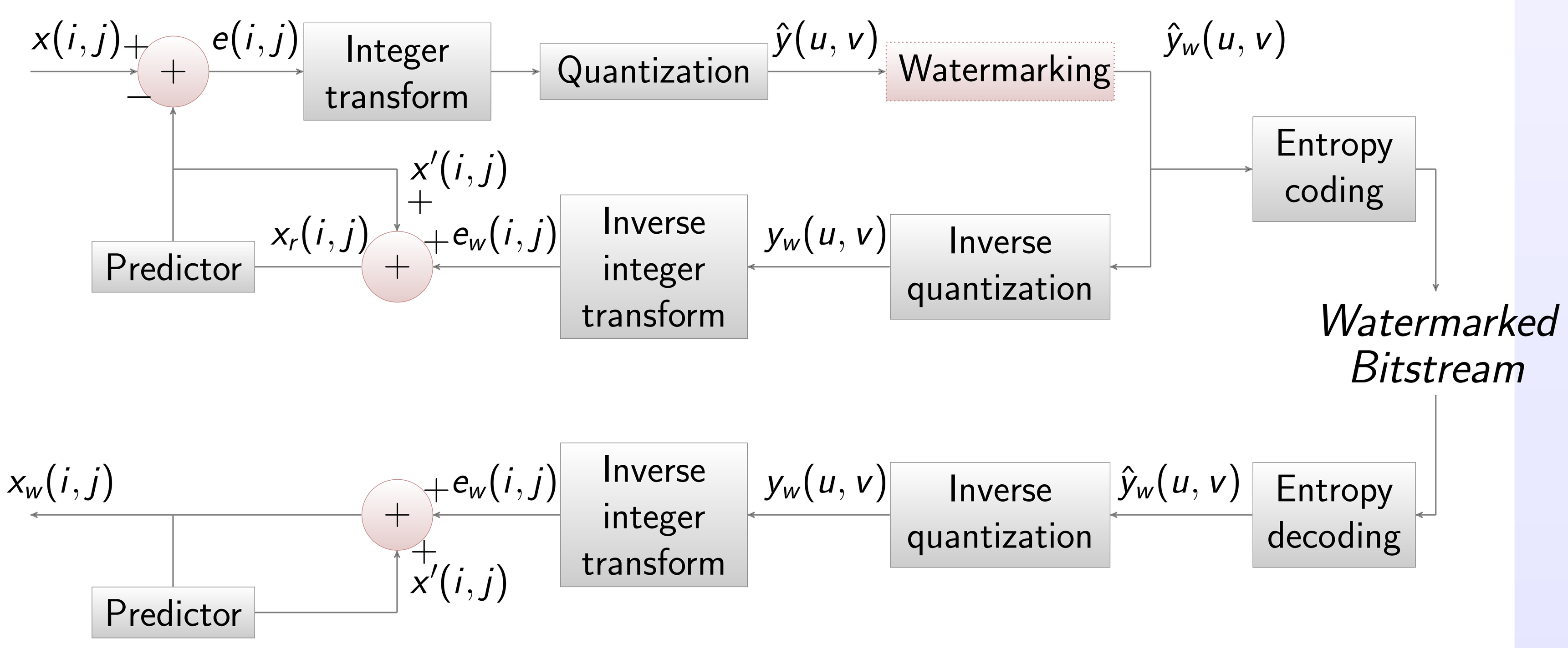
where

$$U(1,1,p) = \sqrt{(1-p)/p} \quad U(1,0,p) = -\sqrt{p/(1-p)},$$
$$U(0,0,p) = \sqrt{p/(1-p)} \quad U(0,1,p) = -\sqrt{(1-p)/p}.$$

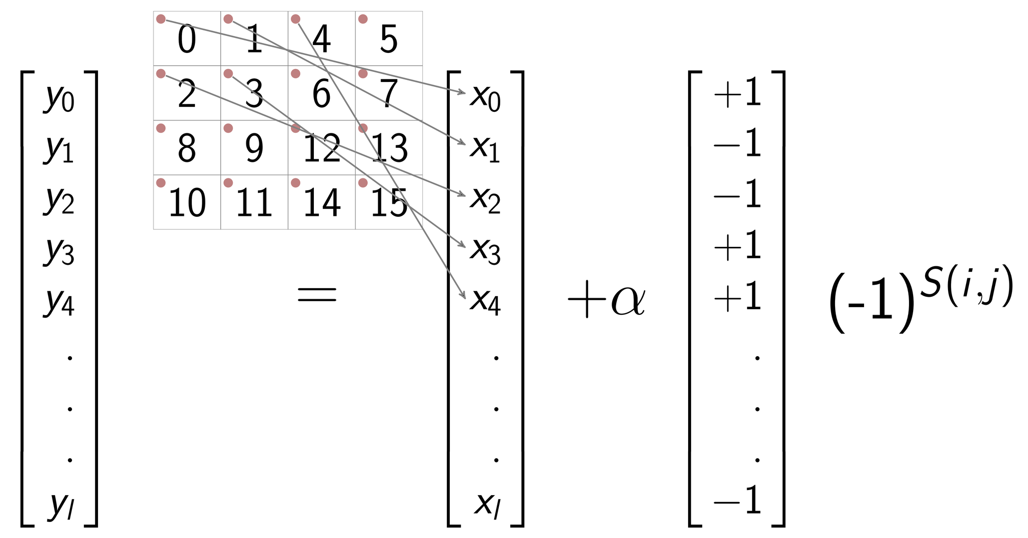
- A_j for accused users may be modeled with a Gaussian centered at $\mu = \frac{2m}{c\pi}$, while A_j for innocent users may be modeled with a Gaussian centered at 0. Accused users (the traitors) have a score above $\mu - \sqrt{m}$ (i-e $\frac{2m}{c\pi} - \sqrt{m}$), where $\sqrt{m}$ is the standard deviation of the Gaussian.

Spread Spectrum Embedding in H.264/AVC

Reconstruction loop is taken into account for Tardos code embedding:



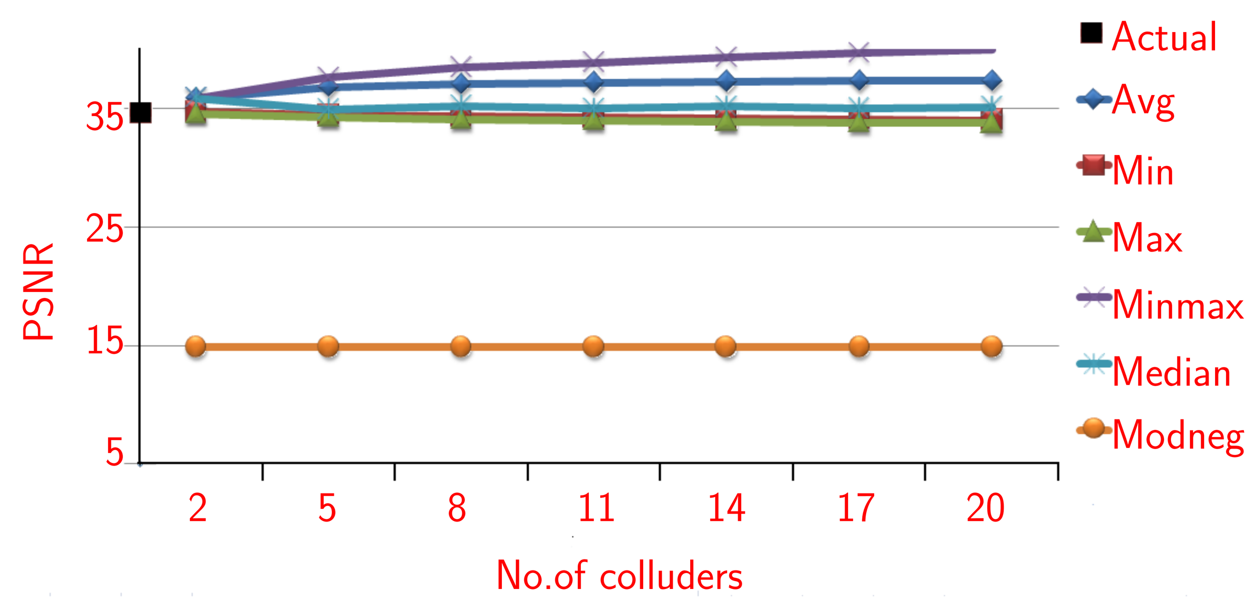
SS embedding is performed on DC coefficients in the following manner:



with $S(i,j)$ is a message bit to be inserted with value 0 or 1.

Experimental Results

- H.264/AVC JSVM 10.2 in AVC mode Intra 4×4 MB mode along with CAVLC.
- Parameters $n = 100$, $\epsilon_1 = 10^{-3}$, $c_0 = 20$, $m = 92104$.
- Payload is 10 bits/frame. Hence 9211 CIF frames (369 seconds of video at 25 fps) are used to embed the code.

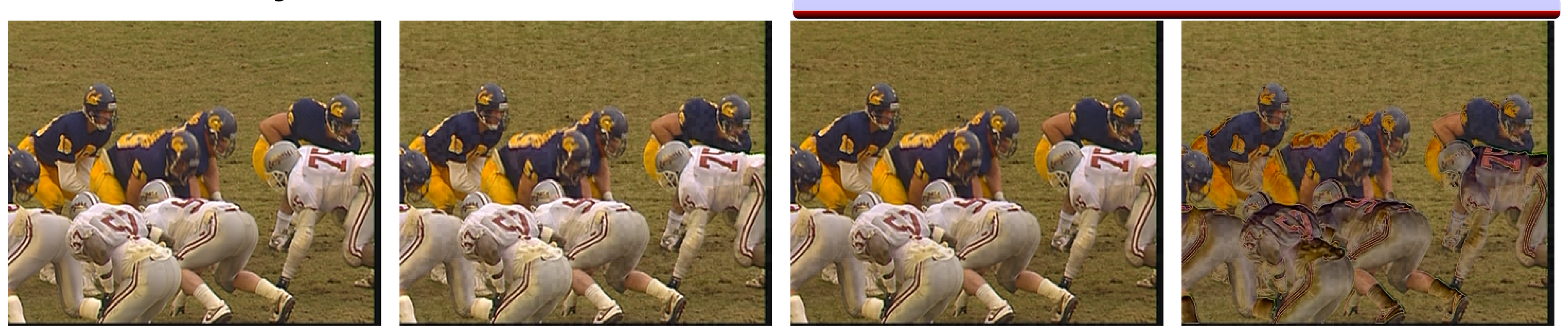


For colluder set S, the j^{th} pixel of colluded video $z(j)$ is:
$z_{ave}(j) = \sum_{k \in S} \lambda_k y_k(j)$,
$z_{min}(j) = \min\{y_k(j)\}_{k \in S}$
$z_{max}(j) = \max\{y_k(j)\}_{k \in S}$
$z_{med}(j) = med\{y_k(j)\}_{k \in S}$
$z_{minmax}(j) = (y_{min}(j) + y_{max}(j))/2$
$z_{modneg}(j) = y_{min}(j) + y_{max}(j) - y_{med}(j)$.

Number of colluders detected

K	Avg	Min	Max	Median	Minmax	Modneg
2	2	2	2	2	2	2
5	5	5	5	5	5	5
8	8	8	8	8	8	6
11	11	10	10	10	10	7
14	14	13	13	13	13	9
17	16	15	16	16	16	10
20	18	18	18	19	18	11

Visual Analysis:



Original — Fingerprinted — Avg(17) — Modneg(8)

Conclusion

- Colluders can be traced in a colluded video having acceptable visual quality.
- Tardos code with SS technique is an efficient video fingerprinting framework.
- The presented results can be improved by using informed watermarking and fine-tuning the parameters.

References

1: ITU-T Rec. H.264 ISO/IEC 14496-10 AVC, *Tech. Rep., Joint Video Team (JVT), Doc. JVT-G050*, March 2003.
2: Gábor Tardos, *Optimal Probabilistic Fingerprint Codes*, in *Proc. ACM symposium on Theory of computing*, New York, NY, USA, 2003.