Mohammad A. Akhaee,
Mohammad J. Saberian,
Soheil Feizi, Farokh Marvasti

ROBUST AUDIO DATA HIDING USING CORRELATED QUANTIZATION WITH HISTOGRAM-BASED DETECTOR

Organized by meconin
Introduction

- Two blind audio watermarking methods using correlated (PPG mapping) quantization for data embedding have been proposed
  - Hard Quantization, Soft Quantization
- Performance of both methods is analyzed by obtaining the radii distribution of PPG points after watermarking
  - Rician distribution, Rayleigh distribution
- Against AWGN, echo, low pass, resampling, MP3
Quantization Watermarking

- $S = (s_1, s_2, \ldots, s_n)$, $s_i \in X$, $Q = (q_1, q_2, \ldots, q_M)$
- Message $m = 0, 1$
- $L_m(s): X \rightarrow Q$
  - $f_m(s, q): X \times Q \rightarrow X$
  - $s_i' = f_m(s_i, L_m(s_i))$
- $L_m(s), f_m$ Requirement
  - Be able to recognize $P_0, P_1$
  - Distortion imposed on the host signal should be controlled within an acceptable range
Suitable Set of Functions

- $L_0(s)$: largest quantization level smaller than $s$
- $L_1(s)$: smallest quantization level larger than $s$

\[
f_m(s_i, L_m(s_i)) = \frac{s_i + dL_m(s_i)}{d + 1} \quad m = 0, 1
\]

$d$ is a design para.

\[
L_0(s) \leq f_0(s, L_0(s)) \leq \frac{L_1(s) + dL_0(s)}{d + 1}
\]

\[
\frac{L_0(s) + dL_1(s)}{d + 1} \leq f_1(s, L_1(s)) \leq L_1(s).
\]

\[
\frac{L_1(s) + dL_0(s)}{d + 1} \leq \frac{L_0(s) + dL_1(s)}{d + 1} \implies d \geq 1
\]
Point-to-point Graph (PPG)

- Putting N samples together in specific order to convert a 1-D signal to n-dimensional counterpart
- Ex: $n = 2, (1, k+1), (2, k+2) \ldots (k, 2k)$
- Small $k \rightarrow$ strong correlation
  Larger $k \rightarrow$ more dispersion
Quantization in PPG Domain

- Quantization can be performed on both radius and angle, in the paper, radii only
- Quantization of the radii of PPG points has three advantages:
  - More transparency in the time domain
  - More robustness against low-pass attacks
  - Higher resistance in noisy environment
Transparency

- Quantization in PPG domain produces correlated noise on the components of points.
- Adjusting PPG mapping index $k$, time domain traces of quantization can be controlled.
- Figure in next page
  - $k = 1$, adjacent samples are quantized similarly great $\rightarrow$ impact on time domain shape
  - $k = 10$, this effect is absolutely transparent in the time domain, guaranteeing the security of the watermarking process.
PPG \((k = 1, \ k = 10)\)

Fig. 1. Time domain shape of a signal with 400 samples for (a) \(k = 1\) and (b) \(k = 10\).
Robustness Against Low Pass

- Radii quantization imposes low-pass distortion watermarked signal
  - The hidden data are inserted in the low-frequency components
- Due to the low-frequency characteristic of hidden data, it’s robust against low-pass attacks
Resistance in Noisy Env.

- 1-D vs. 2-D
  - Bit rate and distortion in both cases must equal
- $\Delta$: Quantization distance $\rightarrow$ distortion $\Delta^2/12$
- Noise $\sim N(0, \sigma^2)$
- The probability of error [20: Digital Comm.]
  - 1-D
  $$P_{e,1-D} = 1 - \left(1 - Q\left(\frac{\Delta}{2\sigma}\right)\right)^2$$
  - 2-D
  $$P_{e,2-D} = 1 - \left(1 - Q\left(\frac{\Delta}{2\sigma}\right)\right)$$

$Q(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\infty} e^{-u^2/2} du$
1-D vs. 2-D

- Tangential noise component $n_2$, affects the error less than $n_1$ in radial direction.

Fig. 2. Decision making region of two samples when (a) they are separately quantized and (b) the PPG radius is quantized. The circle is the effect of noise with a variance of $3\sigma^2$. 
Proposed Methods

- In order to more transparency, only low-pass components are extracted and watermarked.
- Filter the quantized signal with a similar low-pass filter then add high-pass components.

\[
\frac{L_1(s) + dL_0(s)}{d + 1} \leq \frac{L_0(s) + dL_1(s)}{d + 1} \Rightarrow d \geq 1
\]
Hard Quantization Method

- 2 quantization levels, and $d$ tends to infinity
- $g_1 < q_1 < q_2 < g_2$
- Embedding (for each point in a frame)
  - The points with radii outside $(g_1, g_2)$ are kept
  - The points between $[g_1, q_1]$ or $[q_2, g_2]$ are mapped to a circle with radius $g_1/g_2$
  - The points in $[q_1, q_2]$ are quantized to $q_1$ or $q_2$ depends on the value of embedded message for this frame
Hard Quantization Method (Cont.)

- **De-watermarking**
  - PPG points of each frame are obtained
  - Pre-quantization process is performed on the points in \([g_1, g_2]\]
  - Comparing the number of points quantized to the levels \(q_1, q_2\), the data can be obtained

- Adjusting the quantization levels, the robustness and its transparency can be controlled
Soft Quantization Method

- $N$ equidistant quantization levels, $d = 1$
- For a PPG point with radius $R$, to embed “0” or “1”, mapped to the point to $R_0 = (R + q_m) / 2$, $R_1 = (R + q_{m+1}) / 2$
- Embedding
  - Mapped the points to $R_0 / R_1$ depend on message
Synchronization

- Due to each frame contains the message “0” or “1”, the ratio $N_0/N_1$, $N_1/N_0$ will be maximized in accordance to the message bit.
- This synchronization approach suffers from its high complexity cost.
- It can be reduced using wavelet transform and its properties in shifting.
  - Not in the scope of this paper.
Performance Analysis

- Typical PPG point \((x, y)\), after AWGN attack \((n_1, n_2)\) on each component \(\rightarrow A = (x+n_1, y+n_2)\)
  - The distribution of \(|A|\) is Rician
  - The distribution of \(R\) is Rayleigh
    - \(p_{in}\) is input power, affect variance with \(k\)
Rayleigh Distribution

Fig. 4. Rayleigh distribution of the radius of PPG points.
Error Probability (Hard)

\[ P_{F|1} = E\{P_{fst\_reg}(R)\} = \int_{0}^{\infty} P_{fst\_reg}(R) f_R(R; k, P_{in}|m = 1) dR \]

\[ P_{S|1} = E\{P_{snd\_reg}(R)\} = \int_{0}^{\infty} P_{snd\_reg}(R) f_R(R; k, P_{in}|m = 1) dR \]

By Rician Distribution

\[ P_{e|m=1} = \sum_{i_{others}=0}^{N} \binom{N}{i_{others}} (P_{others|1})^{i_{others}} (P_{F|1})^{i_f} (P_{S|1})^{i_s} \]

subject to: \[ \begin{cases} i_{others} + i_f + i_s = N \\ i_f > i_s \end{cases} \]

\[ P_{e|m=0} = \sum_{i_{others}=0}^{N} \binom{N}{i_{others}} (P_{others|0})^{i_{others}} (P_{F|0})^{i_f} (P_{S|0})^{i_s} \]

subject to: \[ \begin{cases} i_{others} + i_f + i_s = N \\ i_f < i_s \end{cases} \]

\[ P_e = \frac{1}{2} (P_{e|m=0} + P_{e|m=1}). \]
Simulation Results

- Simulation Parameter
  - One minute audio signals of various types
    - classic, jazz, rock, vocal with male / female
  - 44100 Hz, 16 bits quantization
  - Frame length = 400
  - For Hard: $g_1, q_1, q_2, g_2$ are 0.05, 0.15, 0.23, 0.33
  - For Soft: $N = 15$
  - Synchronization is assumed to be perfect
BER after AWGN attack
Common Attacks

### TABLE I

**Robustness of the Proposed Methods Against Some Common Attacks**

<table>
<thead>
<tr>
<th></th>
<th>Echo 0.5,100msec</th>
<th>Echo 0.2,100msec</th>
<th>Requantization 16-8</th>
<th>LP 6 kHz</th>
<th>LP 3 kHz</th>
<th>DC attack</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard quant.</td>
<td>11%</td>
<td>8%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10%</td>
</tr>
<tr>
<td>Soft quant</td>
<td>9%</td>
<td>6%</td>
<td>0</td>
<td>0</td>
<td>3%</td>
<td>0</td>
</tr>
</tbody>
</table>

### TABLE II

**BER(%) of the Proposed Methods for Various Sampling Rates (Hz)**

<table>
<thead>
<tr>
<th></th>
<th>44/11/44</th>
<th>44/6/44</th>
<th>44/4/44</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard Quant.</td>
<td>0</td>
<td>0</td>
<td>2.6%</td>
</tr>
<tr>
<td>Soft Quant.</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
### TABLE III
ROBUSTNESS OF THE TWO METHODS AGAINST VARIOUS ATTACKS USING DIFFERENT PPG INDEX PARAMETER $k$

<table>
<thead>
<tr>
<th></th>
<th>Noise (SNR=25)</th>
<th>LP (3 kHz)</th>
<th>Resampling (50%)</th>
<th>Echo (100msec, 20%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PPG Index $k$</strong></td>
<td>2</td>
<td>4</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td><strong>Hard Quant.</strong></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15%</td>
</tr>
<tr>
<td><strong>Soft Quant.</strong></td>
<td>0</td>
<td>0</td>
<td>1%</td>
<td>15%</td>
</tr>
</tbody>
</table>

![Fig. 8](image)

Fig. 8. Number of multiplicity for different sliding shifts in (a) hard quantization and (b) soft quantization.

![Fig. 9](image)

Fig. 9. BER versus (a) the time stretching and (b) MP3 attacks with various rates.
### TABLE IV
BER Performance Comparison for Different Kinds of Attacks With the Technique in [19]

<table>
<thead>
<tr>
<th>Methods</th>
<th>Payload (bps)</th>
<th>AWGN, 15 dB</th>
<th>Resampling, 44/4/44</th>
<th>MP3, 64kbps</th>
</tr>
</thead>
<tbody>
<tr>
<td>[19]</td>
<td>172</td>
<td>7.5%</td>
<td>12.21%</td>
<td>4.34%</td>
</tr>
<tr>
<td>Hard Quant.</td>
<td>176</td>
<td>8.8%</td>
<td>13.1%</td>
<td>10.2%</td>
</tr>
<tr>
<td>Soft Quant.</td>
<td>176</td>
<td>3.9%</td>
<td>4%</td>
<td>0.06%</td>
</tr>
</tbody>
</table>

### TABLE V
BER Comparison of the Proposed Methods With the Technique of [23] for Some Common Attacks

<table>
<thead>
<tr>
<th>Methods</th>
<th>Payload bps</th>
<th>AWGN 20dB</th>
<th>Filtering 8kHz</th>
<th>MP3 64kbps</th>
<th>Resampling 44/22/44</th>
<th>Requantization 16-8</th>
<th>Echo 40%, 100 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>[23]</td>
<td>11</td>
<td>22.8%</td>
<td>4.2%</td>
<td>6.5%</td>
<td>6.4%</td>
<td>11.9%</td>
<td>1.6%</td>
</tr>
<tr>
<td>Hard Quant.</td>
<td>110</td>
<td>2.9%</td>
<td>0%</td>
<td>0.2%</td>
<td>0%</td>
<td>0%</td>
<td>10.2%</td>
</tr>
<tr>
<td>Soft Quant.</td>
<td>110</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>8.1%</td>
</tr>
</tbody>
</table>
THANKS FOR YOUR ATTENTIONS