A CELP-Speech Information Hiding Algorithm Based on Vector Quantization

Ji-Xin Liu and Zhe-Ming Lu, and Hao Luo

Date: 2010/03/30
Speaker: Jhang Hon-Bin
Outline

- Introduction
- Related Works
- The G.729 Vocoder
- The Proposed Method
- Experimental Results
- Conclusions
Introduction (1/3)

User A uses this (Attacker MAC) MAC address

User B uses this (Attacker MAC) MAC address

Broadcasting spoofed MAC address

Attacker
Man in the middle
Introduction (2/3)
Information Hiding Scheme

Vector quantization and G.729 8kb/s speech codec

Index-constrained during the Predictive Two-Stage Vector Quantization procedure of the speech coding
Related Works (1/8)

- Information Hiding Scheme (7 approaches)
  
  - Least Significant bit (LSB) :
    perceptually transparent and high payload

  Content-dependent watermarking scheme in compressed speech with identifying manner and location of attacks.

  This method embeds the watermark, which is generated by referring the line spectrum frequency (LSP) and pitch features from current and succeeding frame into the LSBs of the indices indicating the excitation pulse positions or excitation vectors.
Related Works (2/8)

- Information Hiding Scheme (7 approaches)

  - Phase coding (PE):

    Techniques for data hiding.

    This method works by substituting the phase of an initial audio segment with a reference phase that represents the data. The phase of subsequent segments is adjusted in order to preserve the relative phase between segments.
Information Hiding Scheme (7 approaches)

- Spectrum techniques in transform domain:

  Audio Watermarking for Covert Communication through Telephone System.

  Using spread spectrum method.

Though it is claimed in this paper that this method is robust against mp3, AWGN (Additive White Gaussian Noise), LPF (Low Pass Filtering), down-sampling and companding attacks, the Time Scale Modification (TSM) attack, which will be induced during the speech transmission and A/D (D/A) through the PSTN network, is not considered.
Related Works (4/8)

- Information Hiding Scheme (7 approaches)
  - Echo hiding (EH)
Related Works (5/8)

- Information Hiding Scheme (7 approaches)

  - Analysis by synthesis (ABS)-based:

    ABS-based speech information hiding approach.

    Embeds the secret speech information data bit stream into the carrier speech bit stream.
Related Works (6/8)

- Information Hiding Scheme (7 approaches)
  
  • Feature-based:

  Pitch and Duration Modification for Speech Watermarking.

  A speech watermarking algorithm based on the modification of the pitch (fundamental frequency) feature and duration of the quasi-periodic speech segments.

  This method is robust to low data-rate (5-8kbps) speech coders, while the payload is relatively low (approx. 3bits/sec).
Related Works (7/8)

- Information Hiding Scheme (7 approaches)

  - Combined with coding frames:

    Watermarking Combined with CELP Speech Coding for Authentication.

    This method uses the codebook partition and labeling method to embed watermark bits.
Related Works (8/8)

- Information Hiding Scheme (7 approaches)
  
  - Vector Quantization:
    
    The G.729 speech coding bit-stream by combining the Predictive Two-stage.
The LP filter coefficients are converted to Line Spectrum Pairs (LSP) and quantized using predictive two-stage vector quantization with 18 bits.

The excitation signal is chosen by using an analysis-by-synthesis search procedure in which the error between the original and reconstructed speech is minimized according to a perceptually weighted distortion measure.
The Proposed Method (1/3)

Index-constrained during the predictive two-stage vector quantization procedure of the LSP coefficients.

Each frame speech signals is first quantized by a 10 dimension codebook, then each sub-frame’s LSP residual signal quantized by a 5-dimension codebook.

The embedding process for each watermark bit can be achieved by searching the best matched codeword for each input LSP and LSP residual vector under the constraints that the randomly selected bits of the indices are consistent with the watermark bits to be embedded.
The Proposed Method (2/3)

It is obvious that at least 1 bits of watermark can be embedded into each speech frame.

During the watermark embedding process, the G.729 encoder searches for the codeword that provides the minimum distortion between the input LSP (also residual) vector and the reconstructed vector within the constraint of the index.

Figure 2. Index-constrained information hiding in G.729 Vocoder
Substitute the perceptually unimportant bits in the bit stream.

After the substitution, it only has very small probability that the changed codeword (reconstruction vector) best matches the input vector.
Experimental Results (1/4)

- Embedding Rate:
  
  At least 3 bits are embedded in each speech frame during the predictive two-stage vector quantization procedure.

- Quality Evaluation:
  
  1) waveform compare
  2) objective distortion measure
  3) subjective listening test
Experimental Results (2/4)

All the segSNRs are higher than 19.30.

<table>
<thead>
<tr>
<th>Payload (bits/frame)</th>
<th>Average segSNR(dB)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>female speech signals</td>
<td>male speech signals</td>
</tr>
<tr>
<td>3 bits</td>
<td>19.63</td>
<td>19.45</td>
</tr>
<tr>
<td>4 bits</td>
<td>19.10</td>
<td>19.03</td>
</tr>
<tr>
<td>5 bits</td>
<td>18.20</td>
<td>18.34</td>
</tr>
<tr>
<td>6 bits</td>
<td>14.02</td>
<td>14.17</td>
</tr>
</tbody>
</table>

Thus the embedding rate $R$, which is the ratio of the number of the embedded watermark bits and the bit stream size of one frame, could be given as:

$$R = \frac{3}{80} = 3.75\%.$$
Figure 4. Waveforms of the clear and watermarked speech signals
Comparing these figures, we can find that the signal of the watermarked speech and the speech signal decoded from the clear bit stream are nearly the same.
The selected bits of first-stage and residual vector indices during the Predictive Two-stage Vector Quantization procedure are modulated by the secret information bits.

For the imperceptibility of this scheme, the quality of the watermarked speech signal can be controlled by adaptive embedding through referring to the original and watermarked speech.

The objective and subjective quality test show that this algorithm is perceptually transparent with higher payload than most of the existing method.