An Efficient Transcoding Algorithm for G.723.1 and G.729A Speech Coders


S-W Yoon, S-K Jung, Y-C Park, and D-H Youn

11 / 11 / 09

Adviser  Prof. R. S. Lin
Speaker  J. Y. Wang

Department of Computer Science and Information Engineering
Outline

- Introduction
- Proposed Schemes
- Experiment Simulations
- Conclusions
TABLE 1.
THE FRAME CHARACTERISTICS OF THE CODERS MP-MLQ AND CS-ACELP

<table>
<thead>
<tr>
<th>Coder</th>
<th>Format</th>
<th>Frame size</th>
<th>Frame length</th>
<th>Number of subframe</th>
<th>Time of LPC analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP-MLQ</td>
<td>G.723.1</td>
<td>30 ms</td>
<td>240 samples</td>
<td>4</td>
<td>7.5 ms</td>
</tr>
<tr>
<td>CS-ACELP</td>
<td>G.729 A</td>
<td>10 ms</td>
<td>80 samples</td>
<td>2</td>
<td>5 ms</td>
</tr>
</tbody>
</table>
- G.723.1 (A) and G.729A (B)
The proposed method of transcoding between G.723.1 and G.729A is based of four processing steps:

1. Conversion linear spectral pair (LSP).
2. Conversion pitch interval.
4. Fast Fixed-codebook search (5.3 kbps).
Purpose of Research

- Proposed an efficient method of converting speech codec formats between G.723.1 and G.729A.

- The tandem method has several problems:
  - Quality degradation.
  - High complexity.
  - Longer delay time.

- Fast Search Algorithms
CONVERTING SPEECH CODEC FORMATS FROM G.723.1 TO G.729A

1. Conversion linear spectral pair (LSP).
2. Conversion pitch interval.
3. Adaptive-codebook search.
4. Fixed-codebook search.
CONVERTING SPEECH CODEC FORMATS FROM G.723.1 TO G.729A

\[
p_i'(j) = \begin{cases} 
  p_i(j), & i = 1 \\
  \frac{1}{2}(p_i(j) + p_{i+1}(j)), & i = 2, 1 \leq j \leq 10, \\
  p_{i+1}(j), & i = 3 
\end{cases}
\]

Fig. 3 LSP conversion using linear interpolation
CONVERTING SPEECH CODEC FORMATS FROM **G.723.1** TO **G.729A**

*Fig. 4 LPC spectrum: reference(solid), tandem(dash), transcoding(dot)*
CONVERTING SPEECH CODEC FORMATS FROM G.723.1 TO G.729A (Cont.)

Fig. 5. Open-loop pitch estimation using pitch smoothing (G.723.1 → G.729A).
CONVERTING SPEECH CODEC FORMATS FROM \textbf{G.723.1} TO \textbf{G.729A} (Cont.)

- The closed-loop pitch of G.723.1 is determined as the open-loop pitch of G.729A.
- The pitch smoothing method is applied.

| The distance between the two pitch values |  
| < 10 samples | The closed-loop pitch of G.723.1 is determined as the open-loop pitch of G.729A.  
| >10 samples | The pitch smoothing method is applied.  

Conversion pitch interval
CONVERTING SPEECH CODEC FORMATS FROM G.723.1 TO G.729A (Cont.)

If the pitch difference is larger than 10 samples.

\[ R(k_i) = \sum_{n=0}^{79} s_w(n) \cdot s_w(n - k_i), \]

\[ p_i - 3 \leq k_i \leq p_i + 3 (i = A, B), \]

\[ R'(t_i) = \frac{R(t_i)}{\sqrt{\sum_n s_w^2(n-t_i)}}, \quad i = 1, 2 \]
CONVERTING SPEECH CODEC FORMATS FROM G.723.1 TO G.729A

Proposed Schemes (7/12)

The smoothed open-loop pitch, \( T_{op} \), is determined as:

\[
T_{op} = t_1 \\
R'(T_{op}) = R'(t_1) \\
\text{if } R'(t_2) \geq 0.75 \ast R'(T_{op}) \text{ then} \\
R'(T_{op}) = R'(t_2) \\
(T_{op}) = t_2
\]
CONVERTING SPEECH CODEC FORMATS FROM G.723.1 TO G.729A (Cont.)

Proposed Schemes (8/12)


Fig. 2. Block diagram of transcoding from G.723.1 to G.729A.
CONVERTING SPEECH CODEC FORMATS FROM G.729A TO G.723.1

1. LSP conversion using **linear interpolation**.
2. Open-loop pitch conversion using **pitch smoothing**.
3. Fast adaptive-codebook search (5.3 kbps).
4. Fast fixed-codebook search (5.3 kbps).
Fast Adaptive-codebook search

The adaptive-codebook search in G.723.1 uses a 5th order pitch predictor, and estimates the **pitch delay** and **gain** simultaneously.

A 1st-Order
Pitch Predictor

A 5th-Order
Pitch Predictor

170 -> 85
Fast Fixed-codebook search  G.723.1

5.3 kbit/s

ACELP excitation codebook

<table>
<thead>
<tr>
<th>Sign</th>
<th>Positions</th>
</tr>
</thead>
<tbody>
<tr>
<td>±1</td>
<td>0, 8, 16, 24, 32, 40, 48, 56</td>
</tr>
<tr>
<td>±1</td>
<td>2, 10, 18, 26, 34, 42, 50, 58</td>
</tr>
<tr>
<td>±1</td>
<td>4, 12, 20, 28, 36, 44, 52, (60)</td>
</tr>
<tr>
<td>±1</td>
<td>6, 14, 22, 30, 38, 46, 54, (62)</td>
</tr>
</tbody>
</table>

\[ 8 \times 8 \times 8 \times 8 = 4096 \]

Depth-first tree search

\[ 2 \times \{(8 \times 8) + (8 \times 8)\} = 256 \]
CONVERTING SPEECH CODEC FORMATS FROM G.729A TO G.723.1 (Cont.)

Fig. 7. Block diagram of transcoding from G.729A to G.723.1.
Objective quality evaluation

<table>
<thead>
<tr>
<th></th>
<th>LPC-CD (dB)</th>
<th>PSQM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>G.723.1 Tandem</td>
<td>3.90</td>
<td>3.98</td>
</tr>
<tr>
<td>G.729A Tandem</td>
<td>3.54</td>
<td>3.66</td>
</tr>
<tr>
<td>G.729A Transcoding</td>
<td>3.65</td>
<td>4.17</td>
</tr>
<tr>
<td>G.723.1 Transcoding</td>
<td>3.25</td>
<td>3.86</td>
</tr>
</tbody>
</table>

LPC-CD (LPC Cepstral-Distance)
PSQM (Perceptual Speech Quality Measure)
PSQM (0 (no degradation) ~ 6.5 (highest degradation))
Complexity check using TMS320C6201

<table>
<thead>
<tr>
<th>MIPS</th>
<th>G.723.1 → G.729A</th>
<th>G.729A → G.723.1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tandem</td>
<td>Transcoding</td>
</tr>
<tr>
<td>LPC</td>
<td>6.41</td>
<td>2.36</td>
</tr>
<tr>
<td>Olp</td>
<td>0.94</td>
<td>0.21</td>
</tr>
<tr>
<td>ACB</td>
<td>2.45</td>
<td>2.45</td>
</tr>
<tr>
<td>FCB</td>
<td>4.30</td>
<td>4.30</td>
</tr>
<tr>
<td>Others</td>
<td>4.04</td>
<td>4.04</td>
</tr>
<tr>
<td>Total</td>
<td>18.15</td>
<td>13.37</td>
</tr>
</tbody>
</table>
- Proposed an efficient method of *converting* speech codec formats between *G.723.1* and *G.729A*.

<table>
<thead>
<tr>
<th></th>
<th>Delay</th>
<th>Complexity</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tandem Method</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proposed method</td>
<td><img src="image" alt="trophy" /></td>
<td><img src="image" alt="trophy" /></td>
<td><img src="image" alt="trophy" /></td>
</tr>
</tbody>
</table>
Thanks for your attention!