Temporal Scalability using P-Pictures

Stephan Wenger
stewe@cs.tu-berlin.de

TU Berlin and
University of British Columbia
The Problem

• Interactive Applications: low latency
  – 300 ms one-way delay the absolute tolerable maximum

• Video Coding is based on Inter-Picture Prediction
  – minimum algorithmic delay: two pictures
  – at fixed 10 fps: 200 ms

• Temporal Scalability based on B-pictures is not helpful
  – B-pictures do improve temporal resolution,
  – but NOT the overall latency
(Traditional) B-pictures

Layering scheme
Base layer only
Base + Enhancement layer

<table>
<thead>
<tr>
<th></th>
<th>Frame-rate</th>
<th>Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base layer only</td>
<td>10 fps</td>
<td>200 ms</td>
</tr>
<tr>
<td>Base + Enhancement layer</td>
<td>30 fps</td>
<td>200 ms</td>
</tr>
</tbody>
</table>
P-Picture Scalability

• Problem: Bi-directionally predicted nature of the Enhancement layer
• Solution: use only forward prediction
  – Advantages
    • “Latency” scalability
    • Straightforward implementation
  – Disadvantages
    • Lower coding efficiency (?)
    • Syntax of current video standards does not allow that (?)
P-Picture Scalability

Base, 15 fps

Latency Base: 132 ms

Enh, 30 fps

Latency Base+Enhancement: 66 ms

Base, 10 fps

Latency Base: 200 ms

Enh, 30 fps

Latency Base+Enhancement: 66 ms
Disadvantage 1: Coding efficiency

- QCIF, fixed Quantizer value 13 for base, 16 for enhancement layers, H.263+, bitrates in kbit/s

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Base 10 fps</th>
<th>Base 30 fps</th>
<th>B-picture</th>
<th>P-picture</th>
<th>Overhead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreman</td>
<td>43.8</td>
<td>108.5</td>
<td>73.7</td>
<td>89.5</td>
<td>21%</td>
</tr>
<tr>
<td>Coastguard</td>
<td>60.5</td>
<td>113.0</td>
<td>79.2</td>
<td>86.4</td>
<td>9%</td>
</tr>
<tr>
<td>Paris</td>
<td>58.6</td>
<td>106.6</td>
<td>83.9</td>
<td>119.1</td>
<td>42%</td>
</tr>
</tbody>
</table>
Disadvantage 2: lack of syntax

- True for MPEG 1, MPEG 2, MPEG 4 V. 1
  - But still possible using out-band signaling
  - (e.g. in the corresponding RTP-payload header)

- False for H.263+, MPEG 4 V. 2
  - Reference Picture Selection mode
  - In a P-picture other than the previous picture can be selected as Reference Picture
Reference Picture Selection

Video in

To video multiplex coder
Summary

- P-Picture temporal scalability allows for
  - Scalable latency for real-time systems
  - Users can “buy” lower latency
  - Simple implementation of temporal enhancement layers

- At a cost of
  - 10% to 40% coding efficiency
  - need for the implementation of Reference Picture Selection
  - for older video coding standards: a new RTP payload spec.
Good bye

- Thank you for your attention
- Your questions please...