
Author final version (often known as postprint)

Link to published version (if available):
10.1109/ICASSP.2009.4959684

Link to publication record in Explore Bristol Research

PDF-document

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/pure/about/ebr-terms.html

Take down policy

Explore Bristol Research is a digital archive and the intention is that deposited content should not be removed. However, if you believe that this version of the work breaches copyright law please contact open-access@bristol.ac.uk and include the following information in your message:

• Your contact details
• Bibliographic details for the item, including a URL
• An outline of the nature of the complaint

On receipt of your message the Open Access Team will immediately investigate your claim, make an initial judgement of the validity of the claim and, where appropriate, withdraw the item in question from public view.
Enhanced Spatially Interleaved DVC using Diversity and Selective Feedback

Nantheera Anantrasirichai,
Dimitris Agrafiotis,
Dave Bull
Contents of Presentation

• Introduction
  ⇒ Distributed Video Coding (DVC)

• Proposed schemes
  ⇒ Block-based concealment
  ⇒ Gray code
  ⇒ Diversity scheme
  ⇒ Selective feedback

• Results

• Conclusions
Introduction

- **Distributed Video Coding (DVC)**
  - Very low complex video transmitters
  - Two information theory principles: Slepian-Wolf (SW) and Wyner-Ziv (WZ)
  - Separate encoding, but joint decoding
  - Single source video frames are split into two categories, key frames and WZ frames
Introduction

• Distributed Video Coding (DVC)
  
  • Key frames are intra coded with a conventional encoder and are made available at the decoder.
  
  • WZ frame coding, which may involve transformation, uses quantisation followed by channel coding applied.
  
  • At the decoder, the key frames are used for creating an estimate of the WZ frames (side information)
  
  • The side information is seen as a noisy version of the original coded WZ frame.
  
  • Finally, the received parity bits are used to correct the errors present in this noisy version of the WZ data.
Proposed System

This paper presents a concealment based approach to distributed video coding with…

• Hybrid Key/WZ frames via an FMO type interleaving of macroblocks.
• Spatio-temporal concealment for generating the side information on a MB basis
• Gray code
• Diversity technique – multi-hypothesis SI coding
• Two selective feedback mechanisms: reliability map feedback and updated bit request buffer
Proposed System

- SPI DVC codec
Proposed System Encoder

- Hybrid Key/WZ frames

The first step involves splitting of the current input frame into KEY and WZ groups, in a similar fashion to the dispersed type of flexible macroblock ordering (FMO) specified in H.264.
Proposed System Encoder

- **Key slicegroup coding**

KEY MBs are, horizontally shifted to make a new frame of the same height but half the width of the original.

When spatial prediction is used, further temporal interleaving of the KEY groups of two consecutive frames can be applied to avoid any performance loss relative to full frame KEY coding.
Proposed System Encoder

- **Gray code**

The Gray code can improve the system performance since it has a hamming distance of one which provides higher error resilience than the natural binary code. It also offers a higher accuracy of bit probability estimation.

This can be performed in parallel by a bit-shift and exclusive-or operation. The $n$th Gray code is obtained by computing $n \oplus \lfloor n/2 \rfloor$.

The reverse translation can be given recursively: $b_o = g_o$ and $b_i = g_i \oplus b_{i-1}$ where $b_i$ is the $i$th binary-coded bit ($0$ being the most-significant bit) and $g_i$ is the $i$th gray-coded bit.
Proposed System Decoder

- **Side Information Generation**

Temporal Error Concealment

Each missing WZ block is divided into four sub-blocks. Each sub-block is concealed using two motion vectors of neighbouring KEY blocks and two motion vectors generated using external boundary matching error (EBME) of its own WZ block and adjacent WZ blocks. These four possible sub-blocks are blended using a cosine weighting matrix.
Proposed System Decoder

- **Side Information Generation**

**Spatial Error Concealment**

SEC uses bordering KEY pixels to conceal the missing WZ pixels of each WZ block through
(a) directional interpolation or
(b) bilinear interpolation along detected edges.

Proposed System Decoder

- **Side Information Generation**

Mode Selection

- Evaluating the levels of motion compensated activity (TA) and spatial activity (SA) in the neighbourhood of that block

\[
SA = E[(x - \mu)^2] \quad \text{and} \quad TA = E[(x - x*)^2]
\]

If (SA<TA) and (TA>3), MB is processed by SEC.

\[
TA = \text{mean squared error between the key blocks surrounding the missing block in the current frame and those surrounding the replacement block in the reference frame}
\]

\[
SA = \text{variance of the surrounding key blocks in the current frame}
\]
Proposed System Decoder

- Diversity scheme
  - Multi-hypothesis SI

Multiple SI is generated and defined as $SI_i$, $i \in \{1,...,N\}$, with increasing $i$ denoting the ranking of the SI, from best to worst, according to the MSE of the predicted KEY group associated with each SI group.

Log-Likelihood Ratio

$$LLR^l = \log \frac{\Pr[X^l = 0|SI_1, SI_2, ..., SI_N, Y^{l-1}]}{\Pr[X^l = 1|SI_1, SI_2, ..., SI_N, Y^{l-1}]}$$

where $X^l$ is a bit value at bit-plane $l$ and $Y^{l-1}$ is a reconstructed WZ data $Y$ using decoded bit-planes 1-($l$-1).
• Diversity scheme

→ Multi-hypothesis SI

Method allows usage of the correct SI to compensate for the errors appearing in the others.

WZ Reconstruction

The decoded bit-planes are converted to quantisation symbol $\tilde{Q}$.

The pixel $Y$ is then reconstructed by firstly using the best SI based on its quantisation bins as follows:

$$Y_1 = \begin{cases} 
\Delta Q \cdot (\tilde{Q} + 1) & Q_{SI_1} < \tilde{Q} \\
SI_1 & Q_{SI_1} = \tilde{Q} \\
\Delta Q \cdot \tilde{Q} & Q_{SI_1} > \tilde{Q} 
\end{cases}$$

where $\Delta Q$ is a quantisation step size and $Q_{SI_1}$ is the quantisation bin of the SI.

Next, the pixels of $SI_2$ with $Q_{SI_2} = \tilde{Q}$ are used to replace the clipped pixels (where $Q_{SI_1} \neq \tilde{Q}$).
Proposed System Decoder

• **Diversity scheme**

→ **Multi-hypothesis SI**

Method allows usage of the correct SI to compensate for the errors appearing in the others.

**WZ Reconstruction (cont.)**

The other SI is sequentially applied to the clipped pixels as follows:

\[
Y_p = \begin{cases} 
SI_p & Q_{SI_p} = \tilde{Q} \text{ and } Q_{SI_{p-1}} \neq \tilde{Q} \\
Y_{p-1} & \text{otherwise}
\end{cases}
\]

where \(Y_p\) is the pixel value after the \(p\)th SI has been received, \(p \in \{2, ..., N\}\). Thus the final refined version of the reconstructed WZ data (\(Y = Y_N\)) exhibits better image quality.
Results

Transform domain SPI-DVC performance of Foreman.

- The proposed codec outperforms the frame-based DVC approach.
- The Gray code and multi-hypothesis coding can further reduce the WZ bitrate by up to 15% and improve the PSNR by approximately 1.5 dB.
- The performance is comparable to Intra H.264 for Breakdancers, and superior for the case of Foreman by up to 1.5 dB.
- For the Foreman sequence, the proposed DVC codec also outperforms H.264 Inter coding with zero motion (~0.8 dB at high bitrates).
Proposed Selective Feedback

- SPI DVC codec
Proposed Selective Feedback

- **Simple Encoder Approach**
  
  - At the decoder, a feedback buffer is created for collecting the clipping maps which show pixels with $Q_{SI} \neq \tilde{Q}$ for each decoded bit-plane.
  - The WZ data is divided into $(BP_{max}+1)$ groups, where $BP_{max}$ is the total bit-planes used with the WZ coding.
  - The first priority is defined to the most significant bit-plane. The encoder will then send the parity bits of this group first. The decoder stops request when BER < threshold.

![Diagram of the Proposed Selective Feedback Process](image)
Proposed Selective Feedback

- **Smart Encoder Approach**
  - Grouping of the WZ pixels in fixed size blocks (e.g. 32x32) first takes place.
  - Parity bit requests are then made for each block in raster scan order.
  - Both encoder and decoder keep track of the number of parity bit requests made for each block.
  - This ranking is then employed for deciding the coding order of those blocks in the following frame, starting with the region with the largest number of requests.
Results

- Parity bits of Breakdancers (pixel-domain coding)

<table>
<thead>
<tr>
<th>Bit-plane</th>
<th>without enhancement (kbits/frame)</th>
<th>with enhancement</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simple encoder</td>
<td>Smart encoder</td>
<td>kbits/frame</td>
<td>%</td>
</tr>
<tr>
<td>1</td>
<td>6.28</td>
<td>4.89</td>
<td>5.19</td>
<td>17.3</td>
</tr>
<tr>
<td>2</td>
<td>14.55</td>
<td>12.41</td>
<td>13.32</td>
<td>8.4</td>
</tr>
<tr>
<td>3</td>
<td>19.44</td>
<td>17.07</td>
<td>18.17</td>
<td>6.5</td>
</tr>
<tr>
<td>total</td>
<td>40.27</td>
<td>34.37</td>
<td>36.68</td>
<td>8.9</td>
</tr>
</tbody>
</table>

- The simple encoder approach provides slightly better performance, at the cost of increased feedback overhead.
Conclusions

This paper presents

- Extension of the Key-Wyner/Ziv framework with spatio-temporal interleaving
- Use of spatio-temporal concealment for generating the side information on a MB basis
- Use of a Gray code which increases the accuracy of bit probability estimation
- Use of a diversity scheme which produces more reliable results by exploiting multiple SI generated data.
- Two selective feedback mechanisms which reduce the number of requested parity bits
Thank you