Coding of Non-Rectangular Signals with Block-Based Transforms

Priyanka Das
Lehrstuhl für Bildverarbeitung, RWTH Aachen University
SVCP 2021
Introduction

Geometric Partitioning [Bläser, Schneider, Sauer & Wien, 2018]

Discrete Cosine Transformation Bases for $8 \times 8$ block length. Transformation $C = T_M S T_N^T$

Supported Geometric partitioning in VVC, Gao, Esenlik, Alshina und Steinbach [2020]
Non-rectangular Signal Transformation: SOTA

- Non-rectangular Transformation
  - \(\Delta DC-SADCT\) (Shape Adaptive Discrete Cosine Transformation)

- Regular transform by extending the signal

SADCT Basic Principle, [Sikora & Makai, 1995]: (a) Irregular Shaped Object, (b) Vertical Shift, (c) Vertical DCT Coefficients, (d) Horizontal Shift, (e) Horizontal DCT Coefficients

- Graph Fourier Transform

- Zero Extension
- Mean Extension
- Low Pass Extrapolation (LPE), [Kaup, 1999]
- Projection onto Convex Sets (POCS), [Chen, Reha Civanlar & Haskell, 1994]
- Extension Interpolation (EI), [Yi, Cho, Kim, Kim & Lee, 1998]
- Smart Padding, [Shen, Zeng & Ming Lei Liou, 2001]
Algorithm Overview

- Sparse representation using partitioned transform bases
  - Similar to successive approximation technique in [Kaup & Aach, 1994]
- Find extension using this representation and other part of partitioned transform bases

Das, Horst und Wien [2020]
Overcomplete Representation

- Represent the signal using partitioned DCT bases
- Overcomplete system
- 1-D mapping is an under-determined equation system

\[ s = Ax, \quad (1) \]

- \( s \): 1-D mapping of signal
- \( x \): Coefficients
- \( A \): 1-D mapping of partitioned bases and normalized.

- Sparse solution:

\[ c = \min_x \| x \|_0 \quad \text{subject to } s = Ax \quad (2) \]

Extension Method

Extension

- Scale up \( x \)
- Combine the extension dictionary using \( x \) to find extension signal
- Regular transformation of extended block

Dictionary for Extension, Anti-diagonal Partitioning, \( 8 \times 8 \) Block
Dictionary Analysis

Absolute Correlation: $| \mathbf{a}_i^T \cdot \mathbf{a}_j |$

Coherence: $\mu(\mathbf{A}) = \max_{1 \leq k,j \leq N, k \neq j} | \mathbf{a}_k^T \cdot \mathbf{a}_j |$

$\hat{\mu}(\mathbf{A}) = \max_{1 \leq k,j \leq N, k \neq j, k \in \mathcal{N}(j), j \notin \mathcal{N}(k)} | \mathbf{a}_k^T \cdot \mathbf{a}_j |$

<table>
<thead>
<tr>
<th>Block Length</th>
<th>$\mu(\mathbf{A})$</th>
<th>$\hat{\mu}(\mathbf{A})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$4 \times 4$</td>
<td>0.61</td>
<td>0.20</td>
</tr>
<tr>
<td>$8 \times 8$</td>
<td>0.69</td>
<td>0.26</td>
</tr>
<tr>
<td>$16 \times 16$</td>
<td>0.73</td>
<td>0.29</td>
</tr>
<tr>
<td>$32 \times 32$</td>
<td>0.75</td>
<td>0.30</td>
</tr>
<tr>
<td>$64 \times 64$</td>
<td>0.76</td>
<td>0.30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Block Length</th>
<th>$\mu(\mathbf{A})$</th>
<th>$\hat{\mu}(\mathbf{A})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$4 \times 4$</td>
<td>0.29</td>
<td>0.25</td>
</tr>
<tr>
<td>$8 \times 8$</td>
<td>0.27</td>
<td>0.22</td>
</tr>
<tr>
<td>$16 \times 16$</td>
<td>0.25</td>
<td>0.21</td>
</tr>
</tbody>
</table>

Expected Neighborhood ($\mathcal{N}(x)$): 

Triangular Block: Anti-diagonal Partitioning, Trapezoidal Block: Equidistant Partitioning from Anti-diagonal and Bottom-left Corner

Absolute Correlation for First $4 \times 4$ Elements, Anti-diagonal Partitioning, $8 \times 8$ Dictionary
Experiments

• Partitioning : Anti-diagonal
• Dataset:
  – Sequences : Race Horses (RH), BQMall (BM), Party Scene (PS), Basket Ball Drill(BD) of resolution 832 × 480
  – Correlated Dataset: Partitioned blocks from first 5 frames
  – Decorrelated Dataset: Residuals generated from VTM 3.2 using the Random Access Configuration at QP=22 in intra mode from 65 frames, are grouped by block size
• Quantizer: uniform
• QP: 12, ..., 51
• Estimated Rate: Total amount of information in transform coefficients
• PSNR: In region of interest
Experimental Results

Race Horse, block length = 8 × 8, triangular block

Race Horse Residuals, block length = 8 × 8, triangular block

<table>
<thead>
<tr>
<th>Block Sequence</th>
<th>Correlated Data</th>
<th>Residual Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td>IQP</td>
<td>mQP</td>
</tr>
<tr>
<td>4 × 4</td>
<td>RH</td>
<td>-1.44</td>
</tr>
<tr>
<td></td>
<td>BM</td>
<td>-1.87</td>
</tr>
<tr>
<td></td>
<td>PS</td>
<td>1.64</td>
</tr>
<tr>
<td></td>
<td>BD</td>
<td>-1.13</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>-0.70</td>
</tr>
<tr>
<td>8 × 8</td>
<td>RH</td>
<td>-7.19</td>
</tr>
<tr>
<td></td>
<td>PS</td>
<td>-0.96</td>
</tr>
<tr>
<td></td>
<td>BD</td>
<td>-7.73</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>-5.79</td>
</tr>
<tr>
<td>16 × 16</td>
<td>RH</td>
<td>-11.55</td>
</tr>
<tr>
<td></td>
<td>BM</td>
<td>-10.69</td>
</tr>
<tr>
<td></td>
<td>PS</td>
<td>-1.85</td>
</tr>
<tr>
<td></td>
<td>BD</td>
<td>-13.79</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>-9.47</td>
</tr>
<tr>
<td></td>
<td>BM</td>
<td>-10.22</td>
</tr>
<tr>
<td></td>
<td>PS</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>BD</td>
<td>-17.81</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>-10.20</td>
</tr>
</tbody>
</table>

Bjøntegaard Delta Rate Measurements
Observations

• Better RD performance for
  – Correlated data
  – Larger blocks
  – Coarser quantization
  – Non-rectangular shape closer to rectangle

Limitations

• Transform coefficients for whole block has to be sent
• Spatially scattered coefficients
• Probable Solution:
  – Subset of the dictionary
  – Entropy coding
Subset of the dictionary

• Transform coefficients of only these positions are to be sent
Subset of the dictionary - triangular block

- Dictionary must be complete
- Immediate horizontal and vertical neighbors are highly Correlated - keeping alternate atom
- Trivial effect on estimated RD behavior
- DC and diagonal elements must be kept
- Number of usable subset must be low
- Higher redundancy in low frequency region
- Open Problem: Formalization
  – Spectral characteristics

Results

Race Horse Residuals, block length = 8 × 8, triangular block

Number of Coefficients to be sent
Current Work

- Integrating the basic scheme in VTM 11.0
- Transform coefficients for either of the non-rectangular signal is to be coded based on RD cost
- Extra flags to be sent
  - one flag to indicate whether this scheme is used
  - one flag to indicate which mask is chosen
- LFNST is disabled for this scheme
- Quadtree split for the transform unit where geo partitioning is used, is forced disabled
- Only enabled for DCT-2 - can be extended easily

- Close to completion
Conclusion

• Basic scheme outside the reference Software had encouraging gain over SADCT
• Issues with scattered coefficients could be mitigated by
  – Subset of dictionary
  – Entropy coding (open problem)
• Implementation in the reference software
Thank you for your attention


