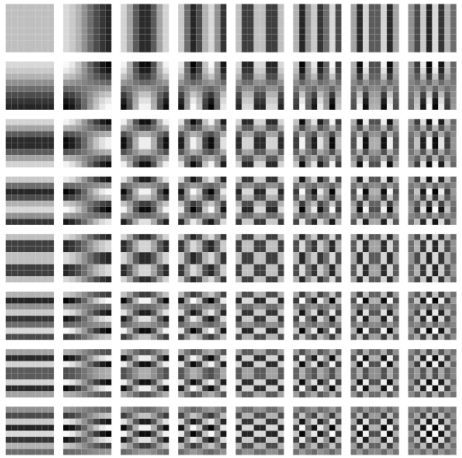
Coding of Non-Rectangular Signals with Block-Based Transforms

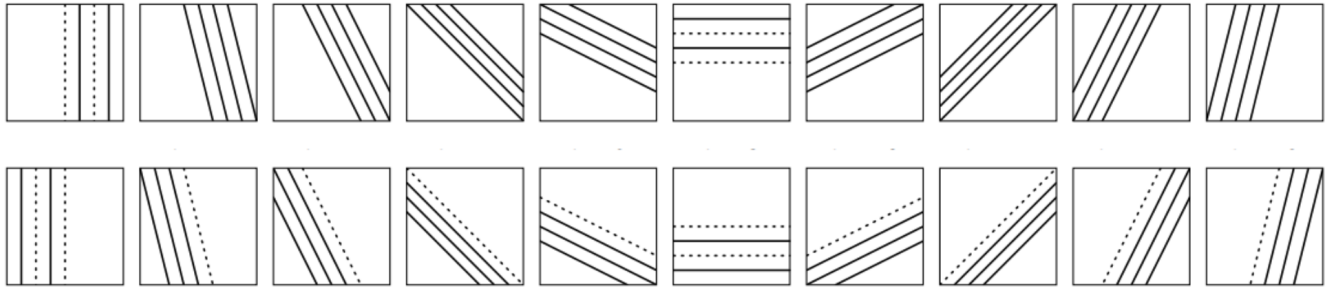
Priyanka Das
Lehrstuhl für Bildverarbeitung, RWTH Aachen University
SVCP 2021

Introduction

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



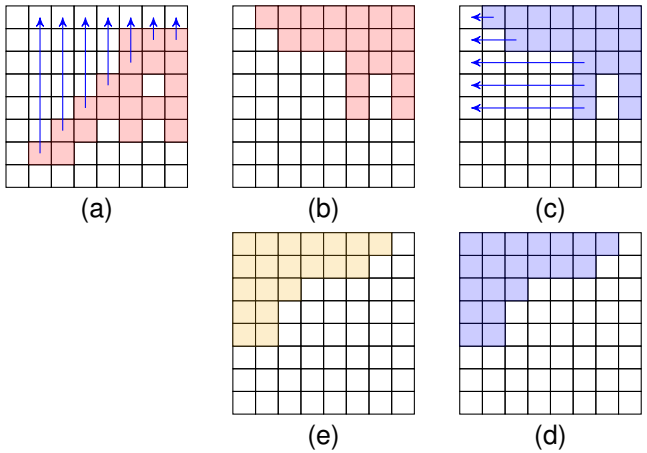
Discrete Cosine Transformation Bases for 8×8 block length. Transformation $\mathbf{C} = \mathbf{T}_M \mathbf{S} \mathbf{T}_N^T$



Supported Geometric partitioning in VVC, Gao, Esenlik, Alshina und Steinbach [2020]

Non-rectangular Signal Tranfsormation: SOTA

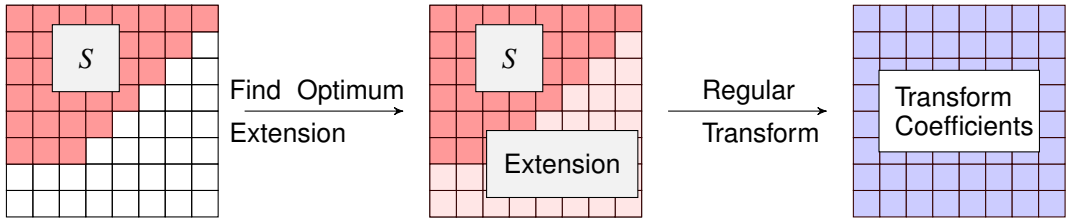
- **Non-rectangular Transformation**
  - **ΔDC-SADCT** (Shape Adaptive Discrete Cosine Transformation)



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

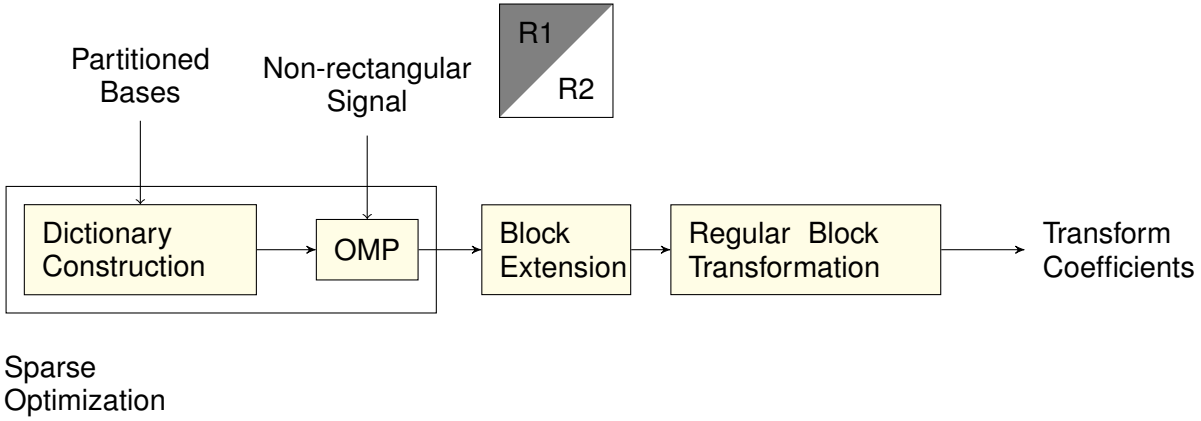
- **Regular transform by extending the signal**



  - 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

$$\mathbf{s} = \mathbf{A}\mathbf{x}, \tag{1}$$

$\mathbf{s}$: 1-D mapping of signal

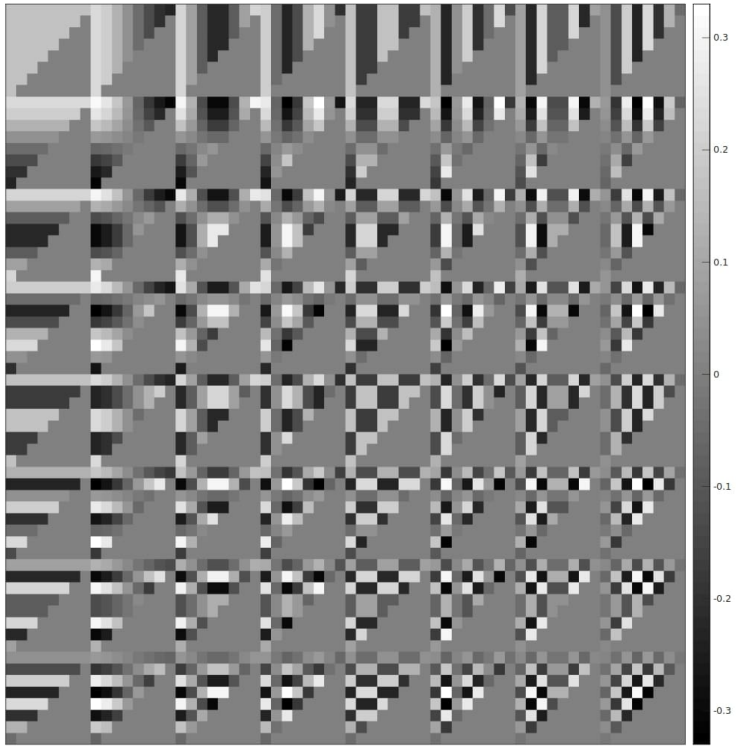
$\mathbf{x}$: Coefficients

$\mathbf{A}$: 1-D mapping of partitioned bases and normalized.

- Sparse solution:

$$\mathbf{c} = \min_{\mathbf{x}} ||x||_0 \text{ subject to } \mathbf{s} = \mathbf{A}\mathbf{x} \tag{2}$$

- NP hard problem - solve using Orthogonal Matching Pursuit [Pati, Rezaiifar & Krishnaprasad, 1993]

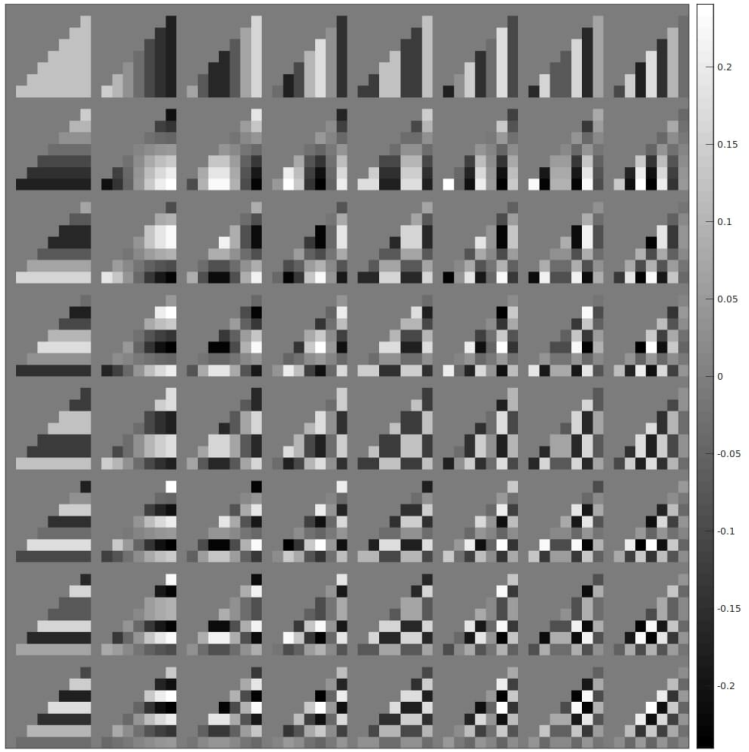


Dictionary for Sparse Coding, Anti-diagonal Partitioning, 8×8 Block

Extension Method

Extension

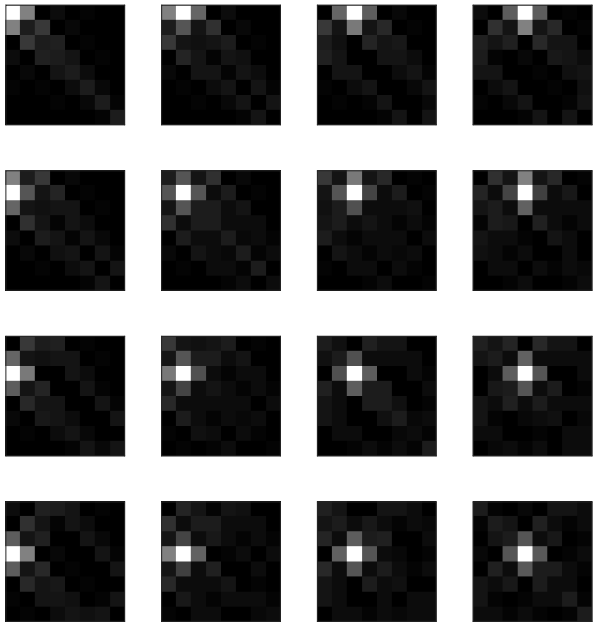
- Scale up $\mathbf{x}$
- Combine the extension dictionary using $\mathbf{x}$ to find extension signal
- Regular transformation of extended block



Dictionary for Extension, Anti-diagonal Partitioning, 8×8 Block

Dictionary Analysis

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



Absolute Correlation for First 4×4 Elements,
Anti-diagonal Partitioning, 8×8 Dictionary

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

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

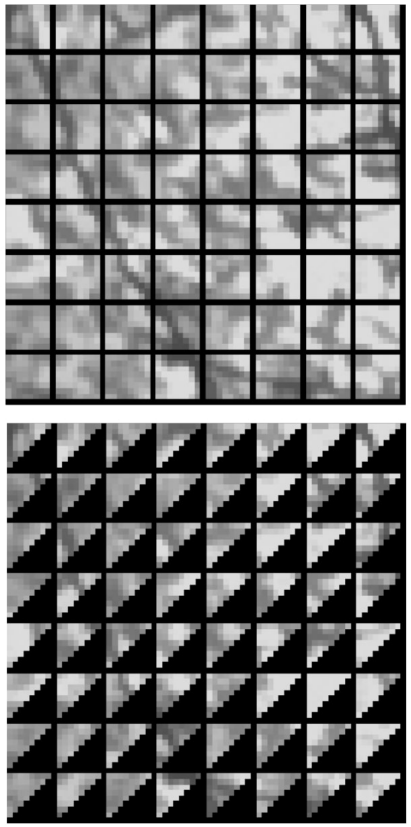
	$\mu(\mathbf{A})$	$\hat{\mu}(\mathbf{A})$
Triangular Block		
Block Length: 4×4	0.61	0.20
Block Length: 8×8	0.69	0.26
Block Length: 16×16	0.73	0.29
Block Length: 32×32	0.75	0.30
Block Length: 64×64	0.76	0.30
Trapezoidal Block		
Block Length: 4×4	0.29	0.25
Block Length: 8×8	0.27	0.22
Block Length: 16×16	0.25	0.21

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

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

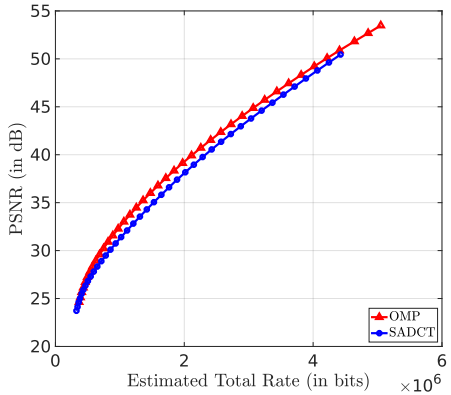
Experiments

- 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

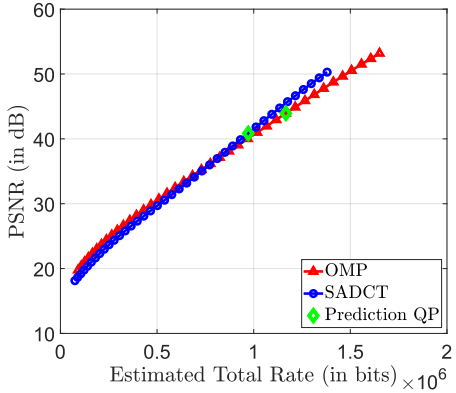


Partitioning of Correlated Data

Experimental Results



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



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

Block Length	Sequence	Correlated Data			Residual Data		
		lQP	mQP	hQP	lQP	mQP	hQP
4×4	RH	-1.44	-1.89	3.57	8.66	0.78	-10.00
	BM	-1.87	-2.19	2.50	8.63	0.54	-9.96
	PS	1.64	-2.87	-1.30	10.88	4.46	-5.66
	BD	-1.13	-0.89	5.92	6.54	-2.98	-13.75
	Average	-0.70	-1.96	2.67	8.68	0.70	-9.84
8×8	RH	-7.19	-11.70	-10.66	3.35	-4.87	-14.08
	BM	-7.29	-11.97	-10.59	3.91	-7.20	-18.99
	PS	-0.96	-9.61	-14.03	7.04	-0.83	-10.75
	BD	-7.73	-12.44	-7.33	3.64	-7.94	-18.48
	Average	-5.79	-11.43	-10.65	4.49	-5.21	-15.58
16×16	RH	-11.55	-19.50	-23.60	-6.46	-16.73	-23.56
	BM	-10.69	-19.47	-22.84	-1.52	-20.32	-39.48
	PS	-1.85	-13.96	-23.94	1.69	-11.86	-24.35
	BD	-13.79	-22.48	-22.22	-0.87	-14.34	-23.58
	Average	-9.47	-18.85	-23.15	-1.79	-15.81	-27.74
32×32	RH	-13.01	-24.20	-31.54	-16.92	-28.64	-30.16
	BM	-10.22	-22.70	-29.06	-11.46	-37.17	-55.45
	PS	0.23	-16.10	-29.80	-2.68	-22.49	-37.73
	BD	-17.81	-29.63	-33.07	-4.31	-19.57	-29.62
	Average	-10.20	-23.16	-30.87	-8.84	-26.97	-38.24

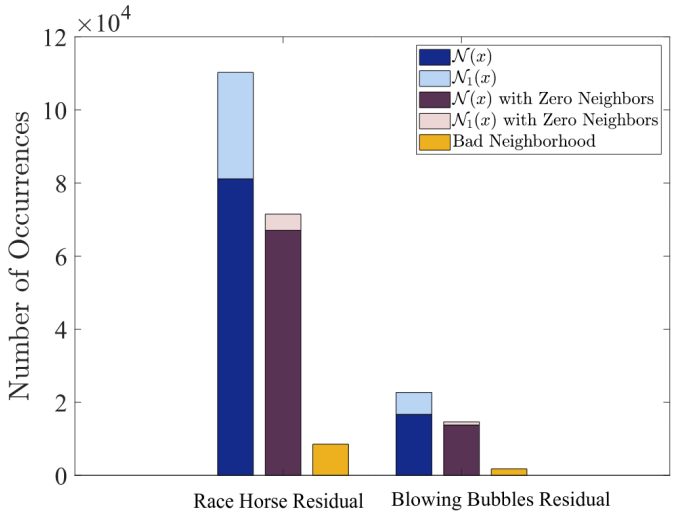
Bjøntegaard Delta Rate Measurements

Observations

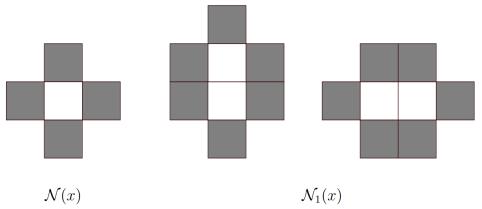
- Better RD performance for
  - Correlated data
  - Larger blocks
  - Coarser quantization
  - Non-rectangular shape closer to rectangle

Limitations

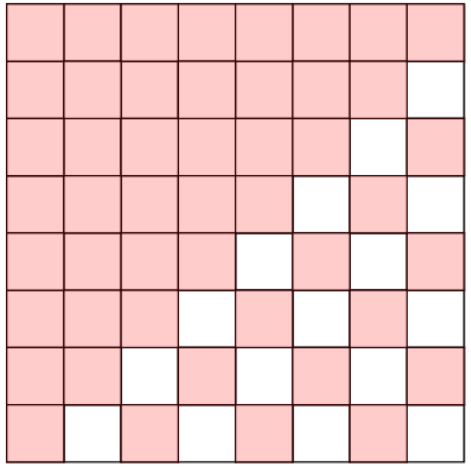
- Transform coefficients for whole block has to be sent
- Spatially scattered coefficients
- Probable Solution:
  - Subset of the dictionary
  - Entropy coding



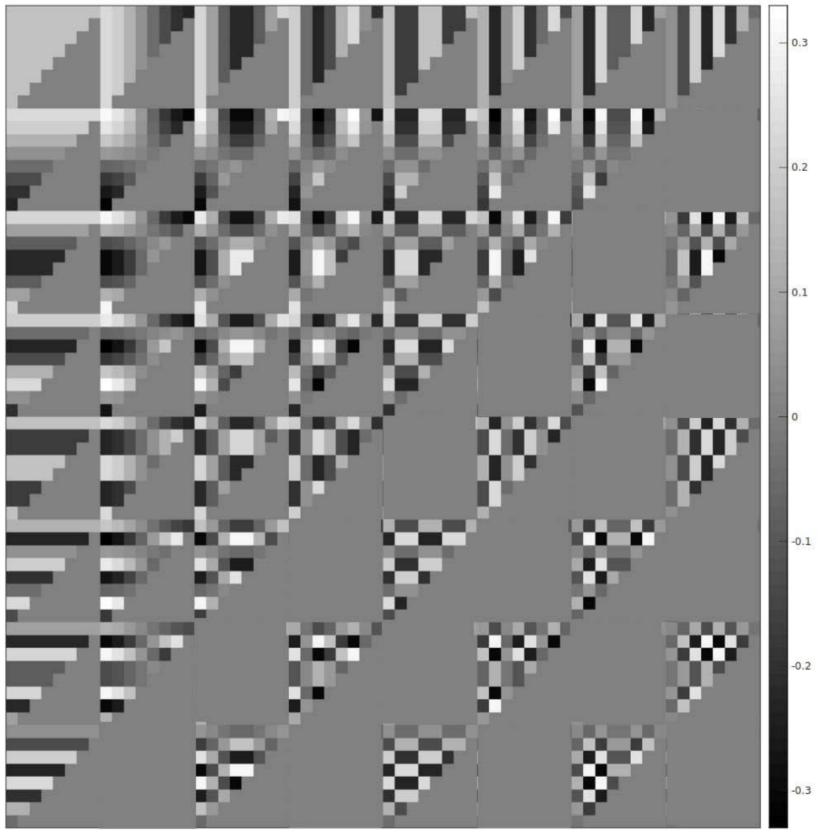
Example of Neighborhood statistics



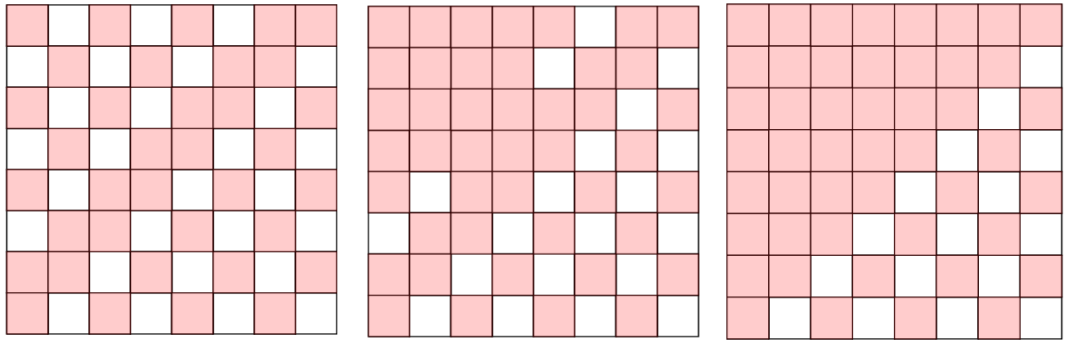
Subset of the dictionary



- Transform coefficients of only these positions are to be sent

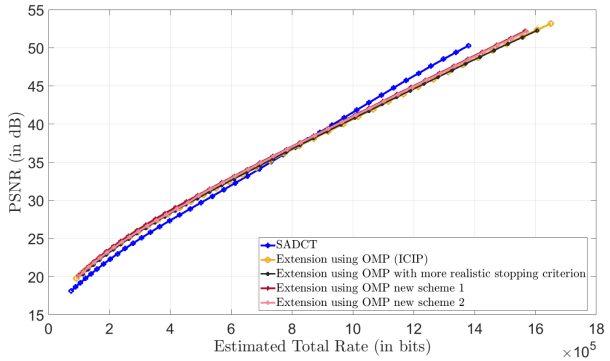


Subset of the dictionary - triangular block

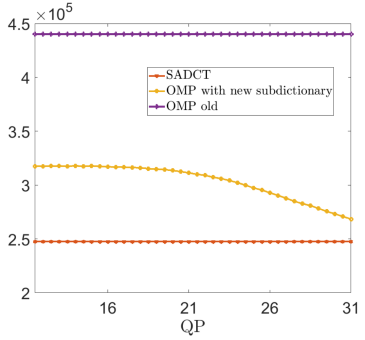


- Dictionary must be complete
- Immediate horizontal and vertical neighbors are highly Correlated - keeping alternate atom
- Trivial effect on estimated RD behavior
- DC and diagonal elements must be kept
- Number of usable subset must be low
- Higher redundancy in low frequency region
- Open Problem: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

Results



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

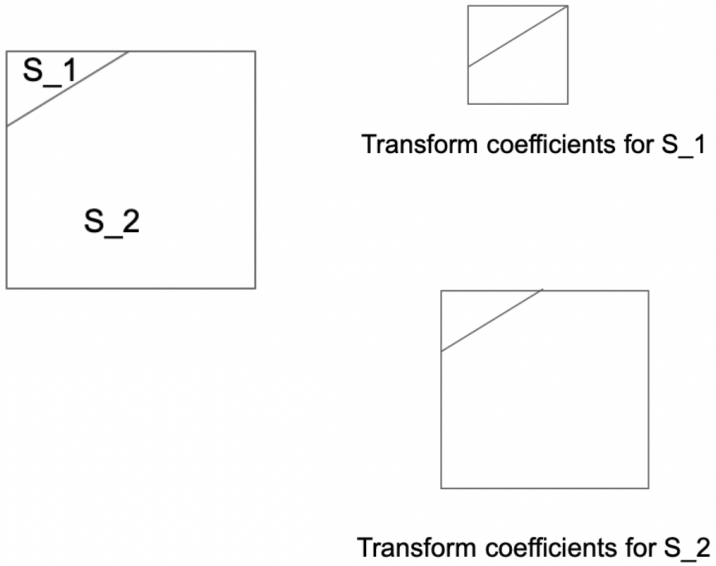


Number of Coefficients to be sent

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



Extension of an exemplary shape

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

Literature I

Bläser, M., Schneider, J., Sauer, J. & Wien, M. (2018, June). Geometry-based partitioning for predictive video coding with transform adaptation. In *2018 picture coding symposium (pcs)* (S. 134-138). doi: 10.1109/PCS.2018.8456238

Chen, H. H., Reha Civanlar, M. & Haskell, B. G. (1994, Nov). A block transform coder for arbitrarily shaped image segments. In *Proceedings of 1st international conference on image processing* (Bd. 1, S. 85-89 vol.1). doi: 10.1109/ICIP.1994.413280

Das, P., Horst, N. & Wien, M. (2020). Coding of non-rectangular signals with block-based transforms. *IEEE International Conference on Image Proicessing*.

Gao, H., Esenlik, S., Alshina, E. & Steinbach, E. (2020). Geometric partitioning mode in versatile video coding: Algorithm review and analysis. *IEEE Transactions on Circuits and Systems for Video Technology*, 1-1. doi: 10.1109/TCSVT.2020.3040291

Kaup, A. (1999, Feb). Object-based texture coding of moving video in mpeg-4. *IEEE Transactions on Circuits and Systems for Video Technology*, *9* (1), 5-15. doi: 10.1109/76.744271

Kaup, A. & Aach, T. (1994, April). Segment-oriented coding of textured images based on successive approximation. In *Proceedings of icsipnn '94. international conference on speech, image processing and neural networks* (S. 197-200 vol.1). doi: 10.1109/SIPNN.1994.344933

Pati, Y. C., Rezaiifar, R. & Krishnaprasad, P. S. (1993, Nov). Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. In *Proceedings of 27th asilomar conference on signals, systems and computers* (S. 40-44 vol.1). doi: 10.1109/ACSSC.1993.342465

Shen, G., Zeng, B. & Ming Lei Liou. (2001, Jan). Arbitrarily shaped transform coding based on a new padding technique. *IEEE Transactions on Circuits and Systems for Video Technology*, *11* (1), 67-79. doi: 10.1109/76.894284

Sikora, T. & Makai, B. (1995, Feb). Shape-adaptive dct for generic coding of video. *IEEE Transactions on Circuits and Systems for Video Technology*, *5* (1), 59-62. doi: 10.1109/76.350781

Yi, J. W., Cho, S. J., Kim, W. J., Kim, S. D. & Lee, S. J. (1998, Jun). A new coding algorithm for arbitrarily shaped image segments. *Signal Processing:Image Communication*, *12*, 231–242.