VVC 비디오 코덱을 위한 변환 커널 유도 방법

산딥 쉬레스따, *이범식
조선 대학교
sandeepshrestha407@chosun.kr, *bslee@chosun.ac.kr (교신저자)

An ANALYTICTRANSFORM KERNEL DERIVATION METHOD FOR VERSATILE VIDEO CODING (VVC)

Sandeep Shrestha, *Bumshik lee
Chosun University

ABSTRACT

In the ongoing standardization of Versatile Video Coding (VVC), DCT−2, DST−7 and DCT−8 are accounted as the vital transform kernels. While storing all of those transform kernels, ROM memory storage is considered as the major problem. So, to deal with this scenario, a common sparse unified matrix concept is introduced in this paper. From the proposed matrix, any point transform kernels (DCT−2, DST−7, DCT−8, DST−4 and DCT−4) can be achieved after some mathematical computation. DCT−2, DST−7 and DCT−8 are the used major transform kernel in this paper.

1. INTRODUCTION

In High Efficiency Video Coding (HEVC) [1], DCT−2 is considered as the vital transform kernel because of its subsampling property for smaller sizes from its larger sized transform kernel [1] and fast−algorithm [2]. In Versatile Video Coding (VVC) [3] standardization, DCT−2, DST−7 and DCT−8 also termed as Multiple Transform Kernels (MTS) [5] is regarded as the vital transform kernel.

During the standardization of VVC, various proposals regarding memory reduction transform kernels were introduced. Compound Orthonormal Transform [6] and Unified Matrix [7] were transform kernel memory related proposals which was considered to study in CE but not adopted in spite of significant result. The Compound Orthonormal Transform is composed of 4−point and 8−point DST−4/DCT−4 and 16−point and 32−point DST−7/DCT−8 embedded into 64−point DCT−2 transform kernel. The proposed transform kernel failed to provide 64−point DCT−2 so that fast algorithm for 64−point DCT−2 could not be applied and five transform kernels are used instead of three. Similarly, in [7] performed mathematical computation which showed no significant gain in AI, RA and LDB for higher resolutions due to the deviation of the DCT−2 transform kernel from the original DCT−2.

In this paper, analytical derivation of transform kernel is performed using common sparse unified matrix which comprises of two parts: Unified DST−3 matrix (**U**) and Grouped DST−7 matrix. The proposed common sparse unified matrix stores only 1648 elements each of 8−bit precision as ROM memory storage which is shown in Fig. 1.

The paper is organized as follows: In Section 2, the proposed common sparse unified matrix is described. In Section 3, experimental results are given and the paper is concluded in Section 4.

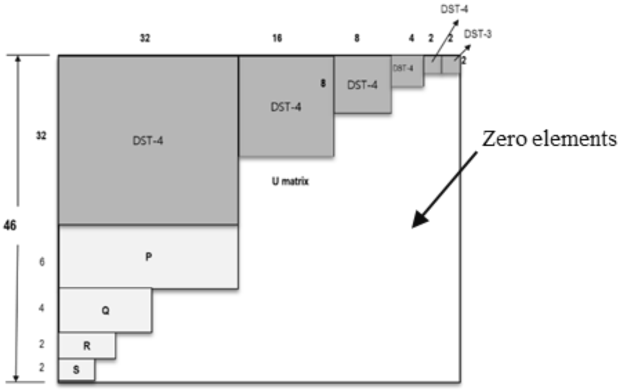


Fig. 1. Common sparse unified matrix

2. COMMON SPARSE UNIFIED MATRIX

As shown in fig. 1, the proposed matrix comprises of two parts: Unified Matrix and grouped DST−7 matrix. The overall size of the proposed sparse unified matrix is 1648 elements each of 8−bit precision.

2.1. UNIFIED DST−3 MATRIX ($\mathbf{U}$)

The proposed Unified matrix ($\mathbf{U}$) stores only 1368 elements out of 64×64 elements and composed up of different point DST−4 transform kernels and a 2×2 DST−3 transform kernel as shown in Fig. 1 and is used to derive DCT−2 transform kernel. To derive DCT−2 transform kernel, initially DST−3 transform kernel is derived using unified DST−3 matrix ($\mathbf{U}$) and five−unit element matrix whose relation can be presented as:

$$\mathbf{S}_{3,64} = \mathbf{UABCDE} \tag{1}$$

where, $\mathbf{S}_{3,64}$ is the 64−point DST−3 matrix and $\mathbf{U}$ represents the 64−point unified DST−3 matrix and A, B, C, D and E represents 64−point unit−element matrices composed of −1, 0 and 1 [8].

Using (1), continuous multiplication with successive unit element matrices gives the larger sizes of DST−3 matrix. After multiplication, the right bottom part of the matrix is selected which gives us required sized DST−3 transform kernel. Finally, DCT−2 can readily be derived using the relation (2).

$$\mathbf{C}_2 = \mathbf{F} \times \mathbf{S}_3 \times \mathbf{S} \tag{2}$$

where, $\mathbf{C}_2$ and $\mathbf{S}_3$ are DCT−2 and DST−3 transform kernels respectively and F and S are a flipping and a sign change matrices [8], respectively.

2.2. GROUPED DST−7 MATRIX

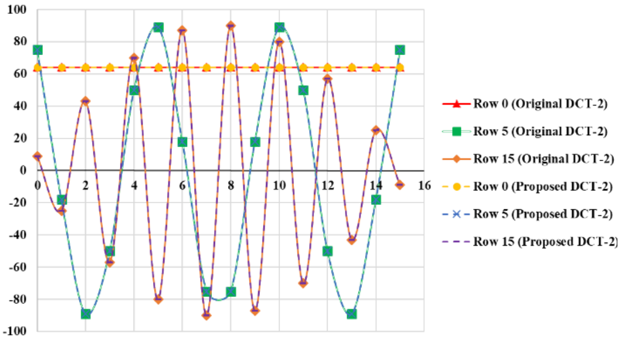
P, Q, R and S in Fig. 1. represent grouped DST−7 matrices which only store few specific rows of different point DST−7 matrix. In grouped DST−7 matrix [8], the specific rows stored are multiplied with the permutation matrix (G) to achieve the full size matrix. The permutation matrix is designed in such a way that it follows the certain repeating element pattern in each group [8]. In the row obtained after multiplication, the first element of the previous row should be equal to the last element of resultant row. The process is not halted until the last element of the first stored row matches with the first element of the last row obtained without considering the sign values. If there exist any row whose first element is negative after multiplication with the permutation matrix, then the sign values of each element of the respective row is changed. [8] provides all the grouping rows that depends on each other and meet the criteria mentioned already. Finally, DCT−8 is derived using the following relation

$$\mathbf{C}_8 = \mathbf{S} \times \mathbf{S}_7 \times \mathbf{F} \tag{3}$$

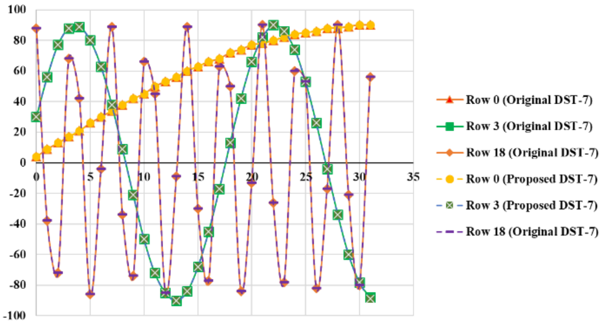
where, $\mathbf{C}_8$ and $\mathbf{S}_7$ are DCT−8 and DST−7 transform kernel respectively and S and F are sign change and flipping matrix respectively.

3. EXPERIMENTAL RESULTS

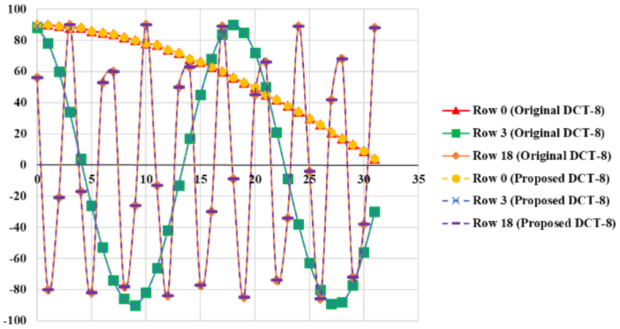
Based on Table.1 results, we can say that the due to significant closeness of the proposed transform kernel signals with the original transform kernels as in fig. 2 −(a), −(b), −(c), there is no loss in the overall result of AI i.e. no loss in the luminance where as 0.02% gain in the chrominance. The experiment was conducted under CTC condition [9] VTM−3.0 anchor and signifies that there is some gain in the higher resolution class sequences.



(a) DCT−2(16−point)



(b) DST−7 (32−point)



(c) DCT−8 (32−point)

Fig. 2. Comparison between the proposed and original kernel

Table.1. Simulation results

	All Intra Main10				
	Over VTM−3.0				
	Y	U	V	EncT	DecT
Class A1	0.00%	−0.08%	0.00%	100%	100%
Class A2	0.00%	−0.04%	0.00%	100%	100%
Class B	0.00%	0.02%	0.06%	100%	100%
Class C	0.00%	−0.04%	−0.07%	100%	100%
Class E	0.00%	0.04%	0.00%	100%	100%
Overall	0.00%	−0.02%	0.00%	100%	100%
Class D	0.00%	0.00%	0.10%	102%	100%

4. CONCLUSION

In this paper, a kernel derivation method for the VVC video codec is proposed. From the limited specific rows all the DCT−2, DST−4, DCT−4, DST−7 and DCT−8 transform kernels can be easily derived with few static memory usage. The proposed method in this paper gives comparable results with the VTM−3.0 anchor with negligible loss which signifies that the proposed transform kernels are significantly close to the original signal in VTM−3.0 anchor.

ACKNOWLEDGEMENT

This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT [NRF−2017R1C1B1008646] also supported by the MIST (Ministry of Science & ICT, Korea, under the National Program for Excellence in SW) supervised by the IITP (Institute for Information & communications Technology Promotion)” (2017−0−00137)

REFERENCES

[1] Gary J. Sullivan , Jens−Rainer Ohm , Woo−Jin Han , Thomas Wiegand, ‘Overview of the High Efficiency Video Coding (HEVC) Standard’, IEEE, pp. 1649-1668, Sep. 2012

[2] Wen−Hsiung Chen , C. Smith , S. Fralick, ‘A Fast Computational Algorithm for the Discrete Cosine Transform’, IEEE Transactions on Communications, vol. Volume: 25, pp. 1004-1009, Sep. 1977.

[3] Chen , Y. Ye, S. H. Kim, ‘Algorithm description for Versatile Video Coding and Test Model 3 (VTM 3),[JVET−L1002]’, JVET of ITU−T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 12th Meeting: Macao, CN, Oct. 2018.

[4] K. Choi, M. Park, C. Kim, ‘CE6: AMT with reduced transform types’, JVET−K0171, 11th JVET meeting, Ljubljana, SI, Jul. 2018.

[5] M. Koo, M. Salehifar, J. Lim, S.H. Kim, ‘CE6−1.11: AMT replacement and restriction’, JVET−K0096, 11th JVET meeting, Ljubljana, SI, Jul. 2018.

[6] X. Zhao, X. Li, S. Liu, ‘CE6: Compound Orthonormal Transform’. JVET−M0496, 13th JVET meeting, Marrakech, MA, Jan−2019.

[7] K. Choi, M. Park, M.W. Park, W. Choi, ‘CE6: Unified matrix for transform’. JVET−M0200, 13th JVET meeting, Marrakech, MA, Jan−2019.

[8] S. Shrestha, A. Kumar, B. Lee, Y. Lee, J. Park ‘CE6−related: MTS kernel derivation for efficient memory usuage’. JVET−M0396, 13th JVET meeting, Marrakech, MA, Jan−2019.

[9] J. Boyce, K. Suehring, X. Li, V. Seregin, ‘JVET common test conditions and software reference configurations for SDR video’. JVET−L1010, 12th JVET meeting, Macao, CN, Oct−2018.