DST-4 와 DCT-4를 위한 DST-3 기반 비디오 압축 변환 커널 유도 방법

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

A DST-3 BASED TRANSFORM KERNEL DERIVATION METHOD FOR DST-4 and DCT-4 IN VIDEO CODING

Sandeep Shrestha, *Bumshik lee
Chosun University

ABSTRACT

In the ongoing standardization of Versatile Video Coding (VVC), DCT-2, DST-7 and DCT-8 are designated as the vital primary transform kernels. Due to the effectiveness of DST-4 and DCT-4 in smaller resolution sequences, DST-4 and DCT-4 transform kernel can also be used as the replacement of the DST-7 and DCT-8 transform kernel respectively. While storing all of those transform kernels, ROM memory storage is considered as the major issue. So, to deal with this scenario, a unified DST-3 based transform kernel derivation method is proposed in this paper. The transform kernels used in this paper is DCT-2, DST-4 and DCT-4 transform kernels. The proposed ROM memory required to store the matrix elements is 1368 bytes each of 8-bit precision.

1. INTRODUCTION

In the ongoing standardization of Versatile Video Coding (VVC) [1], DCT-2, DST-7 and DCT-8 are designated as the major transform kernels which are also known as MTS [2]. DCT-2 fits the residual signal amplitude close to each other whereas, DST-7 and DCT-8 fit the statistics better if the residual signal amplitude increases as a function of the distance from the boundary samples used for prediction [4]. Along with the mentioned transform kernels, DCT-4 and DST-4 can also be used as the transform kernels as they exhibit better coding efficiency for smaller resolution sequences [6]. In general, DST-4 can be used as the replacement of the DST-7 transform kernel and DCT-4 can be used as the replacement of the DCT-8 transform kernel [6].

Storing all those transform kernels are regarded as the major problem. So, to deal with this situation, various ROM memory reduction transform kernels were proposed during the VVC standardization. Compound Orthonormal Transform [7] and Unified Matrix [8] are amongst transform kernel memory-related proposals which was considered to study in core experiment but not adopted in spite of significant result. The Compound Orthonormal Transform is composed of 4-point and 8-poin 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 a fast algorithm for 64-point DCT-2 could not be applied and five transform kernels are used instead of three. Similarly, the proposal presented in [7] uses mathematical computation with no significant gain 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 (DCT-2, DST-4 and DCT-4) is proposed using sparse unified DST-3 matrix (**U**). The proposed methods saves 1368 bytes for 8-bit precision to derive any size and type of transform kernels from a single matrix without any extra computation.

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

2. PROPOSED METHOD

The proposed structure of the sparse unified DST-3 matrix (**U**) can be illustrated as Fig. 1. Based on the Fig. 1, **U** matrix consist of different point DST-4 and a 2×2 DST-3 transform kernel. So, in order to derive DCT-4 matrix of specific block size from specific point DST-4 transform kernel of **U** matrix, the relationship between DST-4 and DCT-4 [6] can be expressed as

$$\mathbf{C_4} = \mathbf{S} \times \mathbf{S_4} \times \mathbf{F} \quad (1)$$

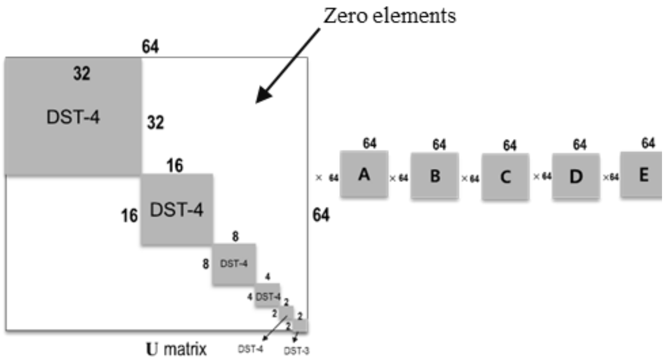


Fig 1. Structure of sparse unified 64−point DST−3 matrix (**U**) and unit−element matrices (A, B, C, D, and E).

where, $\mathbf{C_4}$ and $\mathbf{S_4}$ are DCT−4 and DST−4 respectively and F and S are the flipping and the sign change matrices, respectively [9] [10].

Secondly, for obtaining DCT−2 matrix of specific block size from **U** matrix, firstly, respective block size DST−3 transform kernel is required. To obtain DST−3 transform kernel matrix of specific block size, a sparse unified DST−3 matrix (**U**) and five unit−element matrices A, B, C, D and E is proposed in this paper as shown in Fig 1. The unit−element matrices are 64−point matrixes composed of −1, 0 and 1 [9].

From Fig. 1, the successive multiplication with respective unit element matrices A, B, C, D and E from **U** matrix gives the larger block sized DST−3 matrix and after multiplication, the right bottom part of the matrix is selected which gives the respective block sized DST−3 transform kernel matrix.

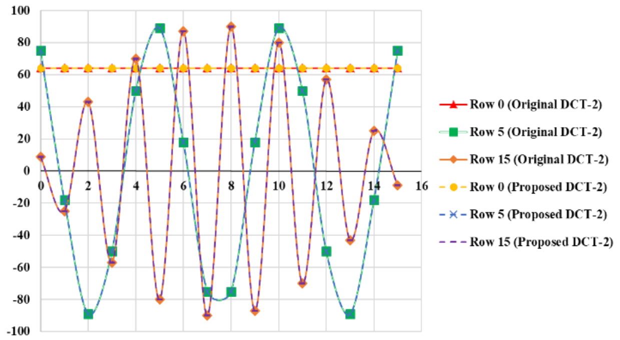
Finally, respective block sized DCT−2 transform kernel matrix can readily derived using the relation [10] as:

$$\mathbf{C_2} = \mathbf{F} \times \mathbf{S_3} \times \boldsymbol{S} \qquad (2)$$

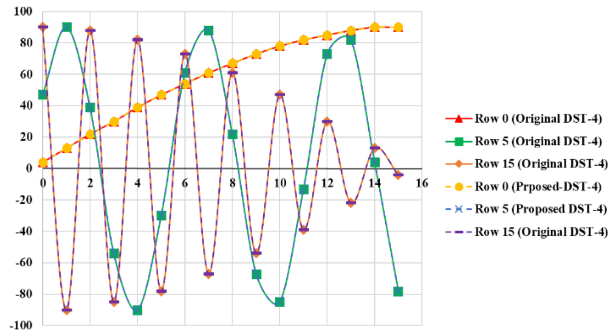
where, $\mathbf{C_2}$ and $\mathbf{S_3}$ are DCT−2 and DST−3 transform kernels respectively and F and S are flipping and sign change matrices [9] [10], respectively.

3. EXPERIMENTAL RESULTS

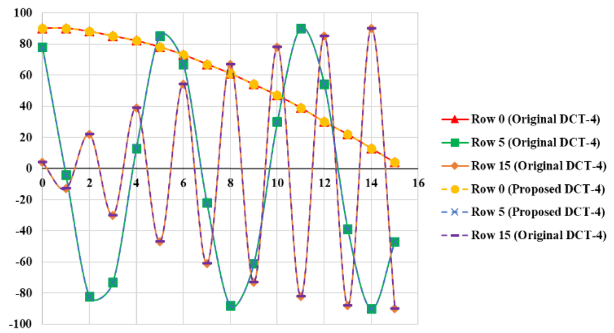
Fig. 2 shows the comparison between the 16−point original transform kernels and the proposed transform kernels. Based on the signal shown, the proposed and original transform kernels perfectly match with each other. Based on signal plot in Fig. 2−(a), −(b), −(c), we can see that in Table. 1 the average loss of 0.17% in the luminance of proposed COT whereas negligible average loss of 0.06% in the proposed method. For chrominance proposed method showed gain of −0.23% whereas COT results showed −0.03% gain which is regarded as not much significant gain. From the memory perspective also proposed method seem to be better than COT. The anchor used for experiment is VTM−3.0 reference software. The number of frames for test is set to 100 frames. Except for the number of frames, the experiments follow the common test conditions (CTC) [11]. The PC for simulation has CPU Intel® Xeon® CPU E3−1275v5 @ 3.60GHz processor, RAM 16 GB with 64−bits Windows 10.



(a) DCT−2 (16−point)



(b) DST−4 (16−point)



(c) DCT−4 (16−point)

Fig 2. Comparisons of the proposed and original kernels

Table 1. Simulation results for the AI configuration over VTM−3.0

Seq.	Proposed			COT		
	Y	U	V	Y	U	V
Class C	-0.04%	-0.35%	0.31%	0.12%	-0.08%	-0.08%
Class D	0.03%	-0.43%	0.04%	0.11%	0.02%	-0.09%
Class E	0.20%	0.09%	-0.07%	0.28%	-0.05%	-0.05%
Average	0.06%	-0.23%	0.09%	0.17%	-0.03%	-0.07%

4. CONCLUSIONS

The proposed method in this paper deals with the ROM memory storage issue by storing few transform kernels. The unified sparse matrix contains a number of zero elements, it is not necessary to save all elements of the matrix, thus resulting in significant memory savings compared to the conventional method. The proposed method shows the outperforming memory saving compared with 33.39% with maintaining compression performance.

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] J. 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.

[2] K. Choi, M. Park, C. Kim, 'CE6: AMT with reduced transform types', JVET-K0171, 11th JVET meeting, Ljubljana, SI, Jul. 2018.

[3] 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.

[4] 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.

[5] J. Chen and E. Alshina, 'Algorithm description for Versatile Video Coding and Test Model 1 (VTM1),[JVET-J1002]'. JVET of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 9th meeting, San Diego, CA, USA, Apr-2018.

[6] Kiyofumi Abe, Tadamasa Toma(Panasonic), Masaru Ikeda, Takeshi Tsukuba(Sony), Karam Naser, Fabrice Le Leannec, Edouard Francois(Technicolor), 'CE6: JVET-L0262: Replacing all DST-7 / DCT-8 by DST-4 / DCT-4 used in MTS(test 6.1.1e)'. JVET of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 13th Meeting: Marrakesh, MA, 09-Jan-2019.

[7] X. Zhao, X. Li, S. Liu, 'CE6: Compound Orthonormal Transform'. JVET-M0496, 13th JVET meeting, Marrakech, MA, Jan-2019.

[8] K. Choi, M. Park, M.W. Park, W. Choi, 'CE6: Unified matrix for transform'. JVET-M0200, 13th JVET meeting, Marrakech, MA, Jan-2019.

[9] S. Shrestha, A. Kumar, B. Lee, Y. Lee, J. Park, 'CE6-related: DST-3 based transform kernels derivation'. JVET-M0397, 13th JVET meeting, Marrakech, MA, 9-18 Jan. 2019.

[10] A. Said, H. Egilmez, V. Seregin, M. Karczewicz, 'Complexity Reduction for Adaptive Multiple Transforms (AMTs) using Adjustment Stages'. JVET-J0066, 10th Meeting, San Diego, US, Apr-2018.

[11] 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.