Joint Collaborative Team on Video Coding (JCT-VC)
of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11
4th Meeting: Daegu, KR, 19-28, January, 2011

Document: JCTVC-D182
WG11 Number: m18939

Title: **Performance report of adaptive DCT/DST selection**

Status: Input Document to JCT-VC

Purpose: Proposal

Author(s) or Contact(s): Atsuro Ichigaya
Shinichi Sakaida

NHK
Science & Technology Research Laboratories
1-10-11, Kinuta, Setagaya-ku, Tokyo, Japan

Tel: +81-3-5494-3354
Email: ichigaya.a-go@nhk.or.jp
sakaida.s-gq@nhk.or.jp

Source: NHK (Japan Broadcasting Corporation)

Abstract

This document describes the coding performance of an adaptive DCT/DST selection method on HEVC test Model (HM) 0.9. In this proposed method, DST is used as an additional transform for intra and inter coding blocks in JCTVC-A122 [1]. DCT and DST are selected adaptively according to input signal characteristics without any training. Adaptive transform selection to the luma signal is implemented in HM0.9 and the performance is evaluated in this document. The coding results shows that the BD-rate reduction is up to 2.7% (0.8% on average) for low delay LC condition.

.

1 Introduction

Alternative transforms are discussed in CE7, which was established as JCTVC-C507 [2]. The target of almost all these proposals, except one, is to only improve intra coding efficiency and use KLT, which is designed using offline training. These approaches require the implementation of a number of transform circuits in a decoder, which increases hardware complexity. Samsung proposed the Mode-Dependent DCT/DST for intra prediction [3]. Their proposed method does not require any training. Moreover, instead of preparing many transform types, this method can make a number of transforms by combining 1D-DCT and 1D-DST type VII. These combinations are used only in intra prediction modes. On the other hand, our proposed method selectively uses 2D-DCT and 2D-DST type II, and the selected transform can be applied to residual signals for both intra and inter blocks with simple signalling.

2 Overview of proposed method

DCT works well in video compression because it has the ability to express DC power by one coefficient. However, since high efficiency predictions are adopted in the new video coding method, DC power in the residual signals decreases. In this case, DCT is not suitable for coding residual signals. On the other hand, DST is suitable for such signals with regards to energy compaction. Moreover, it is not easy for DCT to clearly separate amplitude-frequency and phase-frequency characteristics. Therefore, quantization of DCT components affects signal phase at the same time. This is one of factors of block distortion.

(a) DCT basis (b) DST basis

DST-II has a strong relation to DCT-II, which is the most common DCT type for video coding. It is easy to design from DCT circuits, so our proposal uses DCT and DST adaptively to both intra and inter blocks with signaling. DCT and DST compose DFT by using Euler’s formula. Therefore, these transforms have complementary properties for video coding each other. Both transforms are defined by the following equations.

DCT

$$C(u)=\sqrt{\frac{2}{N}}\gamma(u)\sum_{x=0}^{N-1} f(x)\cos\left(\frac{(2x+1)u\pi}{2N}\right) \qquad \text{u=1, 2,…, N} \qquad (1)$$

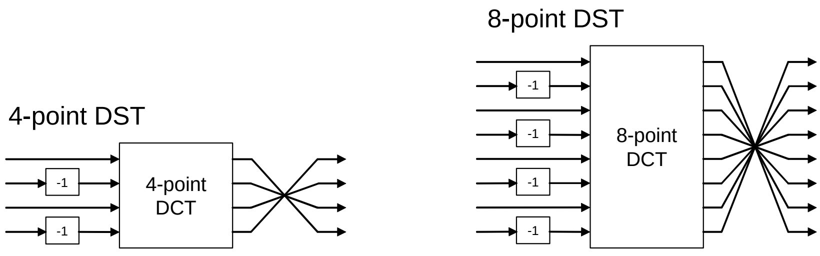
DST

$$S(u)=\sqrt{\frac{2}{N}}\gamma(u)\sum_{x=0}^{N-1} g(x)\sin\left(\frac{(2x+1)u\pi}{2N}\right) \qquad \text{u=1, 2,…, N} \qquad (2)$$

These transforms consist of trigonometric functions. In short, DST can be expressed by DCT with additional elements. The sine function in DST can be re-described with additional theorems.

$$\begin{aligned} S(N-u)&=\sqrt{\frac{2}{N}}\gamma(u)\sum_{x=0}^{N-1} g(x)\sin\left(\frac{(2x+1)(N-u)\pi}{2N}\right) \\ &=\sqrt{\frac{2}{N}}\gamma(u)\sum_{x=0}^{N-1} (-1)^x g(x)\cos\left(\frac{(2x+1)u\pi}{2N}\right) \end{aligned} \qquad \text{u=0, 1,…, N-1} \qquad (3)$$

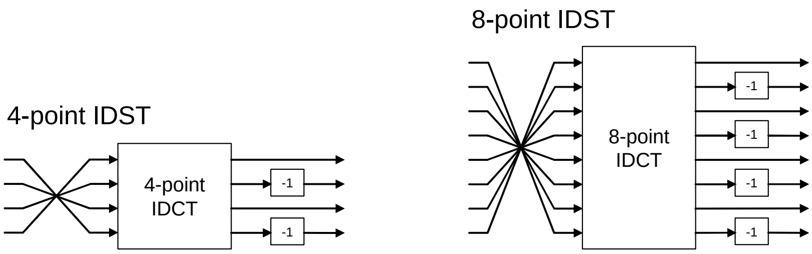
This equation (3) shows that DST can consist of DCT and half the number of inverters. This means DST can be implemented using a DCT circuit as follows.



Similarly, inverse DST can be expressed by inverse DCT with additional elements. The inverse DST is described as follows.

$$\begin{aligned} g(x)&=\sqrt{\frac{2}{N}}\sum_{u=0}^{N-1}\gamma(u)G(N-u)\sin\left(\frac{(2x+1)(N-u)\pi}{2N}\right) \\ &=(-1)^x\sqrt{\frac{2}{N}}\sum_{u=0}^{N-1}\gamma(u)G(N-u)\cos\left(\frac{(2x+1)u\pi}{2N}\right) \end{aligned} \qquad \text{u=0, 1,…, N-1} \qquad (4)$$

The circuit is described as follows.



In short, it is easy to design an NxN DST circuit from an NxN DCT circuit. Our proposed method applies both transforms to residual signals by using RD-Optimization.

3 Coding results

i) Coding condition

The evaluation tests were performed under the common condition [4]. The proposed method is based on HM0.9 and implemented DST modules. According to JCTVC-C405 [5], the block size is defined as 4 x 4 to 32 x 32. Our DST is used for each block size. The test sequences were the same as current formal test sequences. Additionally, cropped Super Hi-Vision (SHV) sequences were used for increasing the variety of test sequences [6] [7].

ii) Coding performance

Table 1 lists the evaluation results of each sequence under all test conditions. We can find BD-Rate reduction without the results of the Intra LC condition. In particular, the results of low delay LC conditions are higher. The achieved improvement is up to a 2.7% bit reduction. These results show that our proposal tends to work better for inter coding than intra coding. For the “Park Scene” in low delay LC condition, DST is selected almost 50% of inter blocks and 30% of intra blocks.

Table 1. Coding performance (BD-Rate Y) of each test sequence

Y BD-rate	IntraHE	IntraLC	RandomHE	RandomLC	LowdelayHE	LowdelayLC
Nebuta	-0.3	-0.3	-0.8	0.0	-0.3	-0.7
SL	-0.3	0.1	-0.3	0.5	-0.6	-0.4
Traffic	-0.3	0.8	-0.4	-0.1	-0.5	-1.2
PeopleOnStreet	-0.7	0.1	-0.3	0.2	-0.4	-0.4
Kimono	-0.3	0.4	-0.6	-0.3	-1.4	-1.8
ParkScene	-1.1	-1.1	-1.2	-1.4	-1.7	-2.7
Cactus	-0.6	-0.3	-0.5	-0.2	-0.8	-1.0
BasketballDrive	-0.2	1.4	0.0	1.0	-0.4	0.2
BQTerrace	-0.3	0.0	-0.7	-0.8	-1.2	-1.9
BasketballDrill	-0.4	0.4	-0.5	0.2	-0.2	-0.4
BQMall	-0.4	0.6	-0.3	0.2	-0.6	-0.5
PartyScene	-0.9	-1.0	-1.2	-1.0	-0.8	-1.2
RaceHorses	-0.7	-0.9	-0.2	-0.4	-0.6	-1.2
BasketballPass	-0.3	0.7	-0.1	0.3	-0.1	0.0
BQSquare	-0.8	-1.1	-1.5	-1.4	-1.0	-1.8
BlowingBubbles	-0.7	-0.8	-0.8	-1.0	-0.8	-1.4
RaceHorses	-0.6	-0.7	-0.3	-0.3	-0.4	-1.2
Vidyo1	-0.2	1.5			-0.5	0.1
Vidyo3	-0.2	1.1			0.0	0.4
Vidyo4	-0.3	1.0			-0.2	-0.2
Average	-0.5	0.1	-0.6	-0.3	-0.6	-0.8

iii) Process time

This proposed method does not depend on any other encoding mode selection nor uses any fast algorithms. Therefore, all blocks have to be evaluated using the RD criterion in both DCT and DST. In intra coding, the encoding time greatly increases because all blocks have to be evaluated for all intra prediction mode decisions in this implementation. This is caused by a number of intra prediction modes. On the other hand, the encoding time of the low delay conditions does not increase. The decoding time does not change.

	Intra	Intra LoCo	RA	RA LoCo	LD	LD LoCo
Enc Time[%]	160%	168%	143%	127%	139%	124%
Dec Time[%]	100%	100%	99%	98%	99%	100%

4 Conclusion

This contribution investigates the performance of adaptive DCT/DST selection. These results show that the proposed method improves performance, especially for inter blocks, although the encoding process time increases due to intra prediction mode decision. We are now developing a fast mode selection algorithm.

Our proposed method has the following features. It does not increase the decoding process, does not require training, and enables easy design of NxN DST circuits from DCT ones. The proposed method can be used together with tools in CE7.

5 References

[1] A. Ichigaya et al., "Description of video coding technology proposal by NHK and Mitsubishi," Doc JCTVC-A122, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T VCEG and ISO/IEC MPEG, Dresden, DE, April. 2010

[2] R. Cohen et al., "Description of Core Experiment 7: Alternative Transforms," Doc. JCTVC-C507, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T VCEG and ISO/IEC MPEG, Guangzhou, CN, October 2010

[3] A. Saxena and F. C. Fernandes,"CE7: Mode-dependent DCT/DST for intra prediction in video coding," Doc. JCTVC-D033, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T VCEG and ISO/IEC MPEG, Daegu, KR, January 2010

[4] F. Bossen, "Common Test conditions and software reference configurations," Doc. JCTVC-C500, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T VCEG and ISO/IEC MPEG, Guangzhou, CN, October 2010

[5] T. K. Tan et al., "Summary of HEVC working draft 1 and HEVC test model (HM)," Doc. JCTVC-C405, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T VCEG and ISO/IEC MPEG, Guangzhou, CN, October 2010

[6] S. Sakaida et al., "7680 x 4320 format test sequences for JCT-VC," Doc. JCTVC-A023, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T VCEG and ISO/IEC MPEG, Dresden, DE, April 2010

[7] A. Ichigaya et al., "Specification of cropped Super Hi-Vision test sequences." Doc. JCTVC-D034, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T VCEG and ISO/IEC MPEG, Daegu, KR, January 2010

6 Patent rights declaration(s)

NHK (Japan Broadcasting corporation) may have IPR relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).