

JVET-N0485

Efficient separable MTS without zero-out

Amir Said, Hilmi Egilmez, Yung-Hsuan Chao,

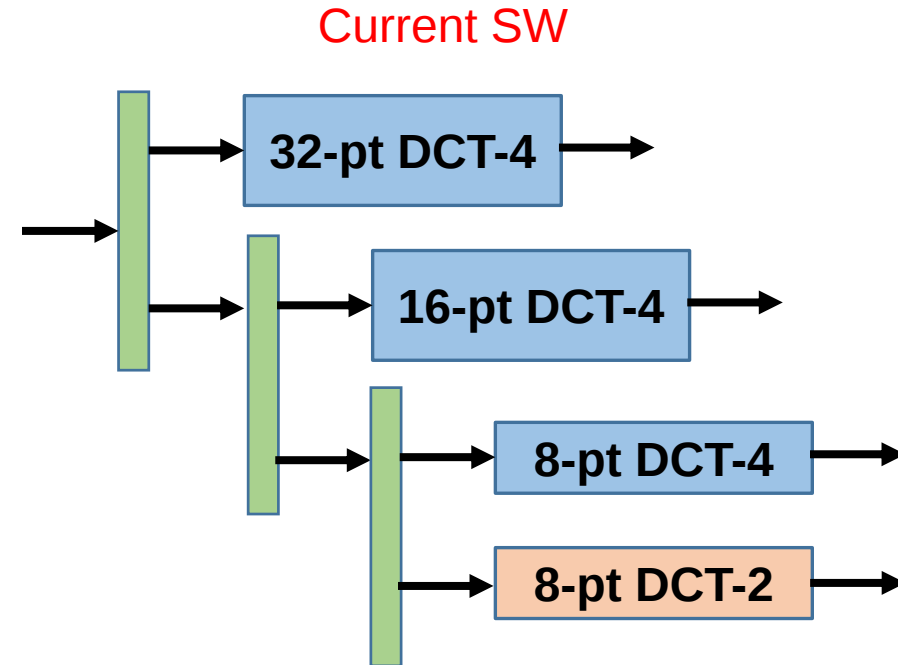
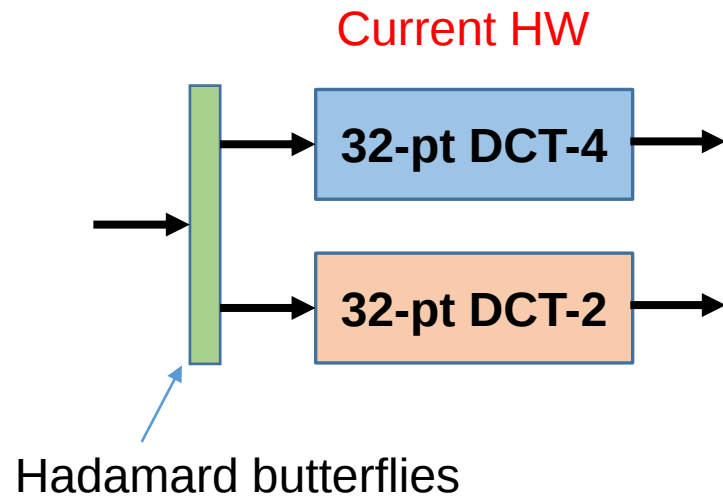
Vadim Seregin, Marta Karczewicz

Qualcomm Inc.

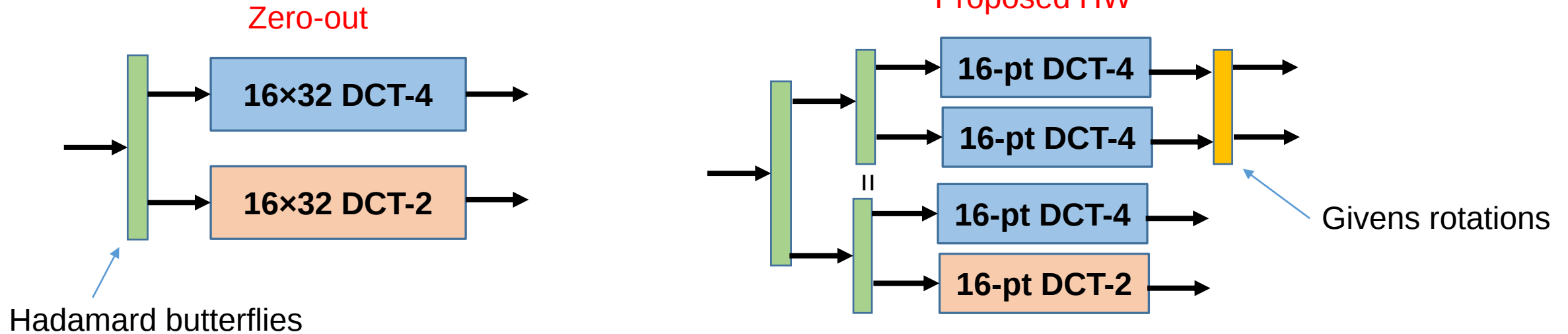
Introduction

- Goal: reduce the complexity of large size transforms (32 and 64-point)
- Current solution: “zero-out” on selected transforms
 - Create two TU types: “full” and “reduced”
 - Affects transform, quantization, and entropy coding functions
 - Lots of “fixes” scattered through the codec
 - Problematic for decisions based on energy remaining after prediction
- This contribution: similar complexity without zero-out
 - Additional factorization of 64-point DCT-2
 - Factorization of 32-point DCT-4 into two 16-point DCT-4s
 - Use transform adjustment (JVET-M0538) to replace matrix-based computation
 - Based on re-using DCT-2 for main computations
 - Single 8×8 matrix for adjustment
 - Can extend MTS to 64-point transforms without increasing encoding/decoding times

“Partial butterfly” DCT-2 factorization

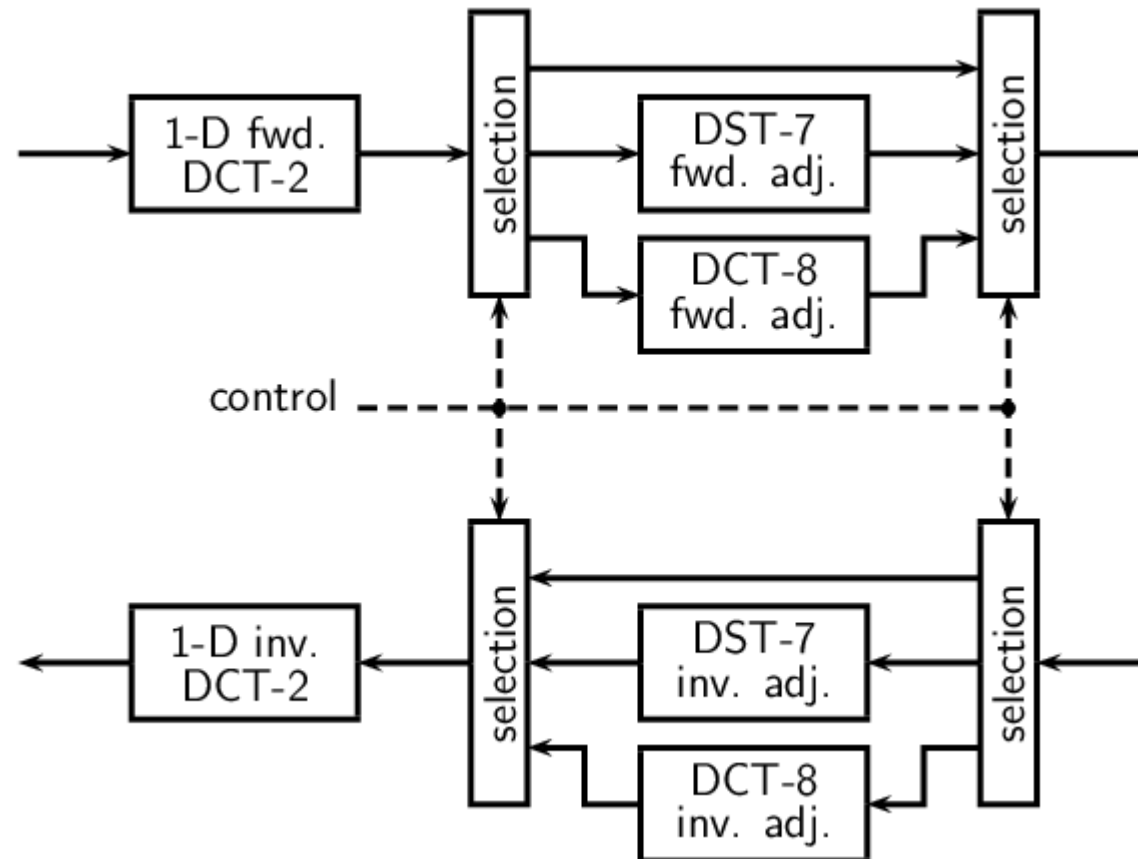


Complexity reduction



- Memory reduction:
 - HW: from 1024 to 544 bytes
 - SW: from 853 to 373 bytes
- Multiplications increase from 1024 to 1088 (6%)
- With 32 parallel multipliers latency increases from 32 to 34 cycles (6%)

Transform adjustment computation (JVET-M0538)



```
const TMatrixCoeff adjustment[8][8] =  
{  
    { 118, -48, -12, -2, -2, -1, -1, 0 },  
    { 38, 108, -54, -15, -8, -6, -3, -2 },  
    { 24, 36, 107, -51, -16, -9, -4, -4 },  
    { 14, 23, 34, 111, -44, -13, -8, -5 },  
    { 10, 16, 19, 29, 114, -38, -14, -8 },  
    { 8, 13, 13, 14, 26, 118, -33, -12 },  
    { 6, 9, 9, 11, 15, 24, 120, -28 },  
    { 5, 8, 8, 8, 11, 14, 23, 124 }  
};
```

Adjustment computation: vector with 8 lowest frequency coefficients multiplied by 8×8 matrix

Complexity for HW implementations

Multiplications per transform coefficient

Memory (bytes)

Size	32	64
DST-7 + DCT-8	64	
DCT-2	341	32
Total	437	

Matrix	DCT-2 + zero-out when $N = 64$				DCT-8/DST-7 + zero-out when $N = 32$			
	$W = 8$	$W = 16$	$W = 32$	$W = 64$	$W = 8$	$W = 16$	$W = 32$	$W = 64$
$H = 8$	8	12	20	18	16	24	20	
$H = 16$	12	16	24	20	24	32	24	
$H = 32$	20	24	32	24	24	32	24	
$H = 64$	20	24	32	24				

DCT-2	DCT-2 double recursion				TA replacing DCT-8/DST-7			
	$W = 8$	$W = 16$	$W = 32$	$W = 64$	$W = 8$	$W = 16$	$W = 32$	$W = 64$
$H = 8$	8	12	20	21	16	24	26	26
$H = 16$	12	16	24	25	24	32	34	34
$H = 32$	20	24	32	33	26	34	36	36
$H = 64$	21	25	33	34	26	34	36	36

*Cross-checker: Orange

Simulation results – CTC*

	All Intra Main10				
	Over VTM-4.0				
	Y	U	V	EncT	DecT
Class A1	-0.61%	-0.15%	-0.22%	99%	94%
Class A2	-0.27%	0.30%	0.17%	99%	98%
Class B	-0.15%	0.12%	0.17%	99%	95%
Class C	-0.06%	-0.14%	-0.09%	100%	93%
Class E	-0.28%	-0.05%	-0.15%	100%	97%
Overall	-0.25%	0.02%	-0.01%	100%	95%
Class D	-0.08%	0.02%	-0.08%	100%	92%
Class F	0.02%	0.02%	-0.10%	100%	95%

	Random Access Main 10				
	Over VTM-4.0				
	Y	U	V	EncT	DecT
Class A1	-0.40%	-0.62%	-0.59%	100%	98%
Class A2	-0.12%	0.04%	0.07%	102%	99%
Class B	-0.18%	-0.13%	-0.43%	102%	99%
Class C	-0.05%	-0.19%	-0.59%	100%	99%
Class E					
Overall	-0.18%	-0.21%	-0.41%	101%	99%
Class D	-0.06%	-0.11%	-0.03%	100%	99%
Class F	-0.06%	0.00%	-0.28%	101%	100%

	Low delay B Main10				
	Over VTM-4.0				
	Y	U	V	EncT	DecT
Class A1					
Class A2					
Class B	-0.18%	-0.44%	-0.71%	102%	96%
Class C	-0.20%	0.30%	-0.25%	98%	95%
Class E	-0.08%	0.87%	0.70%	100%	97%
Overall	-0.16%	0.13%	-0.20%	100%	97%
Class D	-0.08%	-0.05%	-0.66%	98%	95%
Class F	-0.23%	-0.21%	-0.90%	98%	96%

	Low delay P Main10				
	Over VTM-4.0				
	Y	U	V	EncT	DecT
Class A1					
Class A2					
Class B					
Class C	-0.15%	-0.33%	-0.49%	101%	100%
Class E	-0.13%	0.56%	0.22%	103%	100%
Overall					
Class D	-0.04%	0.09%	0.30%	100%	101%
Class F	-0.17%	0.25%	0.37%	104%	103%

Additional results

- Simulations with constrained partitions (VTM config change only)

	All Intra Main10				
	Over VTM-4.0+cfg				
	Y	U	V	EncT*	DecT*
Class A1	-0.76%	0.38%	0.21%		
Class A2	-0.51%	0.52%	0.47%		
Class B	-0.37%	0.33%	0.30%		
Class C	-0.49%	0.17%	-0.01%		
Class E	-0.63%	-0.23%	-0.18%		
Overall	-0.53%	0.24%	0.16%	17%	81%
Class D	-0.36%	-0.12%	-0.06%		

MinQTLumaISlice : 32
MinQTNonISlice : 32
MaxBTDepth : 1
MaxBTDepthISliceL : 1

	All Intra Main10				
	Over VTM-4.0+cfg				
	Y	U	V	EncT*	DecT*
Class A1	-0.82%	0.38%	0.29%		
Class A2	-0.60%	0.41%	0.43%		
Class B	-0.48%	0.21%	0.20%		
Class C	-0.55%	-0.02%	0.02%		
Class E	-0.69%	-0.20%	-0.12%		
Overall	-0.61%	0.15%	0.16%	11%	78%
Class D	-0.25%	-0.14%	-0.09%		

MinQTLumaISlice : 32
MinQTNonISlice : 32
MaxBTDepth : 0
MaxBTDepthISliceL : 0

*Avrg. enc/dec times using VTM-4.0 + CTC as anchor

(See also simulation results in JVET-N0362)

Simulation results – inter MTS enabled

	All Intra Main10				
	Over VTM-4.0 + inter MTS				
	Y	U	V	EncT	DecT
Class A1	-0.61%	-0.15%	-0.22%	99%	96%
Class A2	-0.27%	0.30%	0.17%	99%	98%
Class B	-0.15%	0.12%	0.17%	99%	95%
Class C	-0.06%	-0.14%	-0.09%	100%	93%
Class E	-0.28%	-0.05%	-0.15%	100%	97%
Overall	-0.25%	0.02%	-0.01%	100%	96%
Class D	-0.08%	0.02%	-0.08%	100%	92%
Class F	0.02%	0.02%	-0.10%	100%	95%

	Random Access Main 10				
	Over VTM-4.0 + inter MTS				
	Y	U	V	EncT	DecT
Class A1	-0.49%	-0.30%	-0.32%	111%	99%
Class A2	-0.15%	0.34%	0.21%	110%	99%
Class B	-0.22%	0.16%	-0.19%	111%	100%
Class C	-0.07%	-0.15%	-0.33%	110%	100%
Class E					
Overall	-0.22%	0.02%	-0.17%	111%	99%
Class D	-0.14%	0.01%	-0.13%	108%	98%
Class F	-0.07%	0.08%	-0.23%	110%	99%

	Low delay B Main10				
	Over VTM-4.0 + inter MTS				
	Y	U	V	EncT	DecT
Class A1					
Class A2					
Class B	-0.14%	-0.09%	-0.38%	125%	112%
Class C	-0.17%	0.18%	-0.11%	111%	98%
Class E	0.04%	1.00%	1.29%	119%	101%
Overall	-0.10%	0.27%	0.13%	119%	104%
Class D	-0.07%	-0.10%	0.11%	108%	100%
Class F	-0.13%	0.25%	0.62%	114%	99%

	Low delay P Main10				
	Over VTM-4.0 + inter MTS				
	Y	U	V	EncT	DecT
Class A1					
Class A2					
Class B					
Class C	-0.21%	0.02%	-0.18%	111%	99%
Class E	0.03%	1.30%	0.73%	122%	99%
Overall					
Class D	-0.06%	-0.11%	0.05%	107%	101%
Class F	-0.16%	-0.09%	1.42%	116%	98%

Inter MTS gains on VTM-4.0

	Random Access Main 10				
	Over VTM-4.0				
	Y	U	V	EncT	DecT
Class A1	-0.10%	0.43%	0.47%	109%	100%
Class A2	-0.42%	0.38%	0.50%	111%	101%
Class B	-0.18%	0.55%	0.42%	112%	100%
Class C	-0.06%	0.51%	0.38%	114%	100%
Class E					
Overall	-0.18%	0.48%	0.44%	112%	100%
Class D	-0.08%	0.43%	0.47%	114%	101%
Class F	0.04%	0.30%	0.27%	107%	100%

	Low delay B Main10				
	Over VTM-4.0				
	Y	U	V	EncT	DecT
Class A1					
Class A2					
Class B	-0.32%	1.04%	0.77%	104%	91%
Class C	-0.10%	1.12%	0.72%	118%	101%
Class E	0.09%	1.04%	0.25%	107%	98%
Overall	-0.14%	1.07%	0.62%	109%	96%
Class D	-0.12%	1.00%	0.15%	118%	104%
Class F	-0.08%	0.78%	0.52%	110%	102%

	Low delay P Main10				
	Over VTM-4.0				
	Y	U	V	EncT	DecT
Class A1					
Class A2					
Class B					
Class C	-0.08%	0.96%	0.51%	122%	101%
Class E	-0.09%	0.52%	0.71%	112%	102%
Overall					
Class D	-0.12%	0.90%	0.78%	120%	101%
Class F					

Simulation results – low QP (12, 17, 22, 27)

	All Intra Main10				
	Over VTM-4.0 - low QP			EncT	DecT
	Y	U	V		
Class A1	-0.18%	-0.16%	-0.24%		
Class A2	-0.09%	0.15%	0.10%		
Class B	-0.01%	0.04%	0.05%		
Class C	0.02%	0.06%	0.09%		
Class E	-0.05%	0.14%	-0.13%		
Overall	-0.05%	0.05%	-0.01%		
Class D	-0.02%	0.08%	0.09%		
Class F	0.04%	-0.02%	0.04%		

Advantages

- Hardware:
 - Maintains small worst-case decoder complexity (36 multiplications per coefficient)
 - Significant memory reduction (437 bytes for all DCT-2, plus 32- & 64-pt MTS)
 - Reuse of previous DCT-2 designs
- Software:
 - Increases the throughput significantly (JVET-M0540)
 - No increase of encoder or decoder times for intra MTS
- Adjustment: provides a very good approximation of 32-pt DST7/DCT8
 - Negligible coding gain changes
 - Enables extending MTS to 64-point transforms
- CTC coding gains:
 - Average – AI: 0.25%, RA: 0.18%, LDB: 0.16%
 - Class A1 – AI: 0.61%, RA: 0.40%
 - Class A2 – AI: 0.27%, RA: 0.12%

Recommendations

- Study the proposed method in CE (for intra MTS)
 - Determine additional gains obtainable by larger transforms
- Test:
 - 1) Apply the method to 32-pt and 64-pt DCT-2 and MTS transforms
 - 2) Increase allowed Sub-Block Transform (SBT) dimensions
 - 3) Extend Intra Sub-block Partition (ISP) dimensions