

JCTVC-B057

Performance Report of DPCM-based Memory Compression on TE2

Hirofumi Aoki, Keiichi Chono, Kenta Senzaki,
Junji Tajime and Yuzo Senda
NEC Corporation

Overview

■ Background

■ Motivation

- Problem on memory access overhead

■ Method

■ Experimental results

- Quality
- Complexity

■ Conclusion

Background

- Demand for higher resolution contents
 - HD for mobile and Ultra-HD for home
 - More for the future
- Limited memory for consumer-use devices
 - Limited both on ***size*** and ***bandwidth***
 - Demand for low-cost while handling such contents



Memory compression

Motivation

■ Major memory accesses are caused by:

- Write of CU/MB reconstructed images
- Read of reference pictures for MC/ME
- Read for display output

■ Characteristics of MC memory access

- Frequent
- Random
- Unaligned



Structure for compression should be simple

Overhead caused by unaligned access (1)

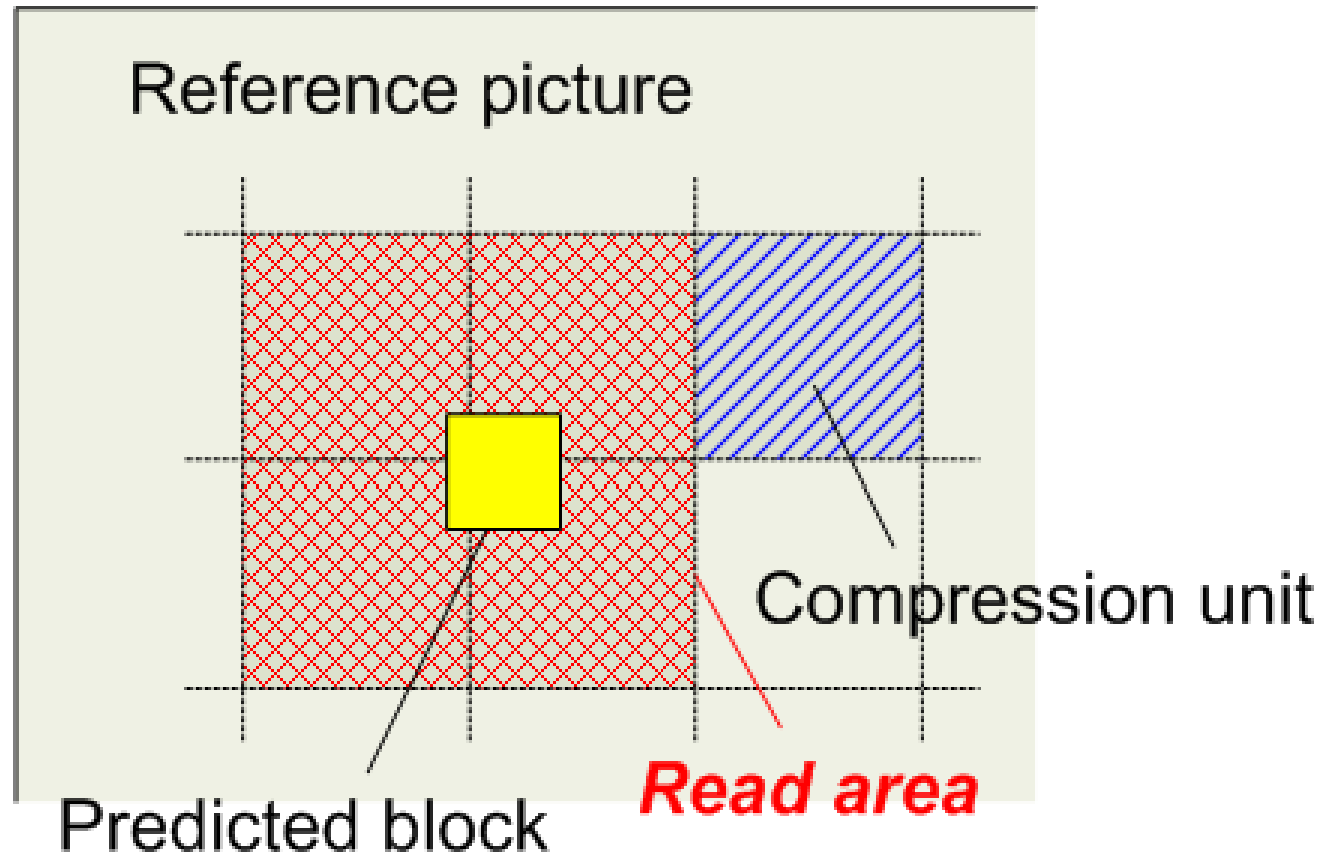
- In MC, Memory access is generally not aligned to memory compression unit
 - Motion vectors are not aligned to fixed pixel boundaries
 - If a MC block is across unit boundaries, codec must read ALL units overlapping the MC block



- To reduce actual memory bandwidth:
 - 2-D structure is not preferable
 - 1-D structure is preferable

Overhead caused by unaligned access (2)

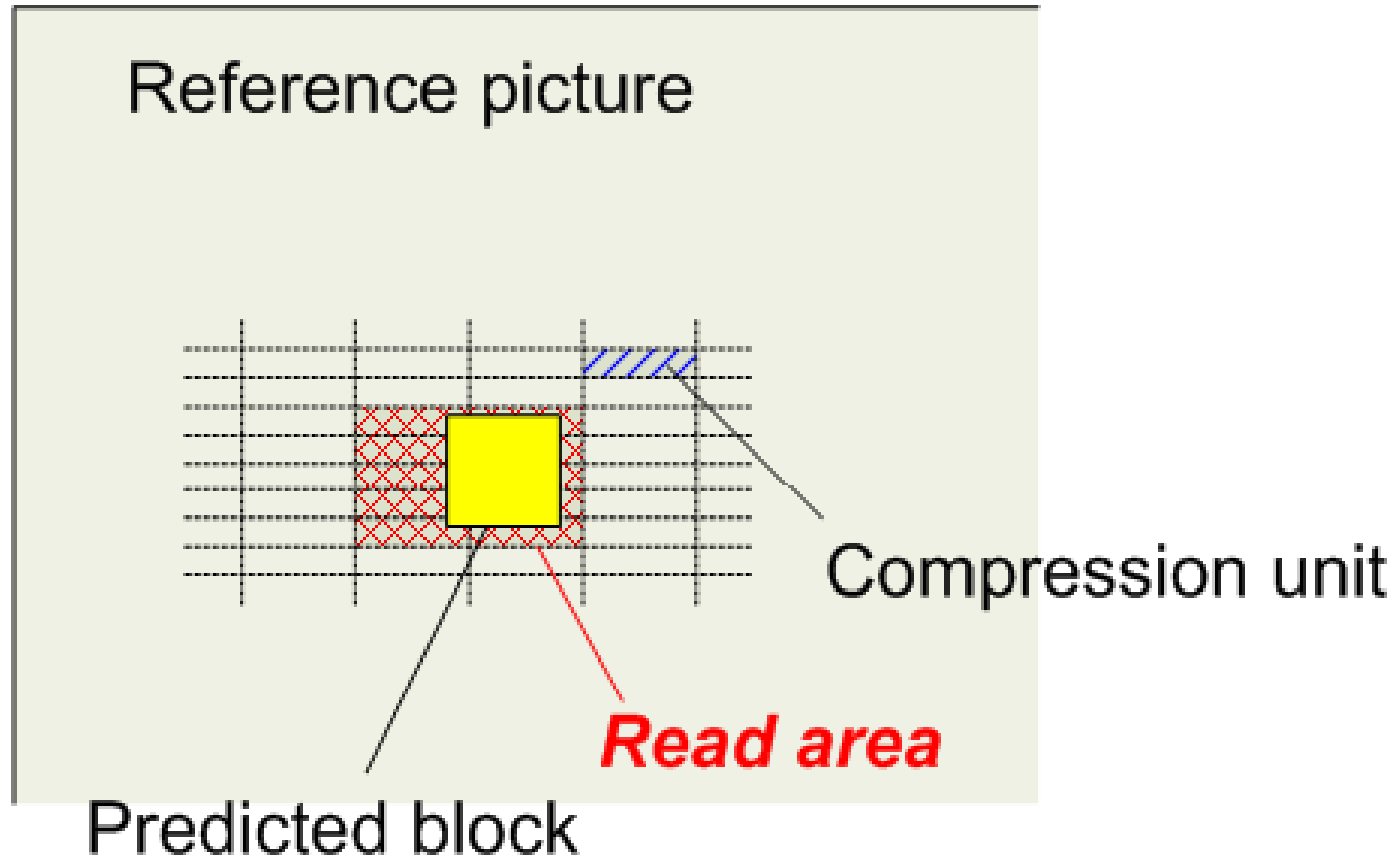
- Considerable overhead with a large unit



Actual memory bandwidth cannot be reduced

Overhead caused by unaligned access (3)

- Overhead reduced with a small unit



Actual memory bandwidth can be reduced

Method

NEC's approach:

1-D DPCM-based memory compression

1-D structure

- Memory bandwidth overhead is reduced

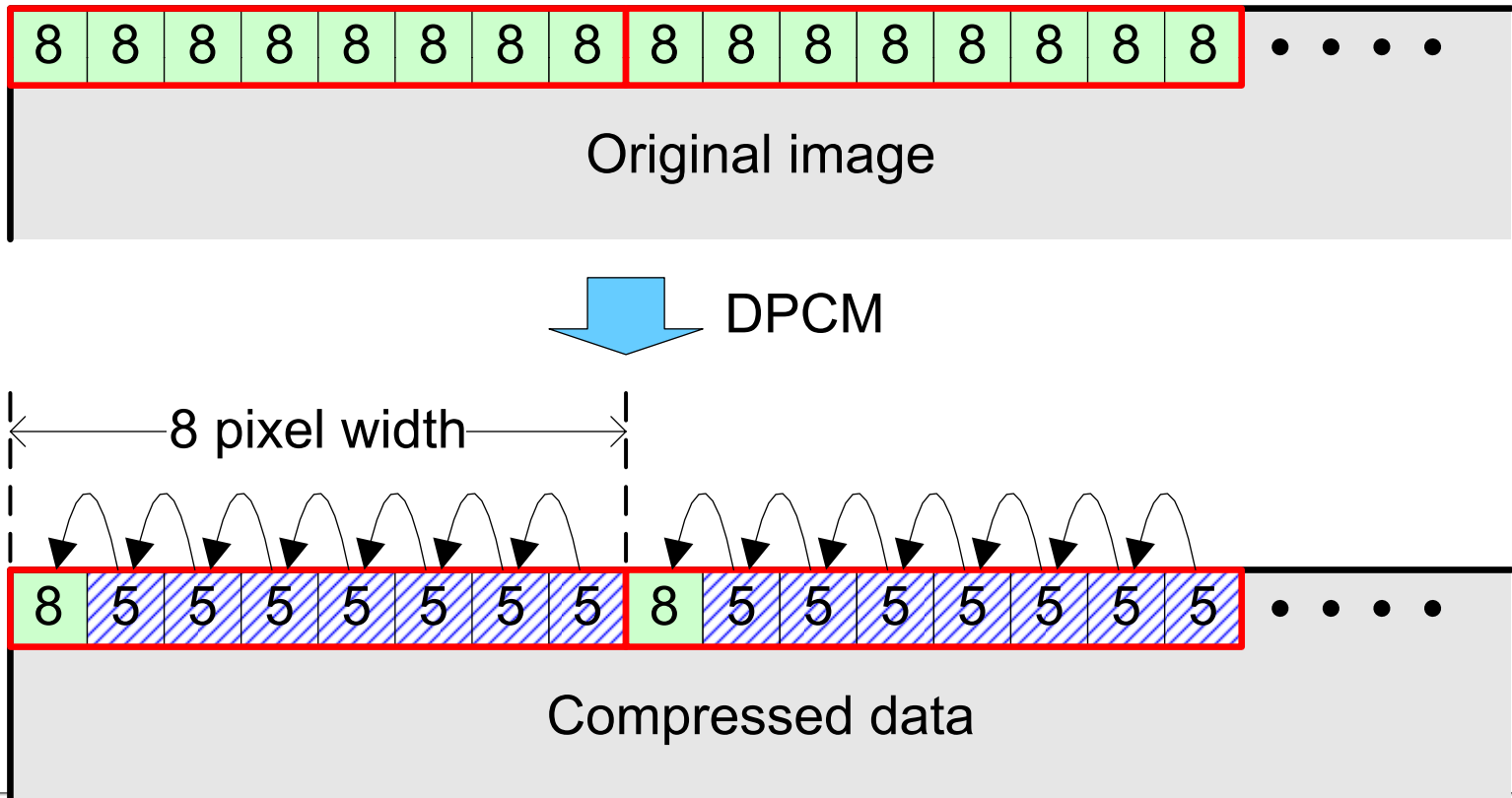
DPCM with nonlinear quantizer

- Decent compression efficiency
- Simple algorithm, in particular, of decompression required for MC
- Ability to guarantee actual bandwidth

Compression process

Compression done in 1-D blocks

- The left-most pixel is first coded with PCM
- The other pixels are then coded with DPCM from left to right, using a nonlinear quantizer



Decompression process

Simply done by:

- Getting a quantization representative value from a look-up table
- Adding the value to a predicted value

$$xd(n) = \begin{cases} y(n); & n = 0 \\ (xd(n-1) + epqr(epqridx(n)) + L) \% L; & n \neq 0 \end{cases}$$

Experiments

Conditions

- CS1: Hierarchical B with periodical I-picture insertion
- CS2: IPPP, where only the first picture is coded as intra

Evaluation measure

- Quality by BD metrics
- Complexity by decoding and encoding time increase

Parameters of the proposed method:

- Base-pixel interval = 8
- IBDI on (12-bit): 7-bit DPCM $\rightarrow (12+7*7)/8 \leq 8$ bpp
- IBDI off (8-bit): 5-bit DPCM $\rightarrow (8+5*7)/8 = 5.4 \leq 6$ bpp

Quality

BD-PSNRs and BD-Rates

IBDI	Anc. type	BD-PSNR dB			BD-Rate %		
		Y	U	V	Y	U	V
Off	CS1	-0.344	-0.092	-0.121	11.552	4.411	6.030
	CS2	-0.241	-0.069	-0.081	7.757	3.580	3.458
On	CS1	-0.020	-0.015	-0.016	0.573	0.686	0.732
	CS2	-0.038	-0.059	-0.073	1.157	3.486	3.898

Subjective quality

- No visible degradation for most cases
- Some ringing artifacts for Vidyo3 CS2 w/o IBDI

Additional results on quality

Memory compression only at decoders

→ Severe degradation in particular for CS2 cases

IBDI	Anc. type	BD-PSNR dB			BD-Rate %		
		Y	U	V	Y	U	V
Off	CS1	-0.372	-0.072	-0.090	12.528	3.321	3.650
	CS2	-1.474	-0.246	-0.369	67.940	18.099	19.387
On	CS1	-0.031	-0.023	-0.031	0.910	1.156	1.405
	CS2	-0.421	-0.797	-0.869	13.376	55.103	59.558

Degradation is suppressible by codec-integrated memory compression

Complexity

- Encoding and decoding time increase %
 - Low complexity because of its simplicity

IBDI	Anc. type	Enc.	Simple dec.	Fast dec.
Off	CS1	-0.32	26.54	4.70
	CS2	-2.67	39.74	14.05
On	CS1	1.41	42.54	1.56
	CS2	3.51	50.45	3.32

Simple dec.: Cross-checked version

Fast dec.: Another implementation with a look-up table

Conclusion

- 1-D DPCM-based lightweight memory compression
 - Invisible quality degradation for most test cases
 - Sufficiently low complexity
-
- Effectiveness of integrating memory compression into codec loop are shown
 - Recommend further study considering on actual memory bandwidth reduction

Empowered by Innovation

NEC