Memory access bandwidth for inter-prediction in HEVC (JCTVC-B086)

Seung-Kyun Oh
Sehoon Yea
July, 2010

LG Electronics

Introduction

- External memory access
  - One of the major design issues
  - Increasing demand for high quality applications such as mobile HD and ultra high-definition (UHD) TV

- Memory bandwidth
  - Defined as the amount of data per unit time as required by the algorithm to be loaded into the data cache from DRAM

- Memory access units
  - Input bitstream store, coding information store, intra prediction, inter prediction, loop-filtering, reference frame store, and display frame read, etc.

---

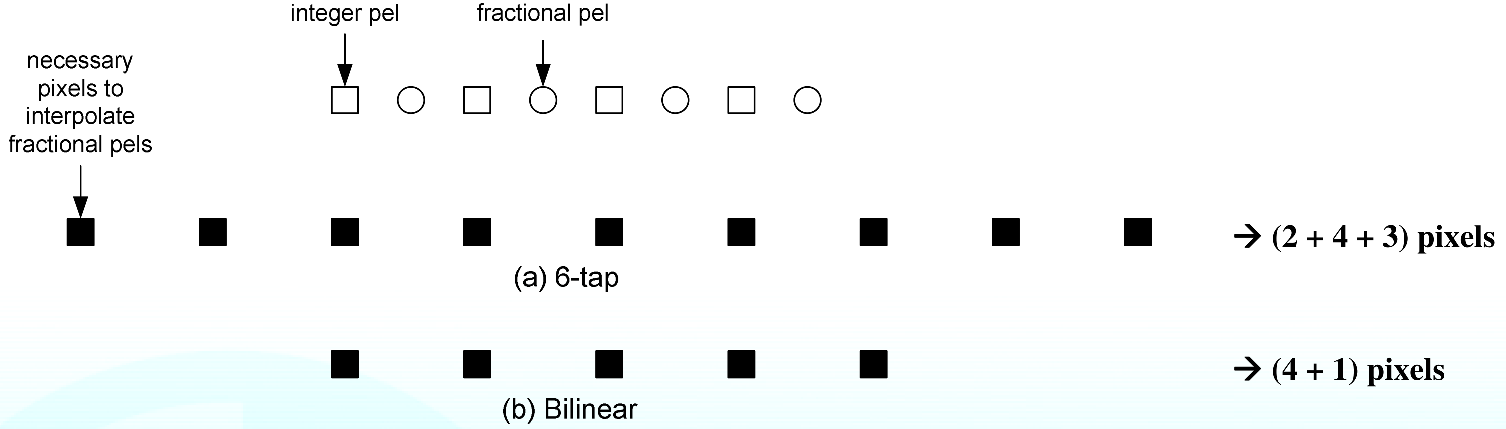
[1] Micheal Horowitz, "Toward Useful Complexity Evaluation Methods", ITU-T SG16/Q6, 33rd VCEG Meeting Shenzhen, China, October 2007, Doc. VCEG-AG19

Introduction

- Motion compensation for inter prediction
  - Exceeds more than half of the total memory access to decode a video bitstream
  - Needs a larger number of pixel date than that of pixels in the block when its motion vector is of fractional-pel
- Real memory bandwidth on a video chip
  - Depends on various factors, such as memory type, video date map, and data cache size as well as the number of transferred pixel data in a unit time
  - More memory bandwidth can be spent by the data quantum of DDR memory and accessing the different banks
- Measure to estimate the memory bandwidth
  - There could be many ways, but the number of pixels to be referenced for motion compensation was used in this contribution

Memory access for motion compensation

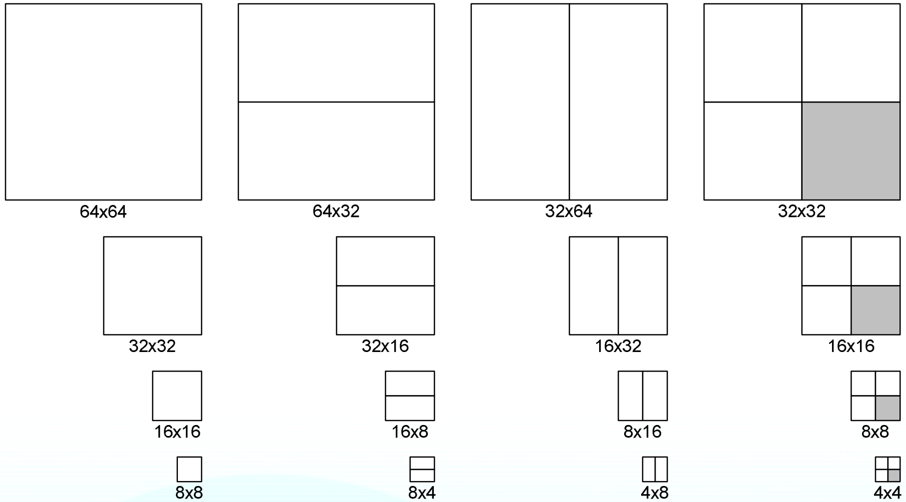
- The number of pixels for motion compensation depends on how many pixels are needed to create a fractional pixel



- Motion vector of a block with a fractional unit
  - A larger block than its size should be read from the frame memory
  - N-tap filter used to interpolate M pixels → (M + N -1) pixels should be read

Memory access for motion compensation

- Extended block partitions in KTA2.6r1



(Y:Cb:Cr 4:2:0 format)

Luma: from 4x4 to 64x64

Chroma: from 2x2 to 32x32

- Overhead ratio of motion compensation

Block Size		Overhead Ratio		Block Size		Overhead Ratio		Block Size		Overhead Ratio		Block Size		Overhead Ratio	
Y	**C**	**Y**	**C**	**Y**	**C**	**Y**	**C**	**Y**	**C**	**Y**	**C**	**Y**	**C**	**Y**	**C**
64x64	32x32	1.16	1.06	32x32	16x16	1.34	1.13	16x16	8x8	1.72	1.27	8x8	4x4	2.64	1.56
64x32	32x16	1.25	1.10	32x16	16x8	1.52	1.20	16x8	8x4	2.13	1.41	8x4	4x2	3.66	1.86
32x64	16x32	1.25	1.10	16x32	8x16	1.52	1.20	8x16	4x8	2.13	1.41	4x8	2x4	3.66	1.86
												4x4	2x2	5.06	2.25

$$(\text{overhead ratio}) = \frac{(\#\text{ of needed pixels for interpolation of a block partition})}{(\#\text{ of pixels of a block partition})}$$

LG Electronics

Memory bandwidth estimate

- Memory bandwidth for motion compensation

(memory bandwidth) = (average number of pixels to be referenced in a frame) x (bit width per pixel) x (frame rate)

- The number of reference pixels for motion compensation of a block depends on the block partition size and its motion vector precision

Experimental results and discussions

- Experimental environments
  - KTA2.6r1 under the JCT-VC TE2 coding conditions with Class B test set

Class B1: Size 1920x1080p 24 fps

Sxx	Name	Duration	QP values for CS1	QP values for CS2
S03	Kimono	72-192 frames	21, 24, 27, 30, 34	22, 25, 28, 32, 35
S04	ParkScene	0-120 frames	24, 27, 30, 33, 36	25, 27, 30, 32, 35

Class B2: Size 1920x1080p 50-60 fps

Sxx	Name	Duration	QP values for CS1	QP values for CS2
S05	Cactus	0-120 frames	25, 27, 30, 32, 35	26, 28, 31, 33, 36
S06	BasketballDrive	0-120 frames	25, 27, 30, 33, 36	27, 29, 32, 35, 38
S07	BQTerrace	0-120 frames	26, 28, 30, 31, 34	28, 29, 30, 32, 34

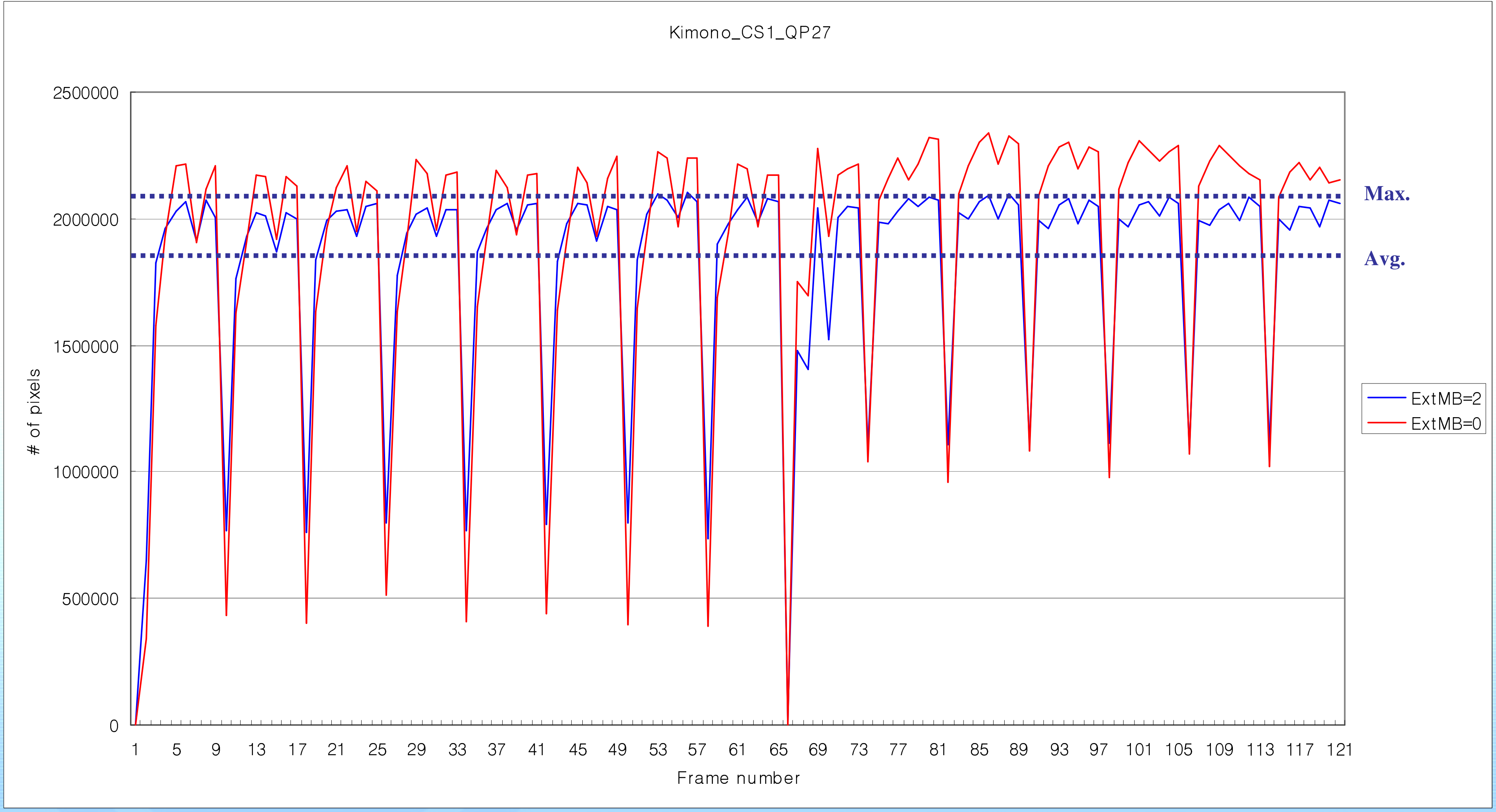
Constraint Set	CS1				CS2			
Sequence Name	BD-Rate (High) [%]	BD-PSNR (High) [%]	BD-Rate (Low) [%]	BD-PSNR (Low) [%]	BD-Rate (High) [%]	BD-PSNR (High) [%]	BD-Rate (Low) [%]	BD-PSNR (Low) [%]
Kimono	-17.30	0.517	-20.69	0.784	-12.98	0.413	-14.73	0.545
ParkScene	-11.53	0.403	-14.51	0.530	-3.15	0.107	-3.76	0.126
Cactus	-10.30	0.262	-11.64	0.357	-7.13	0.193	-7.56	0.232
BasketballDrive	-19.41	0.556	-22.15	0.747	-12.66	0.350	-13.28	0.414
BQTerrace	-17.52	0.282	-22.03	0.438	-10.63	0.214	-13.35	0.280

Experimental results and discussions

Constraint Set	CS1				CS2			
Sequence Name	**QP**	**UseExtMB=0 [Mbytes/s]**	**UseExtMB=2 [Mbytes/s]**	**Bitrate Reduction Rate[%]**	**QP**	**UseExtMB=0 [Mbytes/s]**	**UseExtMB=2 [Mbytes/s]**	**Bitrate Reduction Rate [%]**
Kimono	21	145.59	142.01	-2.46	22	101.26	95.48	-5.71
	24	147.07	140.34	-4.57	25	99.71	91.84	-7.89
	27	147.39	138.34	-6.14	28	98.30	89.25	-9.20
	30	147.93	136.87	-7.48	32	96.17	86.55	-10.01
	34	146.54	135.42	-7.59	35	94.18	85.00	-9.74
ParkScene	24	165.06	152.79	-7.43	25	110.82	104.15	-6.03
	27	161.45	148.72	-7.88	27	108.75	100.53	-7.56
	30	157.00	145.02	-7.63	30	106.21	95.62	-9.97
	33	153.91	142.13	-7.65	32	104.73	92.82	-11.37
	36	151.69	140.33	-7.49	35	102.51	89.50	-12.69
Cactus	25	286.06	279.51	-2.29	26	178.22	177.55	-0.37
	27	285.15	280.71	-1.55	28	176.19	175.40	-0.45
	30	282.78	280.75	-0.72	31	175.10	172.54	-0.89
	33	281.97	280.12	-0.66	33	173.31	170.65	-1.54
	36	280.70	279.32	-0.49	36	170.35	168.22	-1.25
BasketballDrive	25	280.00	266.11	-4.96	27	164.51	169.01	2.73
	27	282.29	263.86	-6.53	29	166.52	167.24	0.43
	30	285.49	262.52	-8.04	32	169.03	166.02	-1.78
	33	287.62	262.47	-8.74	35	173.92	165.35	-4.92
	36	285.59	263.68	-7.67	38	177.59	165.63	-6.74
BQTerrace	26	407.08	370.06	-9.10	28	263.46	244.02	-7.38
	28	411.53	365.66	-11.15	29	261.74	239.24	-8.60
	30	415.13	363.30	-12.49	30	260.34	235.15	-9.68
	31	416.63	363.12	-12.84	32	259.53	227.69	-12.27
	34	425.03	364.13	-14.33	34	257.98	221.35	-14.20

LG Electronics

Experimental results and discussions



Conclusions

- Memory bandwidth
  - One of important issues in developing a video codec
  - Motion compensation is mainly responsible for large memory bandwidth in a video decoder

- In our experimental result
  - Extended MB achieves improvement of coding gain and decrease of memory access rate simultaneously

- For further study
  - Find a more comprehensive way to estimate and reduce the memory access bandwidth
  - Assess proposed coding techniques in view of the balance between the coding performance and the memory bandwidth requirement