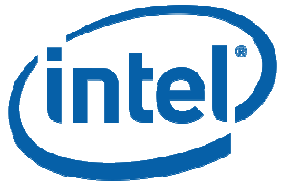


Description of video coding technology proposal by Intel

JCTVC-A106
Yi-Jen Chiu, Lidong Xu
Wenhao Zhang, Hong Jiang
April, 2010

Agenda

- Introduction
- Proposed coding tools
- Performance
- Conclusions

Introduction

- Propose the coding efficiency tools
  - **In-loop Adaptive (Wiener) Loop Filter (ALF):** an average 5.08% for CS1, and of 5.59% for CS2 on KTA2.6r1
  - **Self Derivation Motion Estimation (SDME):** an average 7.0%, up to 21.32%

- Develop the work on KTA2.6r1 as the baseline, with following tools enabled
  - UseAdaptiveFilter = 1
  - UseHPFilter = 1
  - UseAdaptiveLoopFilter = 1

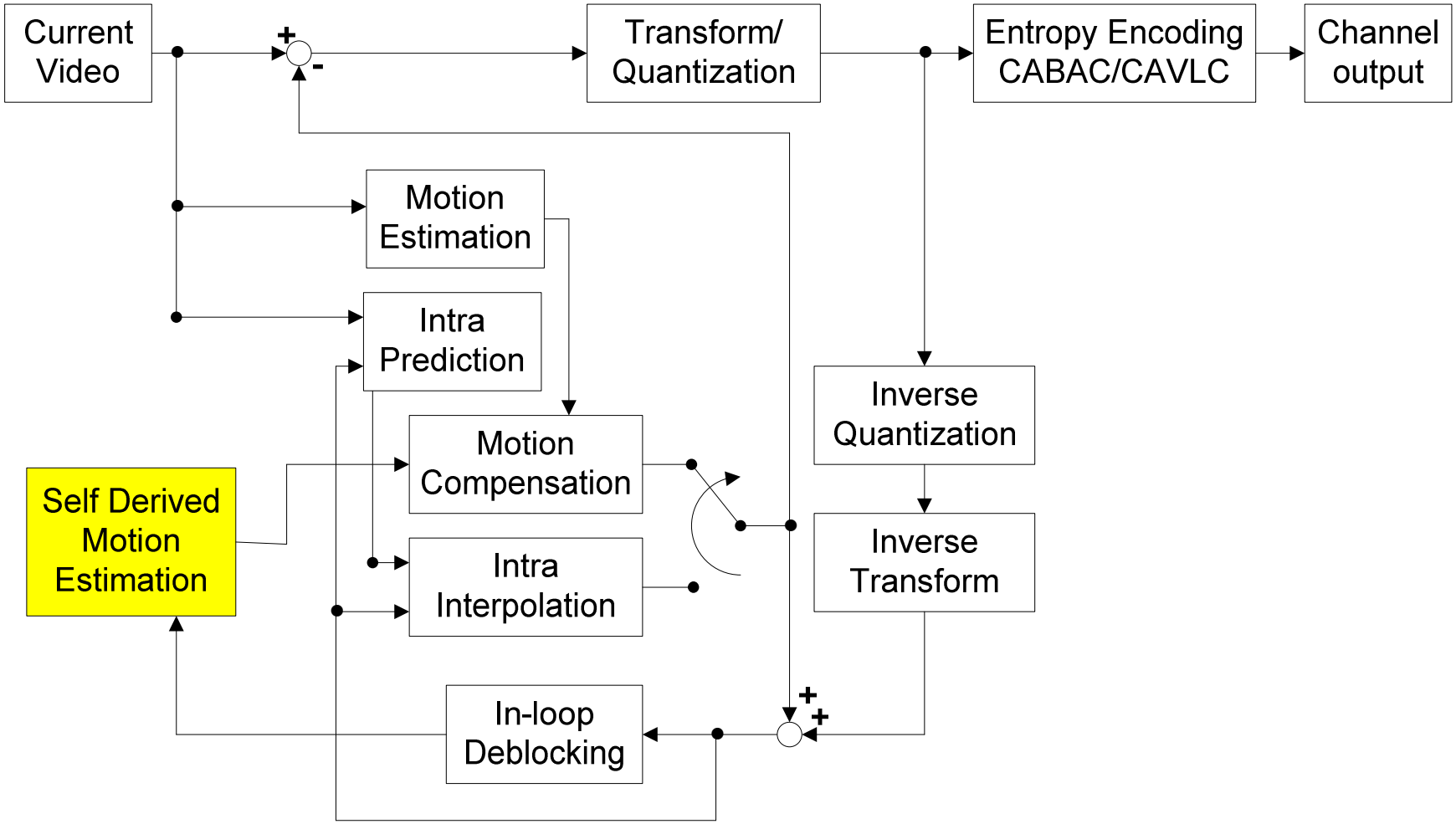
In-loop Adaptive (Wiener) loop filter (ALF)

- Early work was introduced in VCEG meeting at Geneva, April ‘08 (Intel, Toshiba).
- BALF (July’08) & QALF (Oct.’08) (Toshiba)
- A decent software implementation in KTA2.6r1.
- Data shows an overall BD Bitrate improvement of 5.08% for CS1, and of 5.59% for CS2 on KTA2.6r1.
- Draw a lot of attention lately, a couple of proposals to improve further at this meeting.

Self Derivation Motion Estimation (SDME)

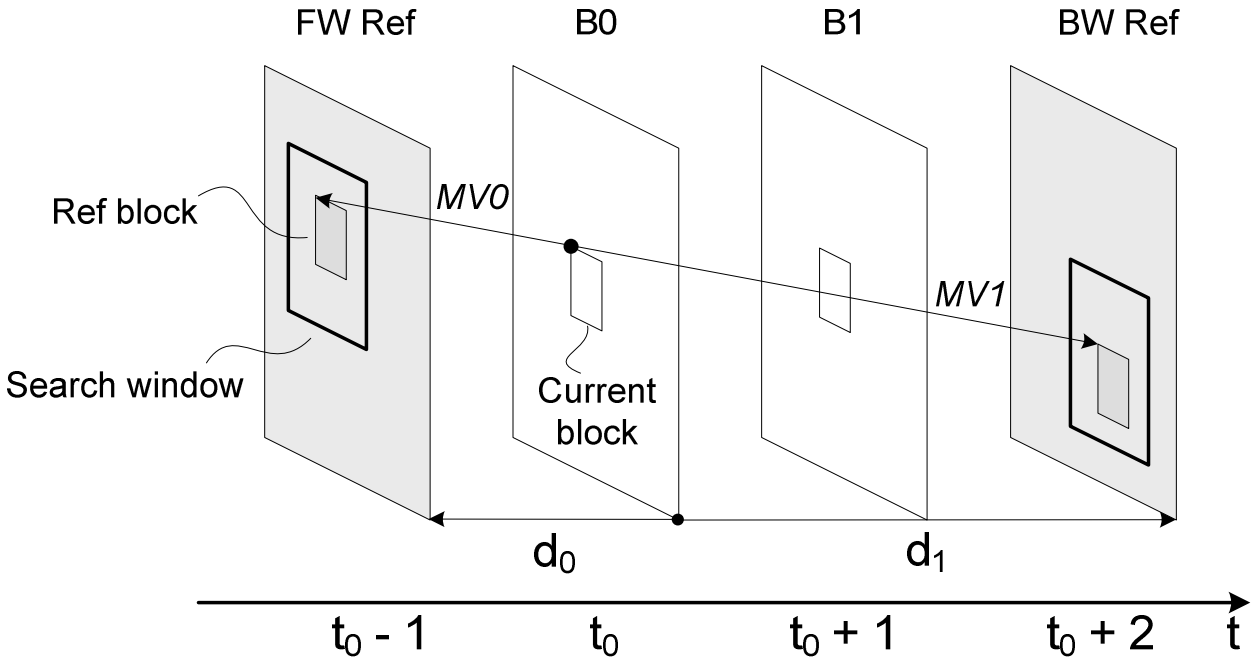
- Early work was introduced for B-picture in VCEG meeting at Yokohama, April'09 (Hitachi).
- Related work on template match technique to derive MV for P-Picture , Oct.'07(RWTH)
- Our work on KTA2.6r1 shows
  - an average 7.0%, up to 21.32% (PeopleOnStreet) in BD Bitrate improvement.
- Related work at this meeting: A108 (JVC), A109 (MediaTek), A111 (Hwawei), A112 (RWTH)

SDME for B frame – Encoder in H.264/AVC



A mirror-based ME for B picture

- ME is performed between the nearest FW & BW reference pictures
- ME aims to find the pair of ref. blocks with minimal SAD
  - Decoder can derive the MV itself -> Save the MV bits



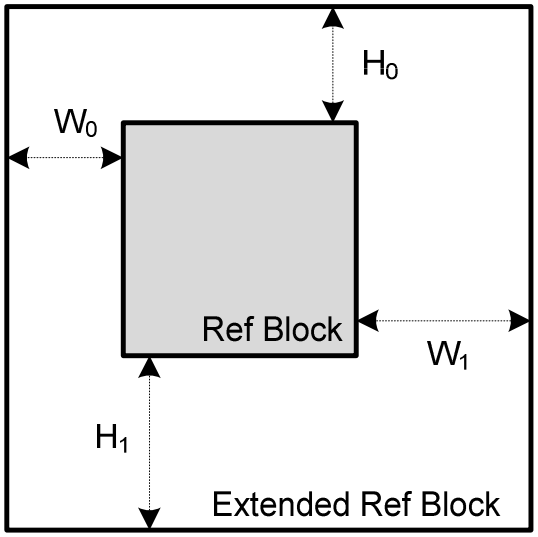
$$MV1 = -\frac{d_1}{d_0} MV0$$

Description of ME scheme

- Integer pixel ME followed by fractional pixel refinement
  - Searching windows are specified in FW reference picture
  - Spiral search pattern, starting from center.
- Integer pixel ME
  - Two search candidate centers, (0,0) and PMV derived by H.264 MV predictor
  - Search Range (SR)
    - Fixed search range (search range = 8)
    - Adaptive search range
- Fractional pixel ME
  - Centered at the best integer MV, search in quarter pixel (7x7 search window)

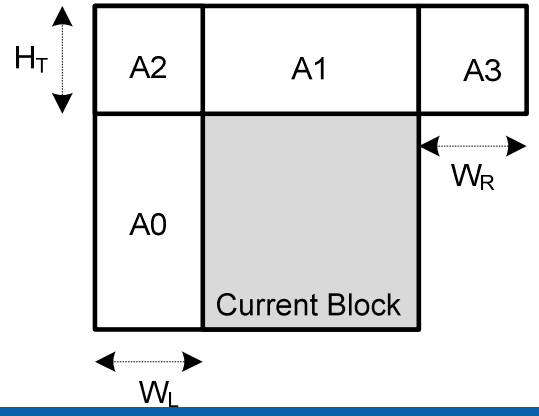
ME cost metric: extended area of block

- The gray ref block is in the same size with target block
- Extended border pixels are considered in SAD metric
  - Can generate more accurate MV for target block
  - Border sizes are experimentally set to 8



ME cost metric: Spatial neighbors

- Four spatial neighbor blocks are considered in ME metric
  - Ai (i = 0, 1, 2, 3) is considered if it is available
  - The neighbor blocks are 8-pixel borders to current block
- Spatial neighbor SAD is included in the ME cost metric
  - SAD between “the spatial neighbor blocks of current target block” and “the weighted average of the spatial neighbor blocks of the two reference blocks”.
  - Weighting factors are decided by the temporal distances to current frame.

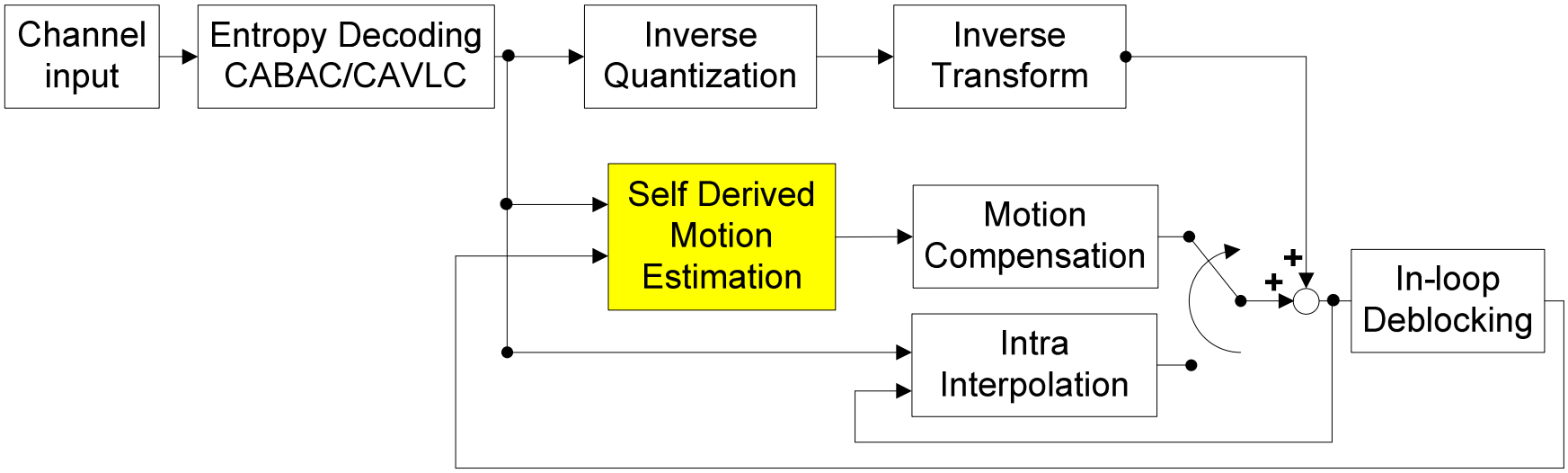


SDME mode

- SDME is applied to the bi-prediction modes of B frames
- A flag is used to signal the MV derivation scheme
  - 0: traditional scheme as in H.264
  - 1: use SDME to derive the MV
  - RDO based mode decision
  - The flag is after the MB mode in syntax table
- SDME currently applies to 16x16 and the sub-partitions
  - Only enable it on 16x16, 16x8 and 8x16 bi-prediction modes currently
    - B_Skip, B_Direct_16x16, B_Bi_16x16
    - B_Ln_Bi_16x8, B_Ln_Bi_8x16, B_Bi_Ln_16x8, B_Bi_Ln_8x16
    - B_Bi_Bi_16x8, B_Bi_Bi_8x16
  - B_Direct_8x8 is replaced with SDME scheme, no flag bit is needed

SDME for B frame – Decoder in H.264/AVC

- Decoder performs exactly the same SDME process to derive the MV as the encoder does.



Compression performance

- Total gains compared to alpha anchor (CS1)

	Sequence	AIF + HPFilter + ALF + SDME					
		BD_PSNR(Full)	BD_BR(Full)	BD_PSNR(High)	BD_BR(High)	BD_PSNR(Low)	BD_BR(Low)
Class A	Traffic	0.6785	-17.3837	0.5691	-15.7194	0.8351	-19.2316
	PeopleOnStreet	1.2287	-22.6805	0.9885	-19.2414	1.6305	-27.636
Class B	Kimono1	0.7741	-20.5587	0.5942	-18.7145	0.9468	-21.9848
	ParkScene	0.5141	-13.2735	0.3778	-10.1795	0.6454	-16.2786
	Cactus	0.5731	-18.1964	0.4216	-16.7214	0.7238	-19.3876
	BasketballDrive	0.5394	-16.1852	0.3762	-13.747	0.6976	-17.9592
	BQTerrace	0.4464	-23.3645	0.3564	-25.3386	0.537	-22.6157
Class C	BasketballDrill	0.7546	-17.0915	0.722	-16.7338	0.7975	-17.2659
	BQMall	1.1707	-22.5261	1.0285	-21.2593	1.3651	-24.5099
	PartyScene	0.9509	-23.1103	0.9268	-21.8189	0.9724	-25.5543
	RaceHorses	0.8631	-19.3436	0.7945	-17.8063	0.9511	-21.4118
Class D	BasketballPass	0.8059	-15.2013	0.7647	-13.6787	0.8726	-18.1336
	BQSquare	1.1251	-27.5481	1.1097	-26.7708	1.1465	-27.9255
	BlowingBubbles	0.8176	-18.1341	0.8776	-18.8646	0.7101	-16.7891
	RaceHorses	0.6916	-12.9866	0.6446	-11.5303	0.7656	-15.7389
Average of Class A		0.9536	-20.0321	0.7788	-17.4804	1.2328	-23.4338
Average of Class B		0.5694	-18.3157	0.4252	-16.9402	0.7101	-19.6452
Average of Class C		0.9348	-20.5179	0.8680	-19.4046	1.0215	-22.1855
Average of Class D		0.8601	-18.4675	0.8492	-17.7111	0.8737	-19.6468
Overall Average		**0.7956**	**-19.1723**	**0.7035**	**-17.8750**	**0.9065**	**-20.8282**

Compression performance

- Total gains compared to beta anchor (CS2)

	Sequence	AIF + HPFilter + ALF					
		BD_PSNR(Full)	BD_BR(Full)	BD_PSNR(High)	BD_BR(High)	BD_PSNR(Low)	BD_BR(Low)
Class B	Kimono1	0.4872	-12.3459	0.4842	-13.3383	0.4854	-11.3542
	ParkScene	0.2963	-8.0121	0.2424	-6.6543	0.3464	-9.2214
	Cactus	0.1682	-5.212	0.1629	-5.8772	0.1742	-4.8376
	BasketballDrive	0.3648	-9.8805	0.3145	-10.1469	0.4106	-9.7017
	BQTerrace	0.3966	-18.0887	0.3504	-21.1355	0.4429	-16.74
Class C	BasketballDrill	0.311	-7.7577	0.2589	-6.7684	0.3702	-8.9576
	BQMall	0.211	-4.3193	0.1885	-4.1184	0.2386	-4.6027
	PartyScene	0.2259	-5.5438	0.1587	-3.9066	0.3292	-9.1167
	RaceHorses	0.1594	-4.049	0.16	-3.9912	0.1564	-4.0764
Class D	BasketballPass	0.0786	-1.5702	0.0557	-1.0762	0.1236	-2.9173
	BQSquare	0.3478	-9.5153	0.3797	-10.5009	0.2437	-7.0714
	BlowingBubbles	0.0198	-0.5113	-0.0295	0.7389	0.0513	-1.3276
	RaceHorses	-0.096	2.3004	-0.0588	1.16	-0.0838	1.8368
Class E	Vidyo1	0.4166	-9.3551	0.3409	-9.1813	0.476	-9.5289
	Vidyo3	0.374	-8.9751	0.3046	-8.2728	0.4993	-10.1356
	Vidyo4	0.3881	-10.0792	0.3691	-11.0442	0.4712	-9.5011
Average of Class B		0.3426	-10.7078	0.3109	-11.4304	0.3719	-10.3710
Average of Class C		0.2268	-5.4175	0.1915	-4.6962	0.2736	-6.6884
Average of Class D		0.0876	-2.3241	0.0868	-2.4196	0.0837	-2.3699
Average of Class E		0.3929	-9.4698	0.3382	-9.4994	0.4822	-9.7219
Overall Average		**0.2593**	**-7.0572**	**0.2301**	**-7.1321**	**0.2960**	**-7.3283**

Compression performance

- SDME gains in CS1 (base is KTA w/o tools enabled)

	Sequence	SDME					
		BD_PSNR(Full)	BD_BR(Full)	BD_PSNR(High)	BD_BR(High)	BD_PSNR(Low)	BD_BR(Low)
Class A	Traffic	0.2676	-7.3282	0.132	-4.0112	0.4438	-10.5127
	PeopleOnStreet	0.7708	-14.7668	0.5277	-10.7419	1.1902	-21.3187
Class B	Kimono1	0.3109	-9.1127	0.153	-5.4509	0.4699	-11.7825
	ParkScene	0.2756	-7.5579	0.1485	-4.2506	0.4047	-10.6407
	Cactus	0.254	-8.9649	0.141	-6.2316	0.3735	-10.8323
	BasketballDrive	0.2678	-8.6727	0.1321	-5.236	0.4027	-10.9594
	BQTerrace	0.1126	-6.8248	0.0618	-5.0742	0.1651	-7.8527
Class C	BasketballDrill	0.1715	-4.2122	0.1076	-2.7316	0.2622	-6.0608
	BQMall	0.4939	-10.7217	0.351	-8.0108	0.7146	-14.2328
	PartyScene	0.306	-8.2267	0.2613	-6.7013	0.3659	-10.5778
	RaceHorses	0.2263	-5.5548	0.1355	-3.3401	0.3596	-8.7786
Class D	BasketballPass	0.1737	-3.5936	0.1204	-2.3505	0.2567	-5.7374
	BQSquare	0.1438	-3.9865	0.0972	-2.657	0.2171	-5.9377
	BlowingBubbles	0.1274	-3.0686	0.1027	-2.3819	0.1695	-4.3508
	RaceHorses	0.1207	-2.4597	0.0736	-1.4234	0.1954	-4.2604
Average of Class A		0.5192	-11.0475	0.32985	-7.37655	0.817	-15.9157
Average of Class B		0.2442	-8.2266	0.1273	-5.2487	0.3632	-10.4135
Average of Class C		0.2994	-7.1789	0.2139	-5.1960	0.4256	-9.9125
Average of Class D		0.1414	-3.2771	0.0985	-2.2032	0.2097	-5.0716
Overall Average		**0.2682**	**-7.0035**	**0.1697**	**-4.7062**	**0.3994**	**-9.5890**

Complexity analysis – Platform

	Constraint set 1	Constraint set 2
CPU	2 Intel Xeon CPU@ 3.16GHz (8 cores)	2 Intel Core CPU@ 3.19GHz (8 cores)
Memory	16GB RAM	8GB RAM
OS	WinXP Professional x64 Edition SP2	WinXP Professional x64 Edition SP2
Encoding	Run 5 tasks simultaneously	Run 5 tasks simultaneously
Decoding	Run 1 tasks simultaneously	Run 1 tasks simultaneously

Complexity analysis – Encoding time

- Encoding time increment by ALF & SDME in CS1
  - ALF: 6.7%, SDME: 111.8%, Total: 118.5%

	Encoding Time Increase of ALF + SDME (%)					
	bitrate_1	bitrate_2	bitrate_3	bitrate_4	bitrate_5	Average
Kimono1	134.627	133.628	131.114	127.534	122.666	129.914
ParkScene	98.636	100.350	99.113	96.061	95.991	98.030
Cactus	97.422	94.764	86.836	80.257	77.552	87.366
BasketballDrive	189.043	188.999	186.901	177.570	165.766	181.656
BQTerrace	125.795	125.271	123.457	119.676	116.331	122.106
BasketballDrill	124.120	121.877	115.236	104.326	100.987	113.309
BQMall	129.644	123.158	117.131	112.960	110.620	118.703
PartyScene	89.920	88.438	85.390	85.236	86.082	87.013
RaceHorses	182.956	180.838	177.986	169.720	171.804	176.661
BasketballPass	126.838	130.893	130.500	126.817	119.937	126.997
BQSquare	85.128	84.359	83.481	82.039	80.514	83.104
BlowingBubbles	85.044	90.814	87.877	85.943	83.522	86.640
RaceHorses	118.805	129.137	135.087	132.835	129.669	129.107
Average	122.152	122.502	120.008	115.460	112.419	118.508

Complexity analysis – Encoding time

- Encoding time increment by ALF in CS2, about 11.5%

	Encoding Time Increase of ALF (%)					
	bitrate_1	bitrate_2	bitrate_3	bitrate_4	bitrate_5	Average
Kimono1	12.846	12.288	11.985	11.262	11.151	11.906
ParkScene	11.428	11.038	10.688	10.249	10.528	10.786
Cactus	12.576	11.000	11.584	11.063	10.819	11.408
BasketballDrive	11.864	11.640	11.083	10.815	10.535	11.187
BQTerrace	13.214	12.999	12.475	11.932	11.454	12.415
BasketballDrill	13.167	12.462	11.762	10.717	10.308	11.683
BQMall	10.633	10.119	9.851	9.819	9.629	10.010
PartyScene	10.941	10.256	9.654	9.606	9.690	10.029
RaceHorses	10.037	9.725	9.743	9.373	9.569	9.689
BasketballPass	12.536	12.079	11.507	11.282	11.069	11.695
BQSquare	14.987	15.396	15.370	14.234	13.522	14.702
BlowingBubbles	10.974	11.205	11.196	10.857	10.915	11.029
RaceHorses	11.657	11.828	11.165	11.353	10.979	11.396
Vidyo1	14.633	13.179	11.865	10.985	10.520	12.236
Vidyo3	15.452	13.992	11.922	11.292	10.406	12.613
Vidyo4	14.477	12.790	11.003	10.763	10.150	11.837
Average	12.589	12.000	11.428	10.975	10.703	11.539

Complexity analysis – Decoding time

- Decoding time increment by ALF & SDME in CS1
  - ALF: 15.0%, SDME: 550.6%, Total: 565.6%

	Decoding Time Increase of ALF + SDME (%)					
	bitrate_1	bitrate_2	bitrate_3	bitrate_4	bitrate_5	Average
Kimono1	882.514	1015.168	1031.005	1022.555	876.091	965.467
ParkScene	501.980	597.920	634.350	583.753	524.418	568.484
Cactus	900.671	992.952	873.125	846.038	768.250	876.207
BasketballDrive	983.383	1079.517	1157.673	1222.960	1093.822	1107.471
BQTerrace	435.621	459.217	409.822	377.153	352.640	406.891
BasketballDrill	289.301	370.381	275.671	193.789	144.358	254.700
BQMall	480.316	480.199	413.377	332.786	271.005	395.537
PartyScene	354.106	408.869	378.598	309.899	262.180	342.730
RaceHorses	677.165	858.524	835.966	794.989	702.398	773.808
BasketballPass	335.618	445.054	463.919	448.771	413.850	421.442
BQSquare	258.023	279.981	254.209	246.813	253.674	258.540
BlowingBubbles	253.205	303.189	275.572	232.754	209.593	254.863
RaceHorses	553.044	742.603	752.600	825.463	756.328	726.008
Average	531.150	617.967	596.607	572.133	509.893	565.550

Complexity analysis – Decoding time

- Decoding time increment by ALF in CS2, about 13.4%

	Encoding Time Increase of ALF (%)					
	bitrate_1	bitrate_2	bitrate_3	bitrate_4	bitrate_5	Average
Kimono1	17.331	18.113	18.777	23.511	27.786	21.104
ParkScene	16.242	18.044	18.532	18.382	19.647	18.169
Cactus	13.480	12.483	11.797	11.750	11.092	12.120
BasketballDrive	14.617	15.344	17.002	18.450	19.303	16.943
BQTerrace	16.011	16.683	14.950	14.390	14.409	15.289
BasketballDrill	11.275	14.266	16.023	16.285	13.113	14.192
BQMall	11.712	14.638	16.595	16.433	11.779	14.231
PartyScene	6.611	7.262	7.245	7.248	7.140	7.101
RaceHorses	15.017	15.651	18.128	19.969	20.722	17.897
BasketballPass	10.218	12.807	16.160	17.004	15.908	14.419
BQSquare	2.592	3.275	4.566	4.275	8.840	4.710
BlowingBubbles	3.174	5.055	5.412	3.325	2.937	3.981
RaceHorses	12.647	17.321	17.722	20.501	22.687	18.176
Vidyo1	16.745	14.505	11.227	6.539	4.193	10.642
Vidyo3	13.650	14.382	14.016	12.770	10.793	13.122
Vidyo4	15.164	15.836	14.832	12.723	6.679	13.047
Average	12.280	13.479	13.937	13.972	13.564	13.446

Conclusions

- Average gain of “SDME+ALF+AIF+HPFilter” is about 19.17% relative to alpha anchor, max 27.93% (BQSquare)
- Average gain of “ALF+AIF+HPFilter” is about 7.06% relative to beta anchor, max 21.14% (BQTerrace)
- The standalone gain on ALF shows an overall BD Bitrate improvement of 5.08% for CS1, and of 5.59% for CS2.
- The standalone gain on SDME shows an average 7.0%,up to 21.32% (PeopleOnStreet in BD_BR_low) on KTA2.6r1.
- The complexity of SDME is currently high in CPU SW-only platform, but it could be a different story on other platforms.
- Propose to form CE groups for SDME & ALF to study these techniques to be considered as the core coding tools.


intel
Leap ahead™