Fast techniques to improve self derivation of motion estimation

JCTVC-B047

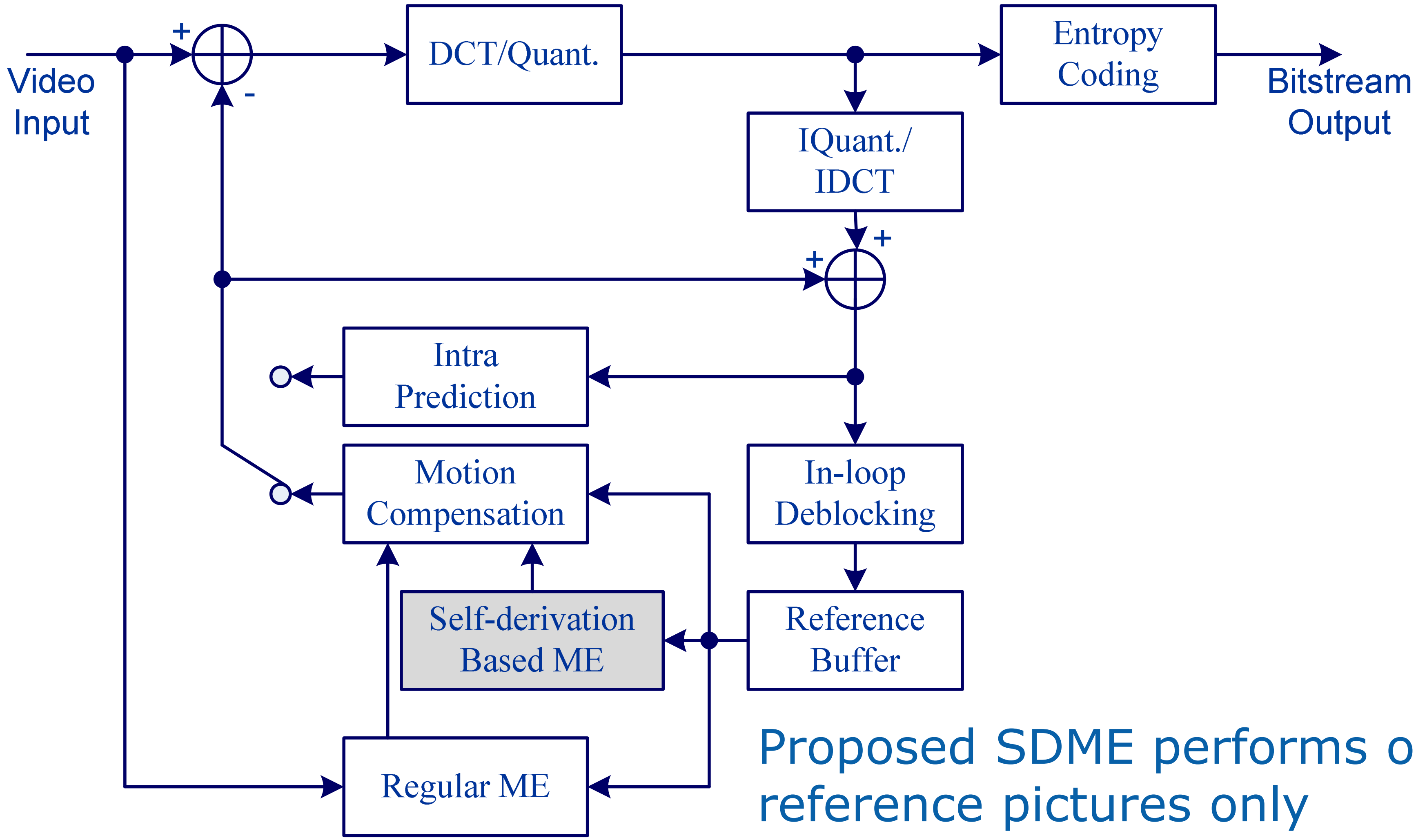
Yi-Jen Chiu, Lidong Xu
Wenhao Zhang, Hong Jiang

July, 2010

Summary of Contribution JCTVC-B047

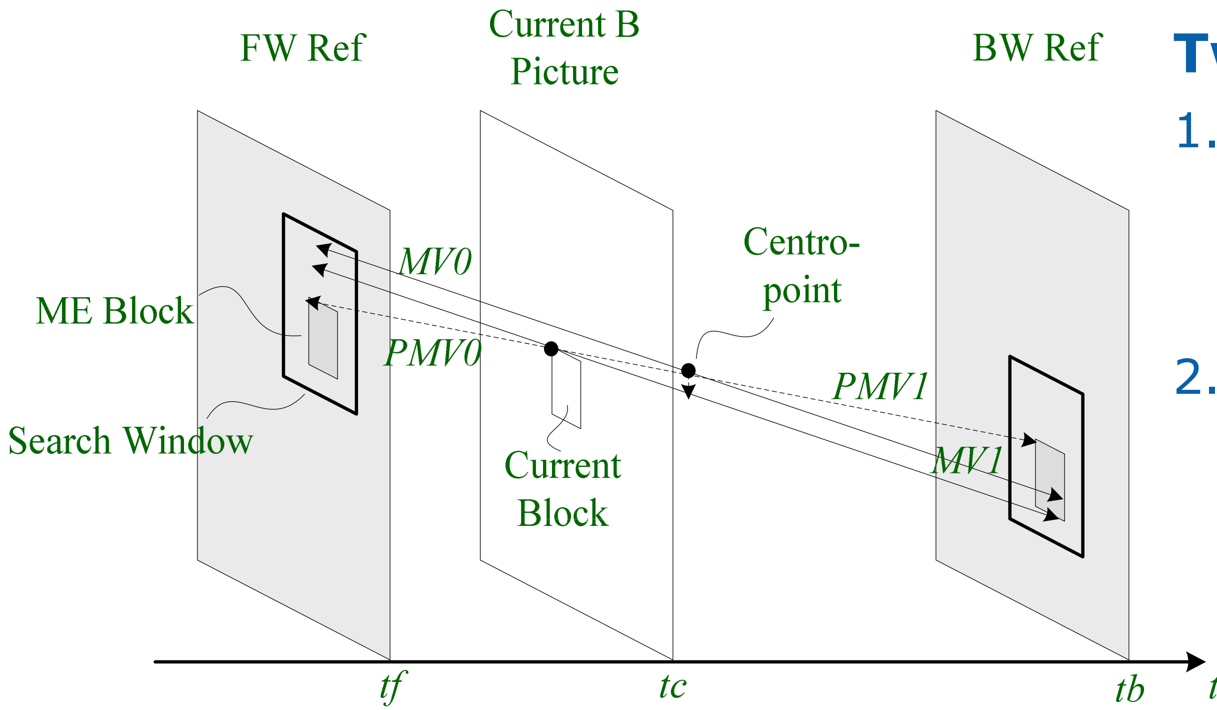
- Fast techniques to improve the self derivation of motion estimation (SDME) presented in JCTVC-A106
  - Derive MVs through ME by capitalizing on the pixels of the previously decoded pictures
  - Candidates based SDME to reduce computational complexity
  - A simple way to derive MV predictor without DMVD dependency of neighboring blocks to realize the parallel implementation of a video decoder.

Video encoder with SDME



Proposed SDME performs on reference pictures only

Mirror-based SDME for B pictures



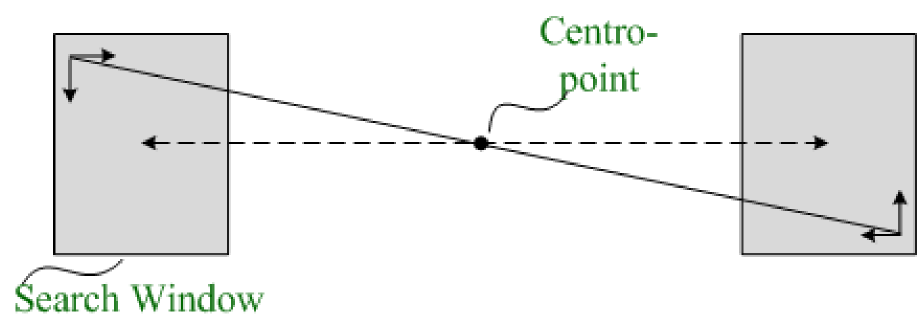
Two assumptions:

1. The inter picture block motions can be treated as a linear trajectory in a short time interval.
2. The current coding block and its spatial neighbor blocks in a stable area have a very similar motion.

(PMV0, PMV1) -- search center

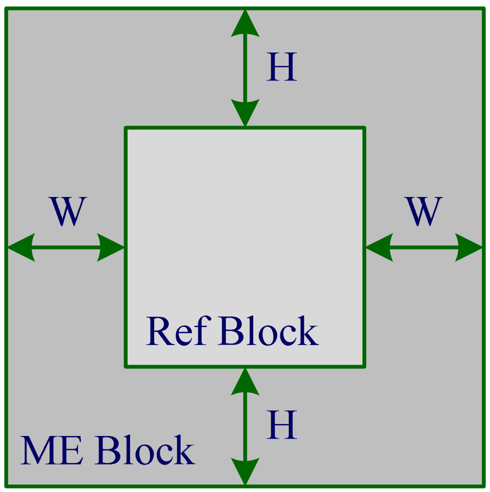
(MV0, MV1) -- SDME output

$$\begin{cases} FMV0 = (MV0 - MV1) * (tc - tf) / (tb - tf) \\ FMV1 = (MV1 - MV0) * (tb - tc) / (tb - tf) \end{cases}$$



Extended ME block

- ME block is bigger than current block



Current Block Size	Ext. ME Block Size
16x16	32x32
16x8	24x16
8x16	16x24
8x8	16x16

- For current blocks bigger than 16x16, extended ME block is NOT used

SDME based coding modes

- SDME is applied to the bi-prediction modes of B pictures
- One flag per bi-predicted coding block signals the ME method, i.e., regular ME or SDME
- The flag is decided based on the RD costs of the two ME methods

Candidates based SDME

- 9 candidates to reduce complexity of SDME

Base MV	**0**	Zero MV
Predicted MVs	1	FW predicted MV pred_fmv
	2	Mirrored MV of the BW predicted MV pred_bmv
	3	(pred_fmv - pred_bmv) / 2
Spatial neighbor MVs Scaled by POC distances	4	MV of left neighbor block *a. If it is FW predicted, use its FW MV mv_fw* *b. If it is BW predicted, use the mirrored MV of its BW MV mv_bw* *c. If it is bi-predicted, use (mv_fw - mv_bw) / 2* *d. If it is intra coded, this candidate will be unavailable*
	5	MV of top neighbor block (*as described in 4*)
	6	MV of top-left neighbor block (*as described in 4*)
Temporal Neighbor MVs Scaled by POC distances	7	MV of FW collocated neighbor block *a. If it has BW MV, scale this BW MV and then mirror it* *b. Otherwise, this candidate will be unavailable*
	8	MV of BW collocated neighbor block *a. If it has FW MV, scale this FW MV* *b. Otherwise, this candidate will be unavailable*

MV prediction in SDME & DMVD mode

- SDME & DMVD demand ME operation at video codec
  - Some blocks in DMVD mode, some blocks in regular mode
- Unbalanced complexity in MV decoding poses a technical challenge in the realization of a parallel friendly decoder
  - If a neighbor block is coded in DMVD mode, the MV prediction of current decoding block needs to wait for the completion of ME.
- Proposal
  - No DMVD neighbor dependency MV prediction
  - Use the regular AVC/H.264 median-based predicted MV to represent the MV of the DMVD coding mode of neighbor blocks.
  - With simple and balanced bitstream parsing at MB layer, all inter-coded blocks can be decoded in parallel

Simulations

- Implemented onto KTA2.6r1
- Evaluated under the CS1 test conditions defined by DMVD TE group
  - Hierarchical coding structure of IbBbBbBbP
  - QPI=27, 30, 34, and 38; QPP=QPI+1; QPB= QPI+2
  - Intra MDDT On, High Precision Filter On, Adaptive Loop Filter ON
  - RDO_Q OFF

Compression performance

- Full search based SDME

	Sequence	UseExtMB = 1		UseExtMB = 0	
		BD_PSNR (dB)	BD_Bitrate (%)	BD_PSNR (dB)	BD_Bitrate (%)
Class A	Traffic	0.198	-5.044	0.391	-9.351
	PeopleOnStreet	0.236	-5.195	0.362	-7.809
Class B	Kimono1	0.300	-9.075	0.588	-15.808
	ParkScene	0.157	-4.631	0.373	-9.994
	Cactus	0.175	-5.686	0.324	-9.730
	BasketballDrive	0.074	-2.601	0.265	-8.309
	BQTerrace	0.048	-2.169	0.177	-7.246
Class C	BasketballDrill	0.037	-0.964	0.109	-2.801
	BQMall	0.162	-3.770	0.237	-5.339
	PartyScene	0.057	-1.384	0.156	-3.672
	RaceHorses	0.023	-0.618	0.121	-3.048
Average of Class A		0.217	-5.120	0.376	-8.580
Average of Class B		0.151	-4.832	0.345	-10.217
Average of Class C		0.070	-1.684	0.156	-3.715
Overall Average		**0.133**	**-3.740**	**0.282**	**-7.555**

Compression performance

- Candidate based SDME

	Sequence	UseExtMB = 1		UseExtMB = 0	
		BD_PSNR (dB)	BD_Bitrate (%)	BD_PSNR (dB)	BD_Bitrate (%)
Class A	Traffic	0.158	-4.023	0.350	-8.382
	PeopleOnStreet	0.200	-4.363	0.321	-6.955
Class B	Kimono1	0.314	-9.483	0.609	-16.390
	ParkScene	0.169	-4.961	0.391	-10.480
	Cactus	0.175	-5.686	0.334	-10.019
	BasketballDrive	0.087	-3.049	0.285	-8.940
	BQTerrace	0.045	-2.060	0.178	-7.261
Class C	BasketballDrill	0.003	-0.088	0.082	-2.112
	BQMall	0.145	-3.348	0.238	-5.362
	PartyScene	0.066	-1.593	0.166	-3.903
	RaceHorses	0.010	-0.289	0.117	-2.932
Average of Class A		0.179	-4.193	0.335	-7.668
Average of Class B		0.158	-5.048	0.359	-10.618
Average of Class C		0.056	-1.330	0.151	-3.577
Overall Average		**0.125**	**-3.540**	**0.279**	**-7.521**

Compression performance

- Candidate + no DMVD neighbor dependency on MV prediction

	Sequence	UseExtMB = 1		UseExtMB = 0	
		BD_PSNR (dB)	BD_Bitrate (%)	BD_PSNR (dB)	BD_Bitrate (%)
Class A	Traffic	0.133	-3.469	0.339	-8.334
	PeopleOnStreet	0.161	-3.523	0.279	-6.127
Class B	Kimono1	0.302	-9.176	0.621	-16.855
	ParkScene	0.138	-4.122	0.376	-10.177
	Cactus	0.151	-4.913	0.321	-9.699
	BasketballDrive	0.049	-1.738	0.226	-7.170
	BQTerrace	0.016	-0.745	0.131	-5.450
Class C	BasketballDrill	-0.023	0.586	0.049	-1.265
	BQMall	0.123	-2.863	0.205	-4.666
	PartyScene	0.044	-1.090	0.151	-3.584
	RaceHorses	-0.026	0.651	0.046	-1.223
Average of Class A		0.147	-3.496	0.309	-7.230
Average of Class B		0.131	-4.139	0.335	-9.870
Average of Class C		0.030	-0.679	0.113	-2.684
Overall Average		**0.097**	**-2.764**	**0.249**	**-6.777**

Encoding and decoding complexities

- Compared to full search based SDME, candidate based SDME can save about 93.9% of the SDME module time
- Compared to the default KTA encoder, full search based SDME increases the total encoding time by about 100%
- Compared to the default KTA encoder, candidate based SDME increases the total encoding time by about 5.58%
- Compared to the default KTA decoder, candidate based SDME increases the total decoding time by about 71.1%

Conclusions

- Compared to CS1 anchor
  - Average gain of full search SDME is about 7.56%, up to 15.81%
    - UseExtMB = 1, the average gain is 3.74%, up to 9.08%
  - Average gain of candidate based SDME is about 7.52%, up to 16.39%
    - For UseExtMB = 1, the average gain is 3.54%, up to 9.48%
    - The SDME time is reduced by 93.9% compared to full search
  - Average gain of “no DMVD dependency” MV prediction scheme on top of candidate based SDME is about 6.78%, up to 16.86%
    - For UseExtMB = 1, the average gain is 2.76%, max 9.18%
  - Notes: Candidate based SDME increases the encoding time by about 5.58% and increases the decoding time by about 71.1%
- The simulation result is verified by Mitsubishi (JCTVC-B069)
- SDME/DMVD can be considered as a potential coding tools for HEVC
- Parallel-friendly implementation of a video decoder should be factored in the design of DMVD tools for HEVC.

Leap ahead™