Collaborative Decoder-side Motion Vector Refinement for Video Coding

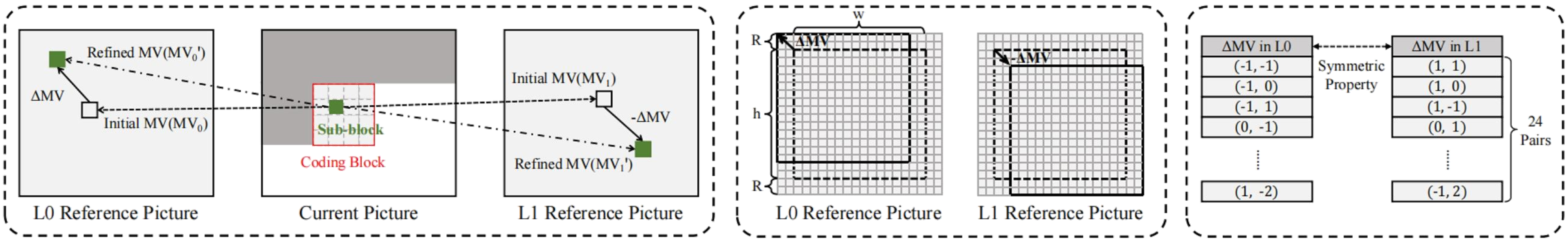
Jialin Li, Zhuoyuan Li, Yao Li, Li Li✉, Houqiang Li✉
University of Science and Technology of China, Hefei, Anhui, China
Institute of Artificial Intelligence, Hefei Comprehensive National Science Center

ISCAS 2025
IEEE INTERNATIONAL SYMPOSIUM ON CIRCUITS AND SYSTEMS
LONDON, UK | May 25-28, 2025

1. Contributions

- Proposing a C-DMVR framework that introduces the sub-block-based collaborative mechanism to optimize the refinement for each sub-block. It includes collaborative distortion calculation (CDC) and collaborative searching strategy (CSS).
- In CDC, the receptive field of the distortion function is enlarged with the collaboration of additional spatial neighbor information to assist the accurate decision of ΔMV.
- In CSS, the coarse-to-fine candidate list derivation scheme is introduced to construct the optimized candidate list with the collaboration of neighbor sub-blocks.
- C-DMVR achieves better R-D performance compared to the AV2 anchor under the random access (RA) configuration.

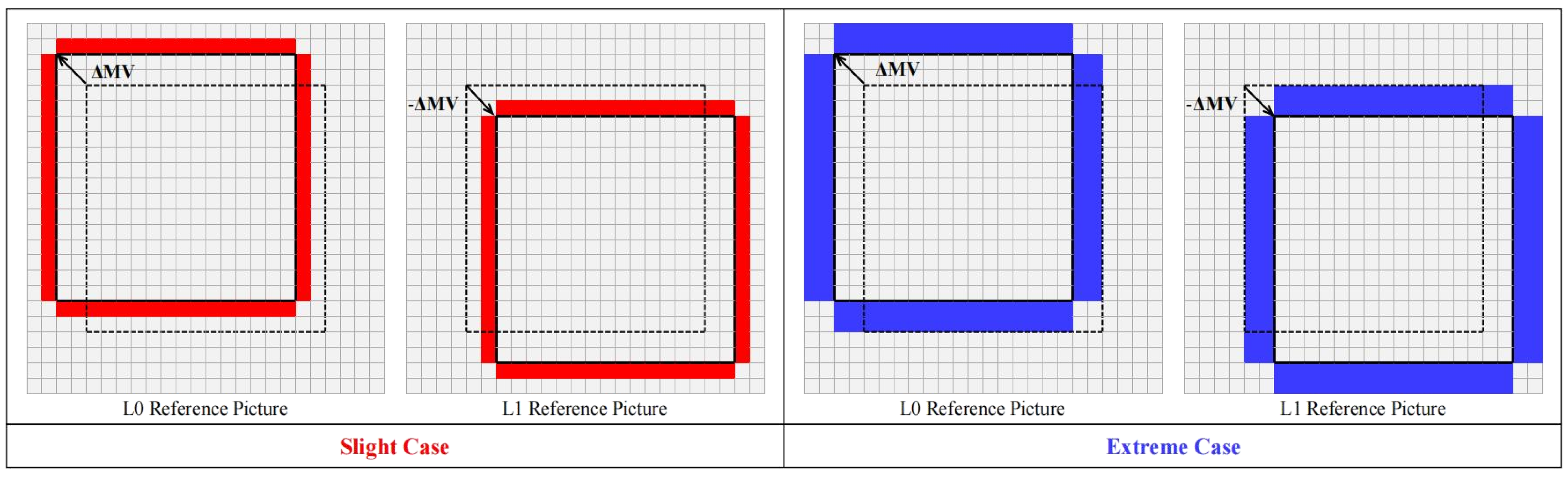
2. Motivation and Analysis



(a) DMVR Process (b) Distortion Calculation (c) Search Strategy

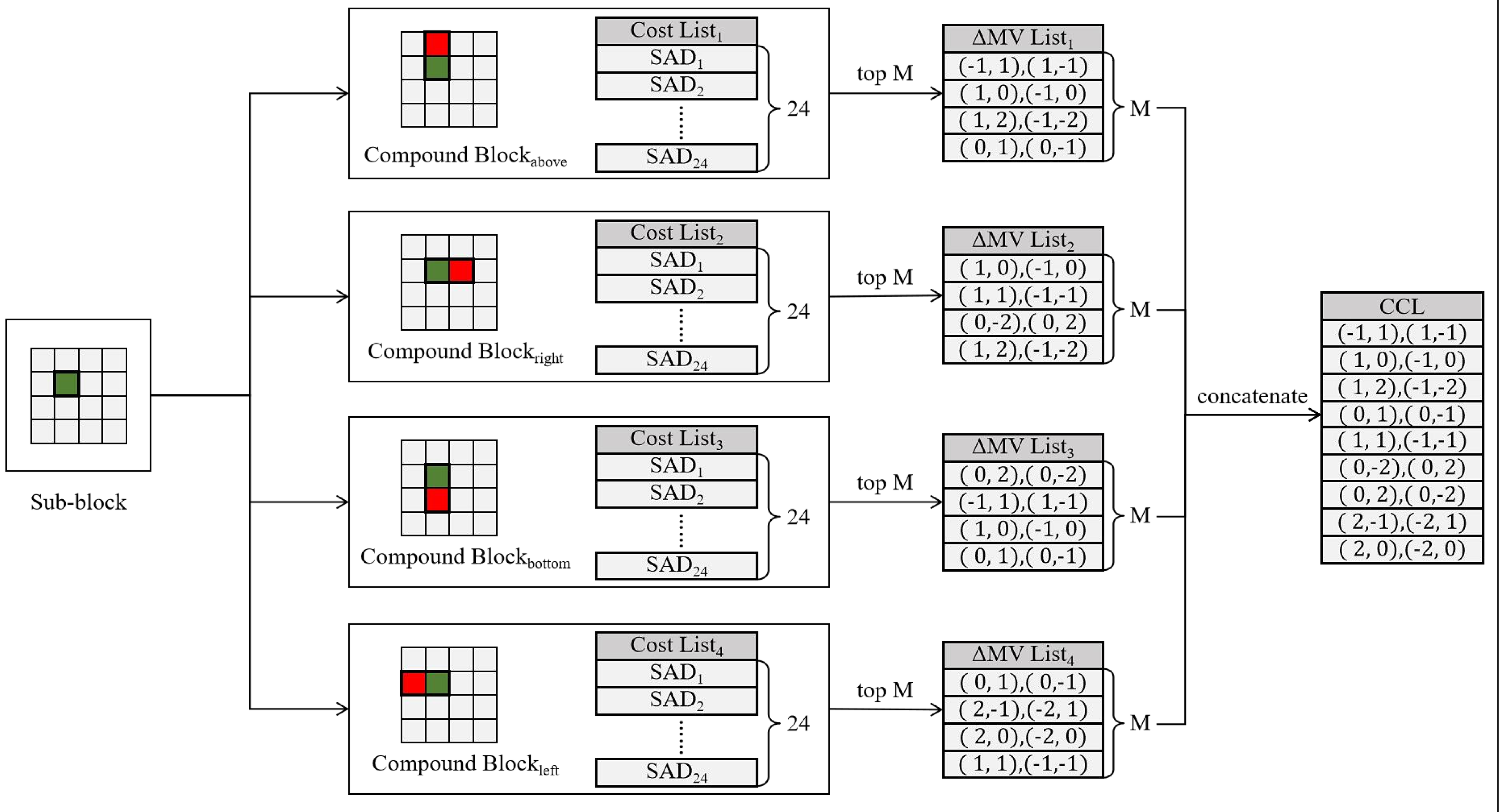
- DMVR assumes that the coding block is divided into sub-blocks, and the refinement is applied to each sub-block independently to search the optimal ΔMV within the fixed searching path. Although the assumption can effectively reduce complexity without extra signal and provide robustness for the searching results, the fixed searching rules limit the motion vector accuracy.
- For the limitation of distortion calculation, only the fixed calculation range of block pairs is considered in distortion function for each sub-block, which leads to a limited receptive field. The limited receptive field may result in an inaccurate decision of ΔMV for different motion situations.
- For the limitation of searching strategy, the motion situation varies from different sub-blocks, so the fixed searching path limits the motion vector accuracy and may result in inaccurate searching results.

3. C-DMVR Framework



I . Collaborative Distortion Calculation (CDC)

- First, the range of spatial neighbor pixels is determined based on the distortion cost of the initial block pair (SAD_0). We assume that the worse the bilateral matching, the more spatial neighbor information needs to be introduced.
- Second, the above distortion function is used to calculate the distortion cost of ΔMVs for each sub-block.



II. Collaborative Search Strategy (CSS)

- In the first stage, the candidate list is constructed with the symmetric property of ΔMV. For each sub-block, SADs of the ΔMVs in the candidate list are calculated and stored in the SAD array for subsequent use.
- In the second stage, the current sub-block and its neighbor sub-block are combined and considered as a compound block. Four compound blocks and the corresponding cost lists can be derived from the SAD array generated in the first stage. For each compound block, the ΔMVs corresponding to the M minimum values in the cost list are selected to construct a pruned ΔMV list. Four pruned ΔMV lists can be obtained and concatenated to construct the CCL, which considers different qualities of the potential ΔMVs for the current sub-block with the collaboration of neighbor sub-blocks.

4. Experiment Results

TABLE I
CODING PERFORMANCE AND RELATIVE COMPLEXITY OF C-DMVR BASED ON AV2 UNDER RA CONFIGURATION

Class	Sequence	Y	U	V
classB (1920x1080)	*MarketPlace*	0.03%	0.21%	0.20%
	RitualDance	-0.15%	-0.33%	-0.17%
	Cactus	-0.10%	-0.40%	-0.37%
	BasketballDrive	-0.21%	-0.64%	0.08%
	BQTerrace	-0.04%	-0.12%	-0.31%
	Avg	**-0.09%**	**-0.26%**	**-0.11%**
classC (832x480)	*BasketballDrill*	-0.01%	-0.36%	0.69%
	BQMall	-0.33%	0.45%	-0.02%
	PartyScene	-0.17%	0.29%	-0.57%
	RaceHorses	-0.06%	-0.31%	0.49%
	Avg	**-0.15%**	**0.02%**	**0.15%**
classD (416x240)	*BasketballPass*	-0.60%	-0.03%	0.64%
	BQSquare	0.03%	0.01%	-0.60%
	BlowingBubbles	-0.03%	-0.95%	1.07%
	RaceHorses	-0.35%	-0.14%	-0.43%
	Avg	**-0.24%**	**-0.28%**	**0.17%**
classE (1280x720)	*FourPeople*	-0.16%	-0.19%	0.23%
	Johnny	-0.33%	-0.63%	-0.32%
	KristenAndSara	-0.18%	-0.47%	-0.41%
	Avg	**-0.22%**	**-0.43%**	**-0.17%**
Overall		**-0.17%**	**-0.23%**	**0.01%**
EncT			104%	
DecT			127%	

TABLE II
CODING PERFORMANCE AND RELATIVE COMPLEXITY OF CDC AND CSS BASED ON AV2 UNDER RA CONFIGURATION

Class	Sequence	CDC			CSS		
		Y	U	V	Y	U	V
ClassB (1920x1080)	*MarketPlace*	-0.07%	0.28%	0.25%	-0.01%	-0.07%	-0.03%
	RitualDance	-0.14%	-0.08%	0.01%	-0.16%	-0.15%	0.12%
	Cactus	0.04%	-0.10%	-0.68%	0.01%	0.19%	-0.35%
	BasketballDrive	-0.04%	-0.53%	0.00%	0.01%	0.18%	-0.30%
	BQTerrace	-0.03%	0.06%	-0.45%	-0.08%	-0.40%	-0.54%
	Avg	**-0.05%**	**-0.07%**	**-0.17%**	**-0.05%**	**-0.05%**	**-0.22%**
ClassC (832x480)	*BasketballDrill*	-0.08%	-0.44%	0.16%	-0.02%	0.00%	0.78%
	BQMall	-0.11%	-0.90%	-0.35%	-0.10%	-0.06%	0.11%
	PartyScene	-0.23%	0.60%	-0.96%	-0.10%	-0.19%	-0.74%
	RaceHorses	-0.01%	-0.75%	-0.38%	0.03%	0.03%	0.32%
	Avg	**-0.11%**	**-0.37%**	**-0.38%**	**-0.05%**	**-0.05%**	**0.12%**
ClassD (416x240)	*BasketballPass*	-0.21%	0.50%	0.28%	-0.08%	0.36%	-0.48%
	BQSquare	-0.02%	-0.94%	0.19%	0.00%	-0.01%	-0.06%
	BlowingBubbles	-0.34%	-1.17%	0.31%	-0.04%	-1.01%	0.48%
	RaceHorses	-0.14%	-0.26%	-0.73%	-0.10%	0.24%	-0.26%
	Avg	**-0.18%**	**-0.47%**	**0.01%**	**-0.06%**	**-0.11%**	**-0.08%**
ClassE (1280x720)	*FourPeople*	0.05%	-0.34%	0.50%	-0.01%	-0.02%	0.42%
	Johnny	-0.40%	-0.62%	-0.59%	-0.24%	-0.53%	0.01%
	KristenAndSara	-0.08%	0.23%	-0.52%	0.01%	0.11%	-0.17%
	Avg	**-0.14%**	**-0.25%**	**-0.20%**	**-0.08%**	**-0.15%**	**0.09%**
Overall		**-0.11%**	**-0.28%**	**-0.19%**	**-0.06%**	**-0.08%**	**-0.04%**
EncT			99%			99%	
DecT			114%			104%	

Email: ljl_sa23218081@mail.ustc.edu.cn