

Hong Kong University of Science and Technology



Towards Low-Latency High-Performance Bidirectional Stereo Image Compression

Zhening Liu, Xinjie Zhang, Jiawei Shao, Zehong Lin, Jun Zhang

Background

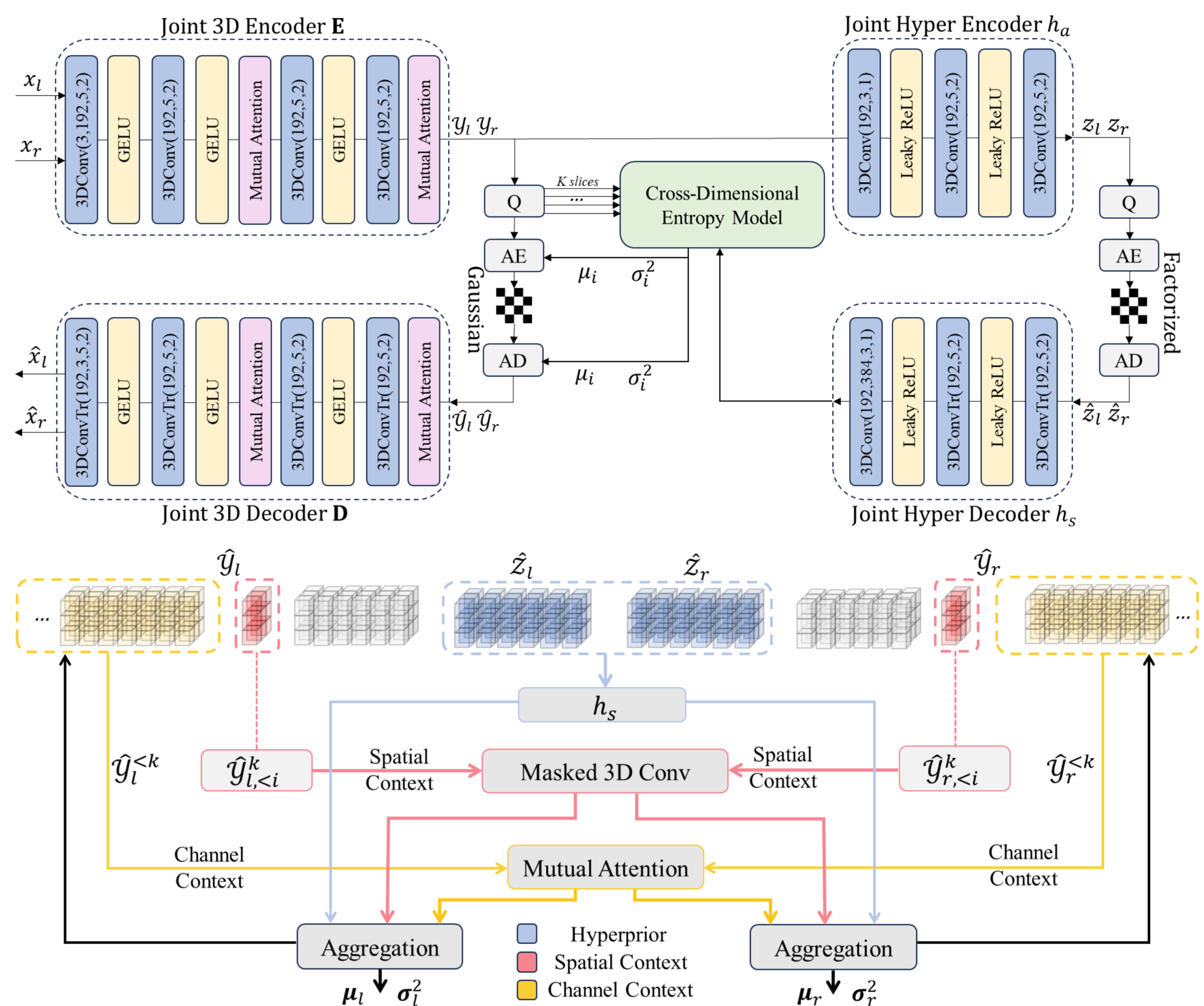
- AR/VR technology and autonomous driving applications increase the demand for stereo vision.
- Towards the era of Big Data 6G scenarios, compressing the stereo images decreases the transmission overhead and benefits communication efficiency.

Motivation

- Previous methods adopt the unidirectional predictive coding, yielding un-balanced compression results.
- Existing entropy models fail to jointly consider the stereo features on multiple dimensions.

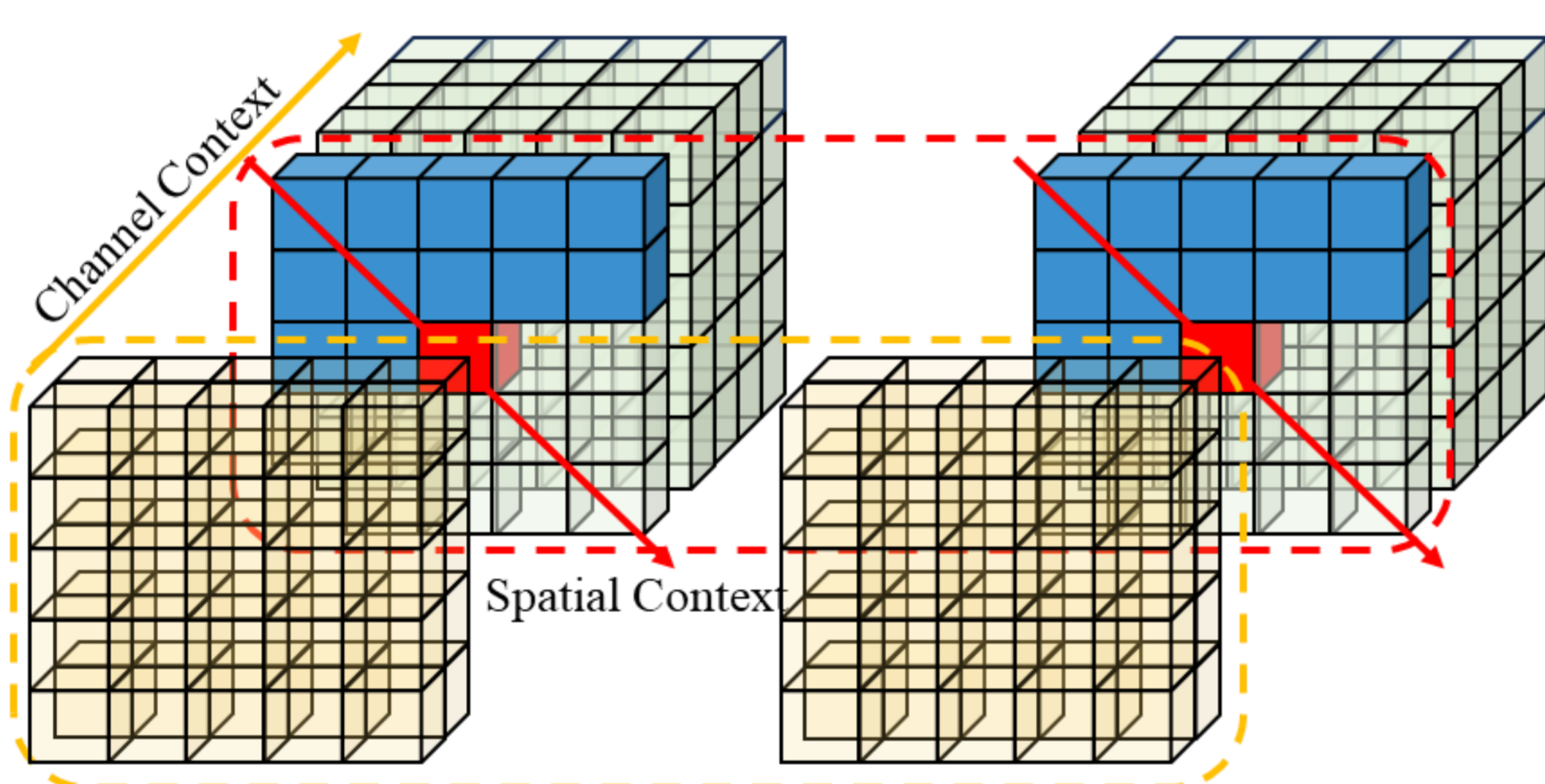
Methodology

- Bidirectional codec, symmetric entropy model.
- A low-latency version fully exploit the spatial context with stereo-checkerboard design.



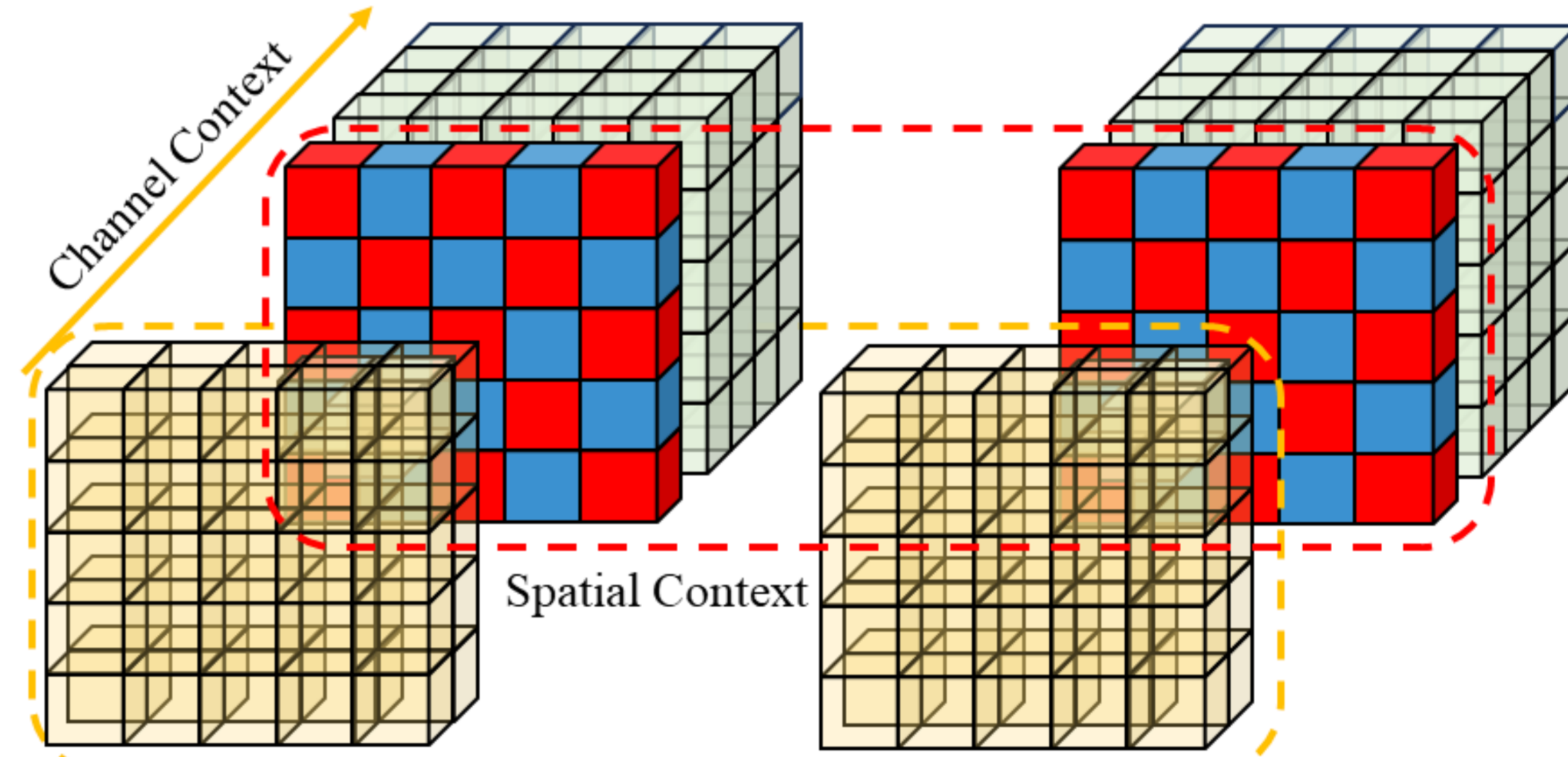
Cross-dimensional Entropy Model

- Aggregates stereo features on spatial and channel dimensions jointly and processed concurrently.



Stereo-Checkerboard

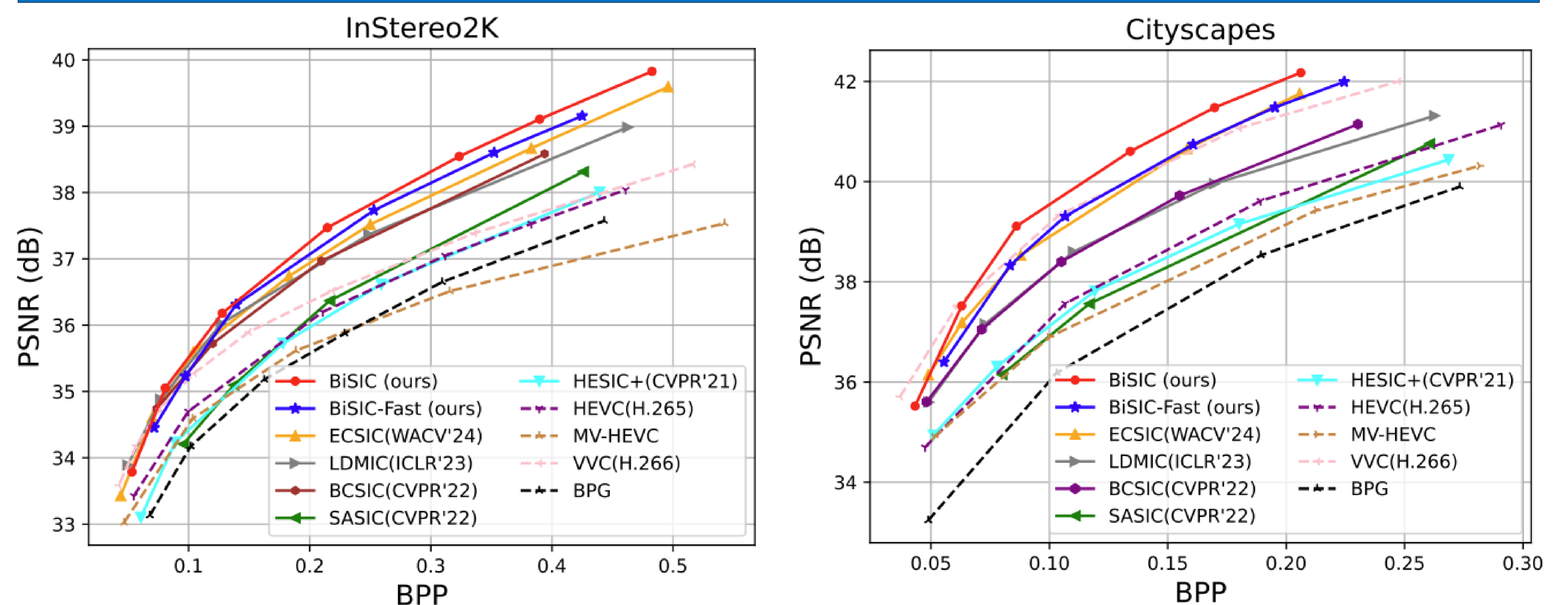
- Enables a two-fold processing towards low-latency, while keeping the neighboring context for high-performance.



Implementation

- End-to-end rate-distortion training on two large-scale stereo image datasets: InStereo2K (in-door scenes) and Cityscapes (out-door scenes).
- Performance is shown in rate-distortion curves, reflecting the quality achieved at the same bit-rate.
- Compared with traditional non-learning codecs and previous learning-based methods.

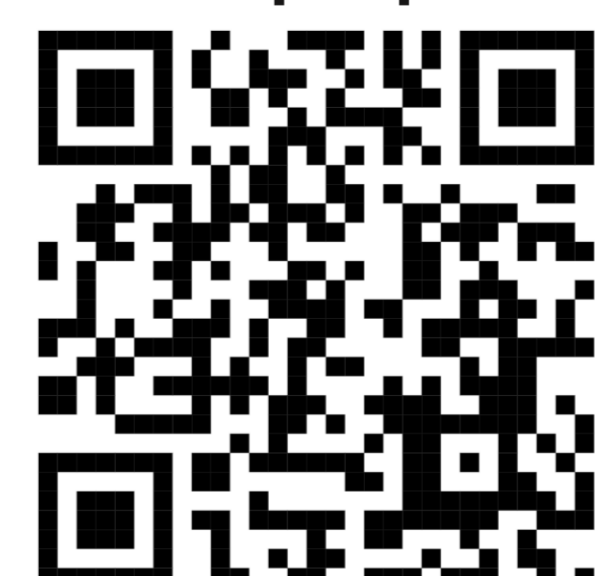
Evaluation Results



Coding Latency

| Methods | SASIC | BCSIC | LDMIC | ECSIC | VVC | Ours | Ours-Fast |
|--------------|-------|-------|-------|-------|--------|-------|-----------|
| Enc-Time (s) | 1.21 | 24.22 | 15.35 | 1.12 | 189.36 | 39.86 | 1.09 |
| Dec-Time (s) | 1.07 | 58.61 | 49.87 | 1.34 | 1.65 | 87.62 | 1.23 |

Full paper



Homepage



Acknowledgment