The 38th Annual AAAI Conference on Artificial Intelligence
FEBRUARY 20-27, 2024 | VANCOUVER, CANADA

中国科学技术大学
University of Science and Technology of China
iVC Lab

(Paper ID: 7069)

Offline and Online Optical Flow Enhancement for Deep Video Compression

Chuanbo Tang, Xihua Sheng, Zhuoyuan Li, Haotian Zhang, Li Li, Dong Liu

cbtang@mail.ustc.edu.cn

University of Science and Technology of China (USTC)

Outline

The 38th Annual AAAI Conference on Artificial Intelligence
FEBRUARY 20-27, 2024 | VANCOUVER, CANADA

中国科学技术大学
University of Science and Technology of China
iVC Lab

- Background
- Method
- Results
- Conclusion

Outline

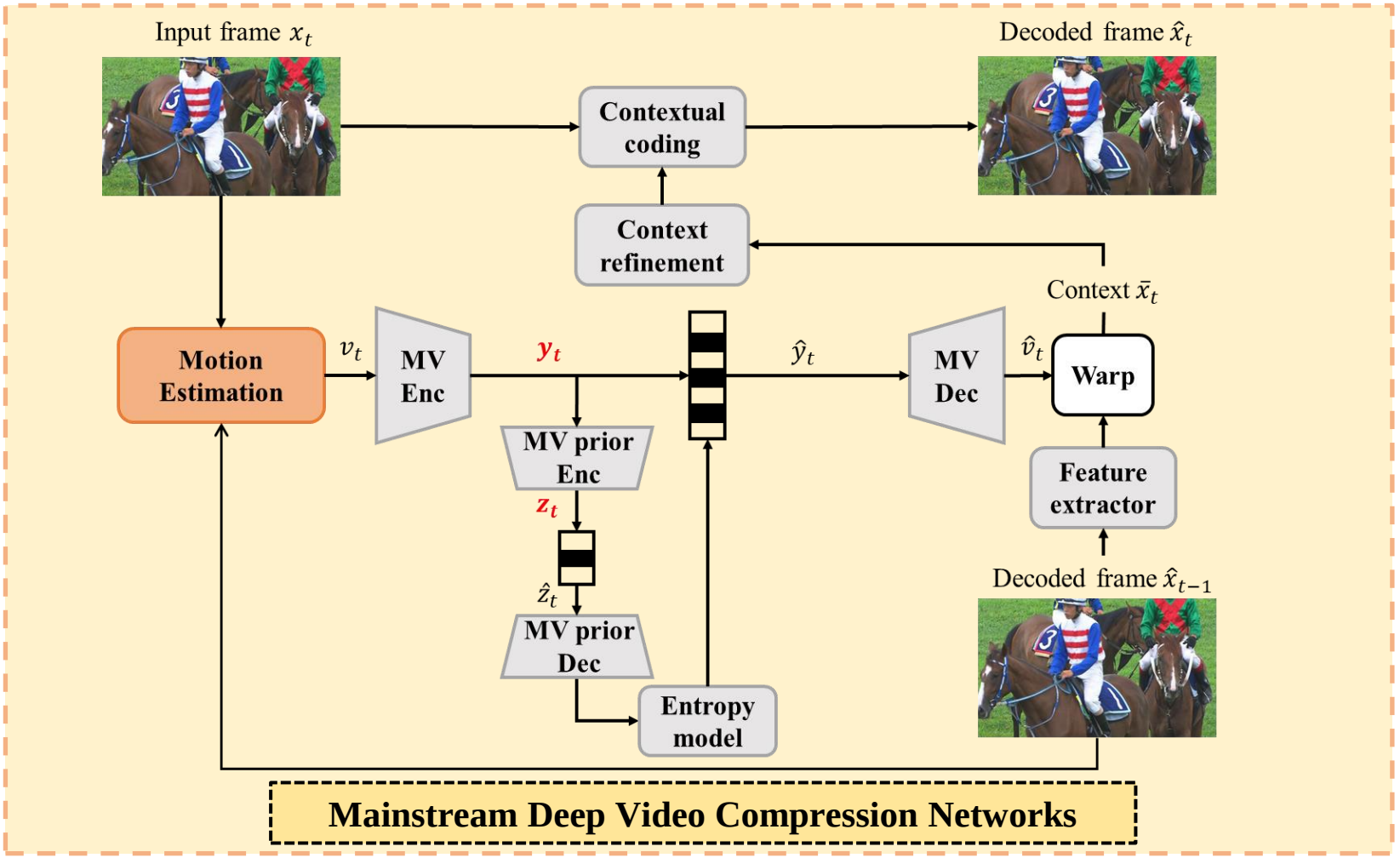
- **Background**
- **Method**
- **Results**
- **Conclusion**

Background

The 38th Annual AAAI Conference on Artificial Intelligence
FEBRUARY 20-27, 2024 | VANCOUVER, CANADA

中国科学技术大学
University of Science and Technology of China
iVC Lab

I. ME：pre-trained optical flow estimation network
II. Motion information is obtained by a simple forward



Background

The 38th Annual AAAI Conference on Artificial Intelligence
FEBRUARY 20-27, 2024 | VANCOUVER, CANADA

中国科学技术大学
University of Science and Technology of China
iVC Lab

Analysis

Optical flows

Trained for inter-frame accuracy only → Lack of rate constrain

Trained on synthesis data → Lack of generalization on real-world data

Motion information in deep video compression → Obtained by a simple forward in inference → Not optimal for different coding sequences

?

Traditional Video Compression

Offline: Designing various hand-crafted inter-frame prediction modes

Offline: Searching the mode achieving best RD performance

Outline

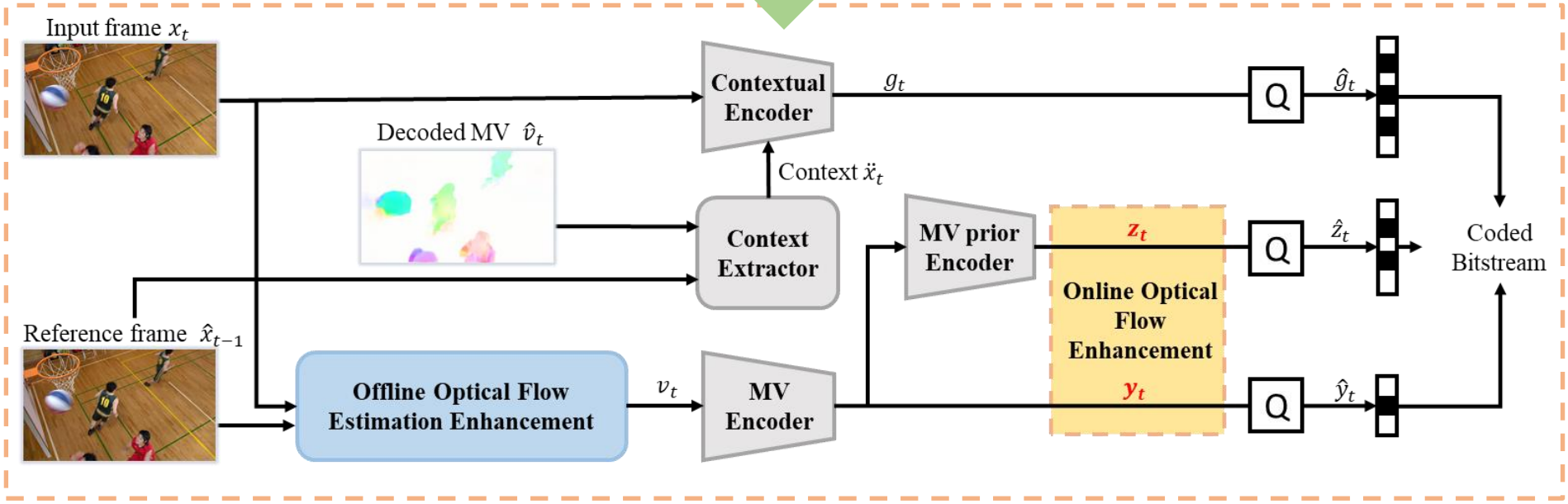
- Background
- **Method**
- Results
- Conclusion

Method

Offline and Online Optical Flow Enhancement:

Offline Enhancement → **Motion estimation optimized for both rate and inter-frame prediction accuracy**

Online Enhancement → **Interframe modules achieve content adaptive**



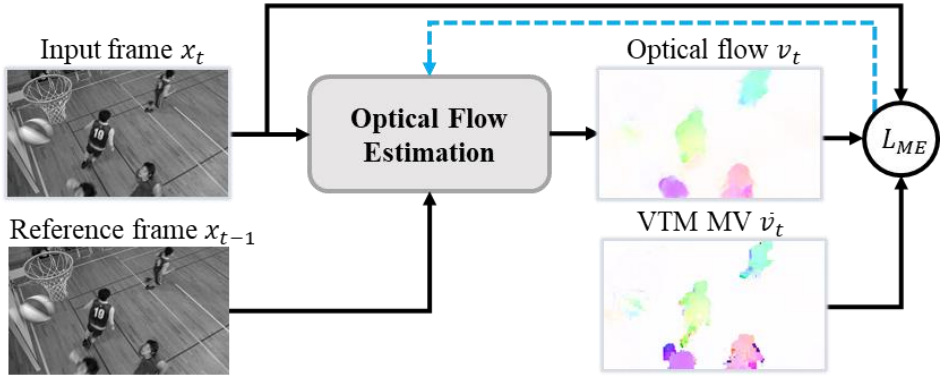
Method

The 38th Annual AAAI Conference on Artificial Intelligence
FEBRUARY 20-27, 2024 | VANCOUVER, CANADA

中国科学技术大学
University of Science and Technology of China
iVC Lab

Offline Enhancement:

➢ Finetuning the pre-trained Spynet supervise by VTM MV $\bar{v}_t$



➢ **Objective 1: Minimizing EPE Loss**

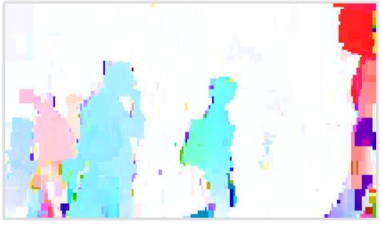
➢ **Objective 2: Minimizing Warp MSE**

$$L_{ME} = \frac{1}{mn}\sum_{i,j}\sqrt{(v_i - \bar{v}_i)^2 + (v_j - \bar{v}_j)^2} + \lambda_{ME} * d(x_t, \bar{x}_t)$$

Objective 1 | Objective 2

(a) Spynet Optical Flow

(b) Enhanced Spynet Optical Flow



(c) VTM MV

(d) Spynet Warp Frame (28.12dB)

(e) Enhanced Spynet Warp Frame (31.32dB)

(f) VTM Warp Frame(33.58dB)

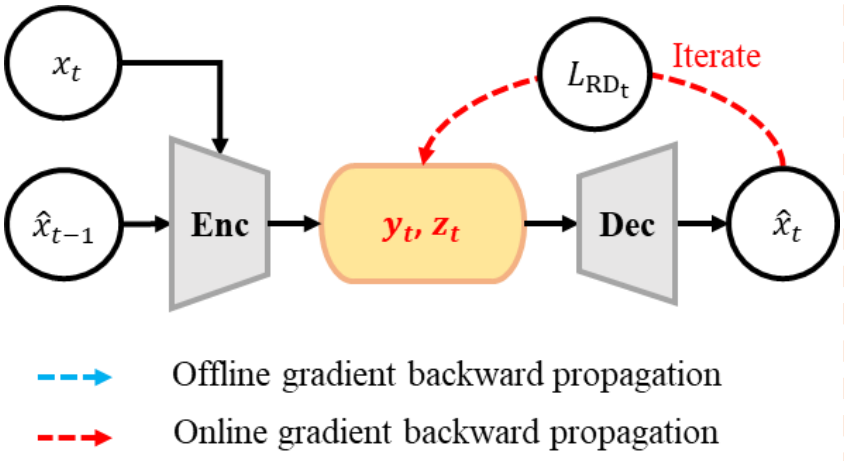
Method

Online Enhancement:

- Online Updating the latent features of the optical flows with a gradient descent-based algorithm.

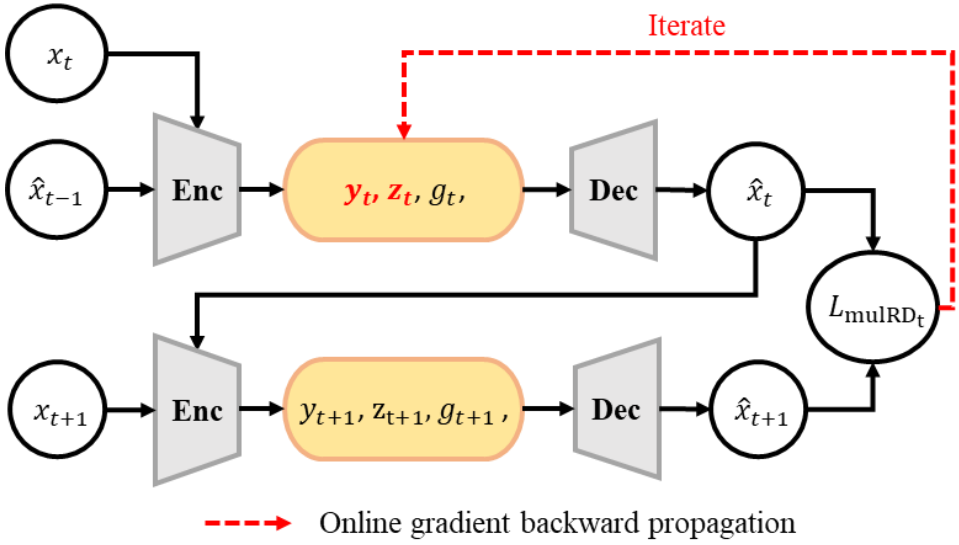
Single-frame level

$$\tilde{L}^{i}_{RD_t} = \lambda d(x_j, \tilde{x}^{i}_{j}) + H(\tilde{y}^{i}_{j}) + H(\tilde{z}^{i}_{j}) + H(\tilde{g}^{i}_{j})$$



Multi-frame level

$$\tilde{L}^{i}_{mulRD_t} = \sum_{j=t}^{W} \alpha_j [\lambda d(x_j, \tilde{x}^{i}_{j}) + H(\tilde{y}^{i}_{j}) + H(\tilde{z}^{i}_{j}) + H(\tilde{g}^{i}_{j})]$$



Outline

- Background
- Method
- **Results**
- Conclusion

Results

The 38th Annual AAAI Conference on Artificial Intelligence
FEBRUARY 20-27, 2024 | VANCOUVER, CANADA

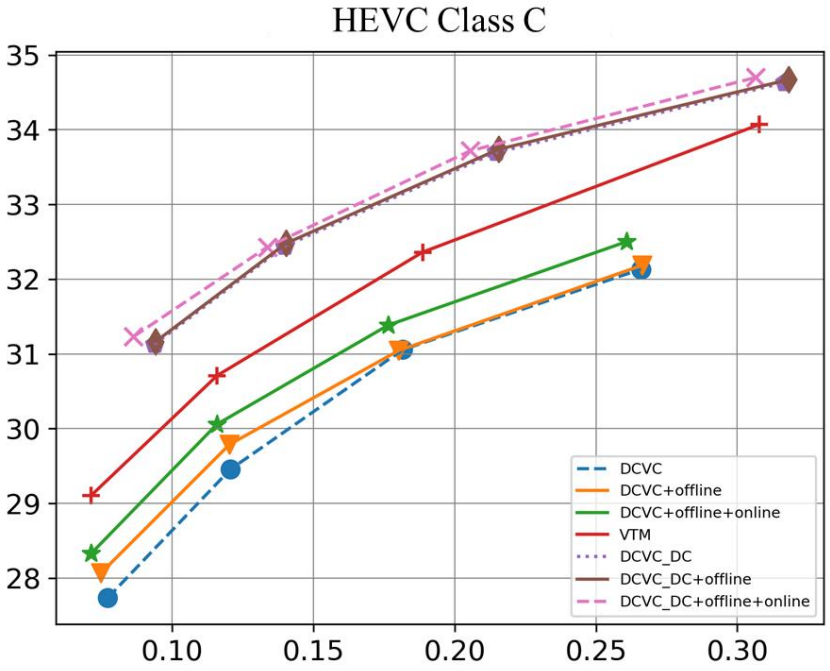
中国科学技术大学
University of Science and Technology of China
iVC Lab

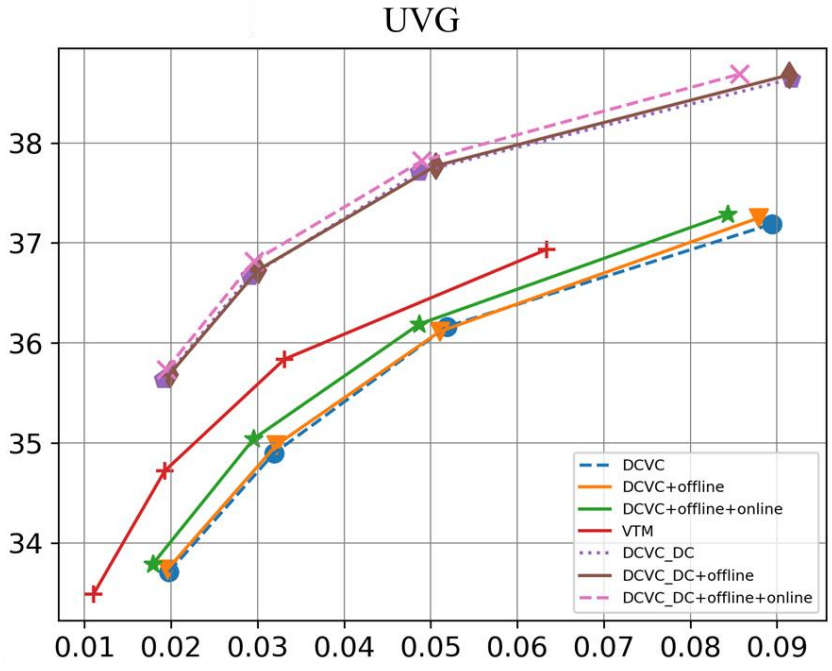
Based on DCVC [1]:

- Offline enhancement: **4.3%** bitrate saving
- Offline and Online enhancement: **13.4%** bitrate saving

Based on DCVC-DC [2]:

- Offline enhancement:**1.1%** bitrate saving
- Offline and Online enhancement: **4.1%** bitrate saving





[1] Li J, Li B, Lu Y. Deep contextual video compression[J]. Advances in Neural Information Processing Systems, 2021, 34: 18114-18125.
[2] Li J, Li B, Lu Y. Neural video compression with diverse contexts[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 22616-22626.

Results

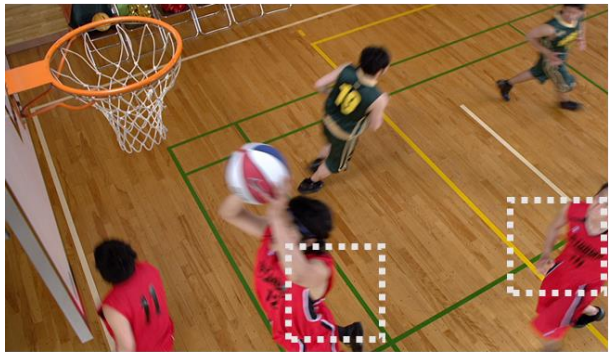
The 38th Annual AAAI Conference on Artificial Intelligence
FEBRUARY 20-27, 2024 | VANCOUVER, CANADA

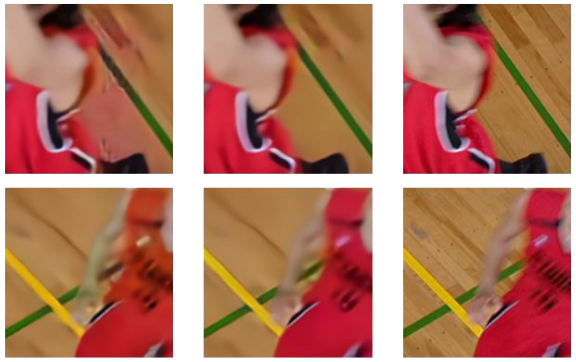
中国科学技术大学
University of Science and Technology of China
iVC Lab

Subjective Results of Reconstructed Frames:

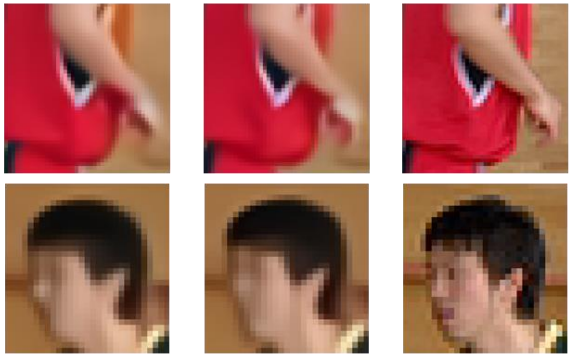
- Higher quality reconstruction using fewer bits than DCVC

Frame No.23
BasketballDrill_832×480





Frame No.22
BasketballPass_416×240



Results

The 38th Annual AAAI Conference on Artificial Intelligence
FEBRUARY 20-27, 2024 | VANCOUVER, CANADA

中国科学技术大学
University of Science and Technology of China
iVC Lab

Effectiveness of the Offline and Online Enhancement on DCVC-DC

	B	C	D	UVG	Average
DCVC-DC	0.0	0.0	0.0	0.0	0.0
DCVC	66.6	79.7	76.7	78.7	75.4
DCVC-DC + offline	-0.7	-1.0	-2.1	-0.4	-1.1
DCVC-DC + offline + online	-2.8	-4.9	-4.6	-4.2	-4.1

Based on DCVC:

- **Offline**: 4.3% bitrate saving
- **Online**: 10.7% bitrate saving
- **Offline and Online**: 13.4% bitrate saving

Ablation study of Offline and Online Enhancement on DCVC

Offline	Online	B	C	D	E	RGB	UVG	MCL	Average
✗	✗	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
✓	✗	-3.0	-5.9	-4.4	-7.9	-0.7	-1.3	-6.7	-4.3
✗	✓	-10.7	-14.3	-11.1	-9.0	-8.5	-10.1	-11.3	-10.7
✓	✓	-12.0	-17.1	-13.1	-15.3	-8.8	-10.5	-16.9	-13.4

Results

The 38th Annual AAAI Conference on Artificial Intelligence
FEBRUARY 20-27, 2024 | VANCOUVER, CANADA

中国科学技术大学
University of Science and Technology of China
iVC Lab

Subjective Results of decoded flows and enhanced flows on DCVC:

(a) DCVC
Decoded Optical Flow
(0.01473bpp)

(b) DCVC + Offline
Decoded Optical Flow
(0.01406bpp)

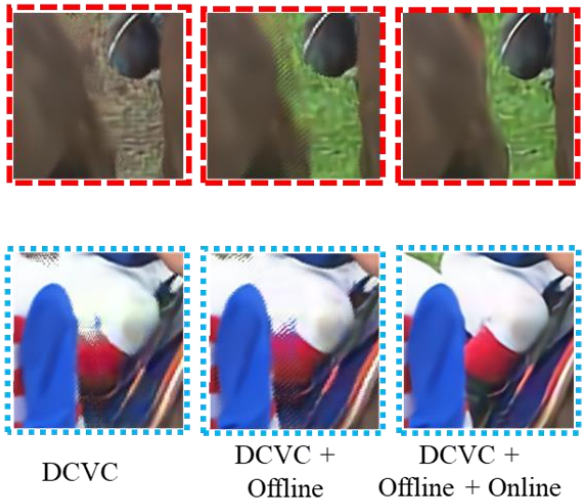
(c) DCVC + Offline + Online
Decoded Optical Flow
(0.01353bpp)

(d) DCVC
Warp Frame (24.41dB)

(e) DCVC + Offline
Warp Frame (24.98dB)

(f) DCVC + Offline + Online
Warp Frame (25.15dB)

(g) Raw Frame



DCVC | DCVC + Offline | DCVC + Offline + Online

Results

Based on DCVC with Offline Enhancement:

- Rate, Distortion, and Encoding Time Tradeoff: $U = 1500$

Ablation study of Offline Updating times *U*

U	C	D	ENC_T C(s)	DEC_T C(s)	ENC_T D(s)	DEC_T D(s)
0	0.0	0.0	2.71	6.94	0.70	1.91
100	-6.1	-5.1	28.15	6.84	10.42	1.90
500	-9.6	-7.9	132.78	6.95	48.99	1.87
1000	-10.8	-8.6	269.20	6.73	92.58	1.89
1500	-11.2	-8.7	388.73	6.86	141.03	1.91
2000	-11.5	-9.1	530.10	6.84	190.64	1.89
2500	-11.6	-9.2	674.54	6.89	239.05	1.88

Based on DCVC with Offline and Offline Enhancement:

- Rate, Distortion, and Encoding Time Tradeoff: $W = 2$

Ablation study of number of frame *W* in the online enhancement

W	C	D	ENC_T C(s)	DEC_T C(s)	ENC_T D(s)	DEC_T D(s)
2	0.0	0.0	518.25	6.84	187.82	1.89
3	-0.5	-0.4	1631.35	6.84	546.99	1.88
4	-0.7	-0.7	2187.58	6.82	683.04	1.88
5	-0.8	-0.8	2706.39	6.87	874.56	1.86

Outline

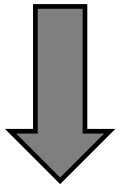
The 38th Annual AAAI Conference on Artificial Intelligence
FEBRUARY 20-27, 2024 | VANCOUVER, CANADA

中国科学技术大学
University of Science and Technology of China
iVC Lab

- Background
- Method
- Results
- **Conclusion**

Conclusion

- **Plug-and-Play Mechanism for Deep Video Compression**
- **SOTA Performance**



Future:

- **Better motion information supervision**
- **Better balance between the RD performance and encoding complexity**
- **…**

Thank you!

Chuanbo Tang

cbtang@mail.ustc.edu.cn

University of Science and Technology of China (USTC)