# Learning for Visual Data Compression

CVPR 2021 Tutorial

# Learned Video Compression

***Guo Lu***

Beijing Institute of Technology, China

June 19, 2021

# Outline

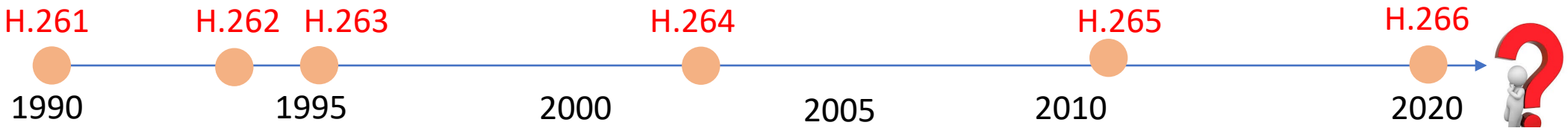
- Background for Video Compression
- End-to-end Learned P-frame Compression
  - Hybrid coding framework
  - RDO techniques
  - Enhanced motion estimation
  - Multiple reference
- End-to-end Learned B-frame Compression
- Learned Autoencoder based Video Compression
- Discussion

# Outline

- Background for Video Compression
- End-to-end Learned P-frame Compression
  - Hybrid coding framework
  - RDO techniques
  - Enhanced motion estimation
  - Multiple reference
- End-to-end Learned B-frame Compression
- Learned Autoencoder based Video Compression
- Discussion

# Background for Video Compression

Traditional codecs rely on classical prediction-transform architecture and hand-crafted techniques.

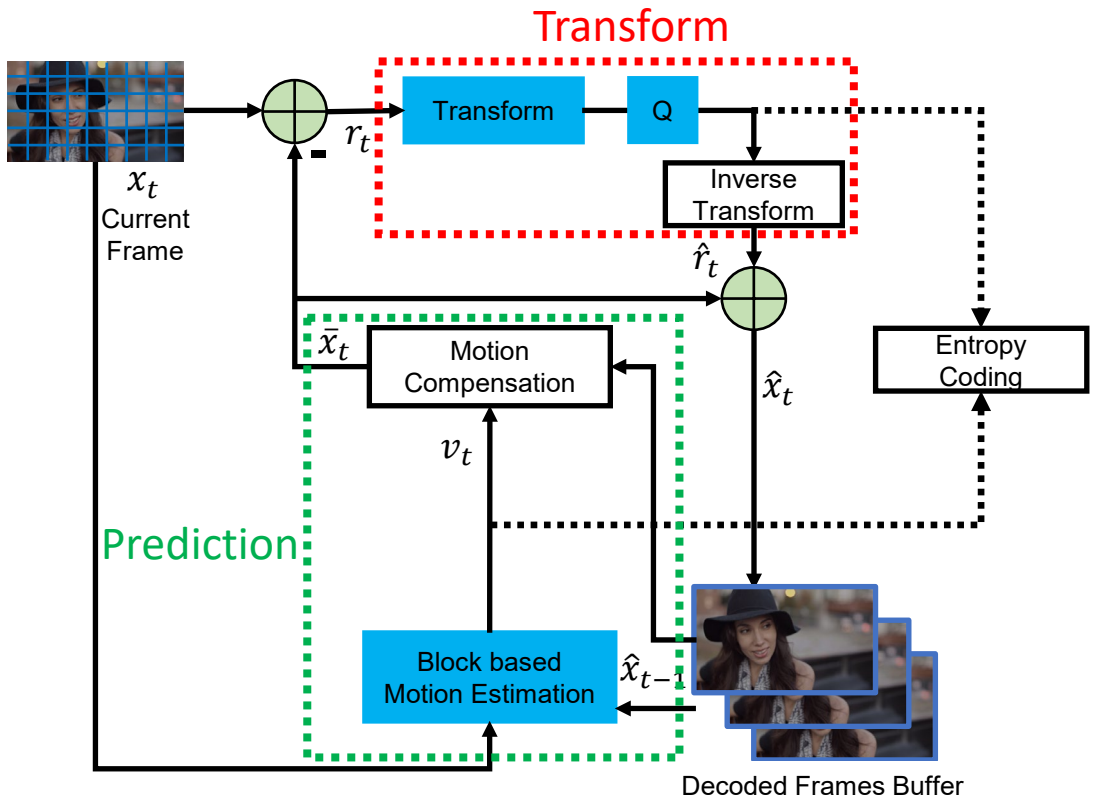


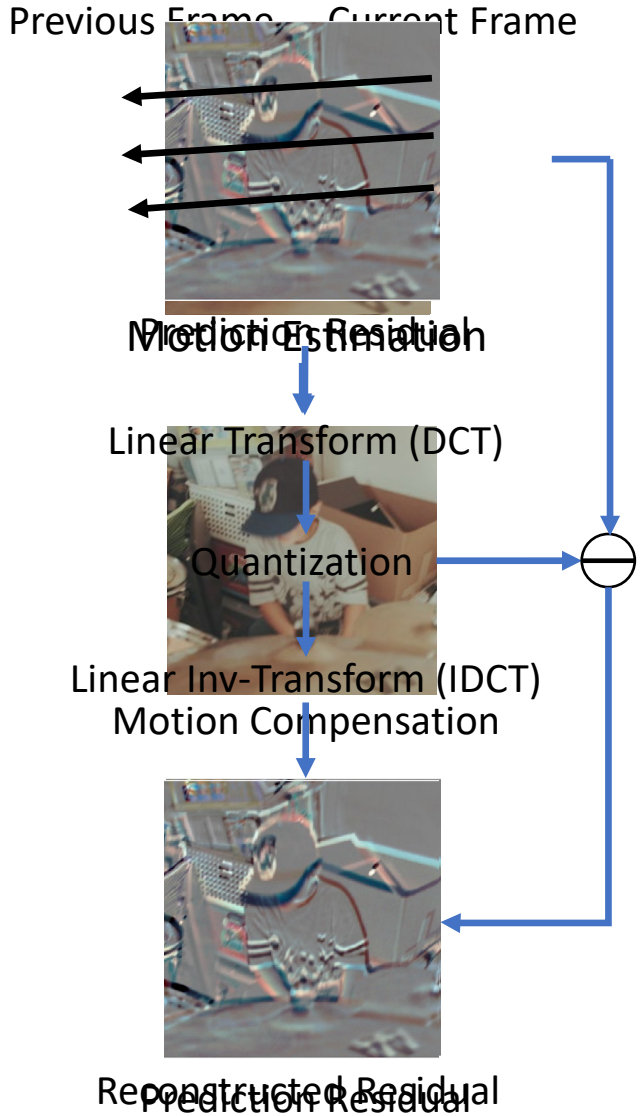

Deep learning has been widely used for a lot of vision tasks for its powerful representation ability.

***What happens when video compression meets deep learning?***

# Background for Video Compression

- Traditional Video Compression



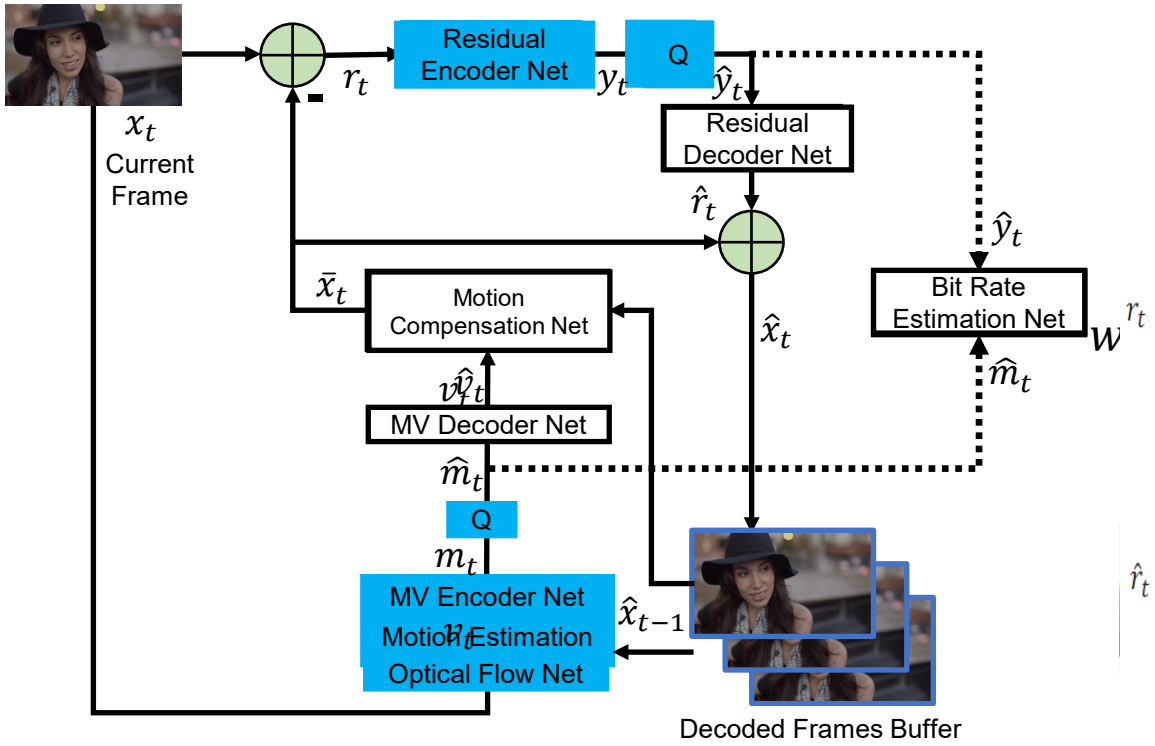



# Outline

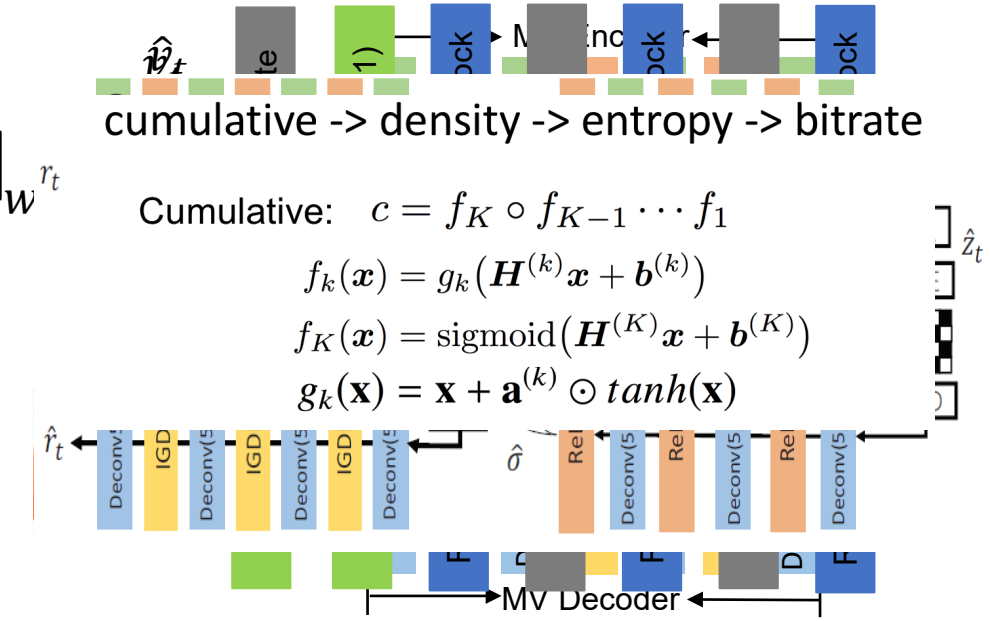
- Background for Video Compression
- End-to-end Learned P-frame Compression
  - Hybrid coding framework
  - RDO techniques
  - Enhanced motion estimation
  - Multiple reference
- End-to-end Learned B-frame Compression
- Learned Autoencoder based Video Compression
- Discussion

# End-to-End Learned P-Frame Video Compression

- The first end-to-end optimized video compression system[1]




$$\min \lambda D + R$$




[1]Guo Lu, et al. "DVC: An end-to-end deep video compression framework", in CVPR(Oral), 2019.

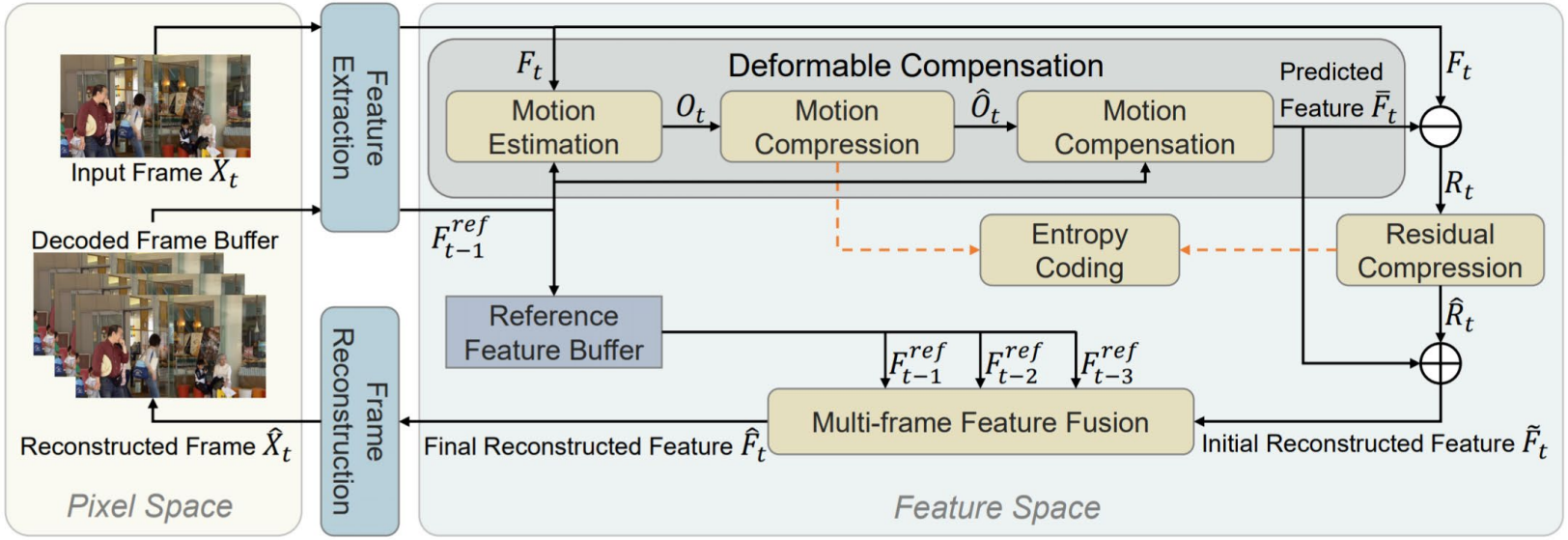
# End-to-End Learned P-Frame Video Compression

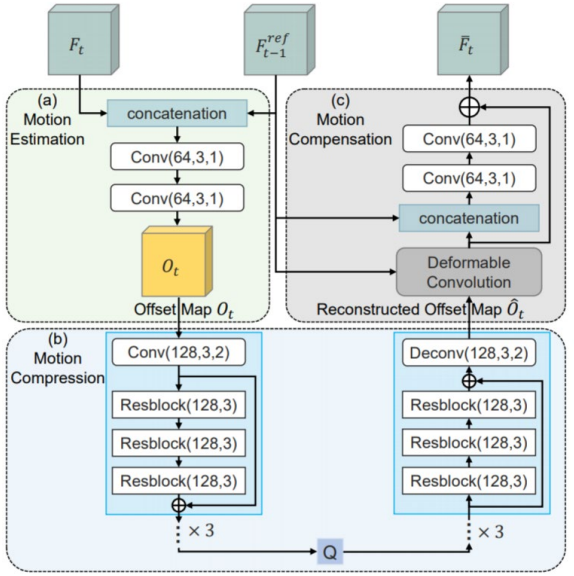
- Pixel-level operations
  - It is difficult to estimate accurate optical flow
  - Pixel-level motion compensation will introduce additional artifacts
  - Pixel residual information is not easy to compress
- Feature-level operations
  - All the operations are in feature space
  - Motion estimation in feature space by using deformable convolution
  - Compress the residual feature instead residual itself

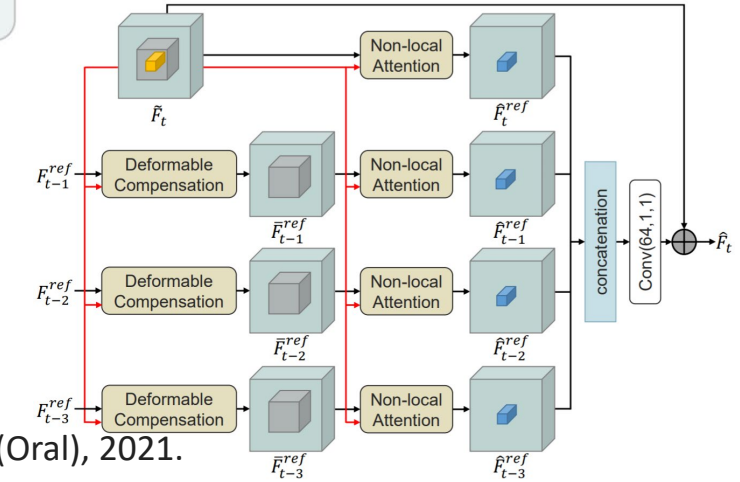
[12] Zhihao Hu, et al. “FVC: A New Framework towards Deep Video Compression in Feature Space”, in CVPR(Oral), 2021.

# End-to-End Learned P-Frame Video Compression

- Hybrid coding framework in feature space



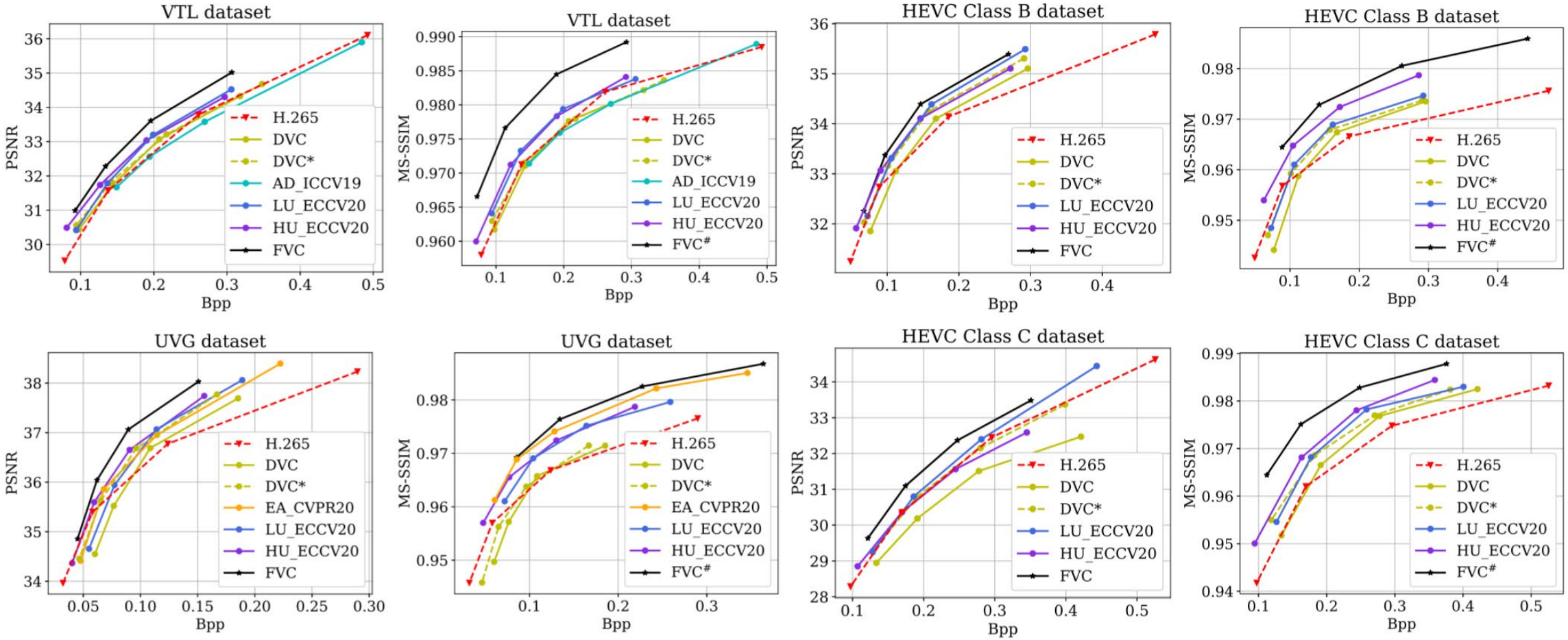




1. Feature Extraction
2. Deformable Compensation
3. Residual Compression
4. Multi-frame Fusion
5. Frame Reconstruction
6. Entropy Coding

[12] Zhihao Hu, et al. "FVC: A New Framework towards Deep Video Compression in Feature Space", in CVPR(Oral), 2021.

# End-to-End Learned P-Frame Video Compression

- Hybrid coding framework in feature space



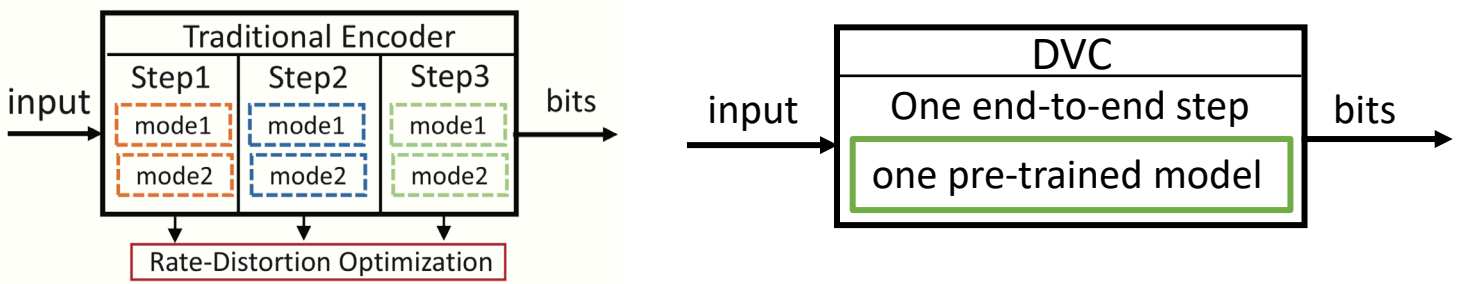

[12] Zhihao Hu, et al. "FVC: A New Framework towards Deep Video Compression in Feature Space", in CVPR(Oral), 2021.

# Outline

- Background for Video Compression
- End-to-end Learned P-frame Compression
  - Hybrid coding framework
  - RDO techniques
  - Enhanced motion estimation
  - Multiple reference
- End-to-end Learned B-frame Compression
- Learned Autoencoder based Video Compression
- Discussion

# End-to-End Learned P-Frame Video Compression

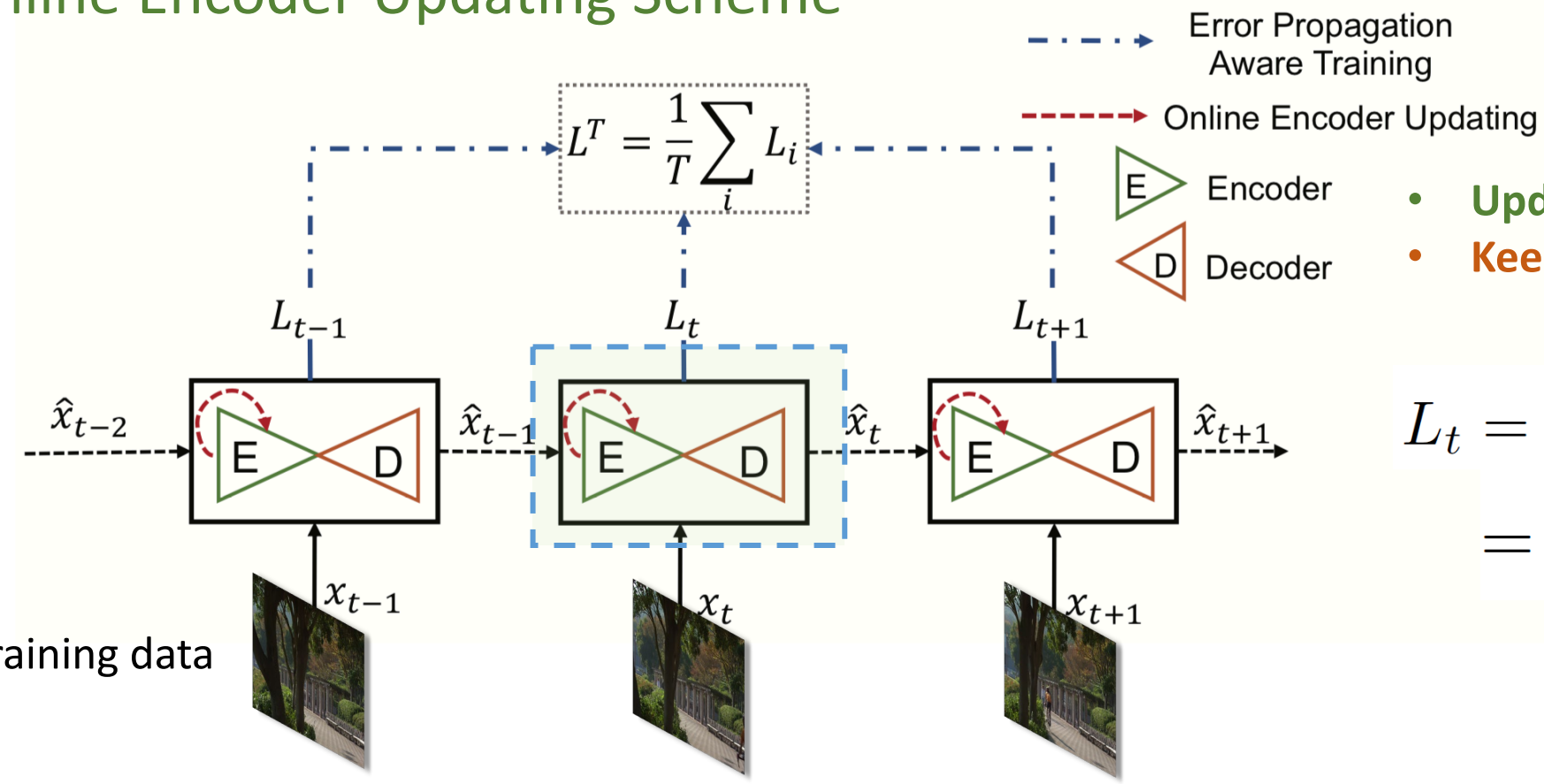
- RDO is the fundamental technique for video compression
  - Choose the optimal mode in the coding stage.
- Learned video compression
  - RDO is ignored in the inference stage
  - the “modes” are fixed after the training stage




- How to apply the RDO technique in learned video compression
  - Directly optimize the encoder
  - Introduce more “modes” and select the optimal mode for each frame

# End-to-End Learned P-Frame Video Compression

## 1. Online Encoder Updating Scheme



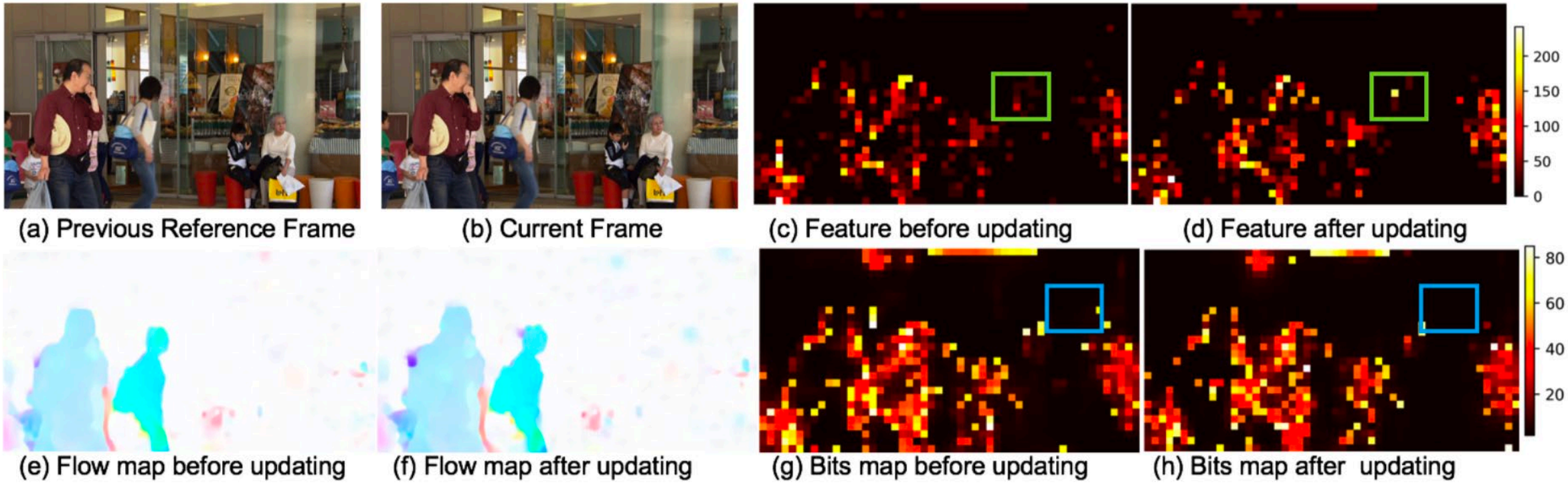

- **Update encoder online**
- **Keep the decoder unchanged**

$$L_t = \lambda D_t + R_t$$
$$= \lambda d(x_t, \hat{x}_t)$$
$$+ [H(\hat{y}_t) + H(\hat{m}_t)]$$

[3]Guo Lu, Chunlei Cai, et al. "Content Adaptive and Error Propagation Aware Deep Video Compression", in ECCV(Oral), 2020.

# End-to-End Learned P-Frame Video Compression

## 1. Online Encoder Updating Scheme



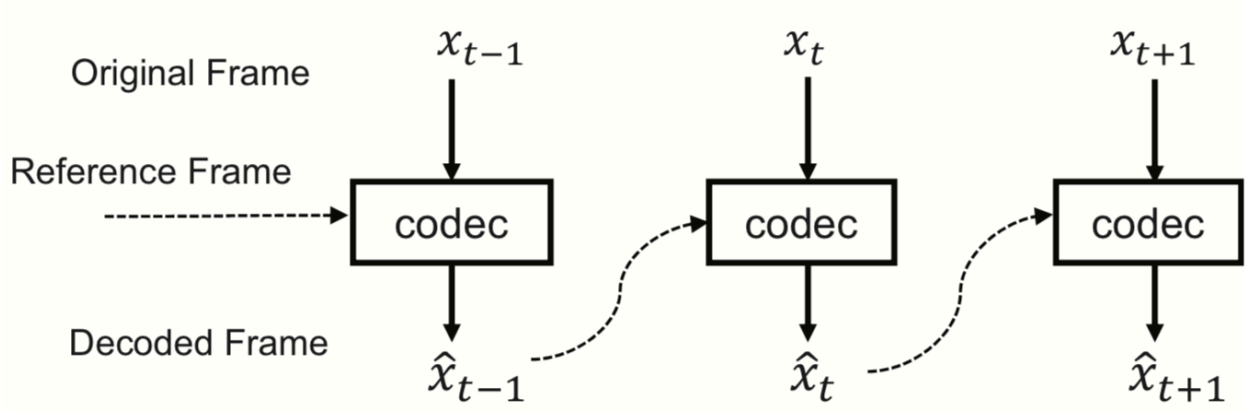

*Improve adaptiveness with same decoding time*

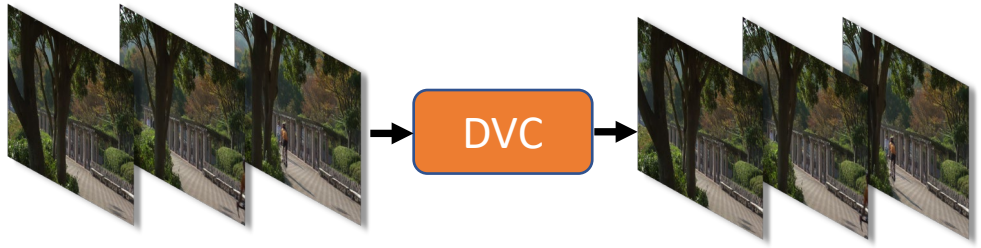
*PSNR: 33.40dB -> 34.13dB  Bpp: 0.056 -> 0.051*

[3]Guo Lu, Chunlei Cai, et al. “Content Adaptive and Error Propagation Aware Deep Video Compression”, in ECCV(Oral), 2020.

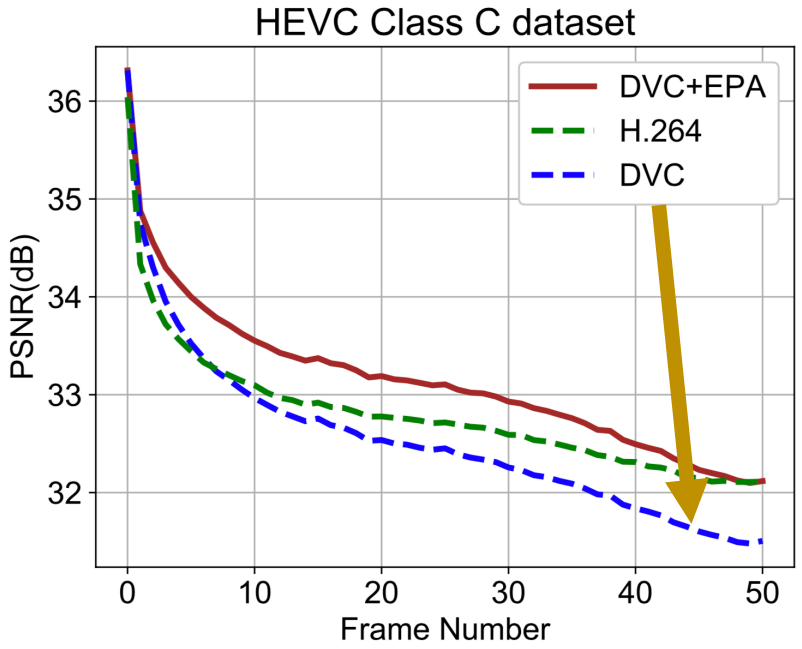
# End-to-End Learned P-Frame Video Compression

## Error propagation in inter predictive coding




**Frame quality keeps decreasing**

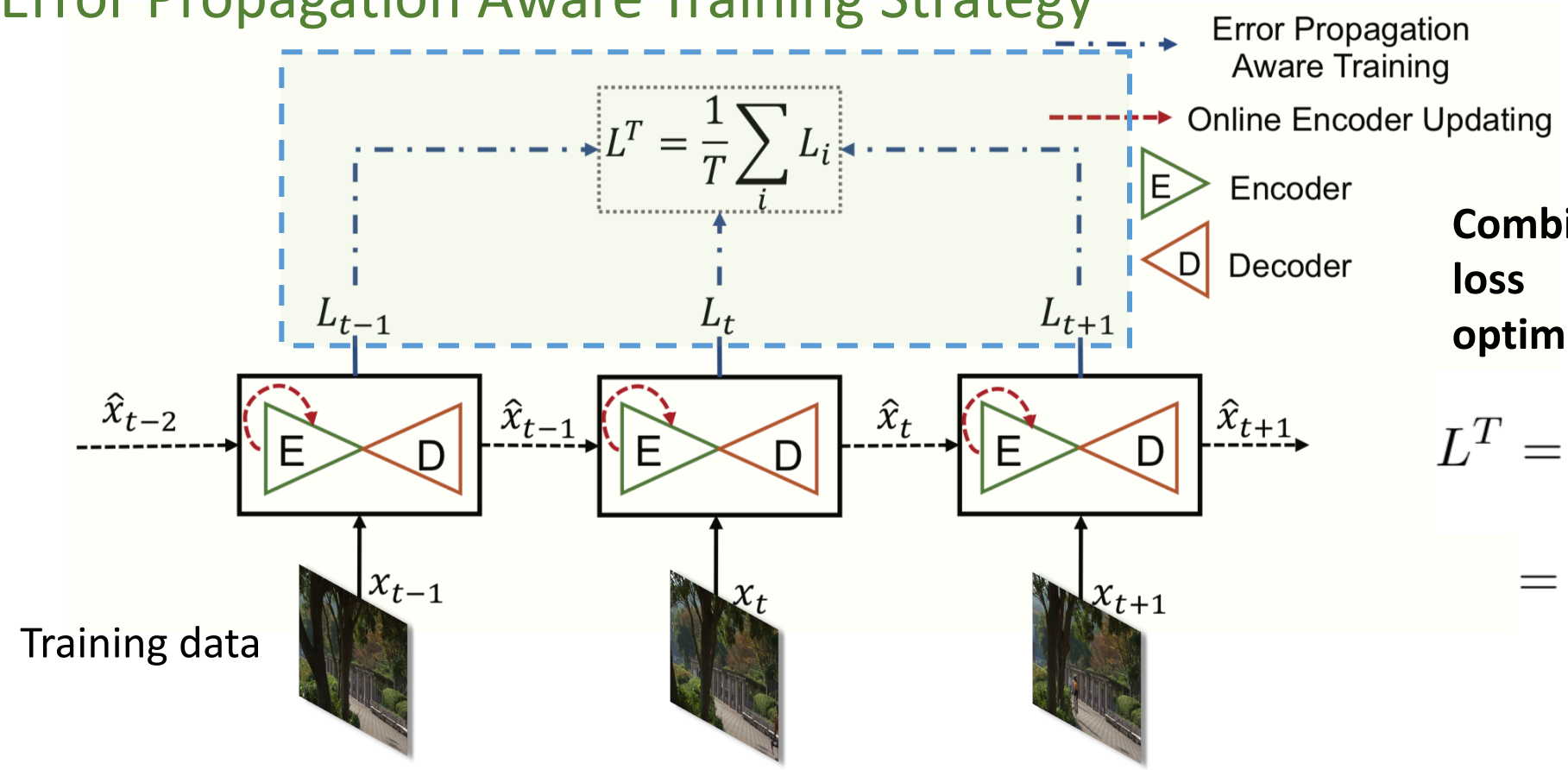



Quality keeps decreasing because **the error propagation is not considered in the training procedure**

[3]Guo Lu, Chunlei Cai, et al. “Content Adaptive and Error Propagation Aware Deep Video Compression”, in ECCV(Oral), 2020.

# End-to-End Learned P-Frame Video Compression

## 2. Error Propagation Aware Training Strategy



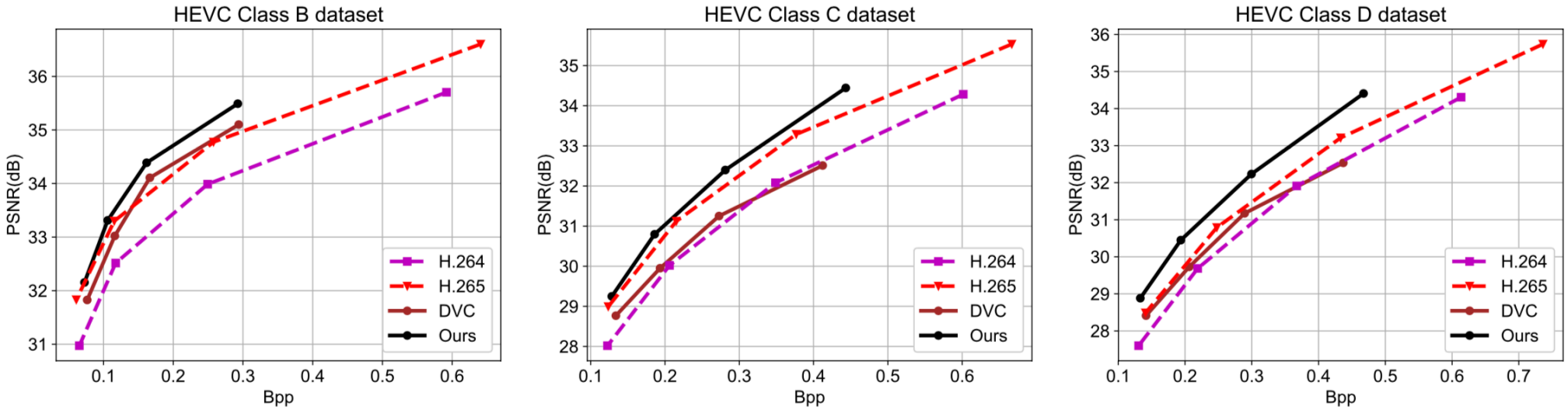

**Combine the rate-distortion loss of several frames to optimize the learned codec**

$$L^T = \frac{1}{T}\sum L_t$$

$$= \frac{1}{T}\sum\{\lambda d(x_t, \hat{x}_t) + [H(\hat{y}_t) + H(\hat{m}_t)]\}$$

[3]Guo Lu, Chunlei Cai, et al. “Content Adaptive and Error Propagation Aware Deep Video Compression”, in ECCV(Oral), 2020.

# End-to-End Learned P-Frame Video Compression




| | BDBR(%) | | | BD-PSNR(dB) | | |
|---|---|---|---|---|---|---|
| Dataset | H.265 | DVC | Ours | H.265 | DVC | Ours |
| Class B | -32.0 | -27.9 | -41.7 | 0.78 | 0.71 | 1.12 |
| Class C | -20.8 | -3.5 | -25.9 | 0.91 | 0.13 | 1.18 |
| Class D | -12.3 | -6.2 | -25.1 | 0.57 | 0.26 | 1.25 |

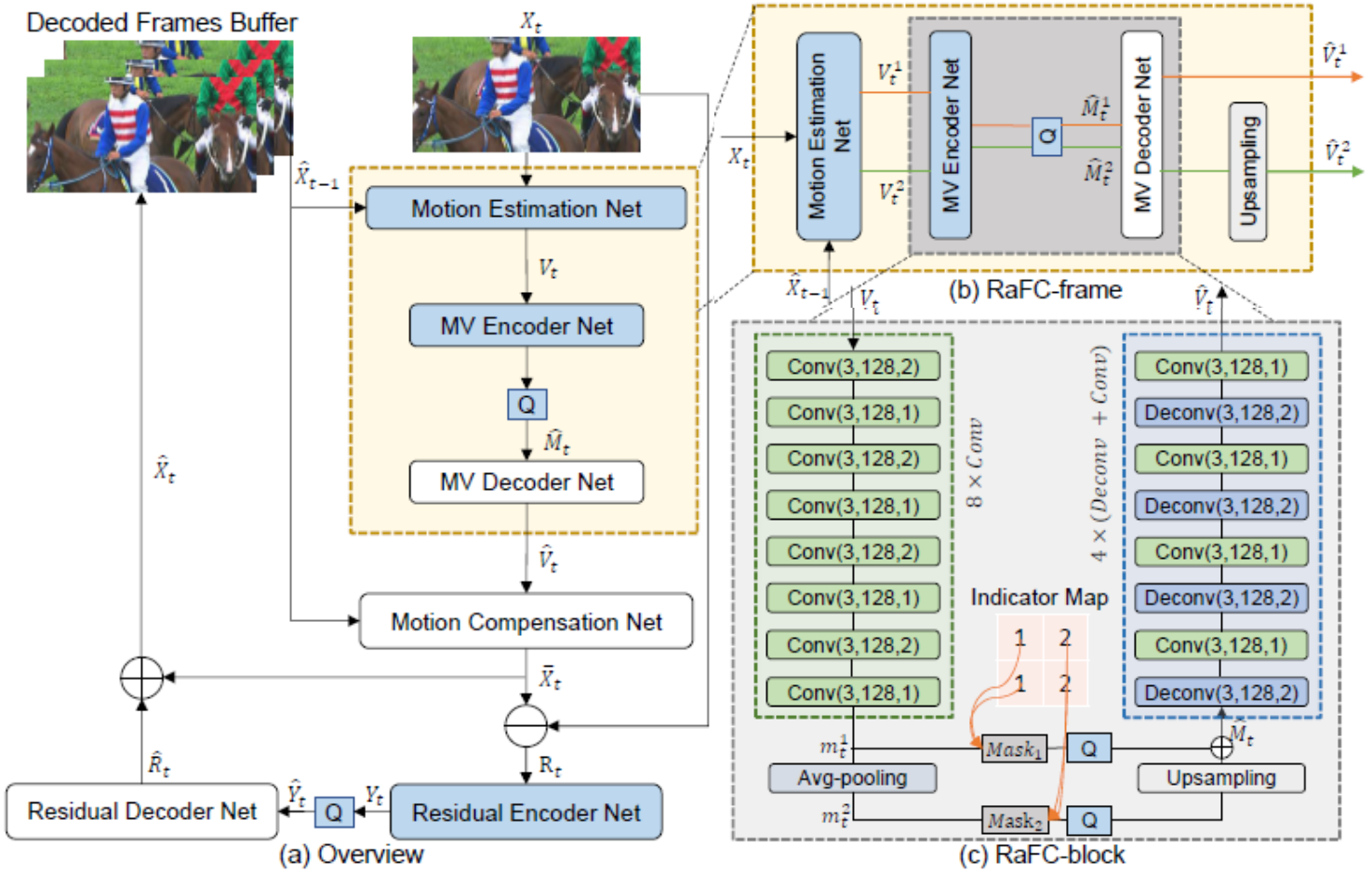
[3]Guo Lu, Chunlei Cai, et al. “Content Adaptive and Error Propagation Aware Deep Video Compression”, in ECCV(Oral), 2020.

# End-to-End Learned P-Frame Video Compression

- Various block sizes are adopted in traditional video compression
  - Large block size for smooth region
  - Small block size for complex region
- Fixed optical flow resolution and motion representation resolutions are used in existing work, like DVC.
  - Generates more modes -> flow resolutions or representation resolutions
  - Choose the optimal mode using RDO

# End-to-End Learned P-Frame Video Compression

- Resolution-adaptive Flow Coding

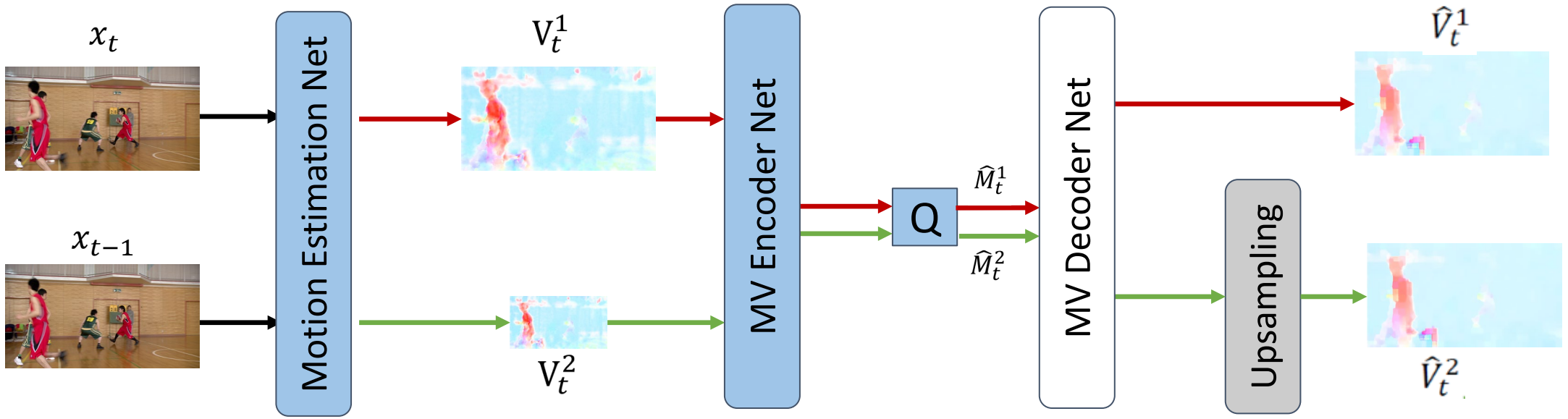



- (a) Overview of the proposed video compression system.
- (b) RaFC-Frame: decides the **Global Optimal Flow Map** resolution for each video frame.
- (c) RaFC-Block: select the optimal resolution for each **Local Block** of motion feature.

[4] Zhihao Hu, et al. "Improving Deep Video Compression by Resolution-adaptive Flow Coding", in ECCV(Oral), 2020.

# End-to-End Learned P-Frame Video Compression

- Resolution-adaptive Flow Coding(Frame-level)

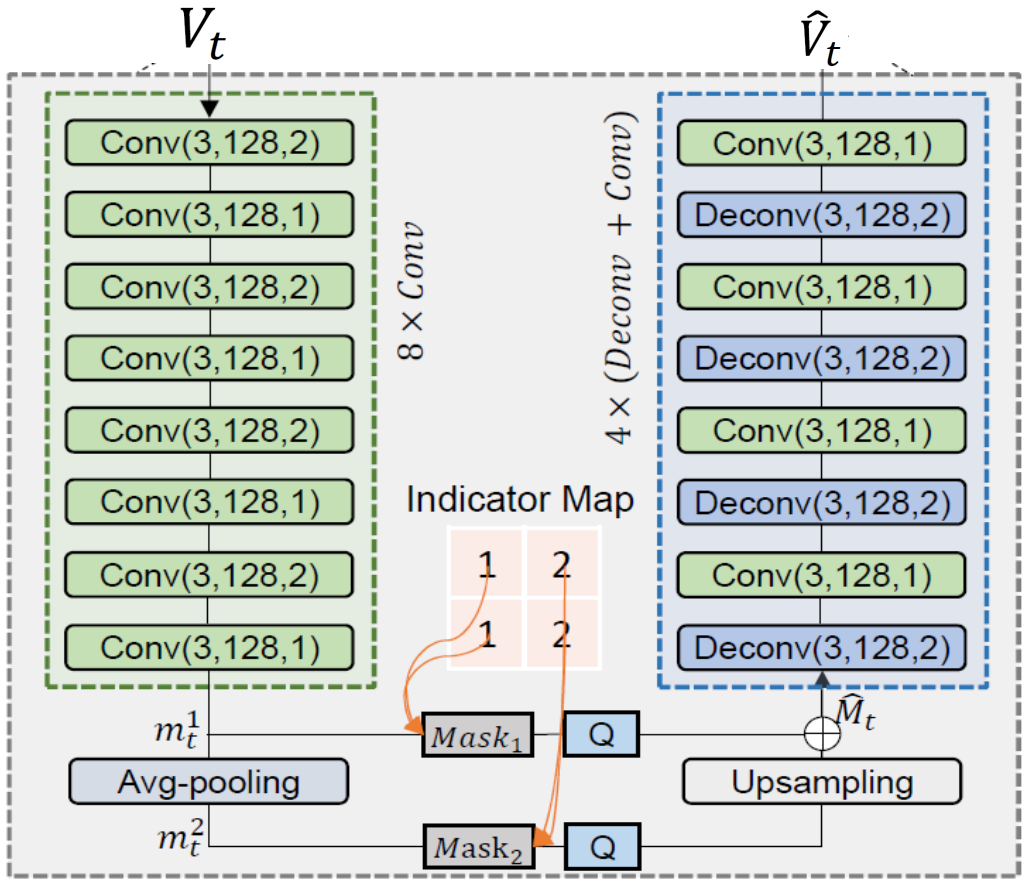



*Generate two optical flow maps with different resolutions*

[4] Zhihao Hu, et al. "Improving Deep Video Compression by Resolution-adaptive Flow Coding", in ECCV(Oral), 2020.

# End-to-End Learned P-Frame Video Compression

- Resolution-adaptive Flow Coding(Block-level)

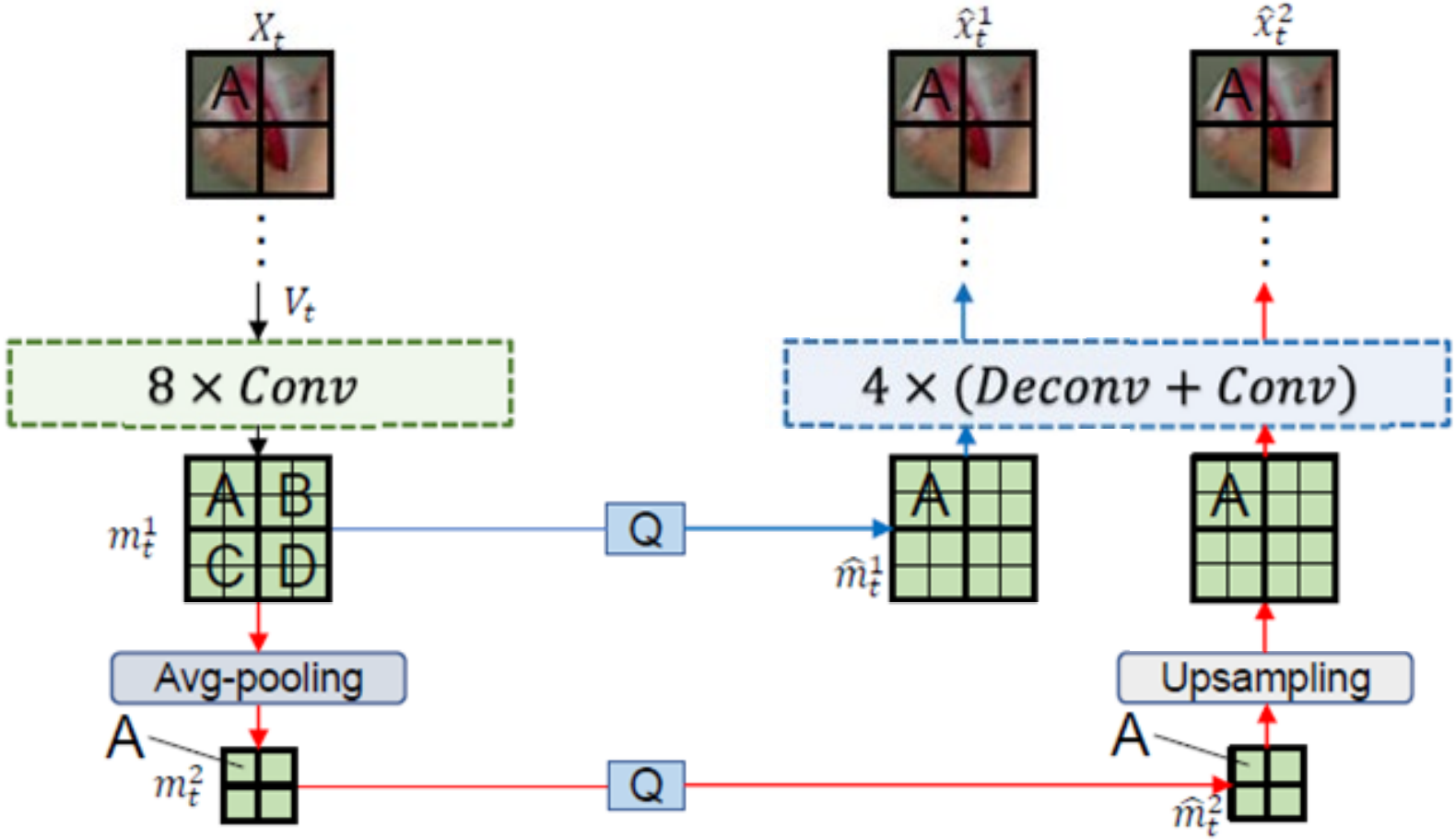



- Generate multi-scale motion features as different modes for each block
- Calculate the RD values of different modes for each block.
- Choose the optimal mode for each block
- Merge these two features

[4] Zhihao Hu, et al. "Improving Deep Video Compression by Resolution-adaptive Flow Coding", in ECCV(Oral), 2020.

# End-to-End Learned P-Frame Video Compression

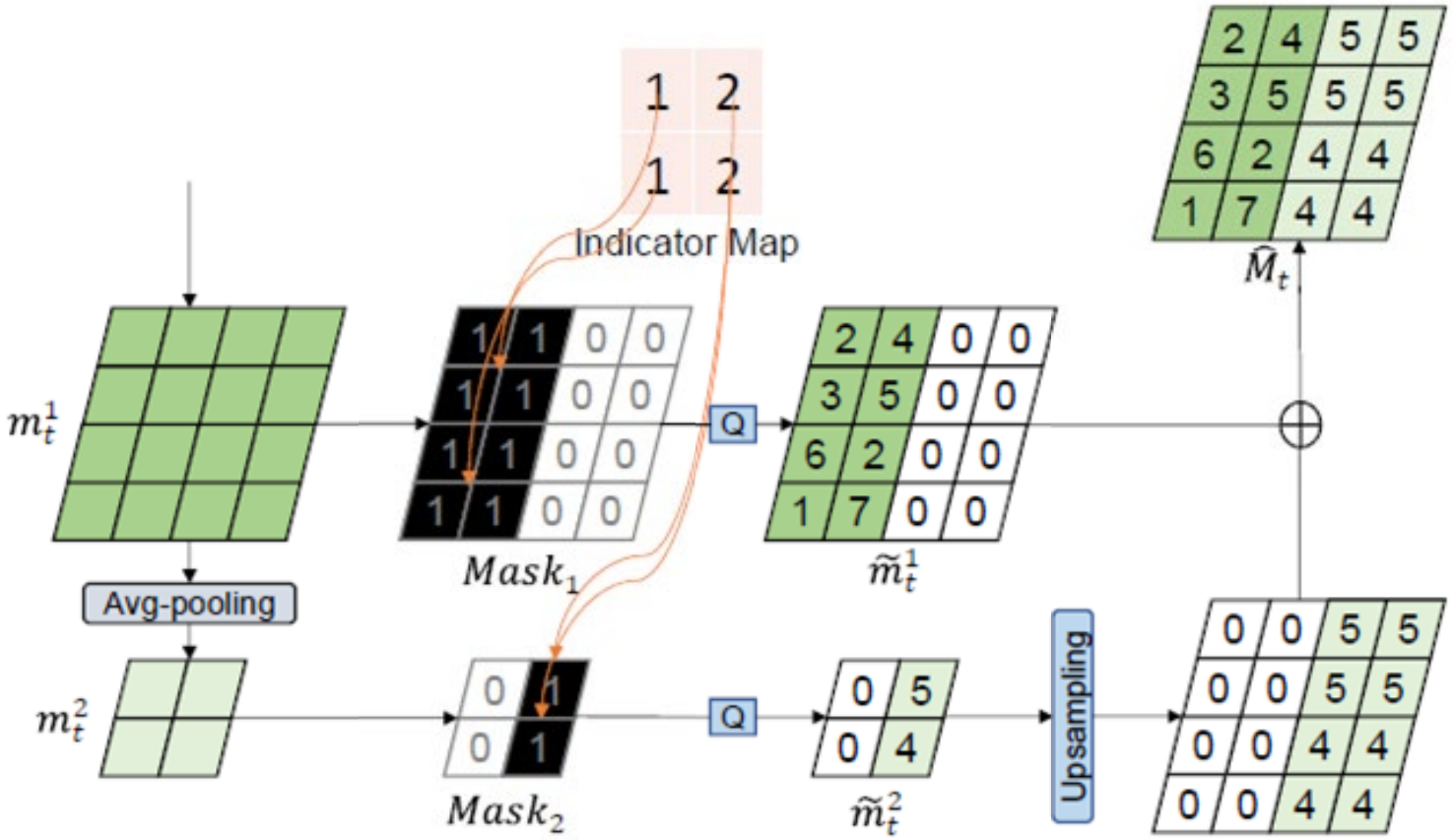
- Resolution-adaptive Flow Coding




[4] Zhihao Hu, et al. "Improving Deep Video Compression by Resolution-adaptive Flow Coding", in ECCV(Oral), 2020.

# End-to-End Learned P-Frame Video Compression

- Resolution-adaptive Flow Coding



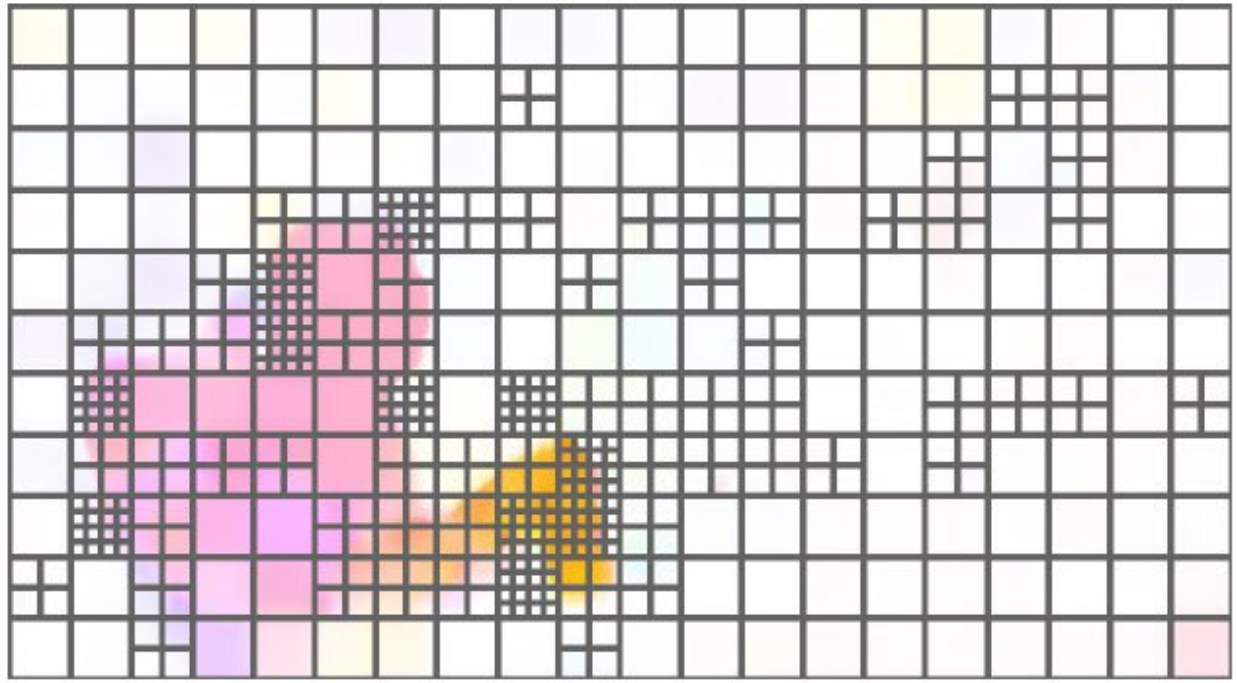

1. Select the feature representations based the indicator map

2. Merge the selected features at the decoder side

[4] Zhihao Hu, et al. "Improving Deep Video Compression by Resolution-adaptive Flow Coding", in ECCV(Oral), 2020.

# End-to-End Learned P-Frame Video Compression

- Resolution-adaptive Flow Coding



Motion features with small resolution are preferred at complex region, *e.g.,* motion boundary
Motion features with large resolution are preferred at smooth region, *e.g.*, background region

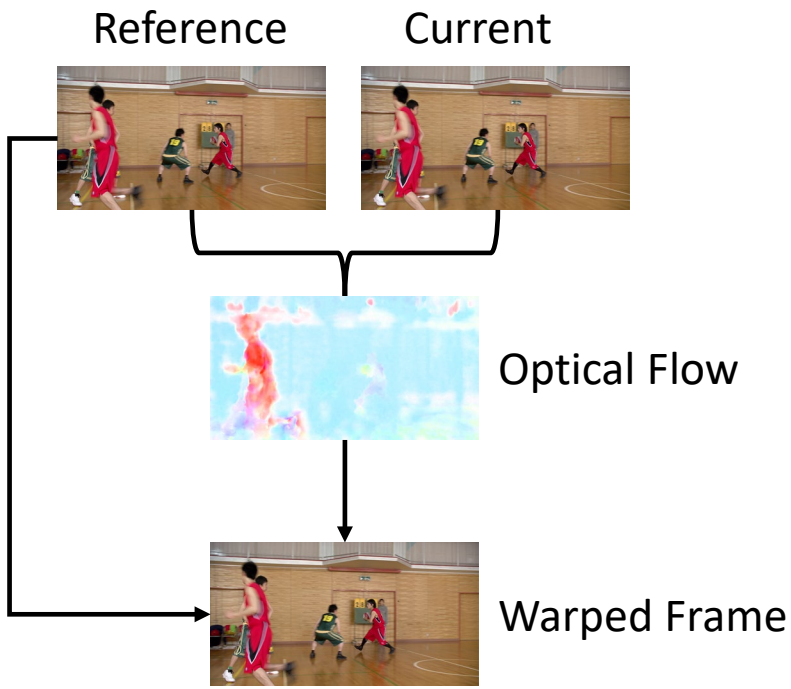
[4] Zhihao Hu, et al. "Improving Deep Video Compression by Resolution-adaptive Flow Coding", in ECCV(Oral), 2020.

# Outline

- Background for Video Compression
- End-to-end Learned P-frame Compression
  - Hybrid coding framework
  - RDO techniques
  - Enhanced motion estimation
  - Multiple reference
- End-to-end Learned B-frame Compression
- Learned Autoencoder based Video Compression
- Discussion

# End-to-End Learned P-Frame Video Compression

## Traditional Motion Compensation Procedure




## Formulations

$$\mathbf{x}' := \text{Bilinear-Warp}(\mathbf{x}, \mathbf{f}) \qquad \text{s.t.}$$
$$\mathbf{x}'[x, y] = \mathbf{x}[x + \mathbf{f}_x[x, y], y + \mathbf{f}_y[x, y]]$$

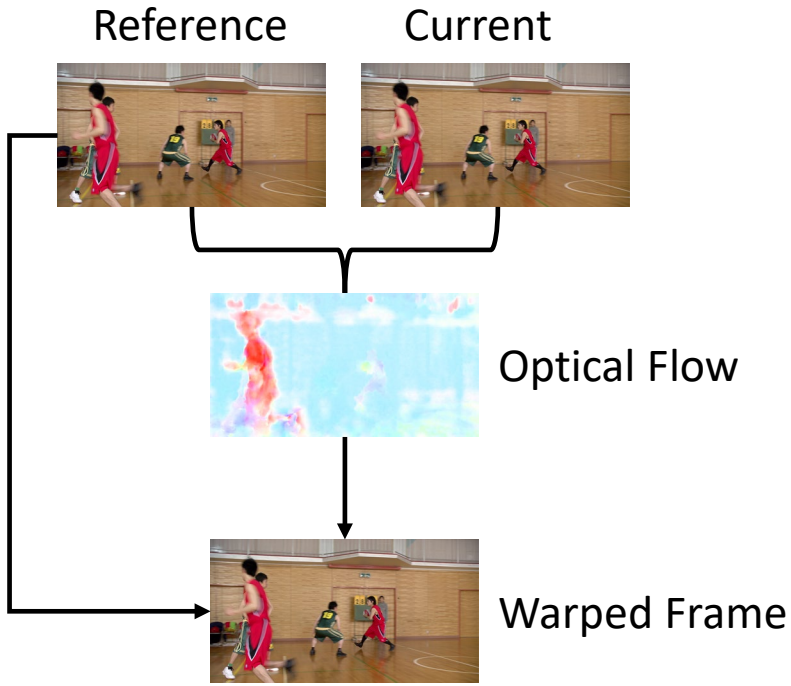
## Limitations

1. Rely on existing network architecture
2. May need pretrain
3. Large residual due to inaccurate warp operation

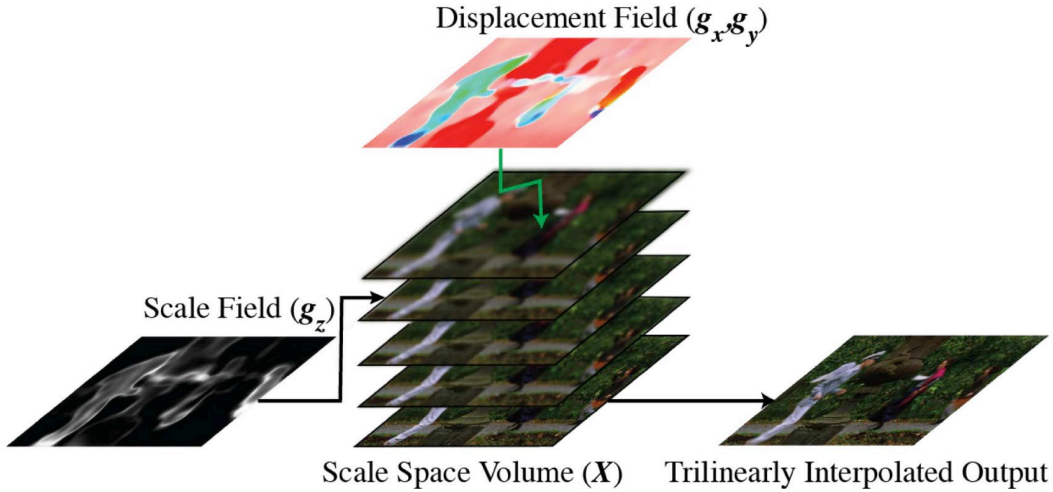
[5] Eirikur Agustsson, et al. “Scale-space Flow for End-to-End Optimized Video Compression”, in CVPR, 2020.

# End-to-End Learned P-Frame Video Compression

## Traditional Motion Compensation Procedure




## Scale-space-warp[5] motion compensation procedure




## Formulations

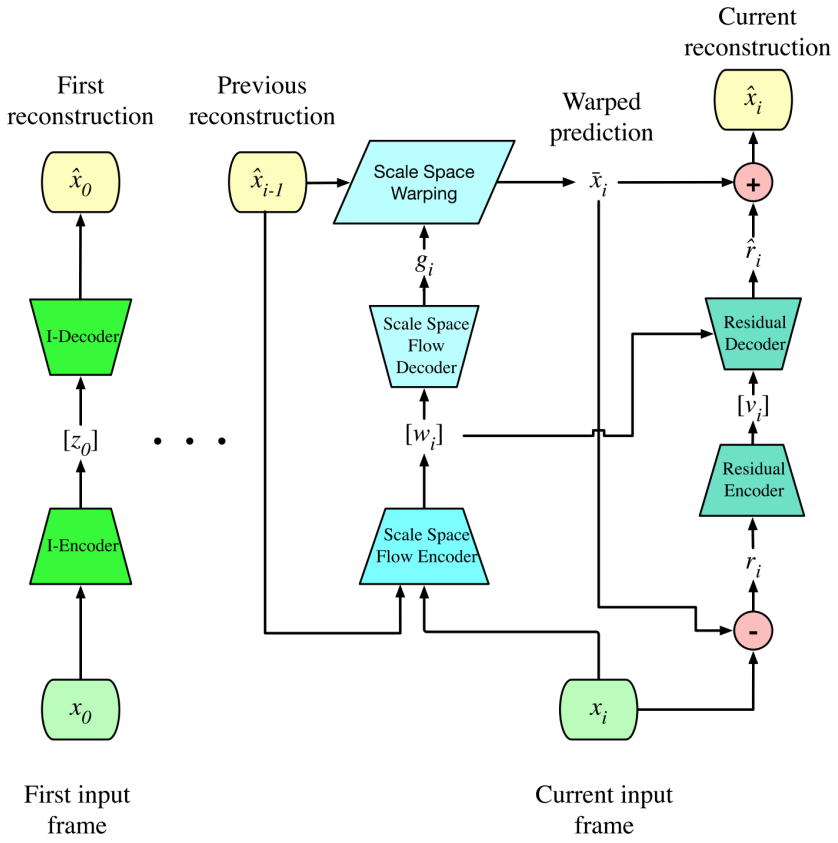
$$\mathbf{x}' := \text{Scale-Space-Warp}(\mathbf{x}, \mathbf{g}) \qquad \text{s.t.}$$

$$\mathbf{x}'[x, y] = \mathbf{X}[x + \mathbf{g}_x[x, y], y + \mathbf{g}_y[x, y], \mathbf{g}_z[x, y]]$$

[5] Eirikur Agustsson, et al. “Scale-space Flow for End-to-End Optimized Video Compression”, in CVPR, 2020.

# End-to-End Learned P-Frame Video Compression

- Overall Architecture



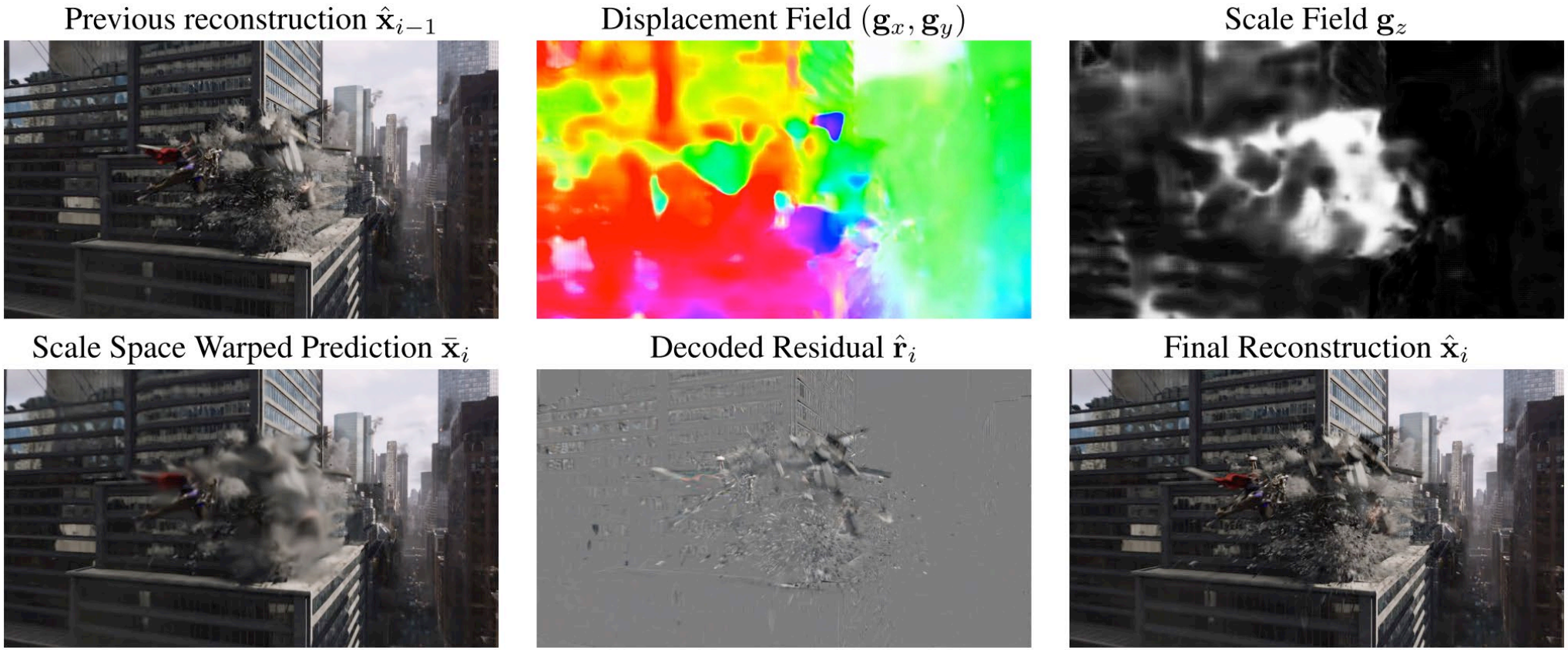

Hybrid Coding Approach:

1. Scale Space Flow Encoder & Decoder
2. Scale Space Warping based Motion Compensation
3. Residual Encoder & Decoder

[5] Eirikur Agustsson, et al. “Scale-space Flow for End-to-End Optimized Video Compression”, in CVPR, 2020.

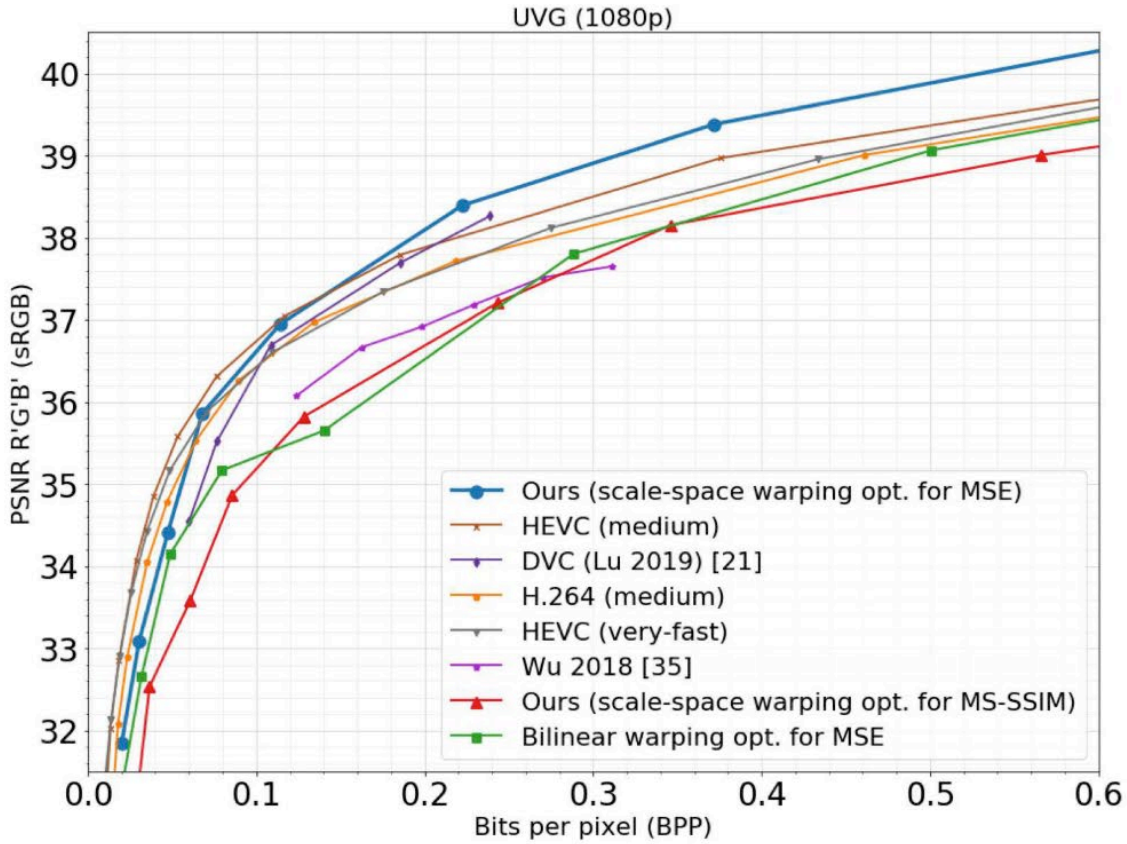
# End-to-End Learned P-Frame Video Compression

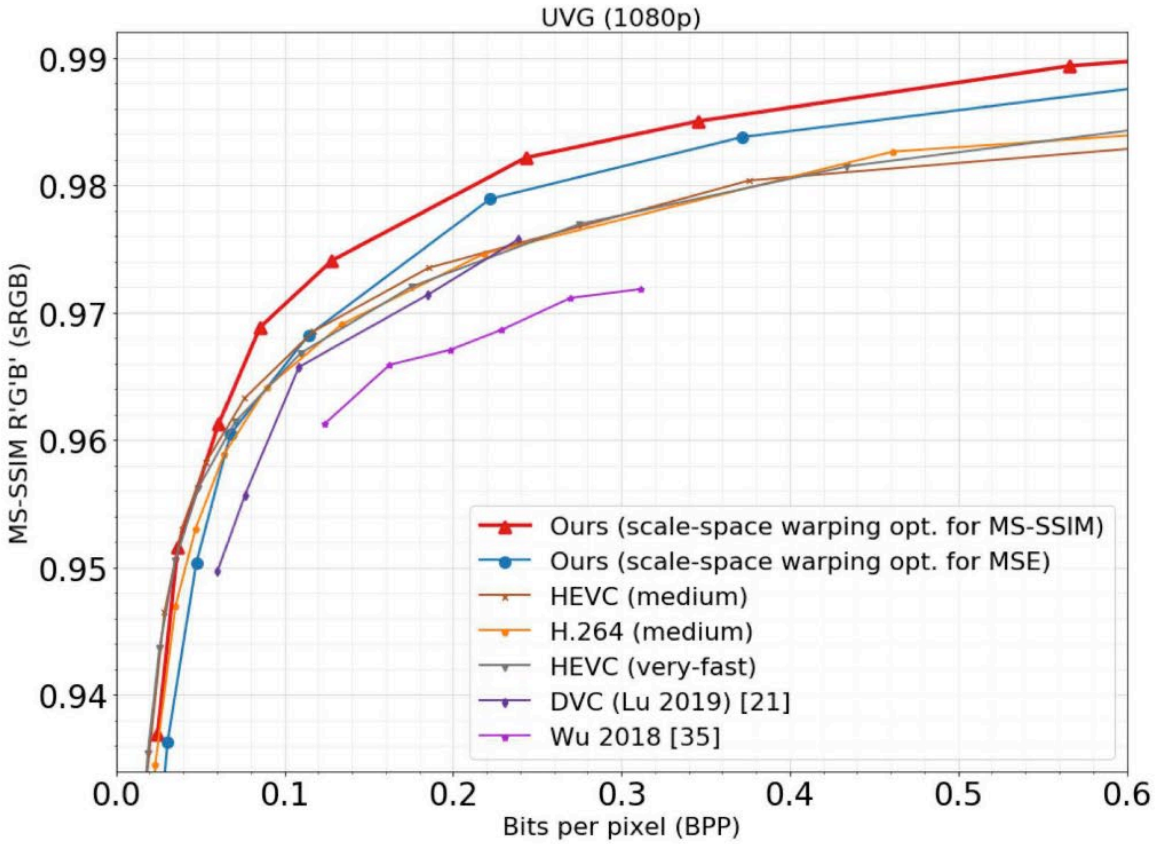
- Scale-space flow visualization




[5] Eirikur Agustsson, et al. “Scale-space Flow for End-to-End Optimized Video Compression”, in CVPR, 2020.

# End-to-End Learned P-Frame Video Compression




[5] Eirikur Agustsson, et al. “Scale-space Flow for End-to-End Optimized Video Compression”, in CVPR, 2020.

# Outline

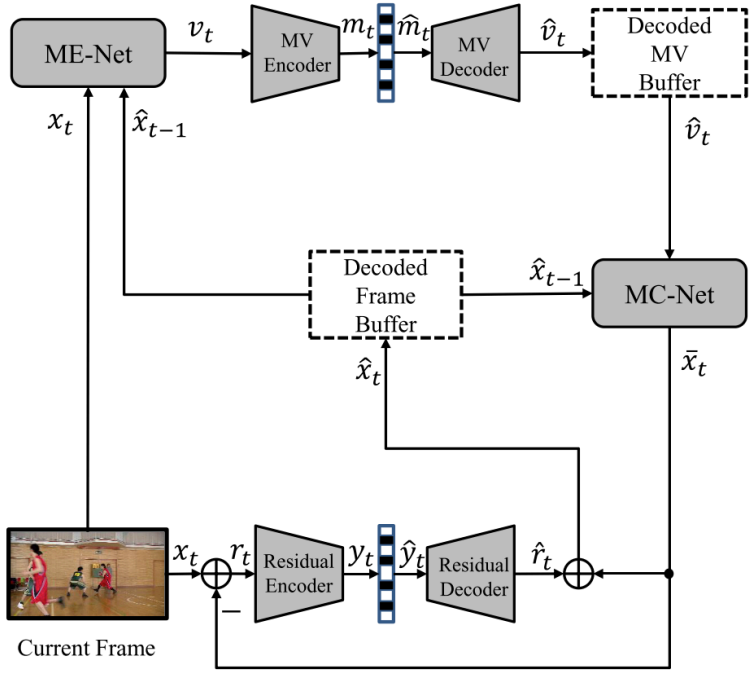
- Background for Video Compression
- End-to-end Learned P-frame Compression
  - Hybrid coding framework
  - RDO techniques
  - Enhanced motion estimation
  - Multiple reference
- End-to-end Learned B-frame Compression
- Learned Autoencoder based Video Compression
- Discussion

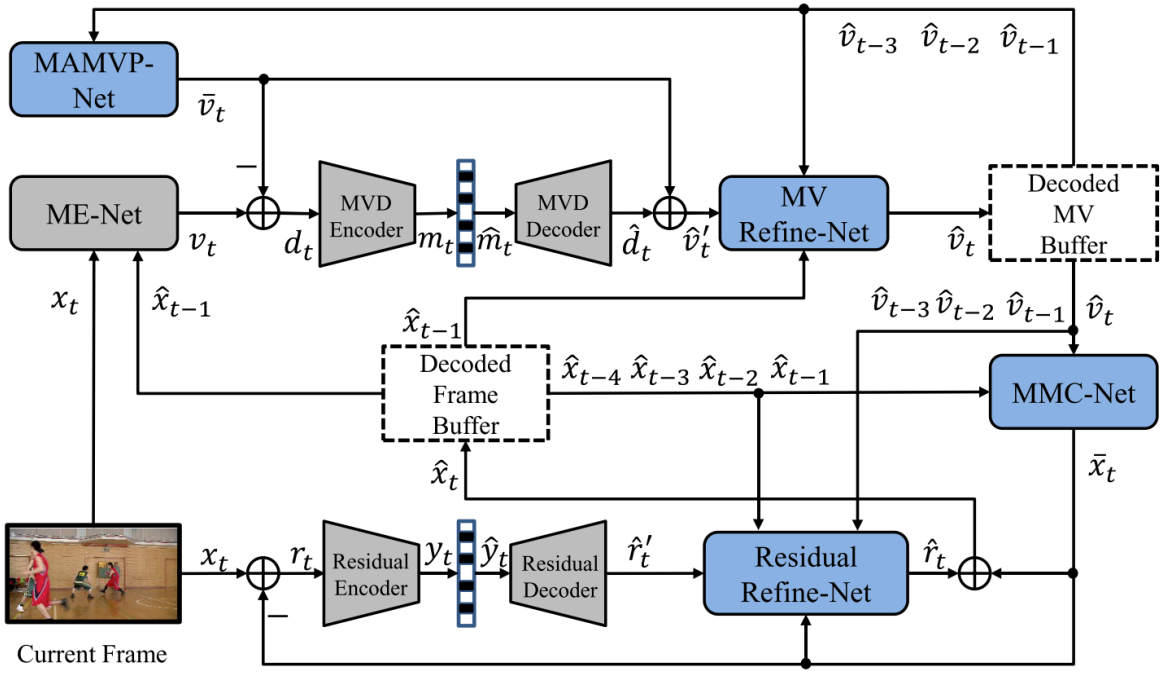
# End-to-End Learned P-Frame Video Compression

- Existing methods use one previous reference frame
- Exploiting multiple reference frames for learned video compression
  - Directly use multiple frames for motion estimation or motion compensation.
  - Explore the long-range temporal information in latent space

# End-to-End Learned P-Frame Video Compression

- Exploiting multiple reference frames for learned video compression

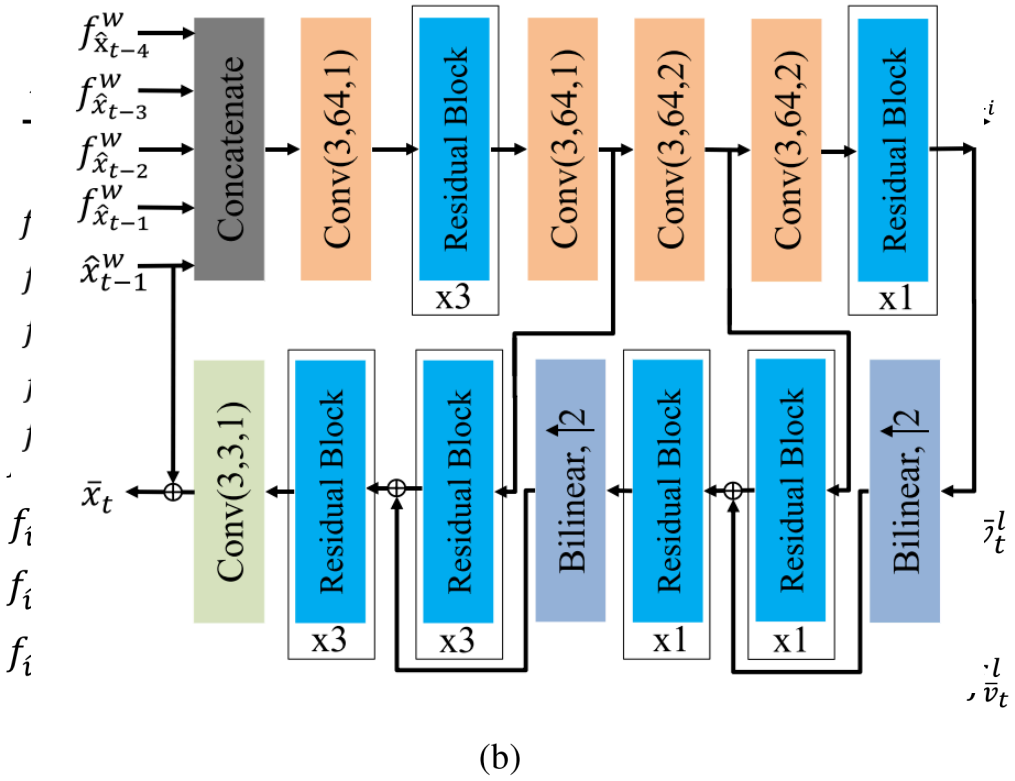


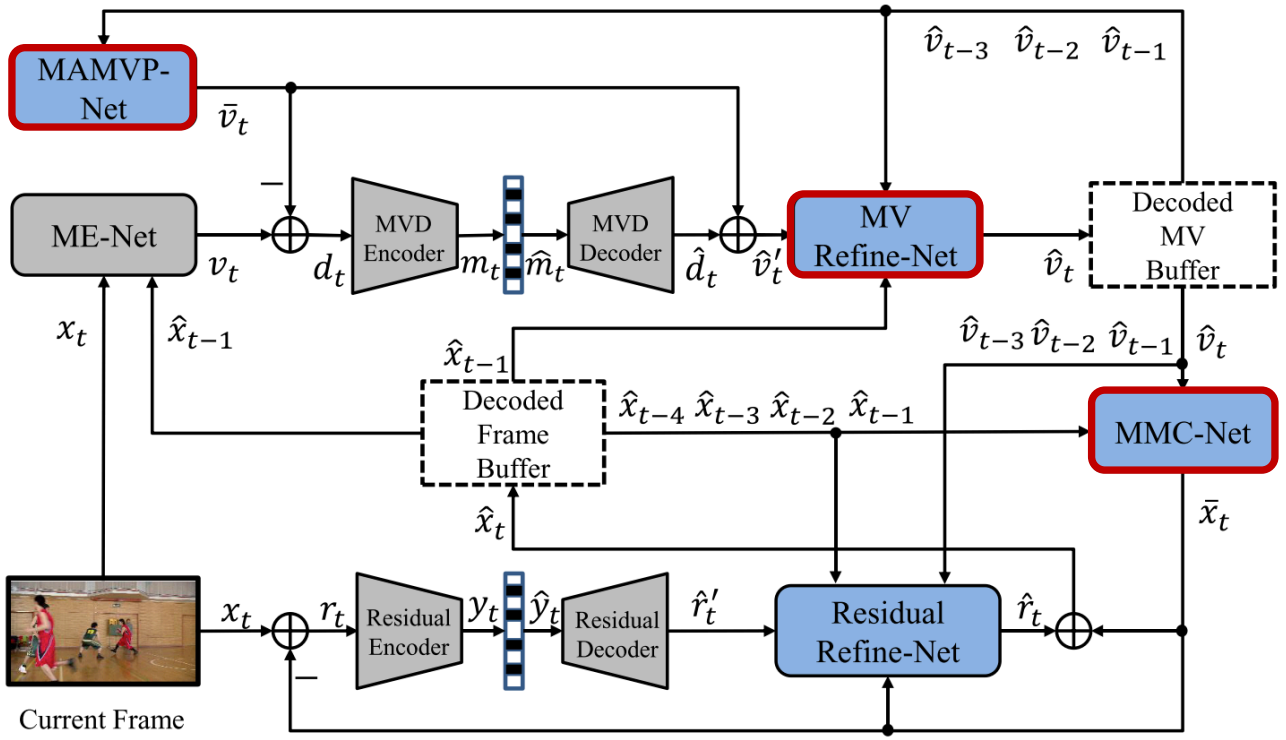

[6] Jianping Lin, et al. “M-LVC: Multiple Frames Prediction for Learned Video Compression”, in CVPR, 2020.

# End-to-End Learned P-Frame Video Compression

- M-LVC: Multiple Frames Prediction for Learned Video Compression



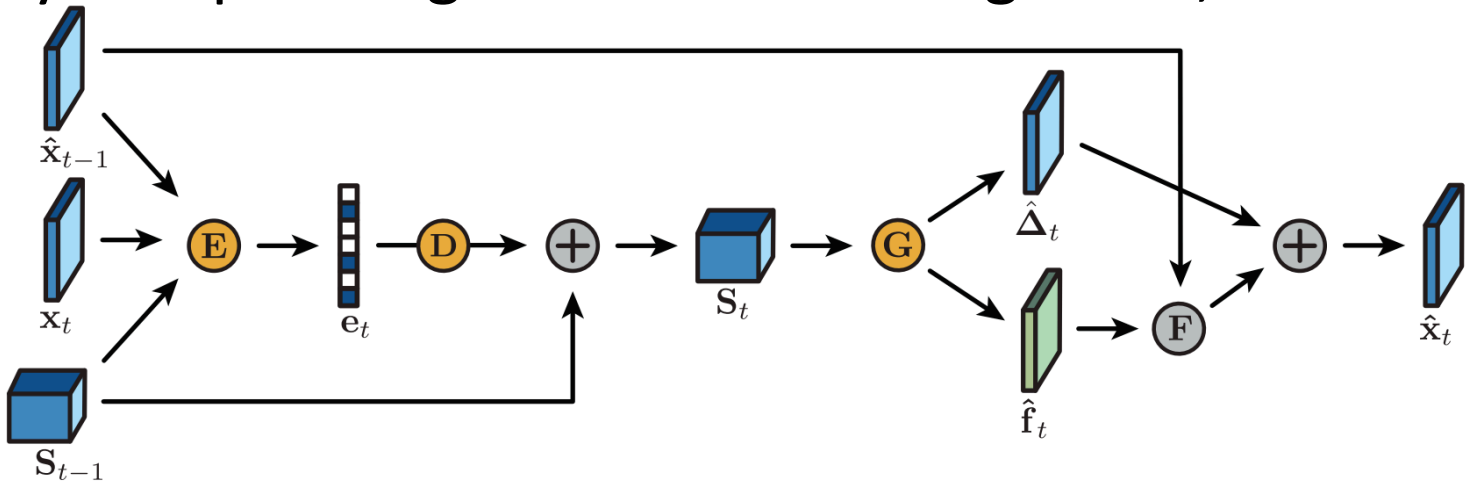

(b)




[6] Jianping Lin, et al. "M-LVC: Multiple Frames Prediction for Learned Video Compression", in CVPR, 2020.

# End-to-End Learned P-Frame Video Compression

- Maintains a state of arbitrary information learned by the model and jointly compressing all transmitted signals[7];




$S_{t-1}$ represents the state from previous time steps and includes the information from both residual and motion.

[7] Oren Rippel, et al. "Learned Video Compression", in ICCV, 2019.

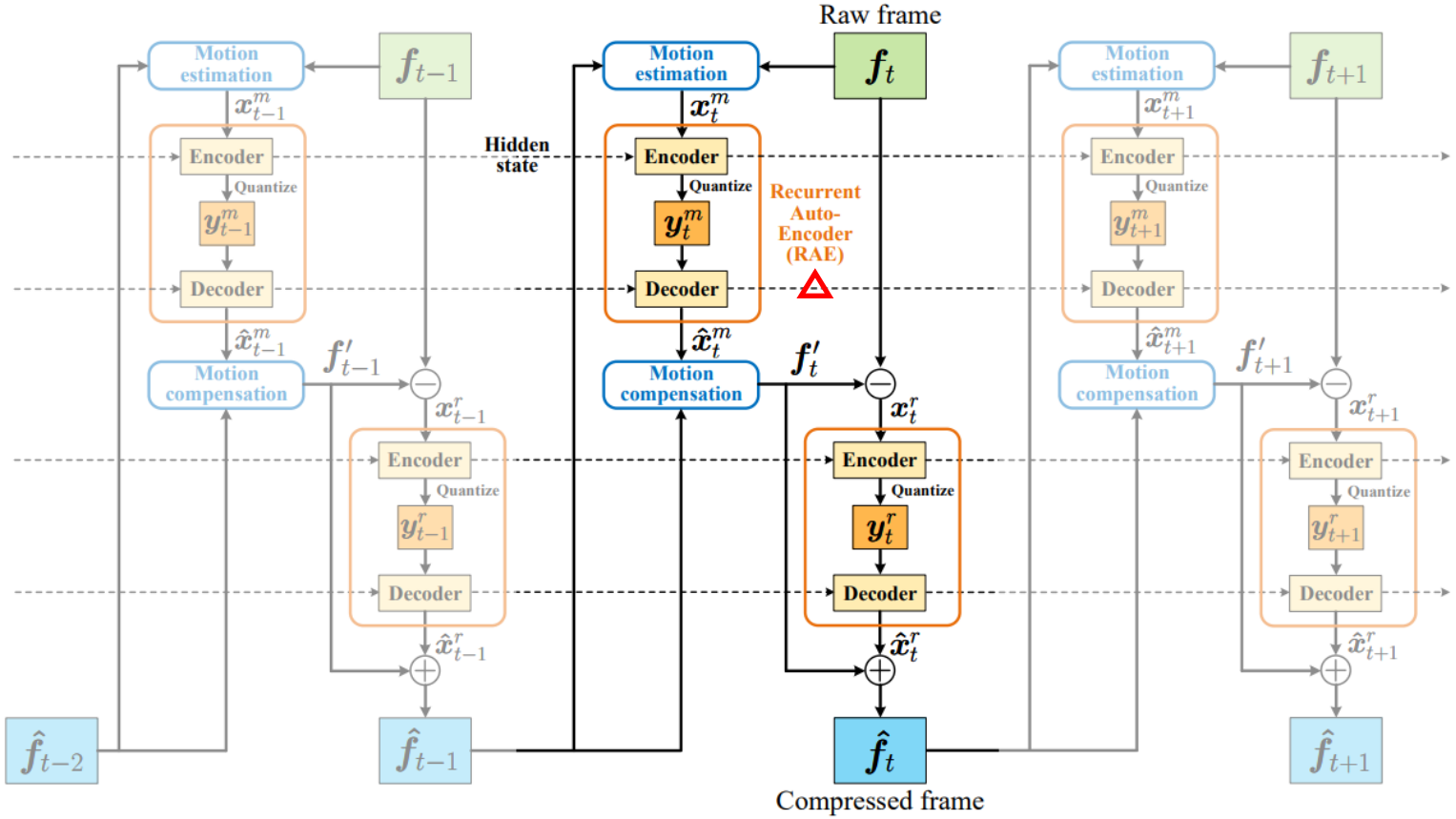
# End-to-End Learned P-Frame Video Compression

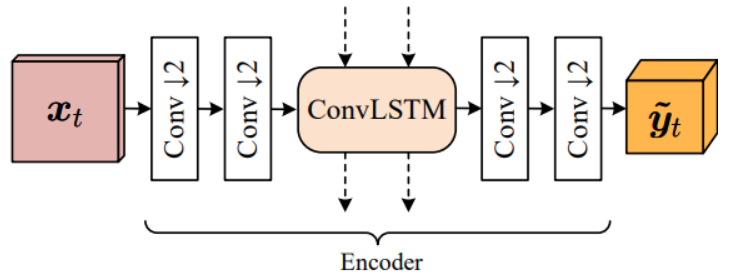
- The latent representations are generated based on limited reference frames;
- Existing work focus on the *independent* context information only;
  - motion compression and residual compression

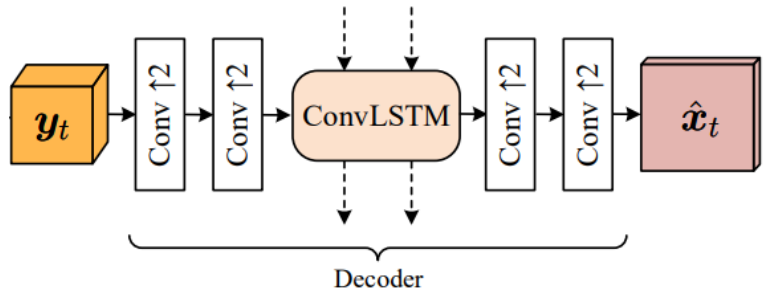
**-> exploiting the temporal redundancy to generate latent representations and more accurate context information**

# End-to-End Learned P-Frame Video Compression

- Implicitly explore temporal information in multiple frames



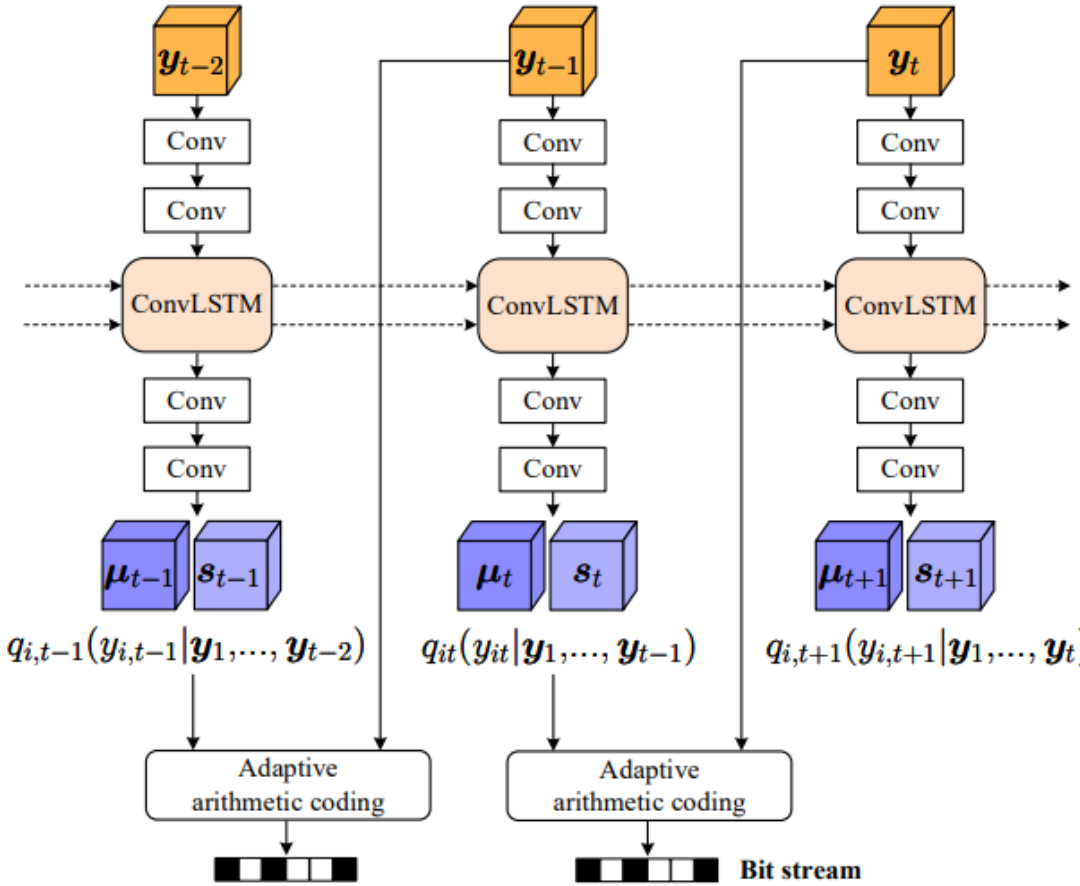




[8] Ren Yang, et al. "Learning for Video Compression with Recurrent Auto-Encoder and Recurrent Probability Model." in submission.

This slide is provided by Ren Yang.

# End-to-End Learned P-Frame Video Compression

- Implicitly explore temporal information in multiple frames




Recurrent Probability Model (RPM)

$$H(p_t, q_t) = \mathbb{E}_{\boldsymbol{y}_t \sim p_t}[-\log_2 q_t(\boldsymbol{y}_t \mid \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{t-1})]$$

$$q_t(\boldsymbol{y}_t \mid \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{t-1}) = \prod_{i=1}^{N} q_{it}(y_{it} \mid \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{t-1})$$

$$q_{it}(y_{it} \mid \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{t-1}) = \int_{y_{it}-0.5}^{y_{it}+0.5} \text{Logistic}(y; \mu_{it}, s_{it}) dy$$

[8] Ren Yang, et al. "Learning for Video Compression with Recurrent Auto-Encoder and Recurrent Probability Model." in submission.

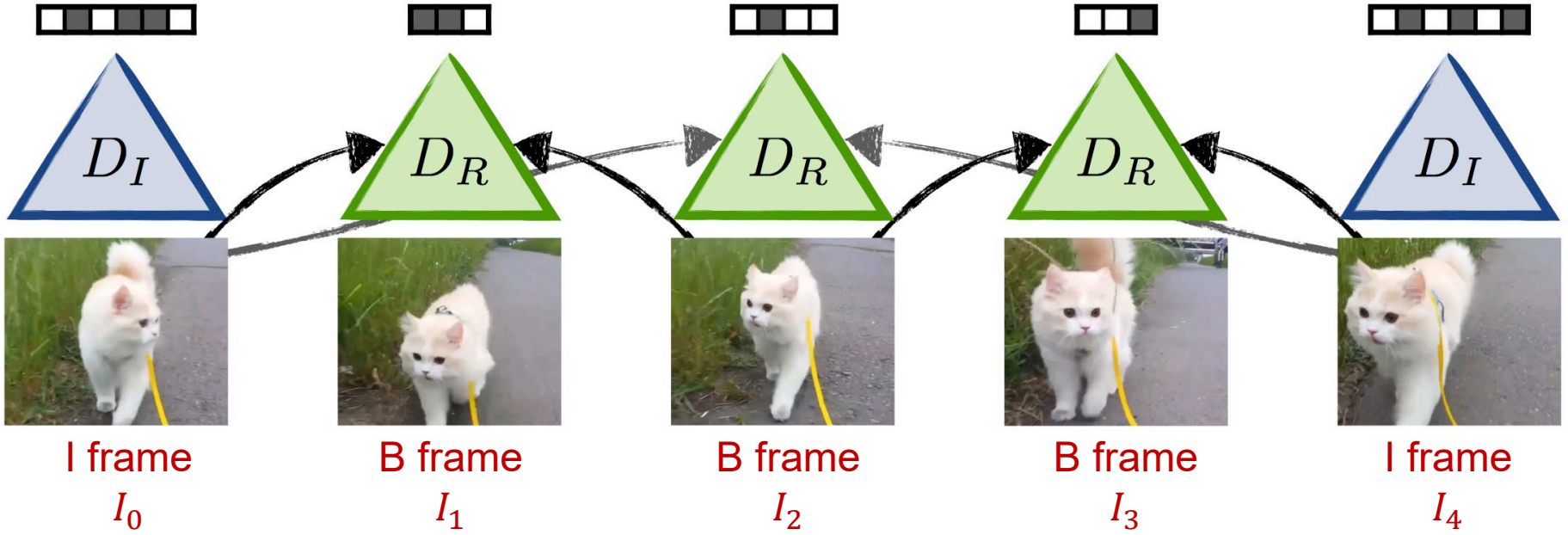
This slide is provided by Ren Yang.

# Outline

- Background for Video Compression
- End-to-end Learned P-frame Compression
  - Hybrid coding framework
  - RDO techniques
  - Enhanced motion estimation
  - Multiple reference
- End-to-end Learned B-frame Compression
- Learned Autoencoder based Video Compression
- Discussion

# End-to-End Learned B-Frame Video Compression

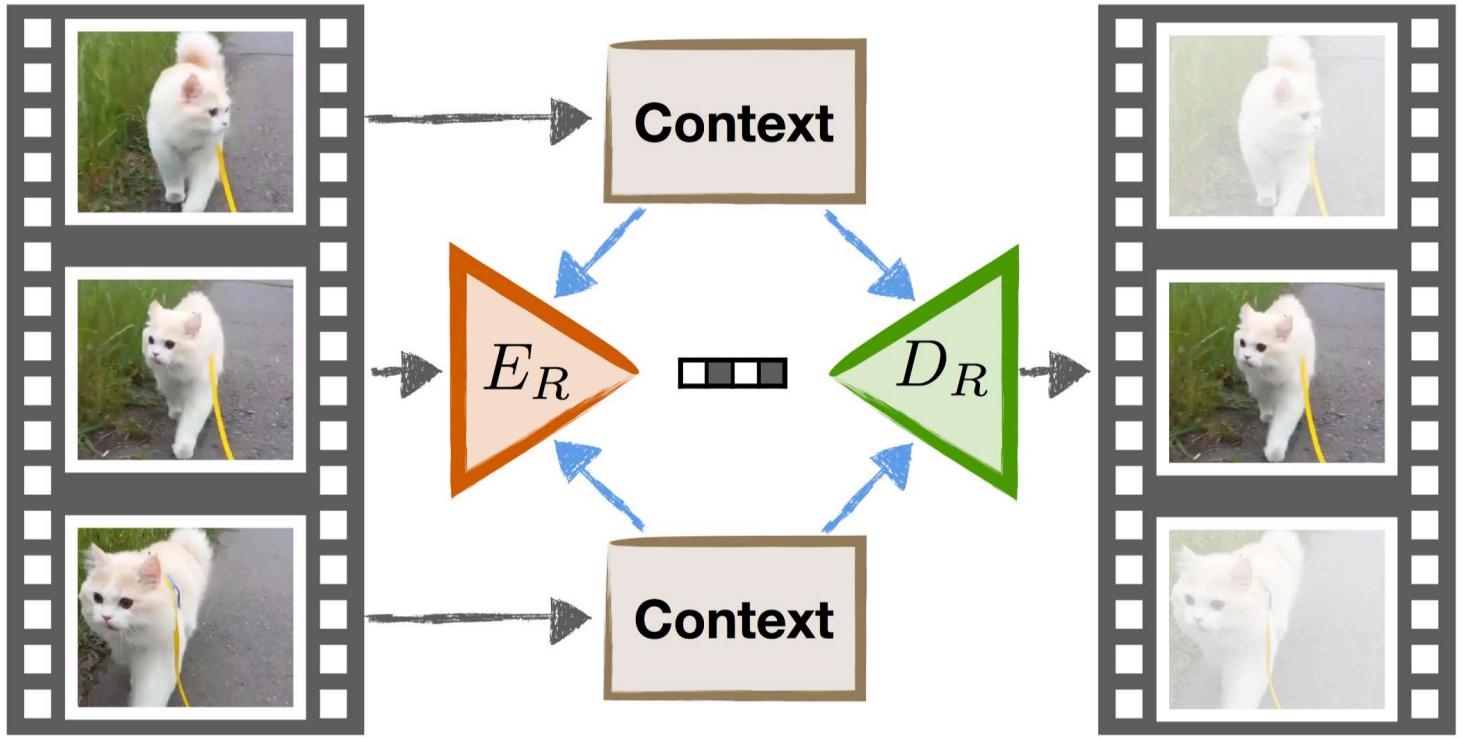
- Frame Interpolation based Video Compression




[9] Chao-yuan Wu, et al. “Video Compression through Image Interpolation,” in ECCV, 2018

# End-to-End Learned B-Frame Video Compression

- Frame Interpolation based Video Compression




1. Extract features from reference images
2. Use block based motion estimation
3. Interpolate the current frame
4. Compress residual using learned image codec
5. Compress motion using traditional image codec

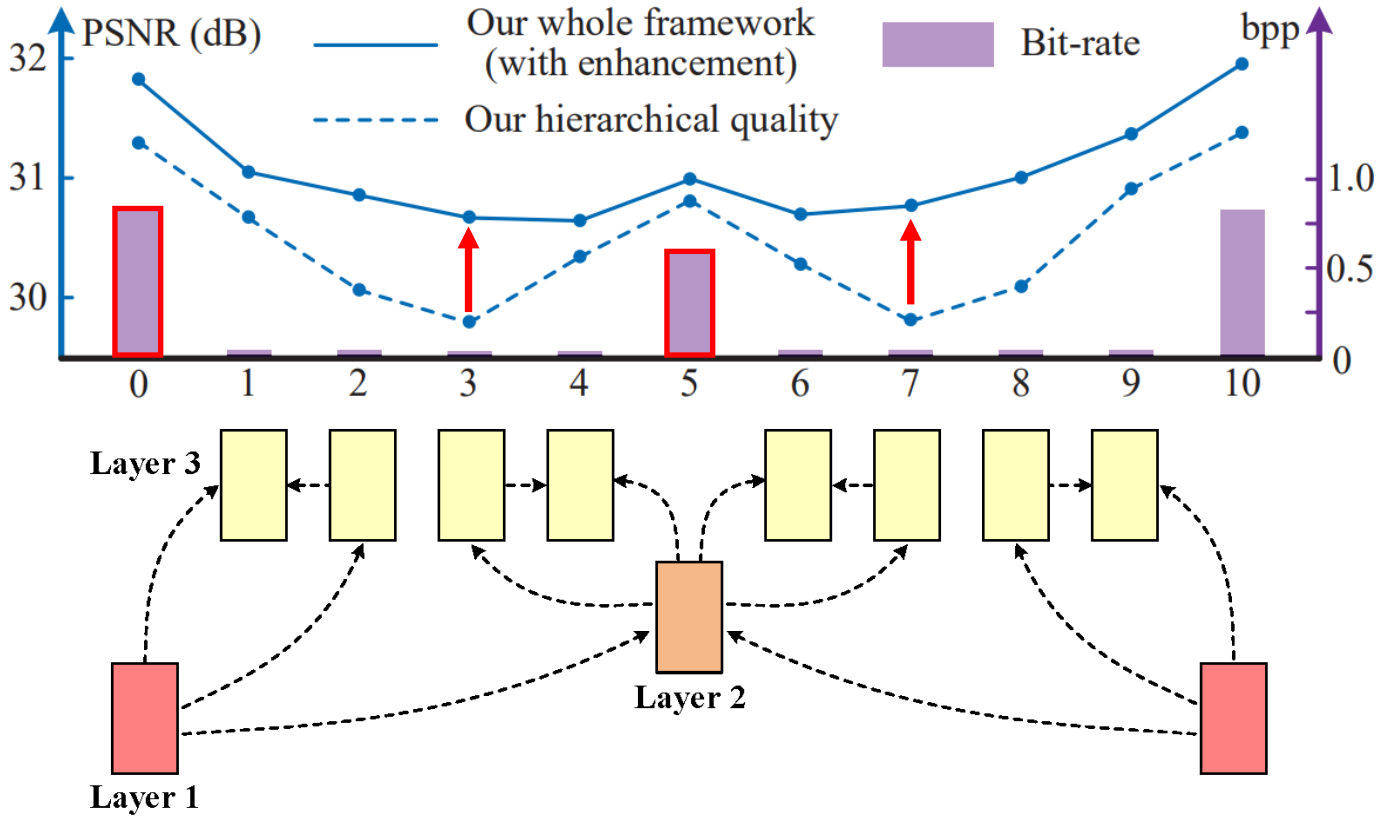
**Limitations:**

1. Not end-to-end optimized
2. Motion compression is not learnt

[9] Chao-yuan Wu, et al. "Video Compression through Image Interpolation," in ECCV, 2018

# End-to-End Learned B-Frame Video Compression

- Hierarchical Learned Video Compression (HLVC) with recurrent enhancement



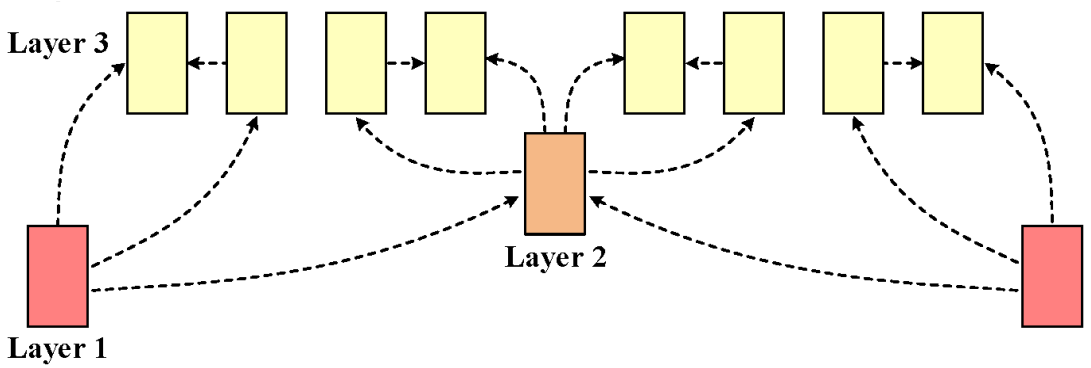

- **At encoder side**, the high quality frames provide high quality references to improve the compression performance of other frames.
- **At decoder side**, the low quality frames can be enhanced by taking advantage of high quality frames without bit-rate overhead. It is equivalent to reducing bit-rate on low quality frames.

[10] Ren Yang, et al. "Learning for Video Compression with Hierarchical Quality and Recurrent Enhancement." in CVPR. 2020.

This slides is provided by Ren Yang

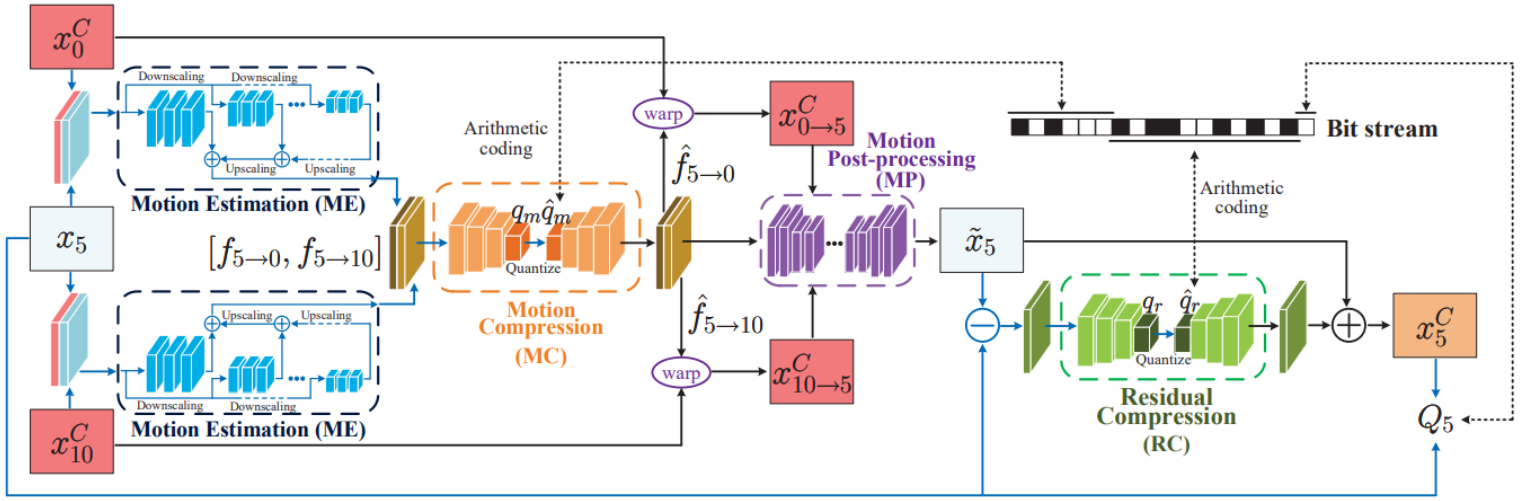
# End-to-End Learned B-Frame Video Compression

- Hierarchical Learned Video Compression (HLVC) with recurrent enhancement




**Layer 1:**
Compressed by BPG for PSNR model, and by Lee *et al.* ICLR 2019 for MS-SSIM model.

**Layer 2:**
Compressed by the proposed Bi-Directional Deep Compression (BDDC) network



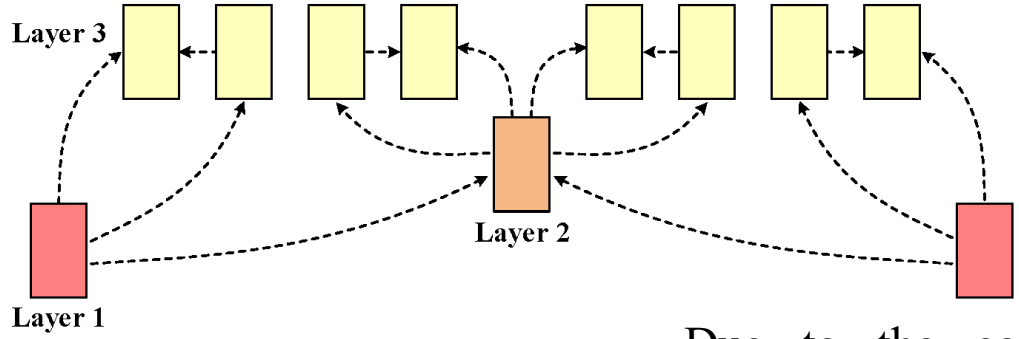

[10] Ren Yang, et al. "Learning for Video Compression with Hierarchical Quality and Recurrent Enhancement." in CVPR. 2020.

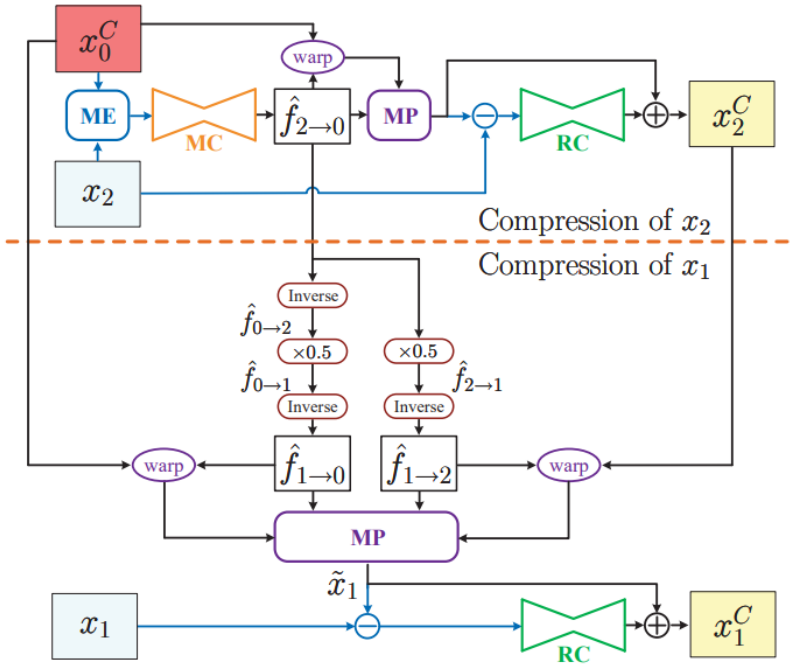
This slides is provided by Ren Yang

# End-to-End Learned B-Frame Video Compression

- Hierarchical Learned Video Compression (HLVC) with recurrent enhancement

**Layer 3:**
Compressed by the proposed Single Motion Deep Compression (SMDC) network



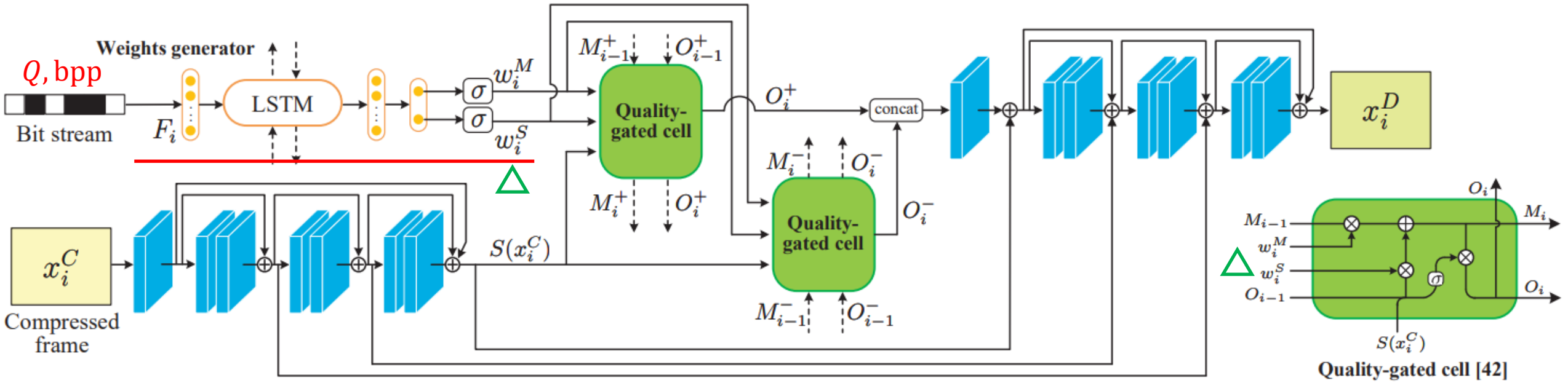

Due to the correlation of motions among multiple neighboring frames, we propose using the motion between $x_0^C$ and $x_2$ to predict the motions between $x_1$ and $x_0^C$ or $x_2$. That is,

$$\hat{f}_{1\to0} = \text{Inverse}(0.5 \times \underbrace{\text{Inverse}(\hat{f}_{2\to0})}_{\hat{f}_{0\to2}}).$$
$$\hat{f}_{0\to1}$$

As such, $x_1$ can be compressed with the reference frames of $x_0^C$ and $x_2$, **without bits consumed for motion map**, thus improving the rate-distortion performance.

# End-to-End Learned B-Frame Video Compression

- Hierarchical Learned Video Compression (HLVC) with recurrent enhancement




[10] Ren Yang, et al. "Learning for Video Compression with Hierarchical Quality and Recurrent Enhancement." in CVPR. 2020.

This slides is provided by Ren Yang

# End-to-End Learned B-Frame Video Compression

- Previous works use separate interpolation network and motion compression module

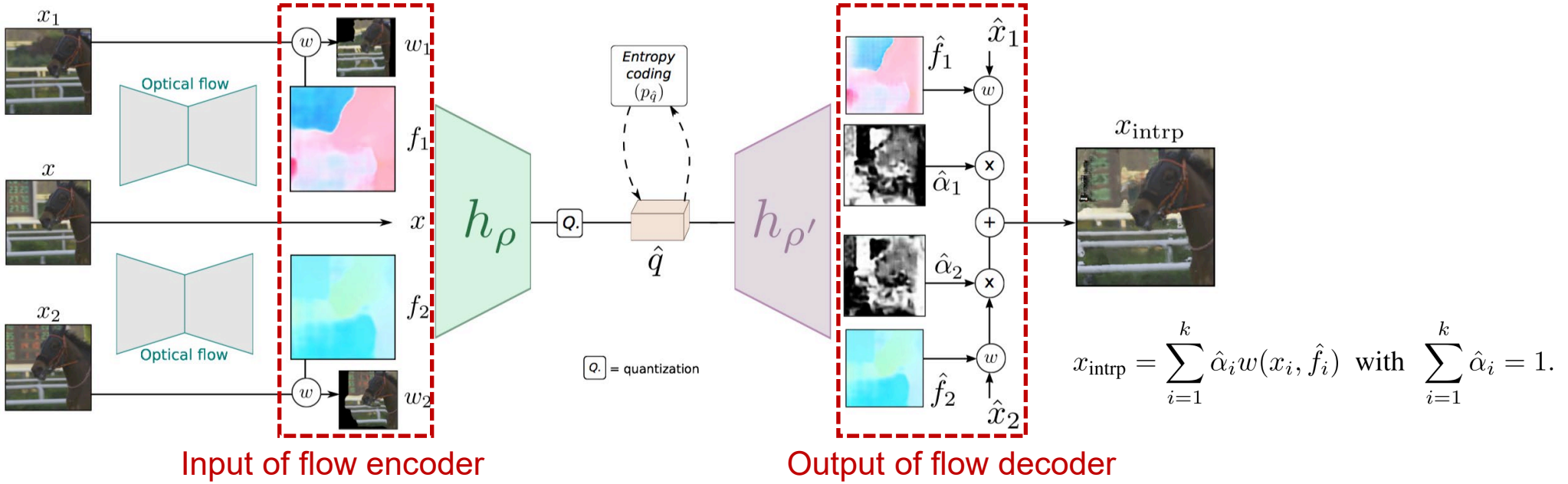
**-> combine interpolation network and motion compression**

- The residual is compressed in the pixel domain and it is a non-trivial task.

**-> Feature space residual compression**

# End-to-End Learned B-Frame Video Compression

- Combine interpolation and flow compression and decode the flow and interpolation coefficients simultaneously.

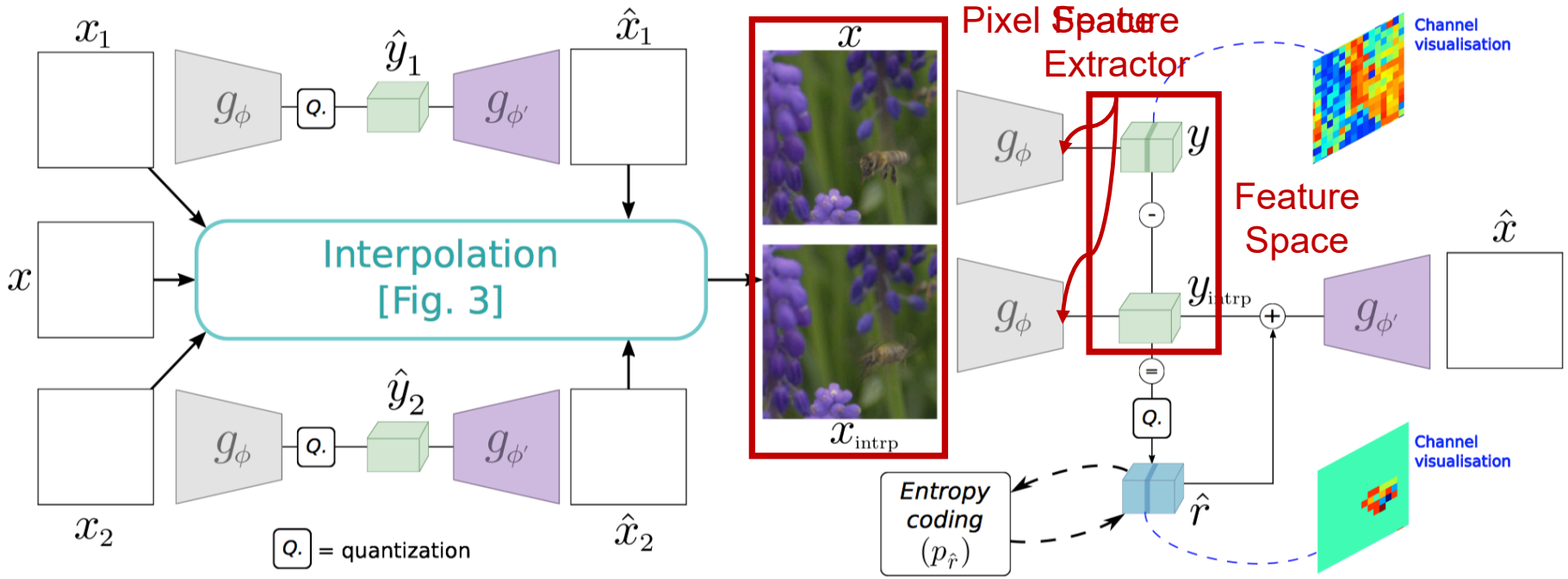



$$x_{\text{intrp}} = \sum_{i=1}^{k} \hat{\alpha}_i w(x_i, \hat{f}_i) \quad \text{with} \quad \sum_{i=1}^{k} \hat{\alpha}_i = 1.$$

[11] Abdelaziz Djelouah, et al. "Neural Inter-Frame Compression for Video Coding." in ICCV. 2019.

# End-to-End Learned B-Frame Video Compression

- Residual Compression in Latent Space




[11] Abdelaziz Djelouah, et al. "Neural Inter-Frame Compression for Video Coding." in ICCV. 2019.

# Outline

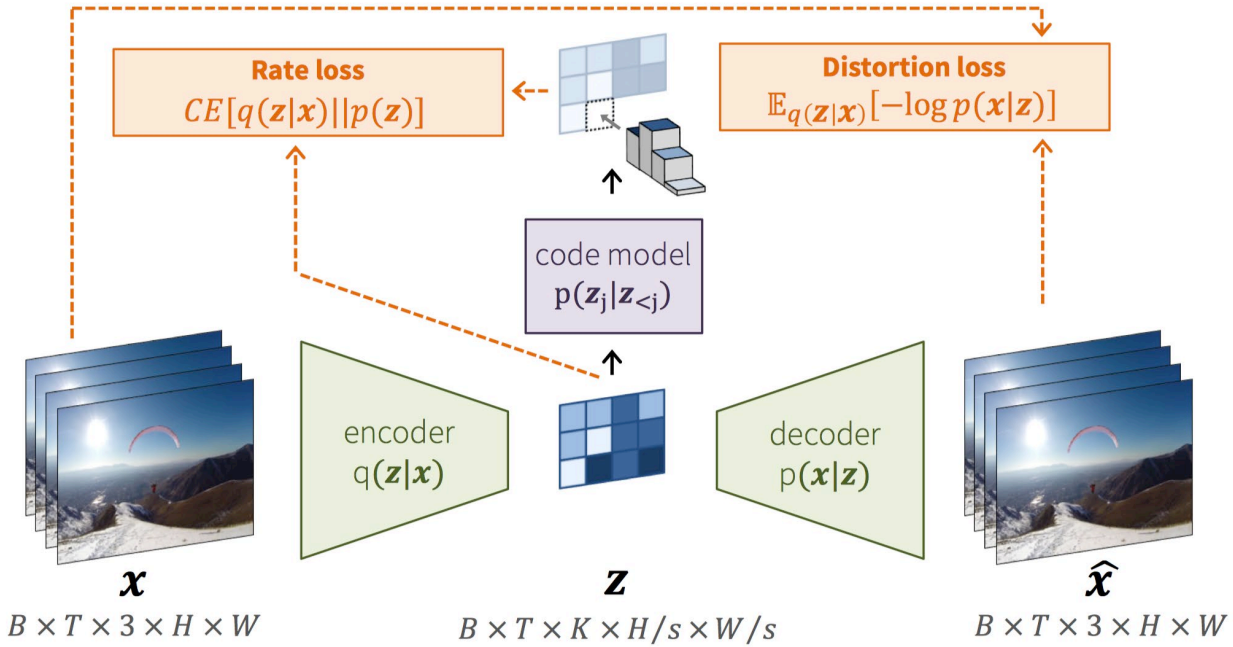
- Background for Video Compression
- End-to-end Learned P-frame Compression
  - Hybrid coding framework
  - RDO techniques
  - Enhanced motion estimation
  - Multiple reference
- End-to-end Learned B-frame Compression
- Learned Autoencoder based Video Compression
- Discussion

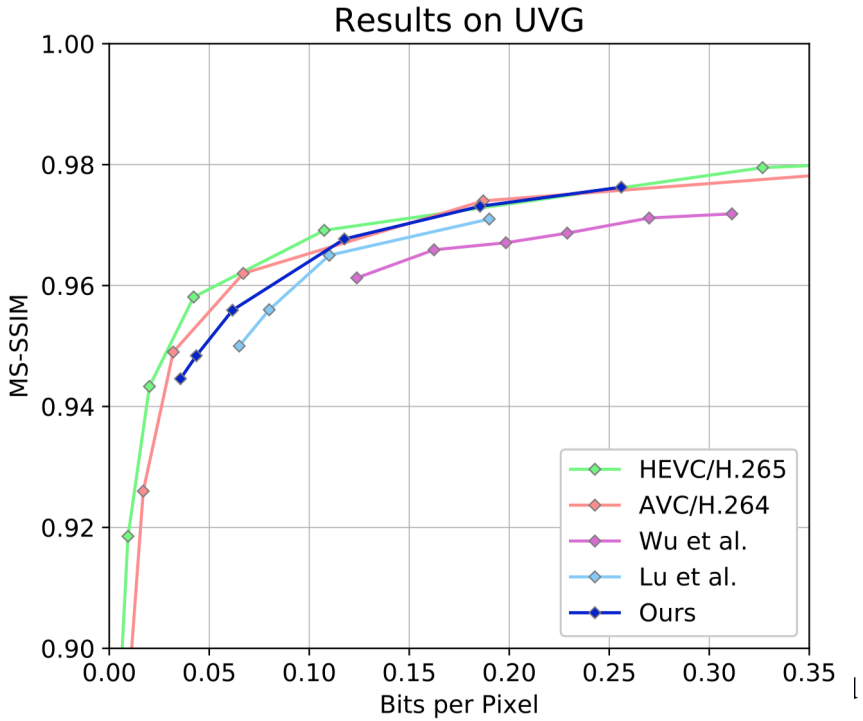
# Learned Autoencoder based Video Compression

- Previous works follow the hybrid coding framework, i.e., motion compensation and residual coding.
- Using optical flow for explicitly motion estimation
- Compressing motion and residual separately

# Learned Autoencoder based Video Compression

- Use 3D autoencoders to compress video frames without explicitly motion estimation.




[12] Amirhossein Habibian, et al. "Video compression with rate-distortion autoencoders ." in ICCV. 2019.

# Outline

- Background for Video Compression
- End-to-end Learned P-frame Compression
  - Hybrid coding framework
  - RDO techniques
  - Enhanced motion estimation
  - Multiple reference
- End-to-end Learned B-frame Compression
- Learned Autoencoder based Video Compression
- Discussion

# Discussion

- Benchmark Results

BDBR Results: Red and Blue represent the best and second best results when compared with x265(placebo)

| Dataset | DVC[1] | DVC++[2] | Djelouah[11] | HLVC[10] | Agustsson[5] | RLVC[13] | Lu[3] | Hu[4] | FVC[12] |
|---|---|---|---|---|---|---|---|---|---|
| Class A | | | | | | | | | |
| Class B | 32.91% | 4.84% | | 15.81% | | -2.32% | 7.45% | 9.81% | -2.35% |
| Class C | 57.54% | 28.60% | | 44.34% | | 26.69% | 24.15% | 33.40% | 11.17% |
| Class D | 52.72% | 23.26% | | 18.18% | | -0.32% | 21.50% | 31.82% | 8.47% |
| Class E | 73.61% | 38.66% | | | | | 49.35% | 33.32% | 38.62% |
| UVG | 28.46% | | 37.78% | | 10.40% | 4.99% | 20.93% | -0.55% | -9.86% |
| MCL | 31.79% | 1.87% | -9.68% | | 12.15% | | 15.72% | 4.82% | -10.90% |
| VTL | 47.11% | | 51.31% | | | | 36.49% | 31.88% | 15.14% |

# Discussion

- Open-Source Project
  - Pytorch Data Compression
    - Learned Image Compression
      - Ballé, ICLR2017
      - Ballé, ICLR2018
      - Minnen, NeurIPS2018
    - Learned Video Compression
      - DVC, CVPR2019
      - Hu[4], ECCV2020(ongoing)
    - Learned Point Cloud Compression
      - OctSqueeze, CVPR2020

https://github.com/ZhihaoHu/PyTorchDataCompression/

# Reference


- [1] Guo Lu, Wanli Ouyang, Dong Xu, Xiaoyun Zhang, Chunlei Cai, and Zhiyong Gao. “DVC: An End-to-end Deep Video Compression Framework.” In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 11006-11015. 2019. **Oral**
- [2] Guo Lu, Xiaoyun Zhang, Wanli Ouyang, Li Chen, Zhiyong Gao, and Dong Xu.“An End-to-End Learning Framework for Video Compression.” IEEE Transactions on Pattern Analysis and Machine Intelligence (T- PAMI), 2020.
- [3] Guo Lu*, Chunlei Cai*, Xiaoyun Zhang, Li Chen, Wanli Ouyang, Dong Xu, Zhiyong Gao. “Content Adaptive and Error Propagation Aware Deep Video Compression. In Proceedings of the European Conference on Computer Vision (ECCV), 2020. **Oral**
- [4] Zhihao Hu, Zhenghao Chen, Dong Xu, Guo Lu, Wanli Ouyang, Shuhang Gu. “Improving Deep Video Compression by Resolution-adaptive Flow Coding.” In Proceedings of the European Conference on Computer Vision (ECCV),2020. **Oral**
- [5] Eirikur Agustsson, David Minnen, Nick Johnston, Johannes Balle, Sung Jin Hwang, and George Toderici. "Scale-Space Flow for End-to-End Optimized Video Compression." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8503-8512. 2020.
- [6] Jianping Lin, Dong Liu, Houqiang Li, and Feng Wu. "M-LVC: Multiple Frames Prediction for Learned Video Compression." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3546-3554. 2020.
- [7] Oren Rippel, Sanjay Nair, Carissa Lew, Steve Branson, Alexander G. Anderson, and Lubomir Bourdev. "Learned video compression." In Proceedings of the IEEE International Conference on Computer Vision, pp. 3454-3463. 2019.
- [8] Chao-Yuan Wu, Nayan Singhal, and Philipp Krahenbuhl. "Video compression through image interpolation." In Proceedings of the European Conference on Computer Vision (ECCV), pp. 416-431. 2018.
- [9] Yang, Ren, Fabian Mentzer, Luc Van Gool, and Radu Timofte. "Learning for video compression with hierarchical quality and recurrent enhancement." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 6628-6637. 2020.
- [10] Abdelaziz Djelouah, Joaquim Campos, Simone Schaub-Meyer, and Christopher Schroers. "Neural inter-frame compression for video coding." In Proceedings of the IEEE International Conference on Computer Vision, pp. 6421-6429. 2019.
- [11] Amirhossein Habibian, Ties van Rozendaal, Jakub M. Tomczak, and Taco S. Cohen. "Video compression with rate-distortion autoencoders." In Proceedings of the IEEE International Conference on Computer Vision, pp. 7033-7042. 2019.
- [12] Zhiaho Hu, Guo Lu, Dong Xu. “FVC: A New Framework towards Deep Video Compression in Feature Space ." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2021.
- [13] Ren Yang, Fabian Mentzer, Luc Van Gool, and Radu Timofte. "Learning for Video Compression with Recurrent Auto-Encoder and Recurrent Probability Model." arXiv preprint arXiv:2006.13560 (2020).

# Q&A