Adaptive Multi-View Live Video Streaming for Teledriving Using a Single Hardware Encoder

Markus Hofbauer

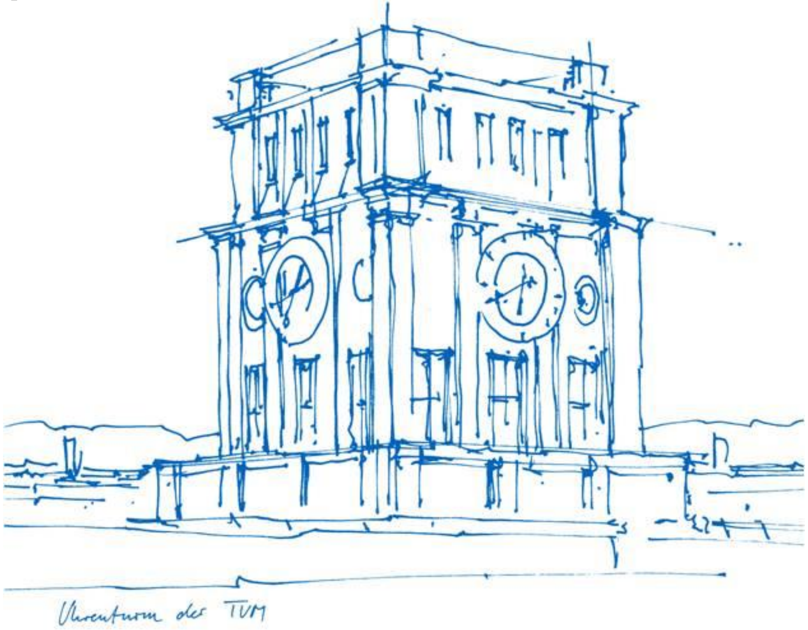
Technical University of Munich

Department for Electrical and Computer Engineering

Chair of Media Technology

Berlin, July 25 – 27, 2021

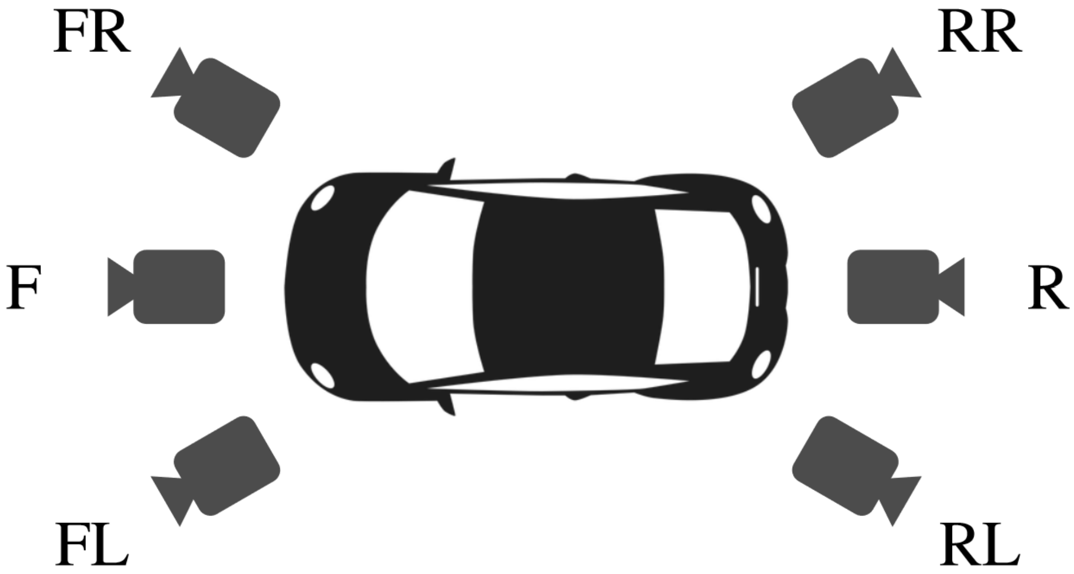
6th ITG/VDE Graduate Summer School on Video Compression and Processing



Motivation – Adaptive Multi-View Live Video Streaming

- Individual adaptation
  - Bitrate
  - Spatial resolution
  - Temporal resolution

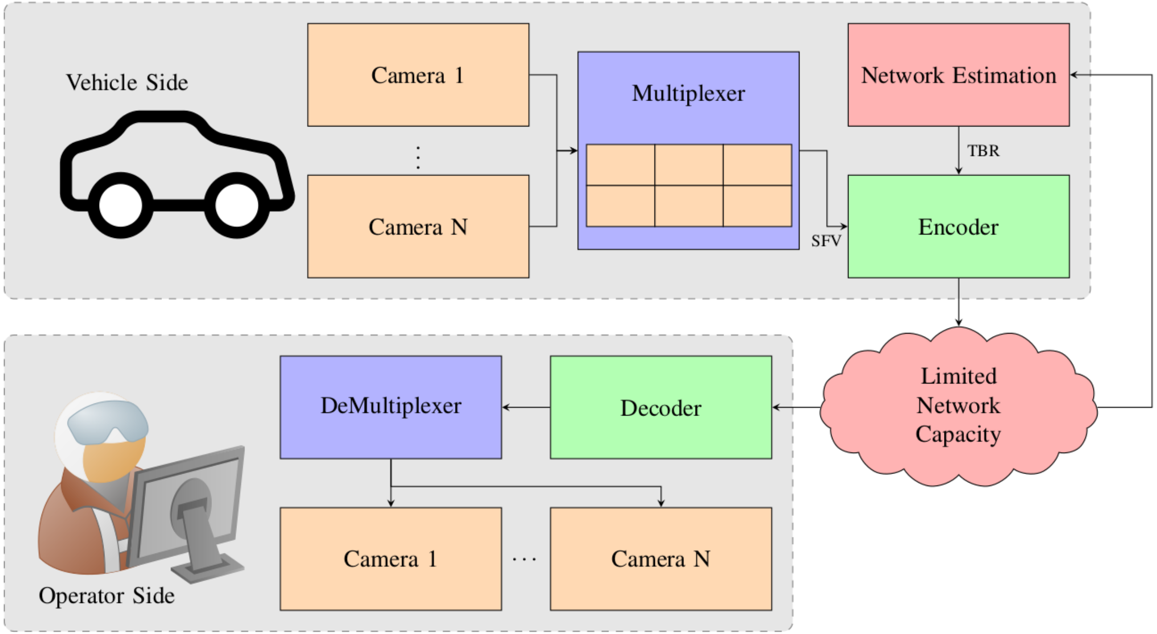
- Multi-view optimization
  - Per view
  - Overall view



State of the Art – Single Encoder Video Transmission

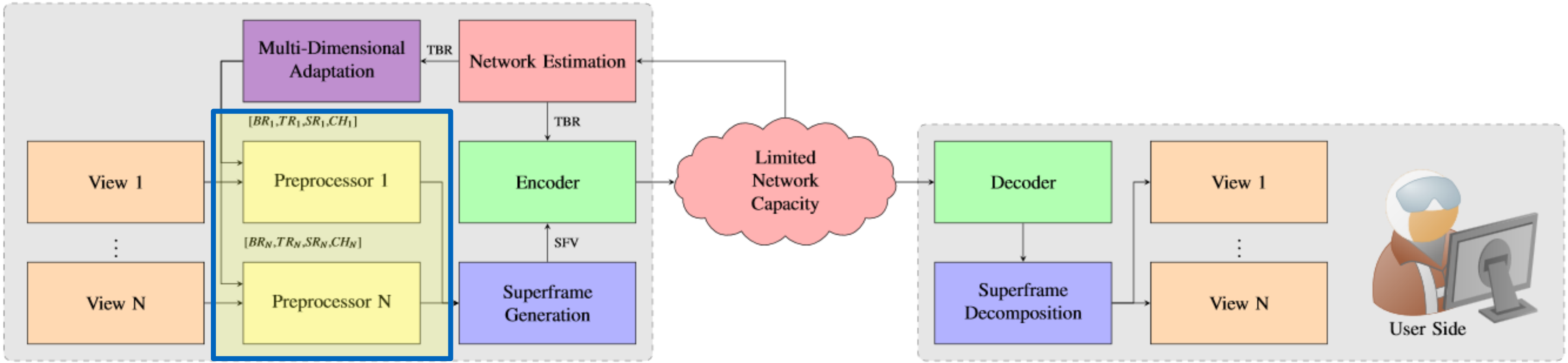
- Limited hardware in autonomous vehicles

- Single encoder
  - Combination into a single superframe
  - Lack of individual adaptation (without direct access to the encoder)



[1] Adaptive Multi-View Live Video Streaming for Teledriving Using a Single Hardware Encoder, Markus Hofbauer, Christopher B. Kuhn, Goran Petrovic, Eckehard Steinbach; IEEE ISM 2020

Concept – Preprocessing Filters

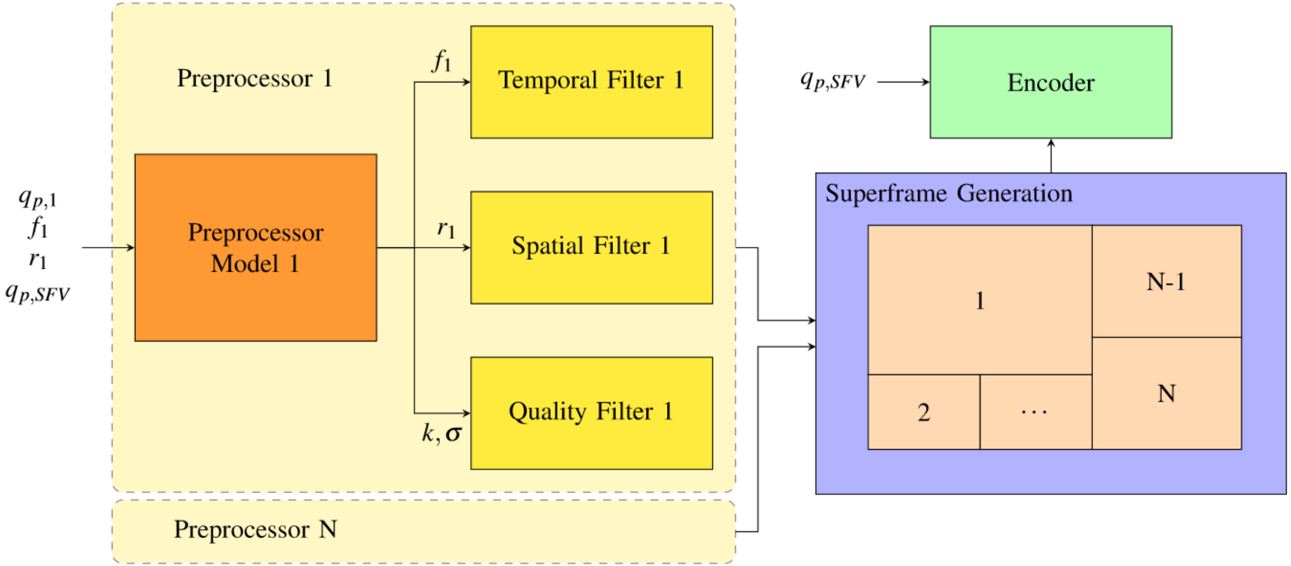


[2] Adaptive Multi-View Live Video Streaming Using a Single Encoder, Markus Hofbauer, Christopher B. Kuhn, Goran Petrovic, Eckehard Steinbach; IEEE Transactions on CSVT, Submitted

- Preprocessing concept
  - Enable individual rate/quality adaptation
  - Multiple camera views
  - Single encoder

Approach – Preprocessing Filters

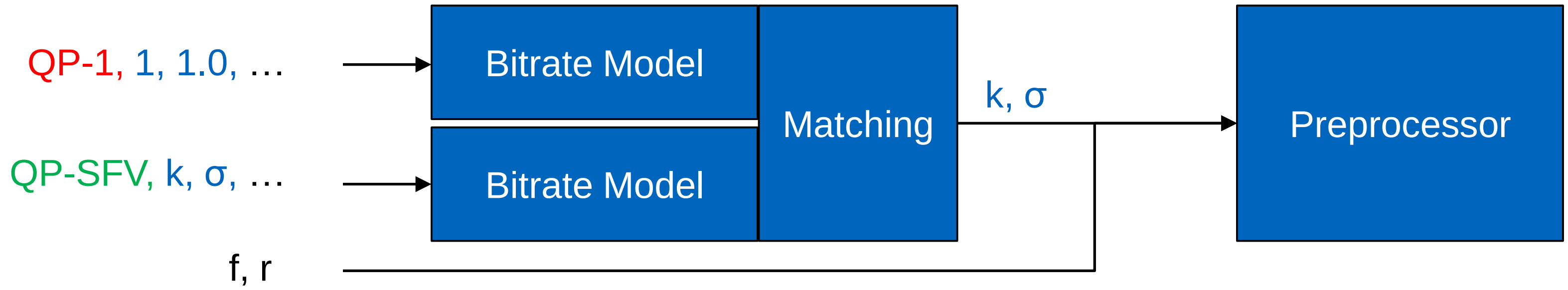
- Temporal filter
  → Frame drops
- Spatial filter
  → Frame downscale
- Quality filter
  → Spatial low pass



[2] Adaptive Multi-View Live Video Streaming Using a Single Encoder, Markus Hofbauer, Christopher B. Kuhn, Goran Petrovic, Eckehard Steinbach; IEEE Transactions on CSVT, Submitted

Approach – Preprocessing Parameter Estimation

- Estimate kernel size k and standard deviation σ using bitrate model
- Workflow with individual encoders: QP, frame rate (f), frame resolution (r) as input for every encoder
- Workflow in case of limited hardware:



- Design bitrate model R to allow solving of
  R(QP-1, …, 1, 1.0) = R(QP-SFV, …, k, σ)

Approach – Bitrate Model Design

- Challenge
  - No direct access to the underlying codec (motion estimation, …)
  - Estimate bitrate with the encoder as black box

- Goal: Estimate the bitrate from
  - Given encoding parameter (QP, f, r, …)
  - Video activity measures [3]
    - Spatial activity SA
    - Temporal activity TA

[3] Rec. ITU-T P.910 : Subjective video quality assessment methods for multimedia applications

Approach – Dataset

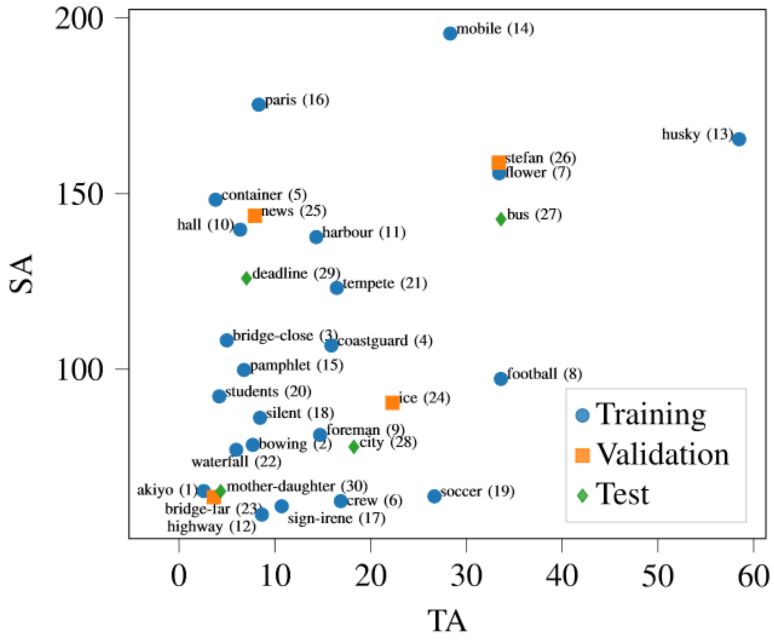
- 30 YUV CIF videos (football, deadline, akiyo, …)
  - CIF as used in most related work

- Data generation with x264
  - Encode all videos with all permutations of encoding parameters (EP)
  - EP: QP, frame rate, GOP length, frame size, Gaussian filter (kernel size & σ)

- Dataset 30 x 554401 = 16.6 million samples

  ➔ SA, TA, EP as input

  ➔ Bitrate as output



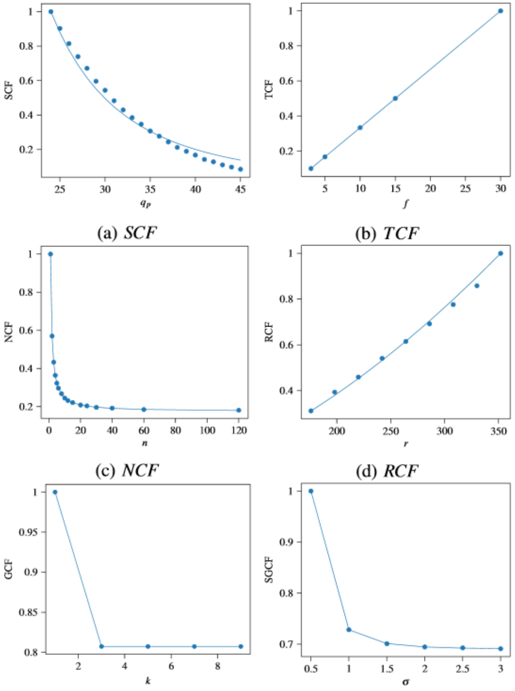
[2] Adaptive Multi-View Live Video Streaming Using a Single Encoder, Markus Hofbauer, Christopher B. Kuhn, Goran Petrovic, Eckehard Steinbach; IEEE Transactions on CSVT, Submitted

Approach – Analytic Bitrate Model

- Extend state-of-the-art approach [3]:

$$R = R_{max} * TCF(f) * SCF(q) * NCF(n) * RCF(r) * GCF(k) * GSDCF(\sigma)$$

- Maximum bitrate R_{max}
- Temporal correction factor $TCF(f)$
- Spatial correction factor $SCF(q)$
- GOP correction factor $NCF(n)$
- Resolution correction factor $RCF(r)$
- Gaussian correction factor $GCF(k)$
- Gaussian standard deviation correction factor $GSDCF(\sigma)$



(a) *SCF* (b) *TCF* (c) *NCF* (d) *RCF* (e) *GCF* (f) *GSDCF*

[3] C. Lottermann, A. Machado, D. Schroeder, Y. Peng, and E. Steinbach, “Modeling the bit rate of H. 264/AVC video encoding as a function of quantization parameter, frame rate and GOP characteristics” in 2014 IEEE ICMEW

[2] Adaptive Multi-View Live Video Streaming Using a Single Encoder, Markus Hofbauer, Christopher B. Kuhn, Goran Petrovic, Eckehard Steinbach; IEEE Transactions on CSVT, Submitted

Approach – Analytic Bitrate Model

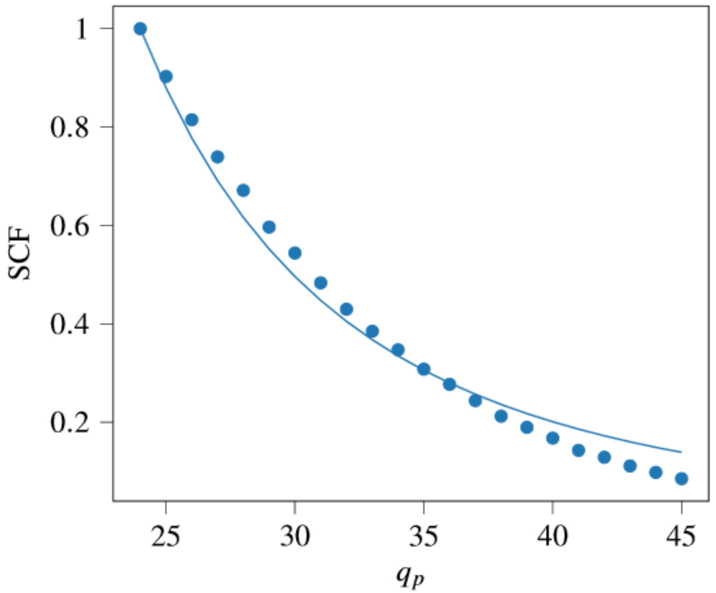
- Model individual correction factor, e.g. SCF

$$SCF(q_p, q_{p,min}) = \left(\frac{q_p}{q_{p,min}}\right)^{-a}$$

- Find model parameter a using least square non-linear fitting

- Model a based on SA and TA
  - Use iterative generalized linear regression method [4]
  - Least square non-linear fitting

$$a_{st} = \alpha_0 * SA + \alpha_1 * SA * TA + \alpha_2$$

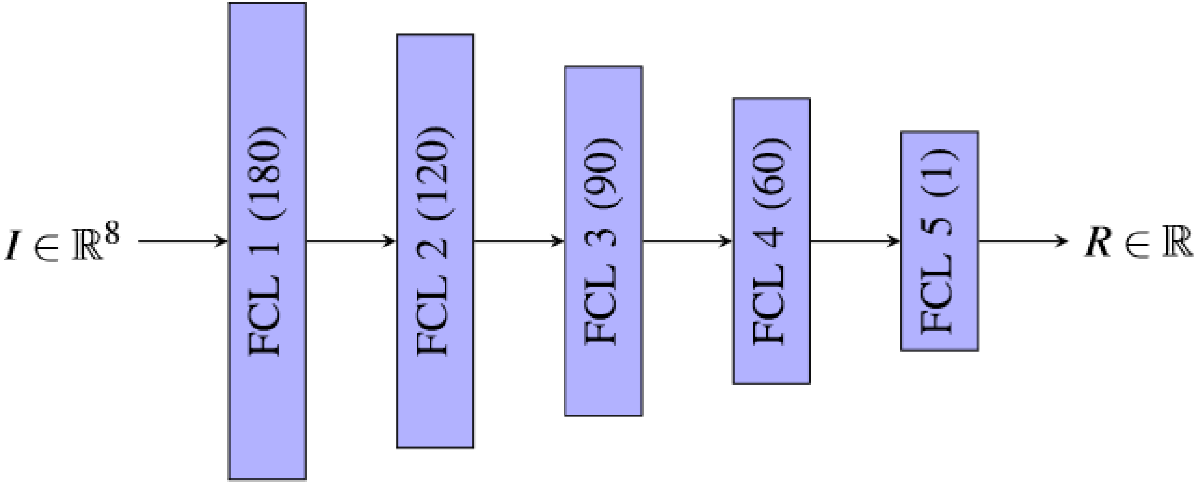


[2] Adaptive Multi-View Live Video Streaming Using a Single Encoder, Markus Hofbauer, Christopher B. Kuhn, Goran Petrovic, Eckehard Steinbach; IEEE Transactions on CSVT, Submitted

[4] P. McCullagh and J. A. Nelder, Generalized Linear Models. CRC Press,1989, vol. 37.

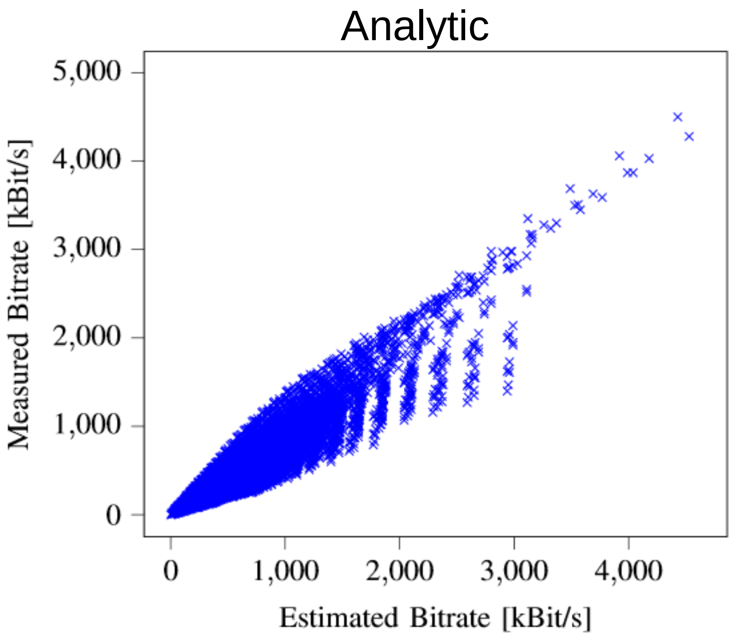
Approach – Machine Learning based Bitrate Model

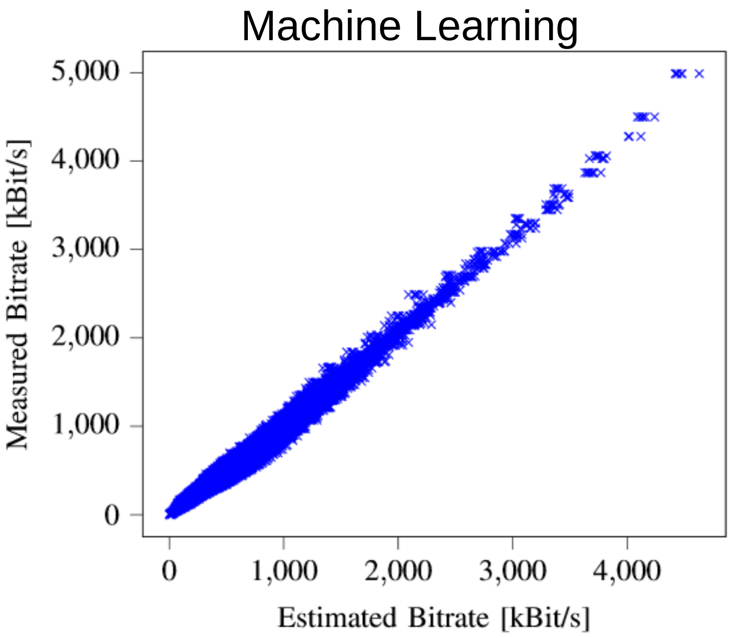
- Multi-layer perceptron
- Five fully connected layers



[2] Adaptive Multi-View Live Video Streaming Using a Single Encoder, Markus Hofbauer, Christopher B. Kuhn, Goran Petrovic, Eckehard Steinbach; IEEE Transactions on CSVT, Submitted

Results – Bitrate Models





[2] Adaptive Multi-View Live Video Streaming Using a Single Encoder, Markus Hofbauer, Christopher B. Kuhn, Goran Petrovic, Eckehard Steinbach; IEEE Transactions on CSVT, Submitted

Results – Bitrate Models

Model	Training Set RMSE	Training Set PC	Test Set RMSE	Test Set PC	All Videos RMSE	All Videos PC
Analytic	0.0179	0.9627	0.0186	0.9564	0.0180	0.9617
ML	**0.0111**	**0.9723**	**0.0078**	**0.9813**	**0.0106**	**0.9721**

[2] Adaptive Multi-View Live Video Streaming Using a Single Encoder, Markus Hofbauer, Christopher B. Kuhn, Goran Petrovic, Eckehard Steinbach; IEEE Transactions on CSVT, Submitted

Results – Application

	Video	Multi	Single	Oracle	Analytic	ML
Settings {QP,k,σ}	bus	{26,-,-}	{26,-,-}	{26,1,0.5}	{26,1,0.5}	{26,1,0.5}
	city	{38,-,-}	{26,-,-}	{26,7,2.5}	{26,9,3.0}	{26,9,3.0}
	deadline	{29,-,-}	{26,-,-}	{26,3,0.5}	{26,3,1.0}	{26,5,0.5}
	mother	{32,-,-}	{26,-,-}	{26,5,2.0}	{26,9,3.0}	{26,5,2.0}
Bitrate [kbit/s]	bus	6145	6145	6145	6145	6145
	city	1353	5225	1355	2282	1179
	deadline	3696	4746	3919	3919	3919
	mother	1098	2088	1097	1806	1317
normalized MAE	∅	-	0.2969	0.0102	0.0731	0.0377

[2] Adaptive Multi-View Live Video Streaming Using a Single Encoder, Markus Hofbauer, Christopher B. Kuhn, Goran Petrovic, Eckehard Steinbach; IEEE Transactions on CSVT, Submitted

Conclusion

- Problem: Limited hardware prevents individual view adaptation for multiple views

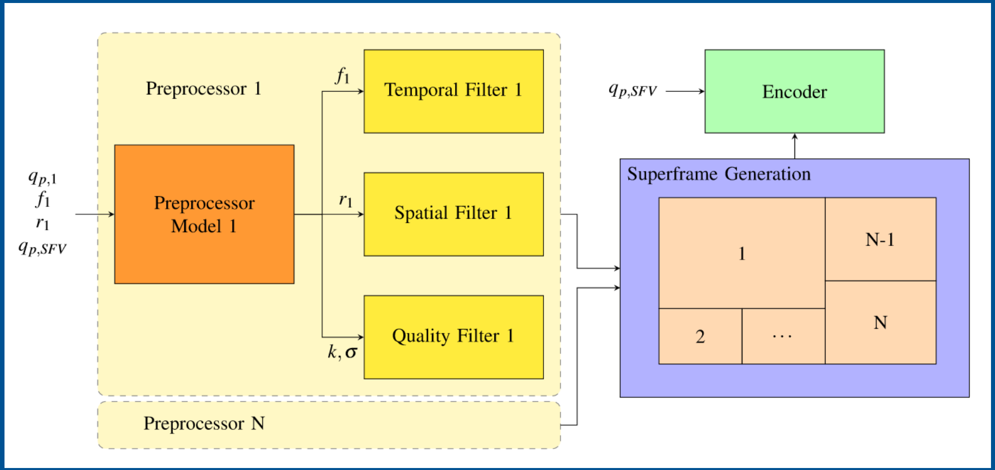
- Solution
  - Preprocessing concept to enable individual view adaptation with a single encoder
    - 3 preprocessing filters to manipulate individual frames
    - Combine preprocessed frames to single superframe
  - Model for estimating the required preprocessing parameter
    - Bitrate model (analytic & ML) for estimating influence of preprocessing filters

- Results
  - Achieve similar bitrate with a single encoder compared to multiple encoders
  - Machine Learning based bitrate model outperforms analytical model

Adaptive Multi-View Live Video Streaming Using a Single Encoder

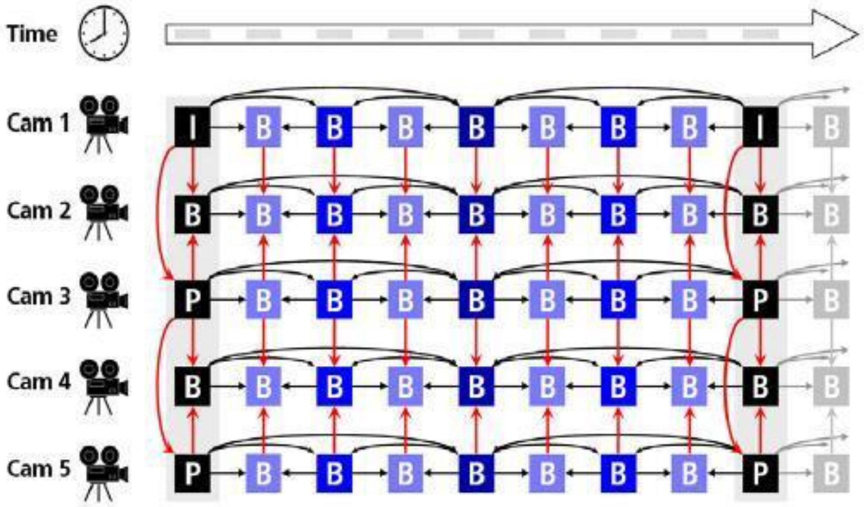
Contact:
markus.hofbauer@tum.de
https://hofbi.github.io

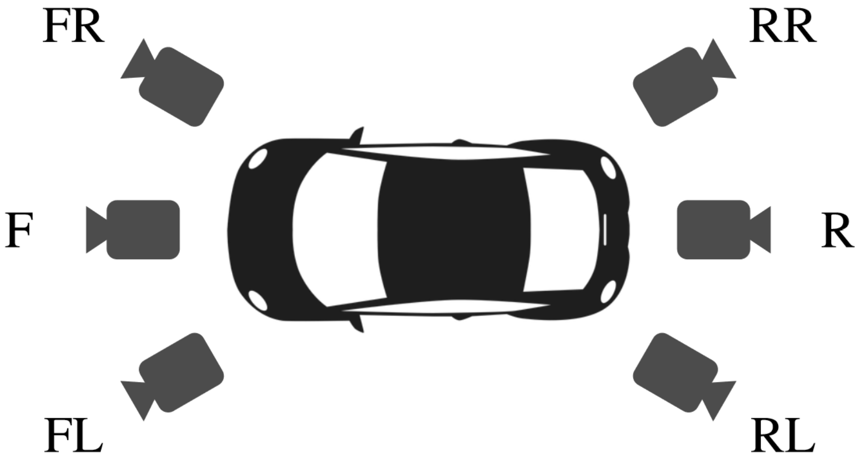
6th ITG/VDE Graduate Summer School on Video Compression and Processing



[2] Adaptive Multi-View Live Video Streaming Using a Single Encoder, Markus Hofbauer, Christopher B. Kuhn, Goran Petrovic, Eckehard Steinbach; IEEE Transactions on CSVT, Submitted

State of the Art – Multi-View in Video Coding Standards





wikipedia.org/wiki/Multiview_Video_Coding

[1] A. Vetro, T. Wiegand, and G. Sullivan, "Overview of the Stereo and Multiview Video Coding Extensions of the H.264/MPEG-4 AVC Standard," Proceedings of the IEEE, vol. 99, pp. 626–642, May 2011.
[2] M. M. Hannuksela, Y. Yan, X. Huang, and H. Li, "Overview of the multiview high efficiency video coding (MV-HEVC) standard," in 2015 IEEE International Conference on Image Processing (ICIP). Quebec City, QC, Canada: IEEE, pp. 2154–2158, Sep. 2015.
[3] G. Tech, Y. Chen, K. Muller, J.-R. Ohm, A. Vetro, and Y.-K. Wang, "Overview of the Multiview and 3D Extensions of High Efficiency Video Coding," IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 1, pp. 35–49, Jan. 2016.