cencro – Speedup of Video Quality Calculation using Center Cropping

Steve Göring, Christopher Krämmer, Alexander Raake

Audiovisual Technology Group, Technische Universität Ilmenau, Germany;
Email: [steve.goering, christopher.kraemmer, alexander.raake]@tu-ilmenau.de

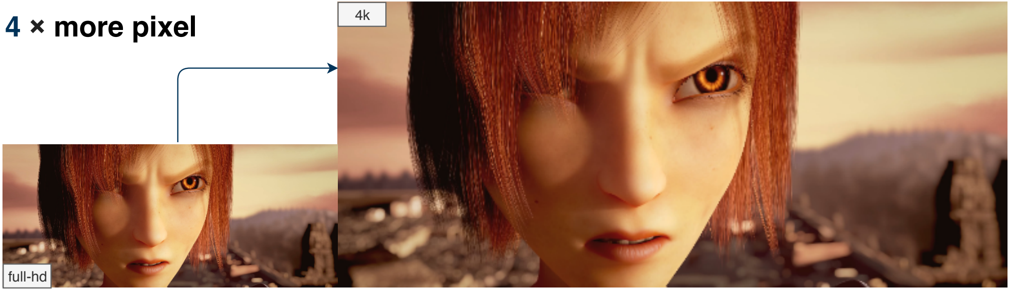
code cencro: `https://git.io/JeR5q`

December 9, 2019

TECHNISCHE UNIVERSITÄT ILMENAU

Motivation

TECHNISCHE UNIVERSITÄT ILMENAU



- increase of video resolution (full-HD, 4K/UHD-1, 8K/UHD-2, ...)
- more processing time for video quality estimation required
- e.g. full-reference methods (VMAF); 45 mins for 4K 10 s sequence

$\rightarrow$ review of video quality models

Video Quality Estimation

TECHNISCHE UNIVERSITÄT ILMENAU

- categorized in no-ref, red-ref, **full-ref** models [13]
  - PSNR, MS-SSIM [16, 15], VIF [14]: fast but not accurate
  - Netflix's VMAF [10, 7]: accurate, slow
  - DNN based models: [2, 4, 3, 5, 6]: accurate and slower
  - brisque+niqe [8, 9]: fast but not always accurate, ...
- subjective tests: cross-lab evaluation *Pinson and Wolf* [11]: Pearson correlations: 0.902 to 0.935 ($\approx 4\%$ error) for MOS,

$\rightarrow$ general principle to speedup full-ref calculations

Video Quality Estimation

TECHNISCHE UNIVERSITÄT **ILMENAU**

- categorized in no-ref, red-ref, **full-ref** models [13]
  - PSNR, MS-SSIM [16, 15], VIF [14]: fast but not accurate
  - Netflix's VMAF [10, 7]: accurate, slow
  - DNN based models: [2, 4, 3, 5, 6]: accurate and slower
  - brisque+niqe [8, 9]: fast but not always accurate, ...
- subjective tests: cross-lab evaluation *Pinson and Wolf* [11]: Pearson correlations: 0.902 to 0.935 ($\approx$ 4% error) for MOS,

$\rightarrow$ general principle to speedup full-ref calculations

Video Quality Estimation

TECHNISCHE UNIVERSITÄT ILMENAU

- categorized in no-ref, red-ref, **full-ref** models [13]
  - PSNR, MS-SSIM [16, 15], VIF [14]: fast but not accurate
  - Netflix's VMAF [10, 7]: accurate, slow
  - DNN based models: [2, 4, 3, 5, 6]: accurate and slower
  - brisque+niqe [8, 9]: fast but not always accurate, ...
- subjective tests: cross-lab evaluation *Pinson and Wolf* [11]: Pearson correlations: 0.902 to 0.935 ($\approx$ 4% error) for MOS,

$\rightarrow$ general principle to speedup full-ref calculations

Video Quality Estimation

TECHNISCHE UNIVERSITÄT
ILMENAU

- categorized in no-ref, red-ref, **full-ref** models [13]
  - PSNR, MS-SSIM [16, 15], VIF [14]: fast but not accurate
  - Netflix's VMAF [10, 7]: accurate, slow
  - DNN based models: [2, 4, 3, 5, 6]: accurate and slower
  - brisque+niqe [8, 9]: fast but not always accurate, ...
- subjective tests: cross-lab evaluation *Pinson and Wolf* [11]: Pearson correlations: 0.902 to 0.935 ($\approx$ 4% error) for MOS,

$\rightarrow$ general principle to speedup full-ref calculations

Video Quality Estimation

TECHNISCHE UNIVERSITÄT ILMENAU

- categorized in no-ref, red-ref, **full-ref** models [13]
  - PSNR, MS-SSIM [16, 15], VIF [14]: fast but not accurate
  - Netflix's VMAF [10, 7]: accurate, slow
  - DNN based models: [2, 4, 3, 5, 6]: accurate and slower
  - brisque+niqe [8, 9]: fast but not always accurate, ...
- subjective tests: cross-lab evaluation *Pinson and Wolf* [11]: Pearson correlations: 0.902 to 0.935 ($\approx$ 4% error) for MOS,

$\rightarrow$ general principle to speedup full-ref calculations

Video Quality Estimation

TECHNISCHE UNIVERSITÄT
ILMENAU

- categorized in no-ref, red-ref, **full-ref** models [13]
  - PSNR, MS-SSIM [16, 15], VIF [14]: fast but not accurate
  - Netflix's VMAF [10, 7]: accurate, slow
  - DNN based models: [2, 4, 3, 5, 6]: accurate and slower
  - brisque+niqe [8, 9]: fast but not always accurate, ...
- subjective tests: cross-lab evaluation *Pinson and Wolf* [11]: Pearson correlations: 0.902 to 0.935 ($\approx$ 4% error) for MOS,

$\rightarrow$ general principle to speedup full-ref calculations

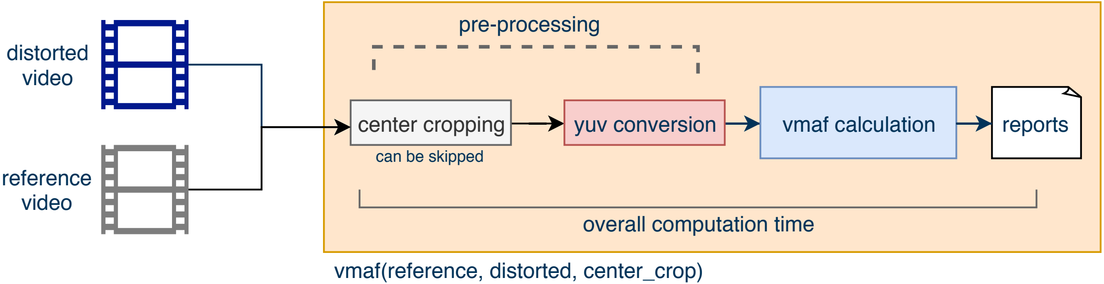
Video Quality Estimation

- categorized in no-ref, red-ref, **full-ref** models [13]
  - PSNR, MS-SSIM [16, 15], VIF [14]: fast but not accurate
  - Netflix's VMAF [10, 7]: accurate, slow
  - DNN based models: [2, 4, 3, 5, 6]: accurate and slower
  - brisque+niqe [8, 9]: fast but not always accurate, ...
- subjective tests: cross-lab evaluation *Pinson and Wolf* [11]: Pearson correlations: 0.902 to 0.935 ($\approx$ 4% error) for MOS,

$\rightarrow$ general principle to speedup full-ref calculations

Our Approach– *cencro*

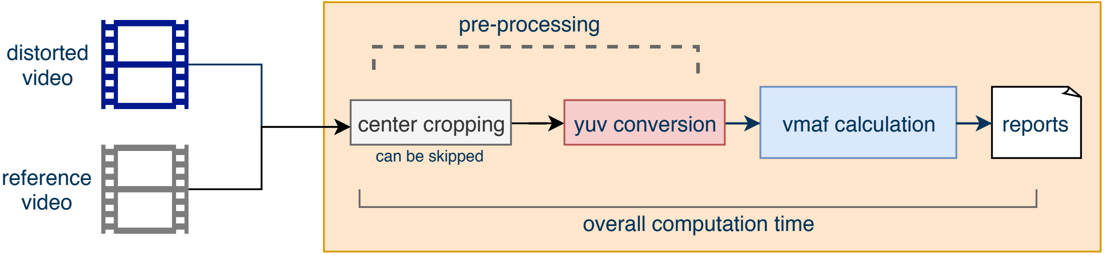
TECHNISCHE UNIVERSITÄT
ILMENAU



- focus: per frame reduction: center cropping of frames → **cencro**
- metric: VMAF; processing time: $t_{all} = t_{dis} + t_{ref} + t_{VMAF}$
- questions: what are suitable center crops? calculation speedup? acceptable error?

Our Approach– *cencro*

TECHNISCHE UNIVERSITÄT ILMENAU

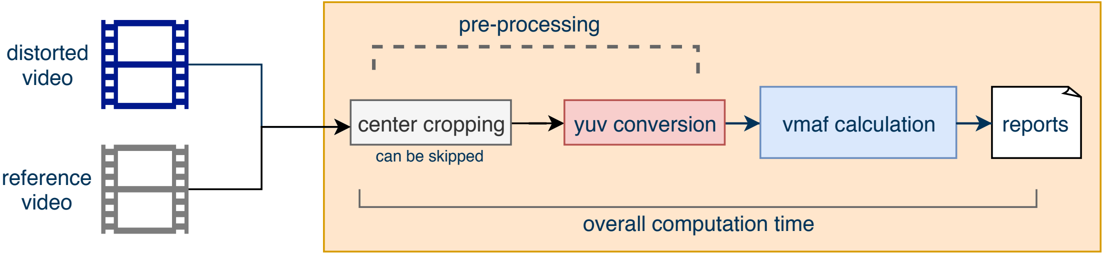


vmaf(reference, distorted, center_crop)

- focus: per frame reduction: center cropping of frames → **cencro**
- metric: VMAF; processing time: $t_{all} = t_{dis} + t_{ref} + t_{VMAF}$
- questions: what are suitable center crops? calculation speedup? acceptable error?

Our Approach– *cencro*

TECHNISCHE UNIVERSITÄT ILMENAU

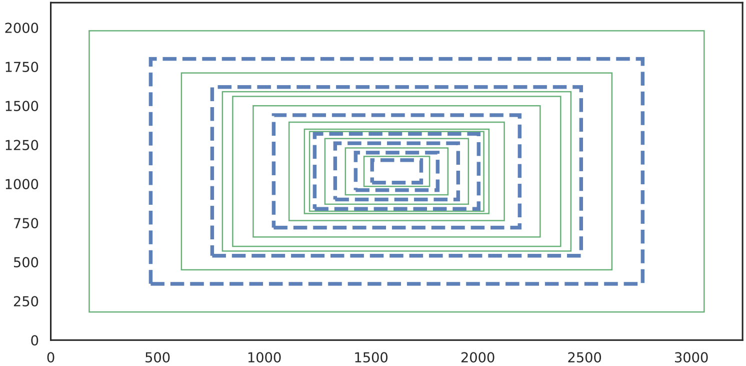


vmaf(reference, distorted, center_crop)

- focus: per frame reduction: center cropping of frames → **cencro**
- metric: VMAF; processing time: $t_{all} = t_{dis} + t_{ref} + t_{VMAF}$
- questions: what are suitable center crops? calculation speedup? acceptable error?

Our Approach

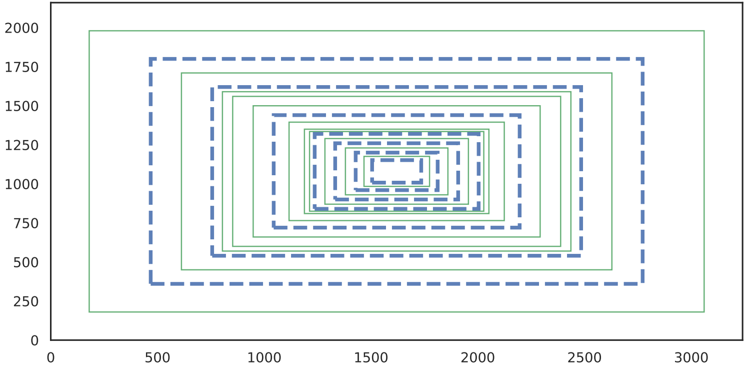
TECHNISCHE UNIVERSITÄT
ILMENAU



- total **18** center crop settings:
  144, 192, **240**, 300, **360**, 420, **480**, 510, 540, 630, **720**, 840, 960, 1020, **1080**, 1260, **1440**, 1800
- widths adapted based on aspect ratio of reference video

Our Approach

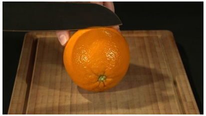
TECHNISCHE UNIVERSITÄT ILMENAU



- total **18** center crop settings:
  144, 192, **240**, 300, **360**, 420, **480**, 510, 540, 630, **720**, 840, 960, 1020, **1080**, 1260, **1440**, 1800
- widths adapted based on aspect ratio of reference video

Evaluation – Datasets





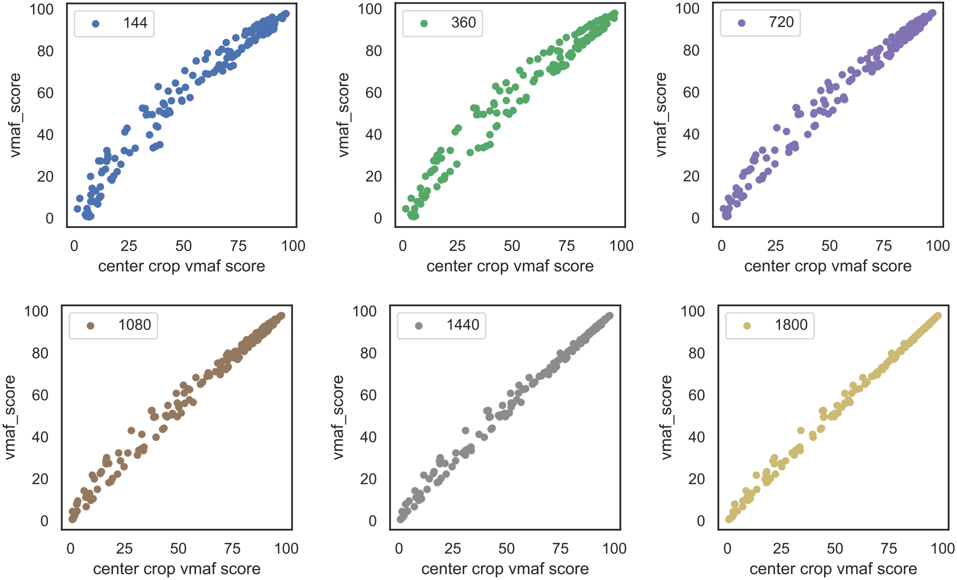

codecs	h264, h265, vp9			
resolutions	360p	720p	1080p	2160p
bitrates in Mbit/s	[0.2,0.75]	[1,2]	[2,7.5,15]	[7.5,15,40]

- 4K videos: subset of AVT-VQDB-UHD-1 [12]; **150 encoded videos**
- + GamingVideoSET [1]; 24 full-HD videos, **90 encoded videos**

Evaluation – Center Cropping – VMAF

TECHNISCHE UNIVERSITÄT ILMENAU

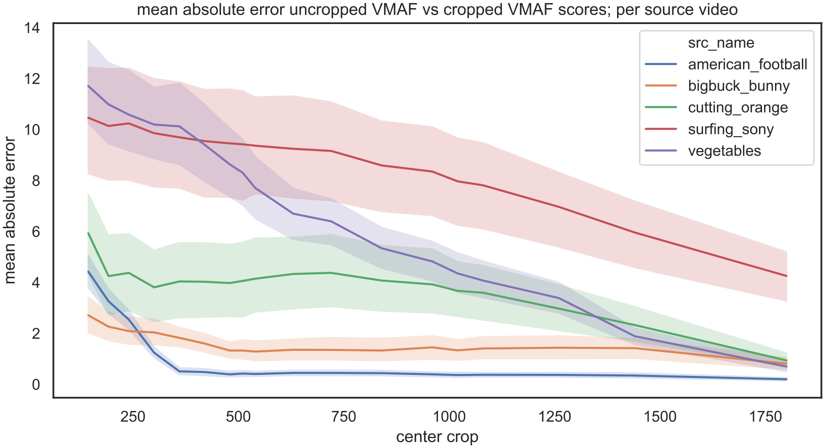
- comparison VMAF score with center cropped VMAF score



Evaluation – Center Cropping – VMAF Error

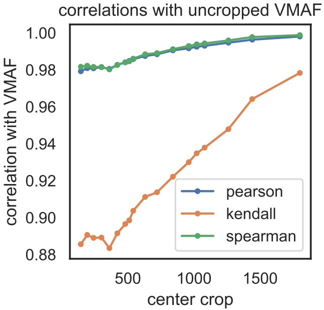
TECHNISCHE UNIVERSITÄT
ILMENAU

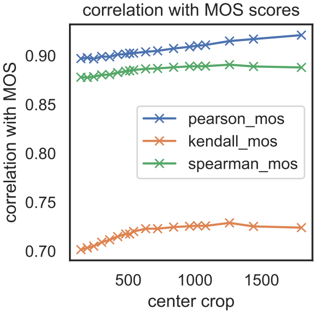
- mean absolute error per source video; VMAF $\in [0, 100]$

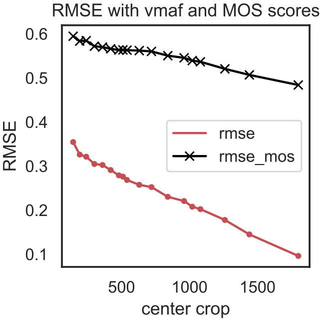


Evaluation – Cencro – Correlations and Error

TECHNISCHE UNIVERSITÄT ILMENAU

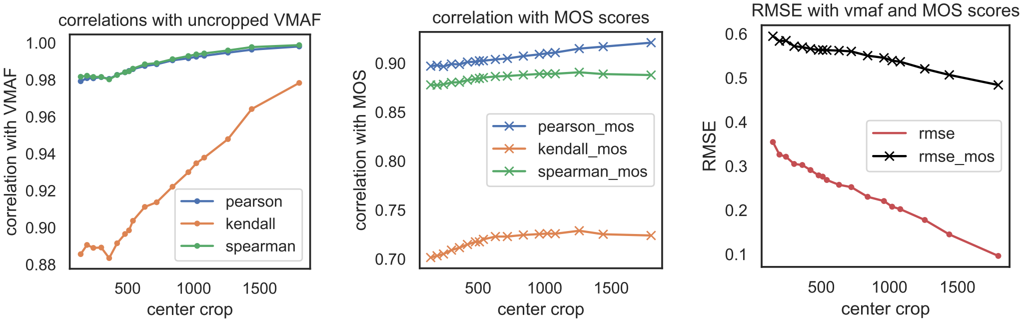






- higher crop → lower RMSE, higher correlation
- error still acceptable (e.g. for 360p cc)

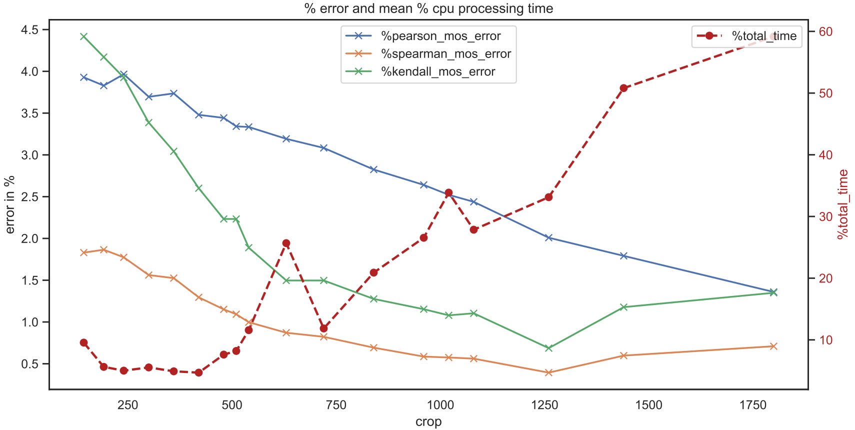
Evaluation – Cencro – Correlations and Error



- higher crop $\rightarrow$ lower RMSE, higher correlation
- error still acceptable (e.g. for 360p cc)

Evaluation – Center Cropping – Error vs Time

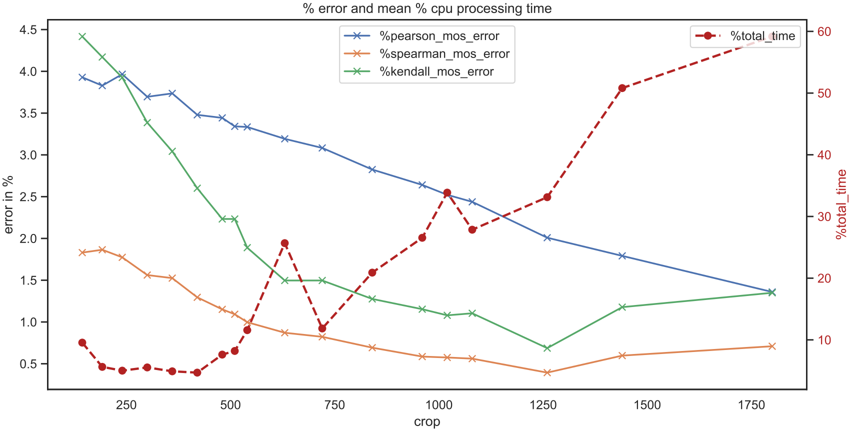
- % error (MOS prediction) vs % t_{all}; 100%time ≈ 45 min



- 360p → **5% cpu time; 4% PCC error**; similar for other FR metrics; Datasets (GamingVideoSET [1])

Evaluation – Center Cropping – Error vs Time

- % error (MOS prediction) vs % t_{all}; 100%time ≈ 45 min



- 360p → **5% cpu time; 4% PCC error**; similar for other FR metrics; Datasets (GamingVideoSET [1])

Conclusion, Summary and Future Work

- overview of **cencro**
  - speedup video quality calculation using center cropping of video frames
  - analysis of possible center crop settings
- error and speedup
  - 360p shows low error in case of 4K, comparable with cross-lab repeatition
  - 360p time saving up to 95%
  - 1080p can save up to 65% time with error of 2.5%
- open and next steps:
  - evaluate on other videos, e.g. 360°; automated optimal crop/pattern selection
  - check reduced-/no-ref metrics

Conclusion, Summary and Future Work

TECHNISCHE UNIVERSITÄT
ILMENAU

- overview of **cencro**
  - speedup video quality calculation using center cropping of video frames
  - analysis of possible center crop settings
- error and speedup
  - 360p shows low error in case of 4K, comparable with cross-lab repeatition
  - 360p time saving up to 95%
  - 1080p can save up to 65% time with error of 2.5%
- open and next steps:
  - evaluate on other videos, e.g. 360°; automated optimal crop/pattern selection
  - check reduced-/no-ref metrics

Conclusion, Summary and Future Work

- overview of **cencro**
  - speedup video quality calculation using center cropping of video frames
  - analysis of possible center crop settings
- error and speedup
  - 360p shows low error in case of 4K, comparable with cross-lab repeatition
  - 360p time saving up to 95%
  - 1080p can save up to 65% time with error of 2.5%
- open and next steps:
  - evaluate on other videos, e.g. 360°; automated optimal crop/pattern selection
  - check reduced-/no-ref metrics

Conclusion, Summary and Future Work

TECHNISCHE UNIVERSITÄT ILMENAU

- overview of **cencro**
  - speedup video quality calculation using center cropping of video frames
  - analysis of possible center crop settings
- error and speedup
  - 360p shows low error in case of 4K, comparable with cross-lab repeatition
  - 360p time saving up to 95%
  - 1080p can save up to 65% time with error of 2.5%
- open and next steps:
  - evaluate on other videos, e.g. 360°; automated optimal crop/pattern selection
  - check reduced-/no-ref metrics

Conclusion, Summary and Future Work

TECHNISCHE UNIVERSITÄT ILMENAU

- overview of **cencro**
  - speedup video quality calculation using center cropping of video frames
  - analysis of possible center crop settings
- error and speedup
  - 360p shows low error in case of 4K, comparable with cross-lab repeatition
  - 360p time saving up to 95%
  - 1080p can save up to 65% time with error of 2.5%
- open and next steps:
  - evaluate on other videos, e.g. 360°; automated optimal crop/pattern selection
  - check reduced-/no-ref metrics

Conclusion, Summary and Future Work

TECHNISCHE UNIVERSITÄT
ILMENAU

- overview of **cencro**
  - speedup video quality calculation using center cropping of video frames
  - analysis of possible center crop settings
- error and speedup
  - 360p shows low error in case of 4K, comparable with cross-lab repeatition
  - 360p time saving up to 95%
  - 1080p can save up to 65% time with error of 2.5%
- open and next steps:
  - evaluate on other videos, e.g. 360°; automated optimal crop/pattern selection
  - check reduced-/no-ref metrics

Conclusion, Summary and Future Work

TECHNISCHE UNIVERSITÄT ILMENAU

- overview of **cencro**
  - speedup video quality calculation using center cropping of video frames
  - analysis of possible center crop settings
- error and speedup
  - 360p shows low error in case of 4K, comparable with cross-lab repeatition
  - 360p time saving up to 95%
  - 1080p can save up to 65% time with error of 2.5%
- open and next steps:
  - evaluate on other videos, e.g. 360°; automated optimal crop/pattern selection
  - check reduced-/no-ref metrics

Conclusion, Summary and Future Work

- overview of **cencro**
  - speedup video quality calculation using center cropping of video frames
  - analysis of possible center crop settings
- error and speedup
  - 360p shows low error in case of 4K, comparable with cross-lab repeatition
  - 360p time saving up to 95%
  - 1080p can save up to 65% time with error of 2.5%
- open and next steps:
  - evaluate on other videos, e.g. 360°; automated optimal crop/pattern selection
  - check reduced-/no-ref metrics

Conclusion, Summary and Future Work

TECHNISCHE UNIVERSITÄT ILMENAU

- overview of **cencro**
  - speedup video quality calculation using center cropping of video frames
  - analysis of possible center crop settings
- error and speedup
  - 360p shows low error in case of 4K, comparable with cross-lab repeatition
  - 360p time saving up to 95%
  - 1080p can save up to 65% time with error of 2.5%
- open and next steps:
  - evaluate on other videos, e.g. 360°; automated optimal crop/pattern selection
  - check reduced-/no-ref metrics

Conclusion, Summary and Future Work

- overview of **cencro**
  - speedup video quality calculation using center cropping of video frames
  - analysis of possible center crop settings
- error and speedup
  - 360p shows low error in case of 4K, comparable with cross-lab repeatition
  - 360p time saving up to 95%
  - 1080p can save up to 65% time with error of 2.5%
- open and next steps:
  - evaluate on other videos, e.g. 360°; automated optimal crop/pattern selection
  - check reduced-/no-ref metrics

Thank you for your attention

TECHNISCHE UNIVERSITÄT ILMENAU

. . . . . . are there any questions?

References I

[1] Nabajeet Barman et al. “GamingVideoSET: a dataset for gaming video streaming applications”. In: *2018 16th Annual Workshop on Network and Systems Support for Games (NetGames)*. IEEE. 2018, pp. 1–6.

[2] Sebastian Bosse et al. “Neural network-based full-reference image quality assessment”. In: *Picture Coding Symposium (PCS), 2016*. IEEE. 2016, pp. 1–5.

[3] Prajna Paramita Dash, Akshaya Mishra, and Alexander Wong. “Deep Quality: A Deep No-reference Quality Assessment System”. In: *arXiv preprint arXiv:1609.07170* (2016).

[4] Prajna Paramita Dash, Alexander Wong, and Akshaya Mishra. “VeNICE: A very deep neural network approach to no-reference image assessment”. In: *Industrial Technology (ICIT), 2017 IEEE International Conference on*. IEEE. 2017, pp. 1091–1096.

References II

[5] Steve Göring and Alexander Raake. "deimeq – A Deep Neural Network Based Hybrid No-reference Image Quality Model". In: *Visual Information Processing (EUVIP), 2018 7th European Workshop on*. IEEE. 2018, pp. 1–6.

[6] Steve Göring, Janto Skowronek, and Alexander Raake. "DeViQ – A deep no reference video quality model". In: *Electronic Imaging, Human Vision Electronic Imaging* (2018).

[7] J. Y. Lin et al. "A fusion-based video quality assessment (fvqa) index". In: *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2014 Asia-Pacific*. Dec. 2014, pp. 1–5.

[8] Anish Mittal, Anush Krishna Moorthy, and Alan Conrad Bovik. "No-reference image quality assessment in the spatial domain". In: *IEEE Transactions on Image Processing* 21.12 (2012), pp. 4695–4708.

References III

[9] Anish Mittal, Rajiv Soundararajan, and Alan C Bovik. "Making a "completely blind" image quality analyzer". In: *IEEE Signal Processing Letters* 20.3 (2013), pp. 209–212.

[10] Netflix. *Netflix VMAF*. URL: https://github.com/Netflix/vmaf (visited on 12/08/2018).

[11] Margaret H Pinson and Stephen Wolf. "Comparing subjective video quality testing methodologies". In: *Visual Communications and Image Processing 2003*. Vol. 5150. International Society for Optics and Photonics. 2003, pp. 573–582.

[12] Rakesh Rao Ramachandra Rao et al. "AVT-VQDB-UHD-1: A Large Scale Video Quality Database for UHD-1". In: *2019 IEEE ISM*. Dec. 2019, pp. 1–8.

References IV

TECHNISCHE UNIVERSITÄT ILMENAU

[13] Muhammad Shahid et al. "No-reference image and video quality assessment: a classification and review of recent approaches". In: *EURASIP Journal on Image and Video Processing* 2014.1 (Aug. 2014), p. 40.

[14] Hamid R Sheikh and Alan C Bovik. "Image information and visual quality". In: *IEEE Transactions on image processing* 15.2 (2006), pp. 430–444.

[15] Zhou Wang, Eero P Simoncelli, and Alan C Bovik. "Multiscale structural similarity for image quality assessment". In: *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on*. Vol. 2. IEEE. 2003, pp. 1398–1402.

[16] Zhou Wang et al. "Image quality assessment: from error visibility to structural similarity". In: *IEEE transactions on image processing* 13.4 (2004), pp. 600–612.