On Evaluation of Video Quality Metrics:
an HDR Dataset for Computer Graphics Applications

Martin Čadík*, Tunç O. Aydın,
Karol Myszkowski, and Hans-Peter Seidel

Outline

- **Quality Assessment**
  - In Computer Graphics

- **Proposed Dataset**
  - Reference-test video pairs
  - LDR-LDR, HDR-HDR, LDR-HDR

- **Example Evaluation**
  - 4 VQMs

- **Conclusion and Future Work**

FR Quality Assessment

Reference **Test**

Rate the Quality
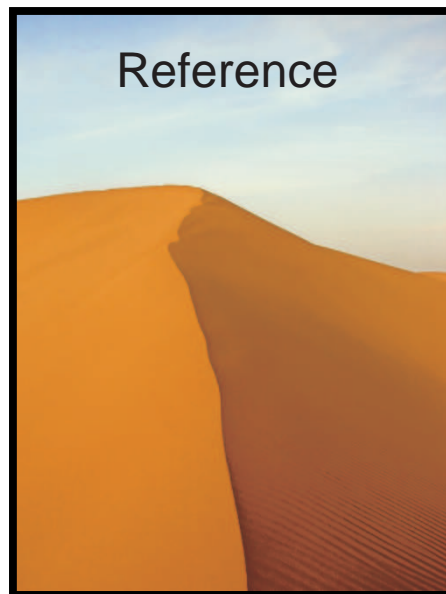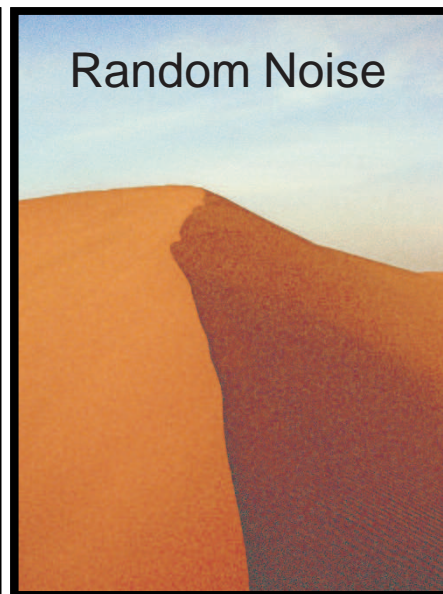
Subjective Studies: + Reliable - Time Consuming

Martin Čadík, mcadik@mpii.de
SPIE HVEI'11, San Francisco

On Evaluation of Video Quality Metrics:
an HDR Dataset for Computer Graphics Applications
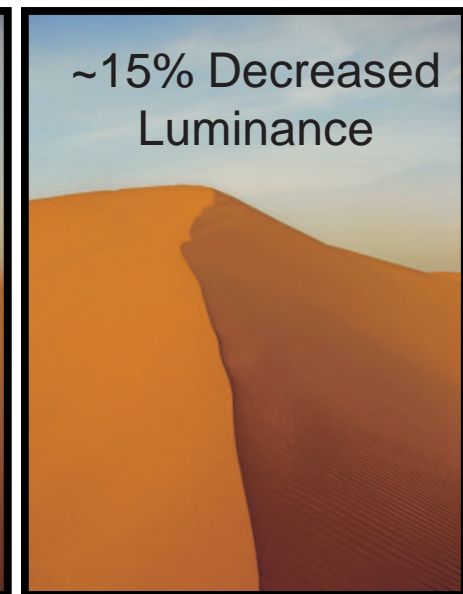
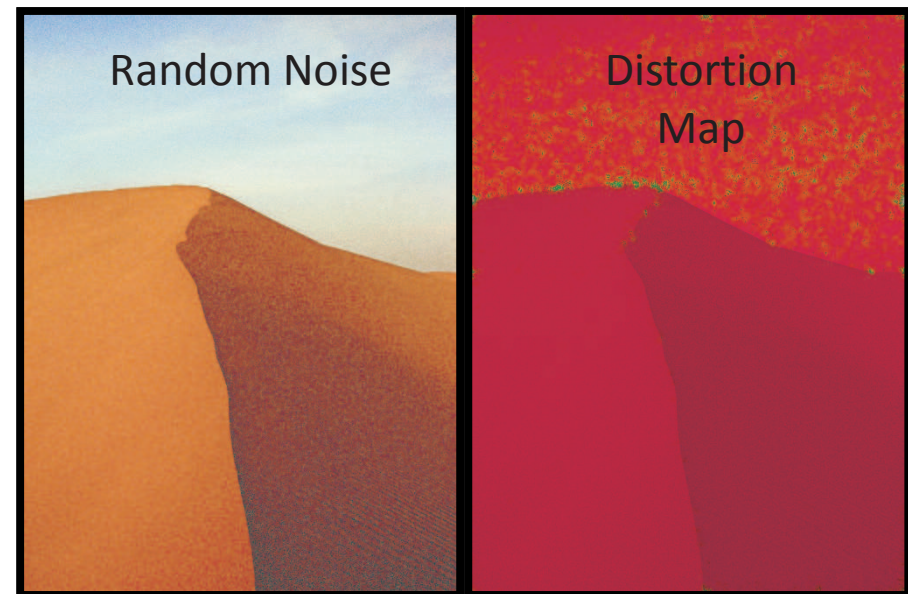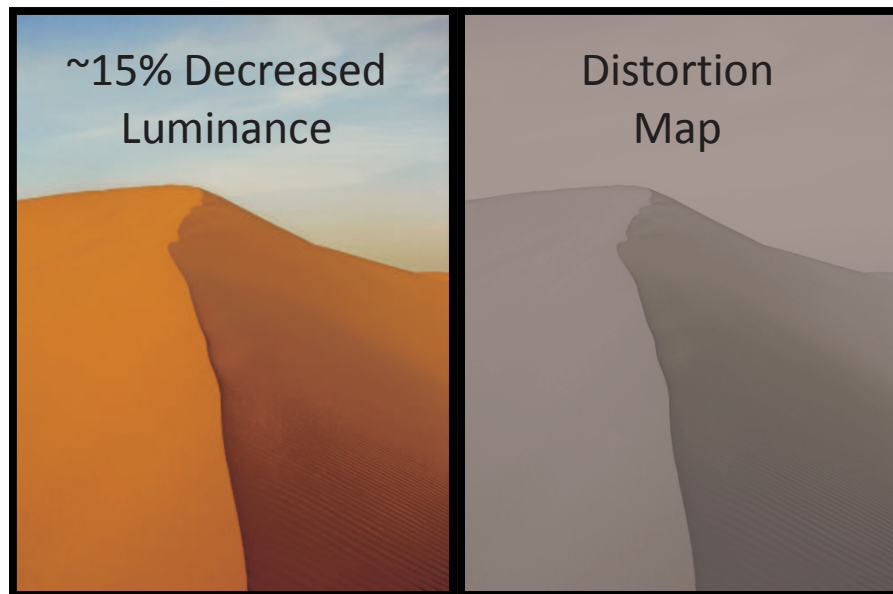# Simple Metrics

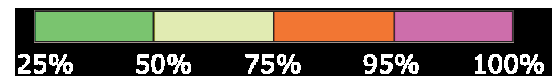| Reference	Random Noise	Blur	~15% Decreased Luminance
	MSE = 280	MSE = 280	MSE = 280 !

Advanced (HVS Based) Metrics

~15% Decreased Luminance

Distortion Map

Random Noise

Distortion Map

Probability of Detection:

25%   50%   75%   95%   100%

Validation of Metrics

- Input data + Subjective responses = dataset

- Datasets
  - Simpler evaluations
  - Reproducible evaluations

  - Should be publicly available
  - Should comprise typical artifacts

Available Video Datasets

- VQEG FRTV Phase 1 [VQEG '00]

- LIVE video db [Seshadrinathan et al. '09]

- LDR videos only

- Focus on compression/transmission related artifacts

- Subjective responses: only overall quality rating

VQA in Computer Graphics

- User studies take time & require expertise in psychophysics

  - Usually only informal studies

  - Objective metrics help in this

  - Objective metrics can be applied in algorithms

- CG

  - LDR, HDR, LDR/HDR
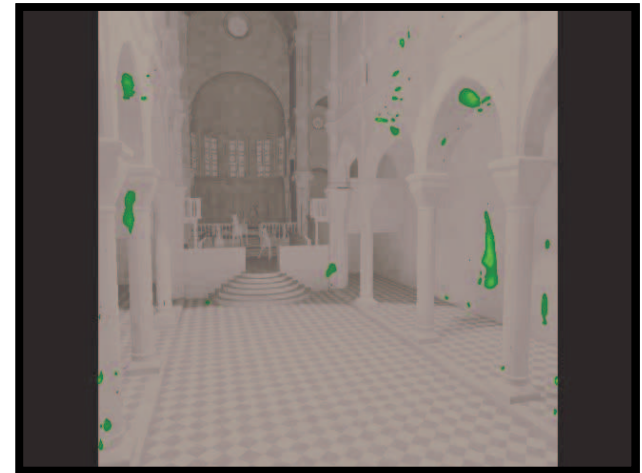
  - Localized distortion maps

  - Specific artifacts

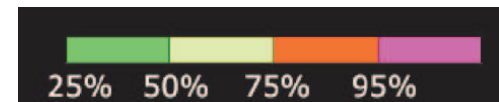# Evaluation of Rendering Methods
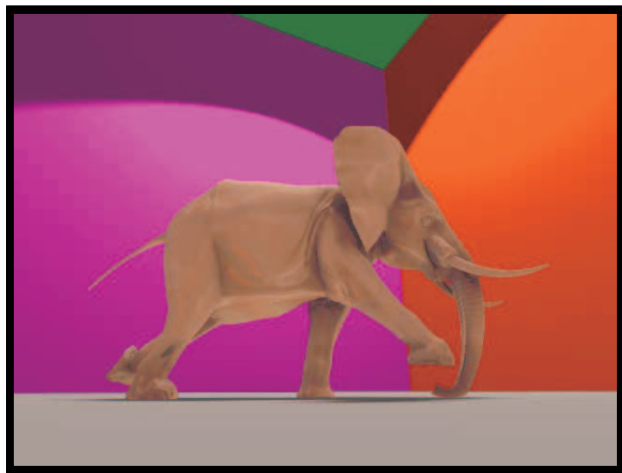

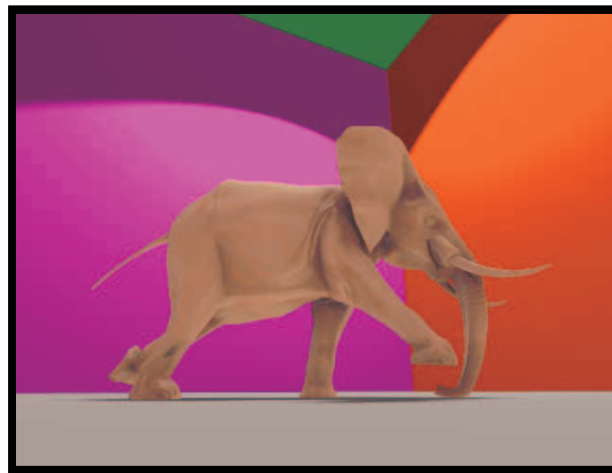With temporal filtering
[Herzog et al. 2010]

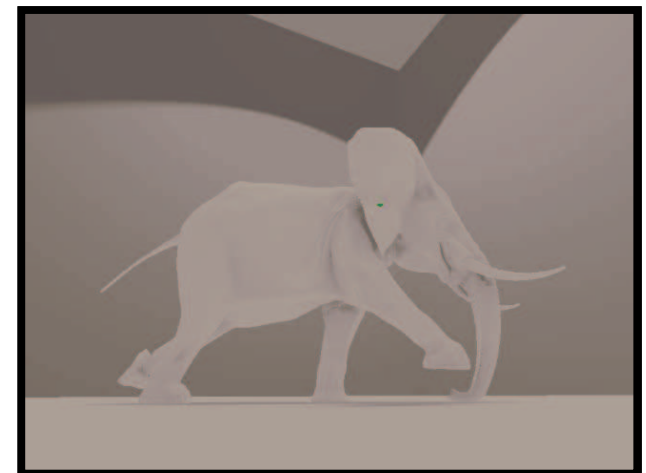No temporal filtering

Predicted distortion map

25%   50%   75%   95%

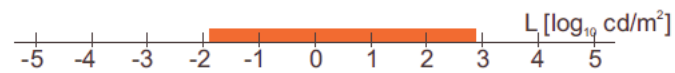# Evaluation of Rendering Qualities

High quality

Low quality

Predicted distortion map

25%   50%   75%   95%
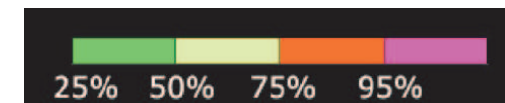
HDR Video

HDR video → Tone mapping → LDR video

Evaluation of HDR Compression

Medium Compression
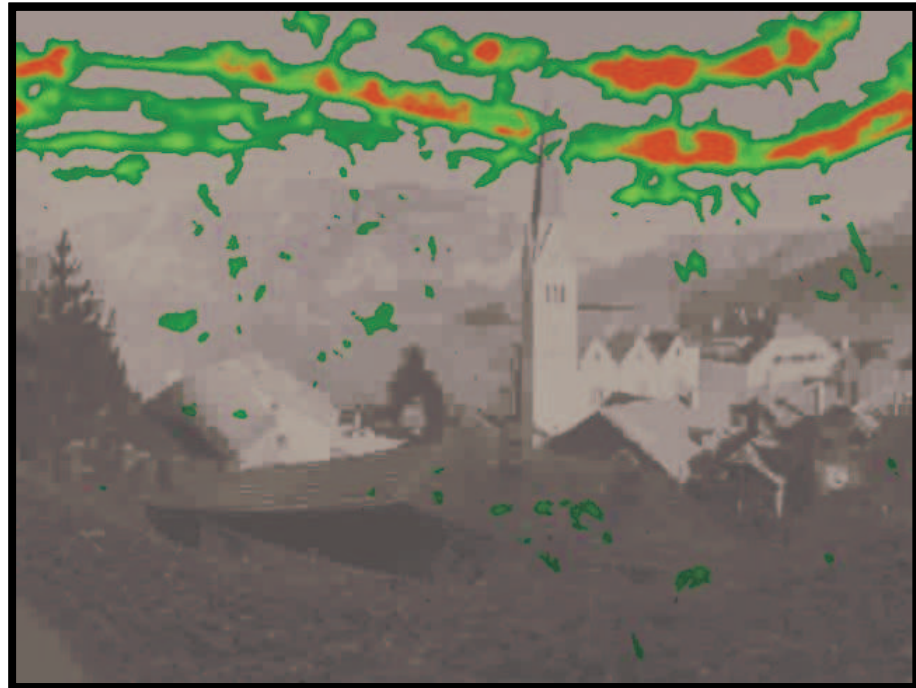
High Compression
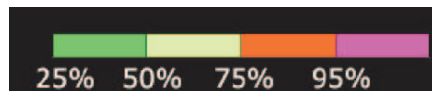
25% 50% 75% 95%

Martin Čadík, mcadik@mpii.de
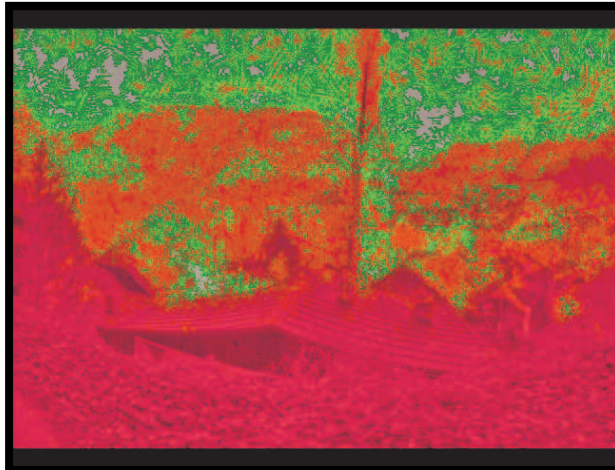SPIE HVEI'11, San Francisco

On Evaluation of Video Quality Metrics:
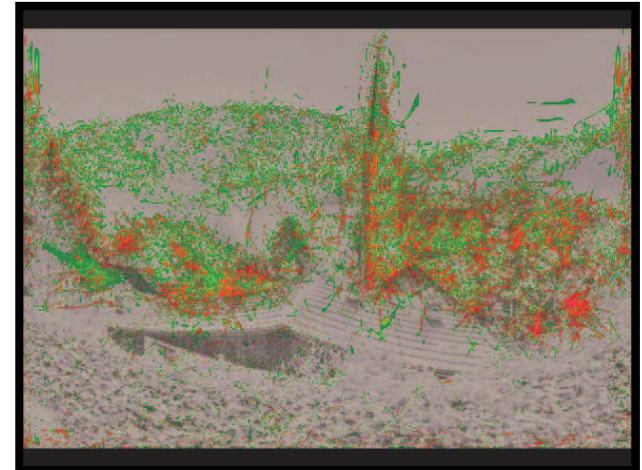an HDR Dataset for Computer Graphics Applications

Evaluation of Tone Mapping


[Fattal et al. 2002]

Detail Amplification

Detail Loss

No suitable dataset available

Proposed Dataset

- LDR-LDR, HDR-HDR, HDR-LDR

- Including subjective distortion maps

- "2.5D videos"

- Temporal noise, HDR video compression, tone mapping

Martin Čadík, mcadik@mpii.de
SPIE HVEI'11, San Francisco

On Evaluation of Video Quality Metrics:
an HDR Dataset for Computer Graphics Applications
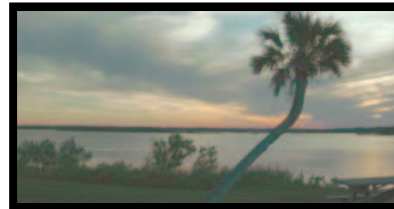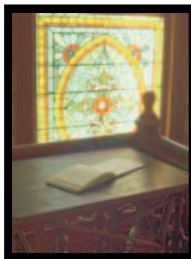
Proposed Dataset (cont.)

- 9 stimuli (reference-test video pairs)
  - 1 LDR-LDR, 2 HDR-LDR, 6 HDR-HDR
  - 60 frames each, 24 fps

- Sub-, Near-, Supra-threshold distortions

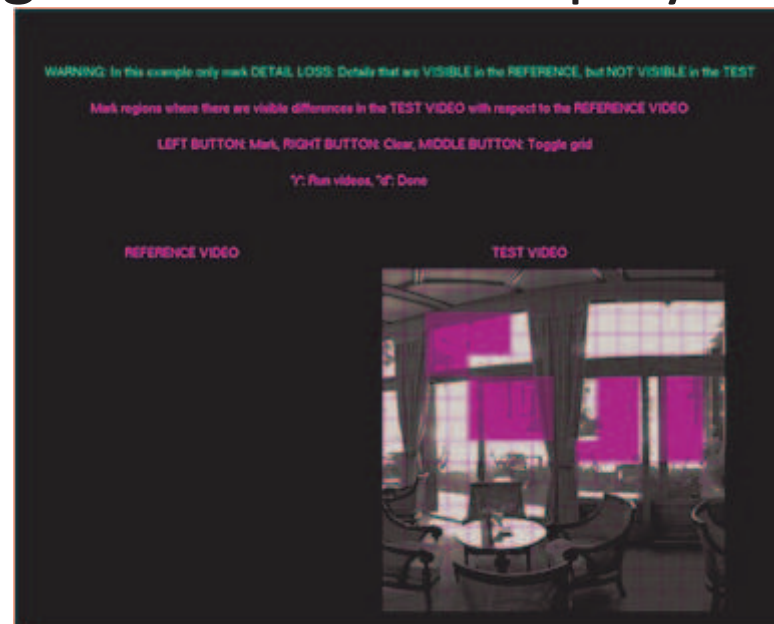- http://www.mpi-inf.mpg.de/resources/hdr/quality
- ~900MB

Subjective Data Acquisition

- 16 subjects, calibrated Brightside DR37-P display

(1) Show videos side-by-side
on a HDR Display

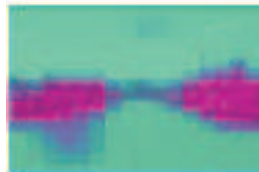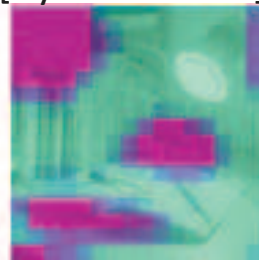(2) Subjects mark regions
where they detect differences

Martin Čadík, mcadik@mpii.de
SPIE HVEI'11, San Francisco

On Evaluation of Video Quality Metrics:
an HDR Dataset for Computer Graphics Applications
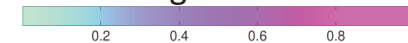
Example Evaluation

| Subj. Response | DRIVQM [Aydin et al.'10] | PDM [Winkler'05] | HDRVDP [Mantiuk al.'05] | DRIVDP [Aydin et al.'08] |

Average distortion

0.2 0.4 0.6 0.8 1

Martin Čadík, mcadik@mpii.de
SPIE HVEI'11, San Francisco

On Evaluation of Video Quality Metrics:
an HDR Dataset for Computer Graphics Applications

Example Evaluation Results - CC

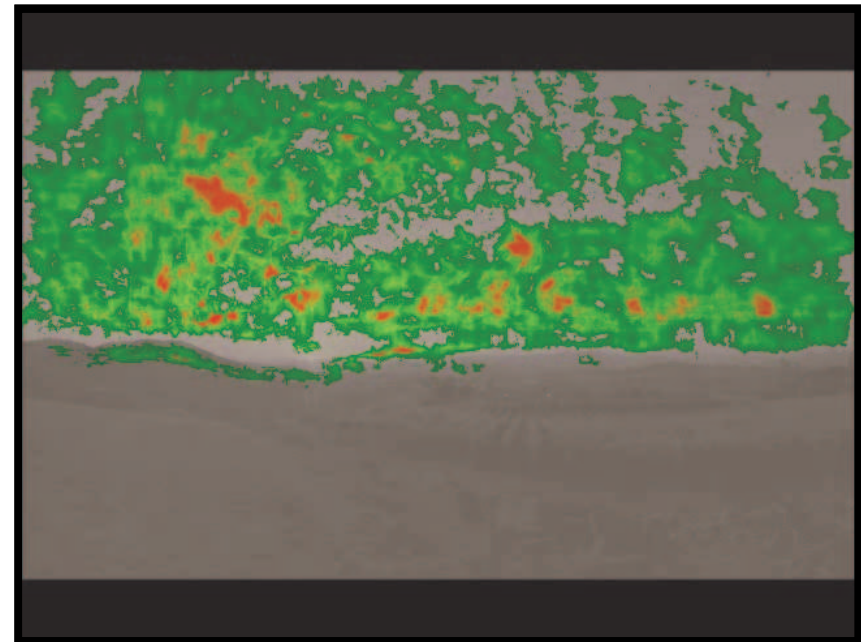| Stimulus		DRIVQM	PDM	HDRVDP	DRIVDP
1		0.765	-0.0147	0.591	0.488
2		0.883	0.686	0.673	0.859
3		0.843	0.886	0.0769	0.865
4		0.815	0.0205	0.211	-0.0654
5		0.844	0.565	0.803	0.689
6		0.761	-0.462	0.709	0.299
7		0.879	0.155	0.882	0.924
8		0.733	0.109	0.339	0.393
9		0.753	0.368	0.473	0.617
Average		0.809	0.257	0.528	0.563

Comparison: Test Scene

- HDR Scene tone mapped with [Pattanaik 2000]

- Spatio-temporal distortion
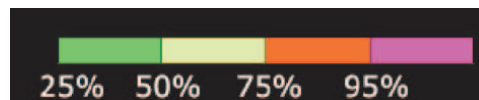 - Random pixel noise filtered with a Gaussian.
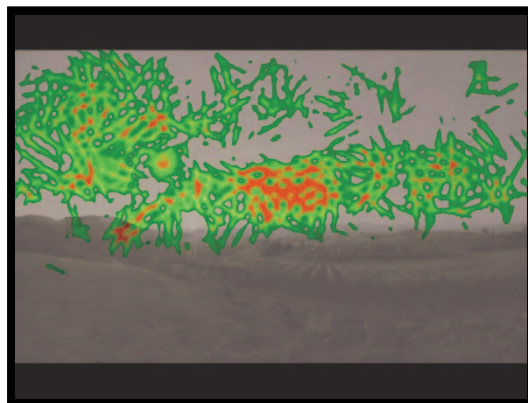
Metric Comparison LDR-LDR

DRIVQM

PDM [Winkler 2005]

25% 50% 75% 95%

Martin Čadík, mcadik@mpii.de
SPIE HVEI'11, San Francisco

On Evaluation of Video Quality Metrics:
an HDR Dataset for Computer Graphics Applications
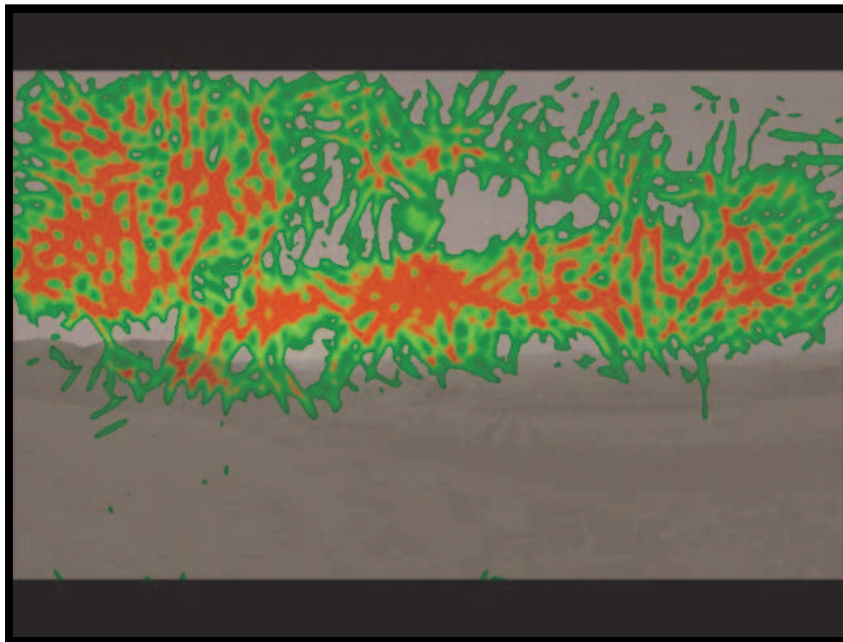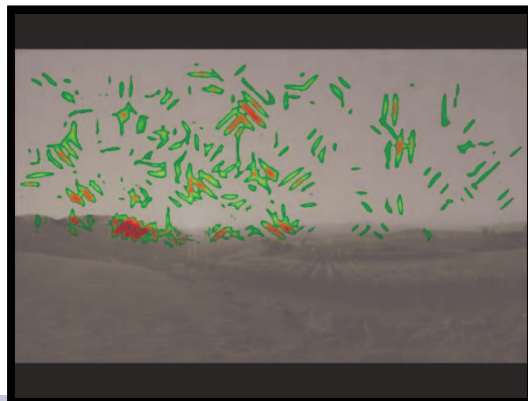
Metric Comparison HDR-HDR

DRIVQM
[Aydin et al.
2010]

PDM
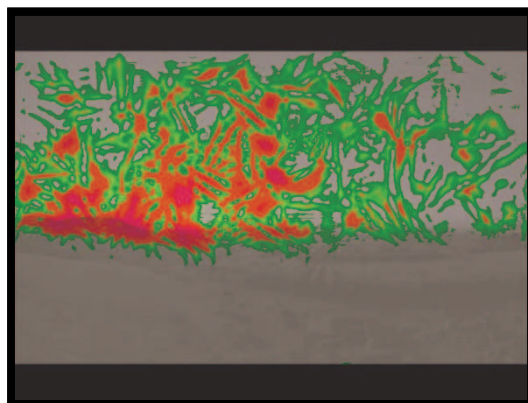[Winkler 2005]

HDRVDP
[Mantiuk
et al 2005]

DRIVDP
[Aydin et al
2008]

Metric Comparison HDR-LDR

DRIVQM
[Aydin et al.
2010]

PDM
[Winkler 2005]
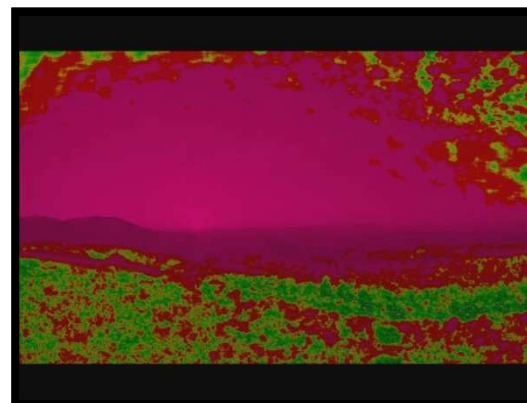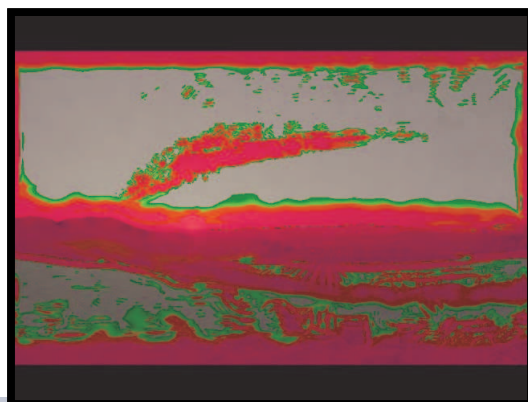
HDRVDP
[Mantiuk
et al 2005]

DRIVDP
[Aydin et al
2008]

Martin Čadík, mcadik@mpii.de
SPIE HVEI'11, San Francisco

On Evaluation of Video Quality Metrics:
an HDR Dataset for Computer Graphics Applications

DRIVQM Online

- http://drivqm.mpi-inf.mpg.de
- Uploaded videos stored only during processing

Martin Čadík, mcadik@mpii.de
SPIE HVEI'11, San Francisco

On Evaluation of Video Quality Metrics:
an HDR Dataset for Computer Graphics Applications
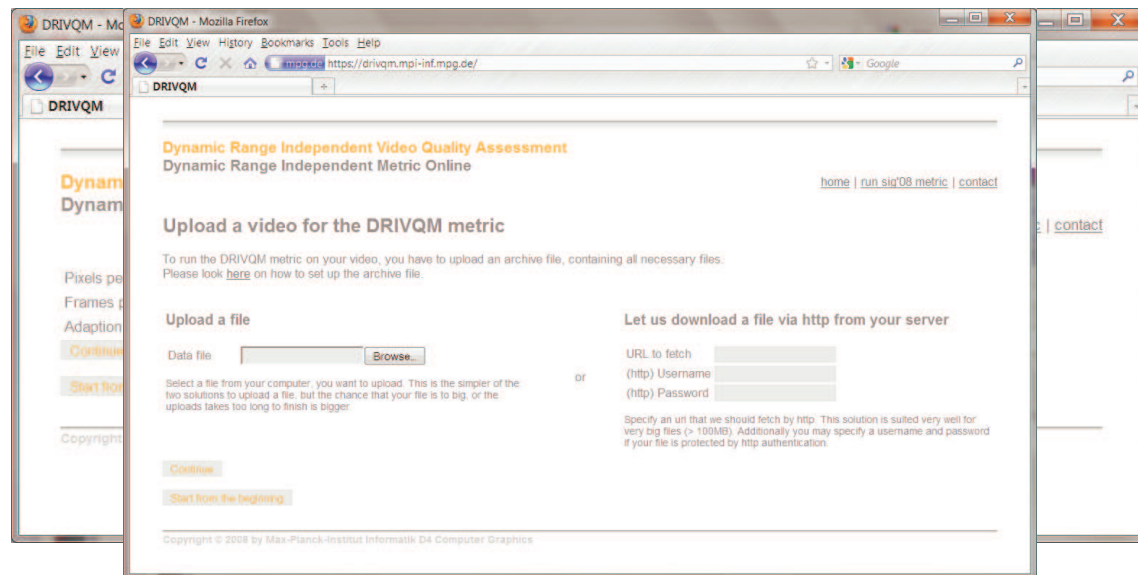
Conclusions

- **Dataset for evaluation of VQMs**
  - LDR-LDR, HDR-HDR, HDR-LDR
  - CG-related artifacts
  - Locally-assessed distortions (distortion maps)

- **Future work**
  - More "advanced" CG artifacts
  - Real videos (how to get subjective distortion maps?)

<http://www.mpi-inf.mpg.de/resources/hdr/quality>

THANK YOU.