HOCHSCHULE DER MEDIEN

Content Aware Quantization: Requantization of High Dynamic Range Baseband Signals Based on Visual Masking by Noise and Texture

2016 IEEE International Conference on Image Processing (ICIP)

Jan Froehlich[1,2], Guan-Ming Su[3], Scott Daly[3],
Andreas Schilling[1], Bernd Eberhardt[2]

[1] Stuttgart Media University, [2] University of Tübingen, [3] Dolby Laboratories Inc.

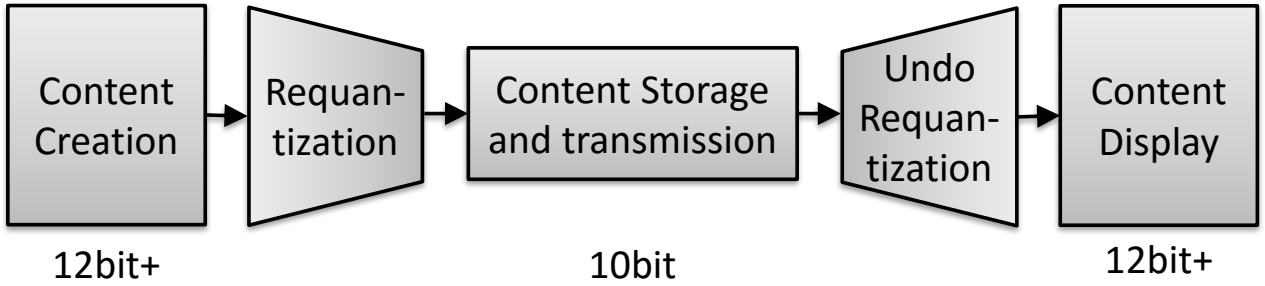
Outline of the Talk

HOCHSCHULE DER MEDIEN

1. Motivation
   - HDR Ecosystem
2. Fundamental concepts
   - Noise & Texture vs. needed quantization step
3. Methods
   - Prediction Kernel
   - Calibration
4. Intended Limitations
   - Flare
5. Results
   - Qualitative
   - Quantitative
6. Conclusion

Motivation

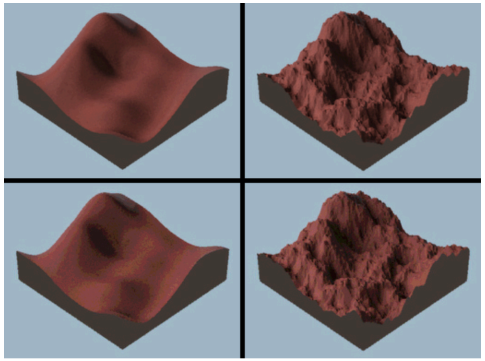
- 0.0005-10000 cd/m^2 zero noise HDR imagery needs 11-12 bits of tonal resolution per color channel for visually lossless quantization*
- Most current video file formats, compression codecs and transmission interfaces are limited to 10 bits of tonal resolution in their mainstream flavors



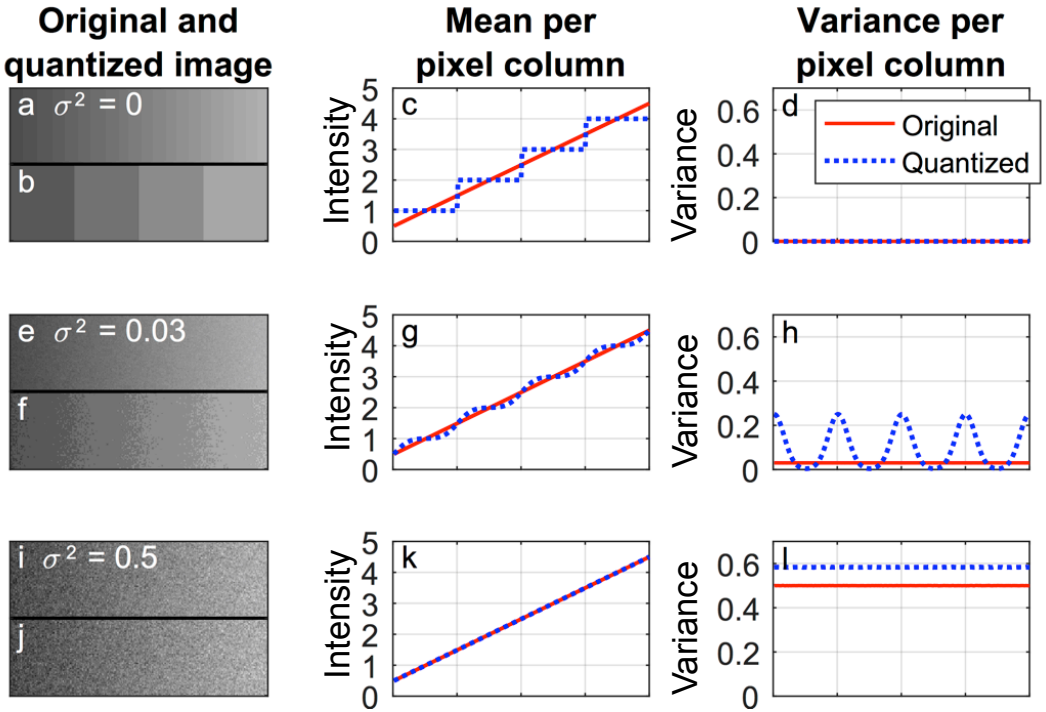
*SMPTE ST.2084 / ITU Rec. BT.2020 / S. Miller, M. Nezamabadi and S. Daly, "Perceptual Signal Coding for More Efficient Usage of Bit Codes," *Annual Technical Conference & Exhibition, SMPTE 2012*, Hollywood, CA, USA, 2012, pp. 1-9.

Fundamental Concept

- Exploit masking of quantization artifacts by noise and texture



Ferwerda, James A., et al. "A model of visual masking for computer graphics." *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 1997.



Methods:
Quantization Study Pattern and Parameters

HOCHSCHULE DER MEDIEN

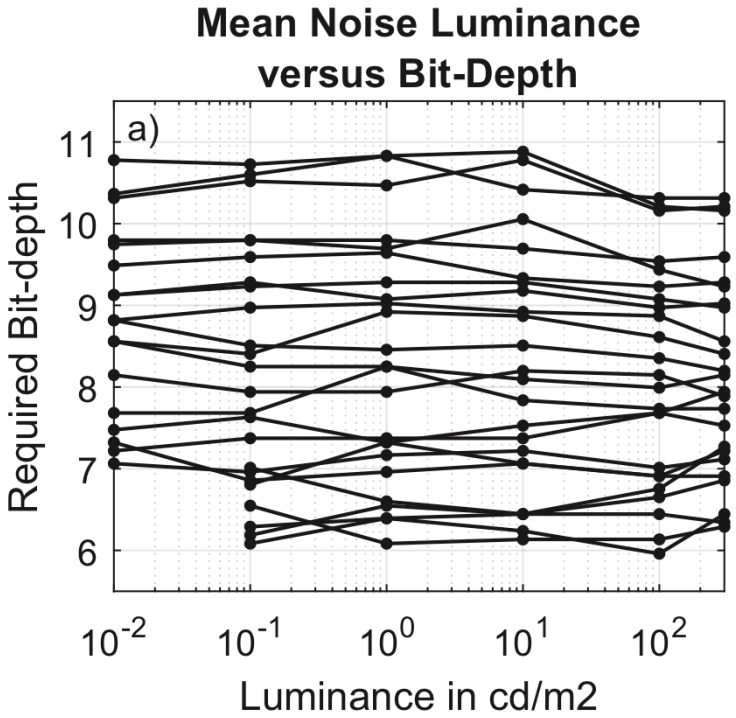
- Use smooth gradients with different slope and orientation as most critical pattern for quantization artifact visibility
- Noise parameters varied:

Mean Luminance	0.01, 0.1, 1, 10, 100, 300cd/m^2
Temporal frequency	0fps (still image), 24fps
Spatial bandwidth	20, 10, 5 cycles per degree
Amplitude	0, 1, 2, 4, 8, 16, 32, 64 standard deviation σ in 12 bit code-values
Quantization (tonal resolution)	q = 5 to 12 for 2^q code values to encode the full PQ range

Quantization Study Results: Luminance

HOCHSCHULE DER MEDIEN

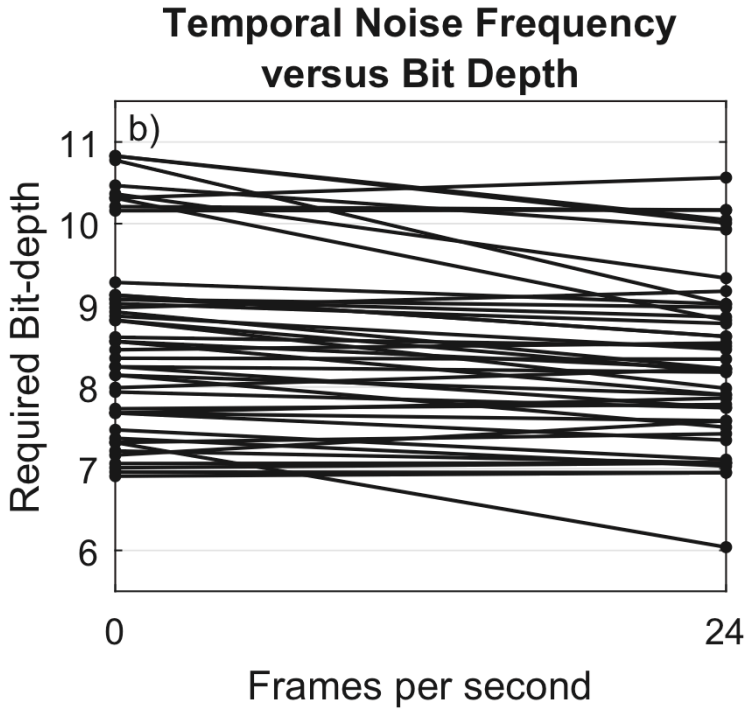
- No strong correlation between luminance and required bit-depth.
- This also confirms the perceptual uniformity of the 'Perceptual Quantizer' (PQ) encoding curve which has been found to deliver a better match to low amplitude visibility compared to previous models like 'log' or 'gamma'.



Dots connected by lines only vary in the parameter on the x-axis

Quantization Study Results: Temporal Frequency

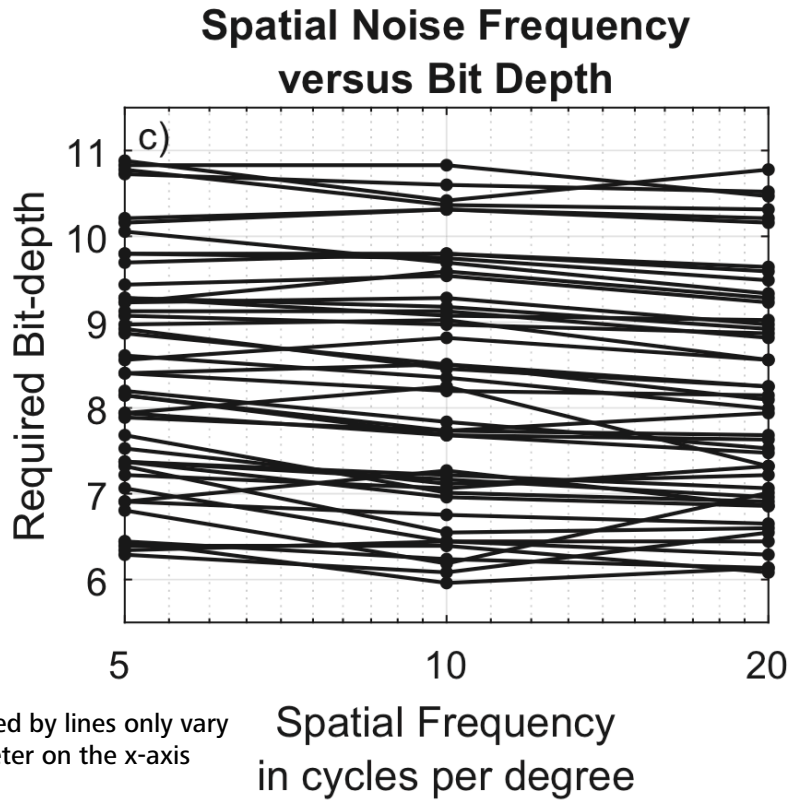
- No strong correlation between the temporal frequency of noise or texture and the required bit-depth.
- Only static images and 24 frames per second were studied.



Dots connected by lines only vary in the parameter on the x-axis

Quantization Study Results: Spatial Frequency

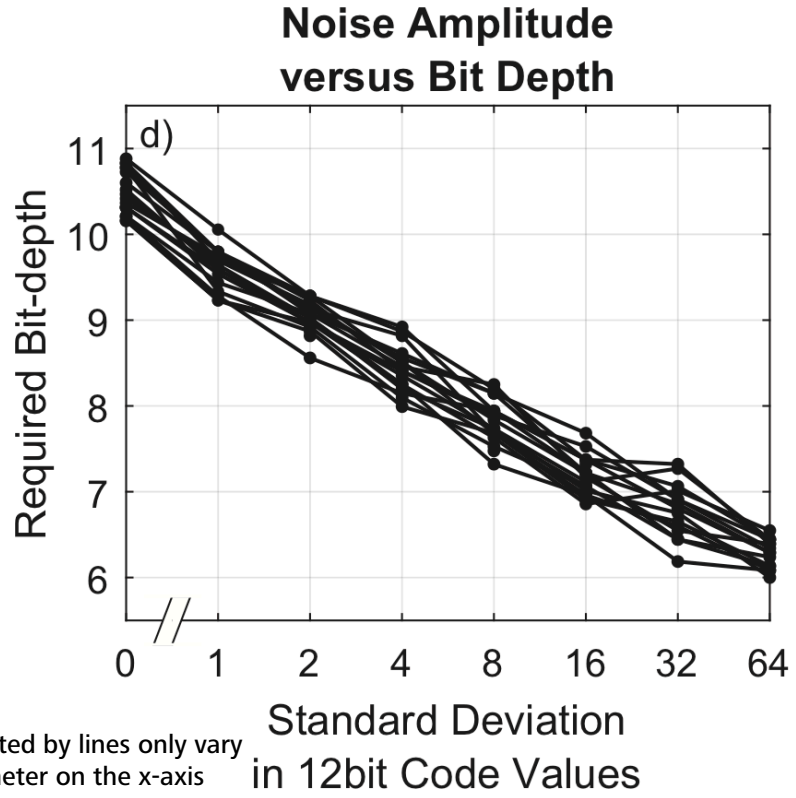
- No strong correlation between the spatial frequency of noise or texture and the required bit-depth.



Dots connected by lines only vary in the parameter on the x-axis

Quantization Study Results: Noise Amplitude

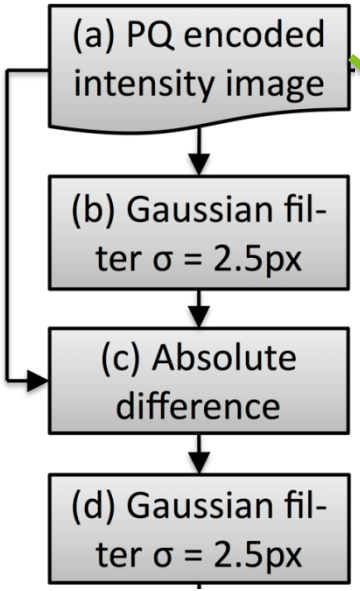
- *Strong* correlation between the amplitude of the noise or texture and the needed bit depth.
- We designed a re-quantization method to exploit this relationship between noise amplitude and needed quantization.
- Our method (CAQ) can reduce bit-depth requirements for HDR images.



Dots connected by lines only vary in the parameter on the x-axis

Methods: Content Adaptive Quantization (CAQ) Block Diagram

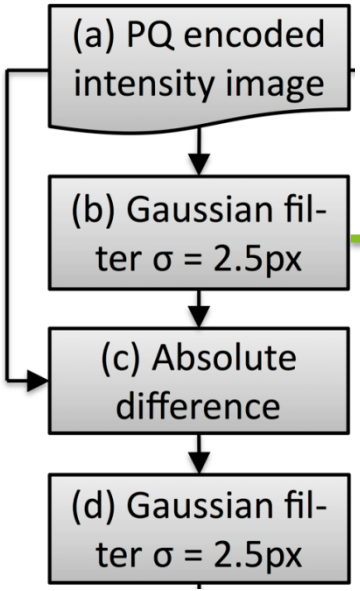
- Calculate intensity image and convert to PQ domain.



Carousel Image from ‚HdM-HDR-2014' data set

Methods: Content Adaptive Quantization (CAQ) Block Diagram

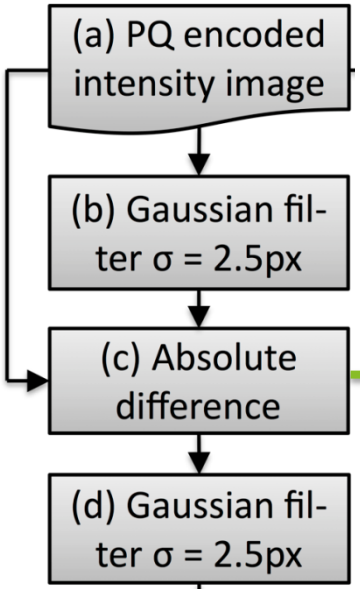
- Low pass filter.

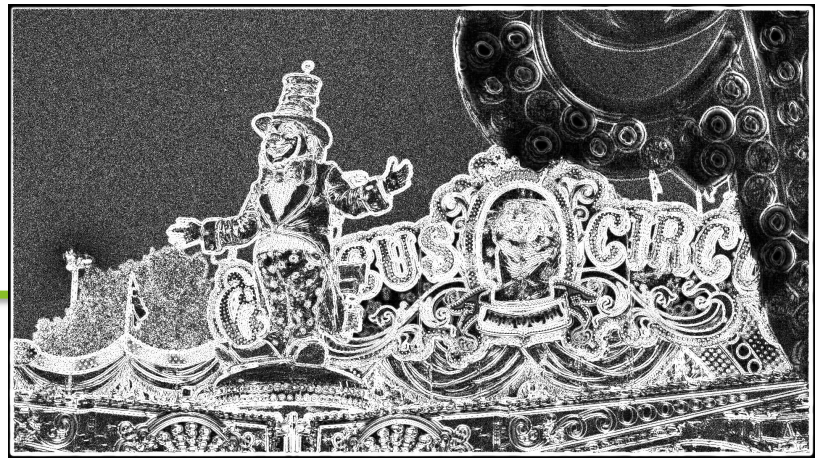


Carousel Image from ‚HdM-HDR-2014' data set

Methods: Content Adaptive Quantization (CAQ) Block Diagram

- Calculate sum of absolute differences (SAD).

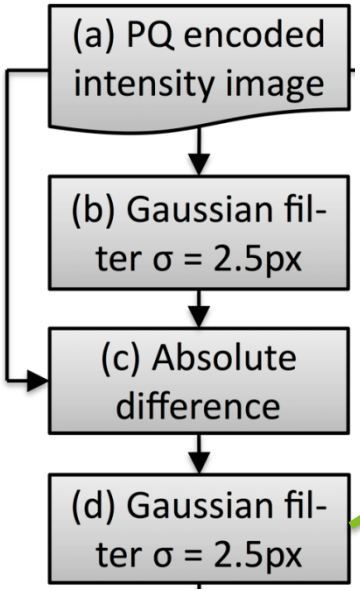


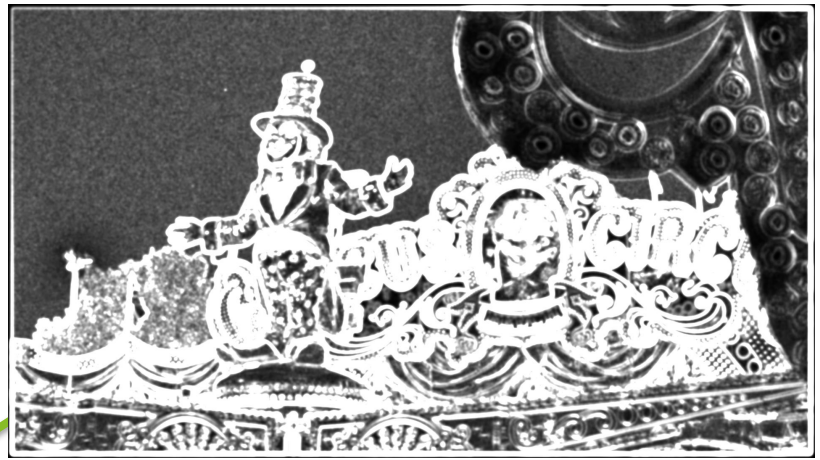


50* amplified for visualization

Methods: Content Adaptive Quantization (CAQ) Block Diagram

- Low pass filter again to increase robust-ness and simulate local masking of the human visual system.



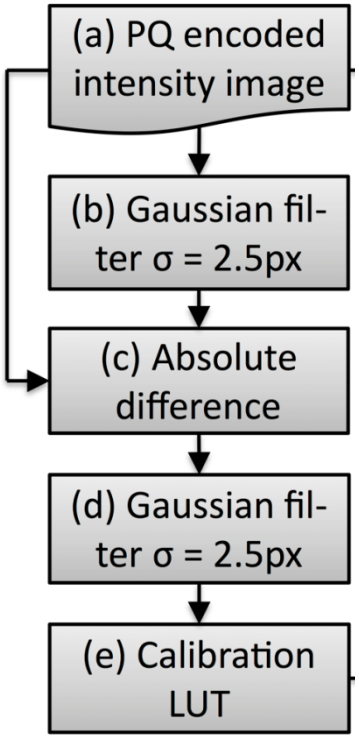


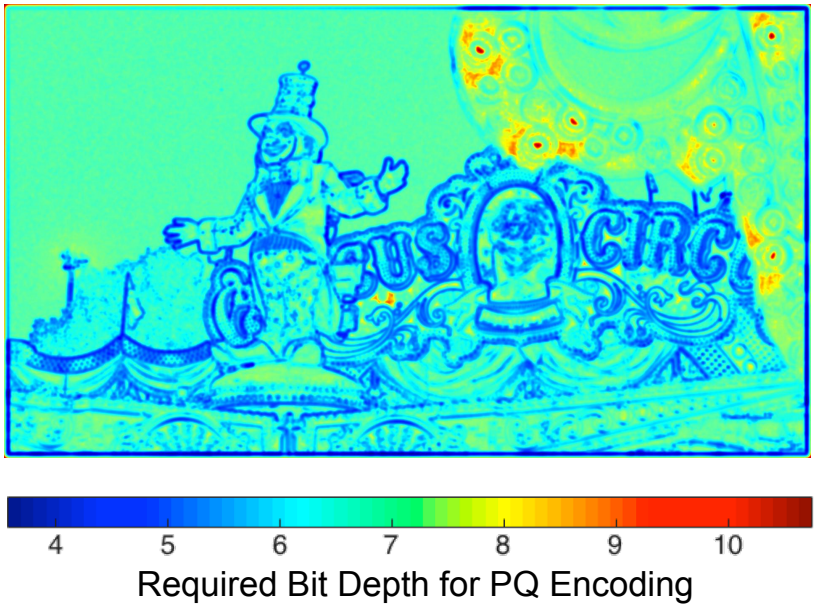
50* amplified for visualization

Methods: CAQ Block Diagram

HOCHSCHULE DER MEDIEN

- Apply calibration look up table to obtain needed bit depth values per pixel from map of presence of high frequencies



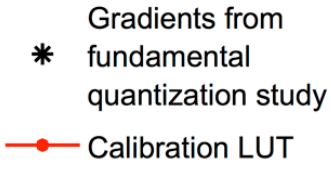


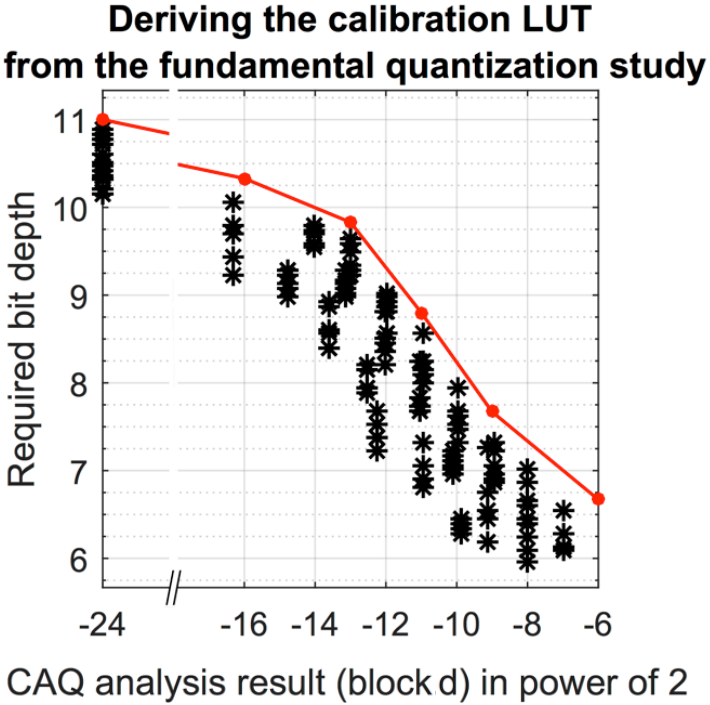
Methods: CAQ Block Diagram - Calculation of the Calibration LUT

HOCHSCHULE DER MEDIEN

- Calibration lookup table predicts the needed bit-depth for each value from the high-pass-filter
- Trained on the images of the fundamental quantization study
- Each CAQ filter result from block (d) is assigned the minimum needed bit-depth for visually lossless quantization

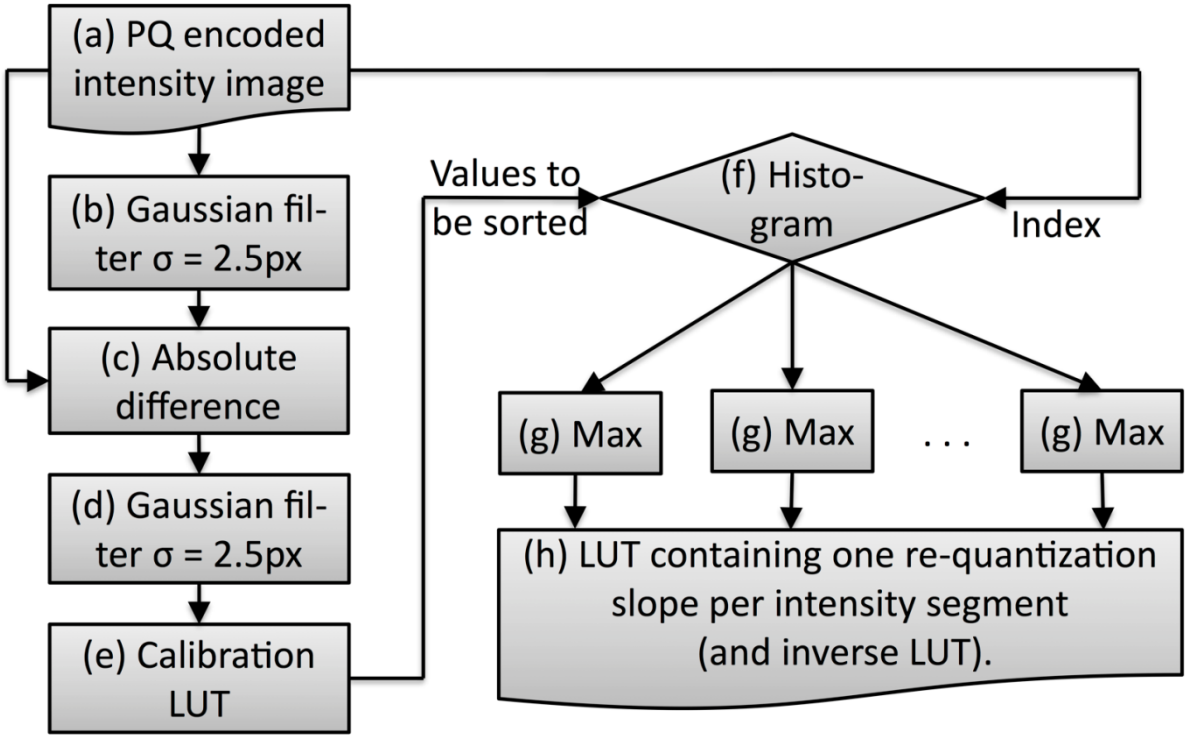
Training set: Fundamental quantization study pattern





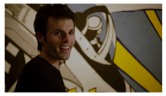
Methods: CAQ Block Diagram

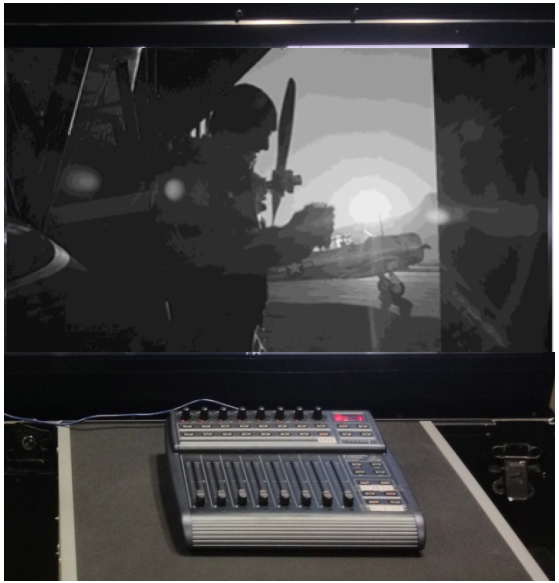
- Typical images contain more noise in the dark areas (photon shot noise).
- The spatial quantization map can be used to calculate a luminance dependent re-quantization LUT.
- This LUT reduces the needed code values per ***intensity range*** opposed to the spatial map from block (e).
- Calculation can be done per frame or shot.



Results and Verification: Test Sequences for the Verification Study

HOCHSCHULE DER MEDIEN

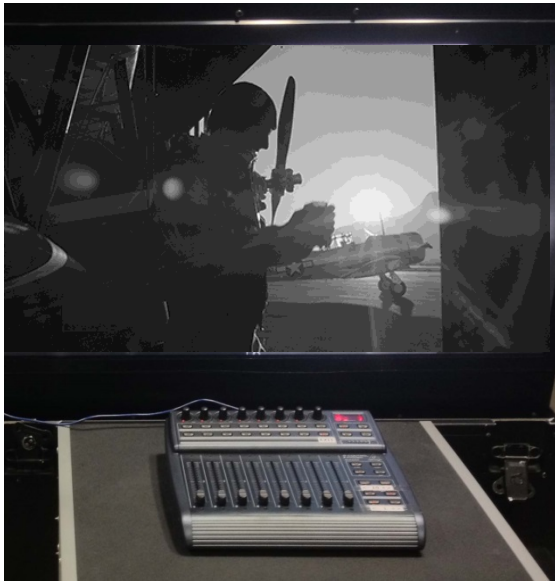
Scene Name *Camera*	Acquisition Medium	Image Description	Thumb-nail
a) Hangar *ARRI Alexa*	ARRIRAW 2.8K	View from Hangar into the sun with a pilot's silhouette.	
b) Fantasy Flight *ARRI Alexa*	ARRIRAW 2.8K	Man standing in front of a painting.	
e) 2009 Kids Film *Toon Boom Harmony*	CG Animation, Rendering, 2K	Dark animated jungle illuminated by fireflies	
f) 2006 A-Movie *Panavision Mill. XL2*	Kodak Vision2 250D, 500T	Sorcerer on stage illuminated by blue searchlights	
g) Flirting with Fire *ARRI Alexa*	Phantom 4K Flex	Explosive Flame / Fireball	
h) Showgirl *2*ARRI Alexa*	2*ProRes 4:4:4 HD	Girl illuminated by directional stagelight	



Results and Verification: Verification Study

- *Surround:* dark room (minimum veiling glare)
- *Display Resolution:* 1920 x 1080 pixel
- *Viewing distance:* 1 picture height
- *Study participants:* 8 expert viewers who perform image evaluation tasks every day.
  - All Participants had 20/20 vision, 3 without eyesight correction, 4 with glasses, 1 with contact lenses
- *Study Task:* Method of adjustment exploiting the pop-out effect of motion when phase-shifting quantization.

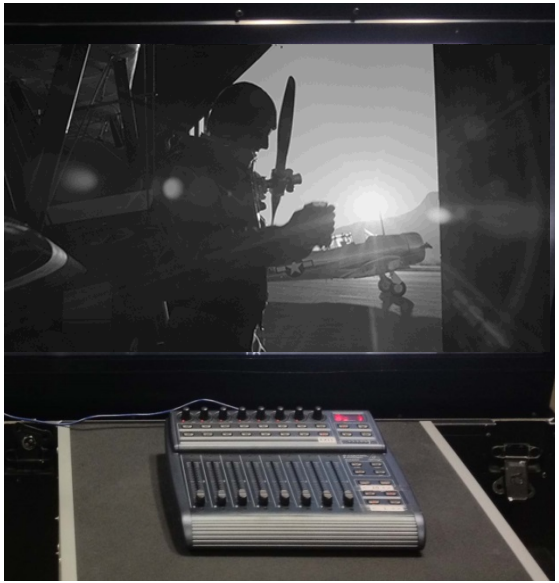
  (Spatio-temporal worst case scenario)



Results and Verification: Verification Study

- *Surround:* dark room (minimum veiling glare)
- *Display Resolution:* 1920 x 1080 pixel
- *Viewing distance:* 1 picture height
- *Study participants:* 8 expert viewers who perform image evaluation tasks every day.
  - All Participants had 20/20 vision, 3 without eyesight correction, 4 with glasses, 1 with contact lenses
- *Study Task:* Method of adjustment exploiting the pop-out effect of motion when phase-shifting quantization.

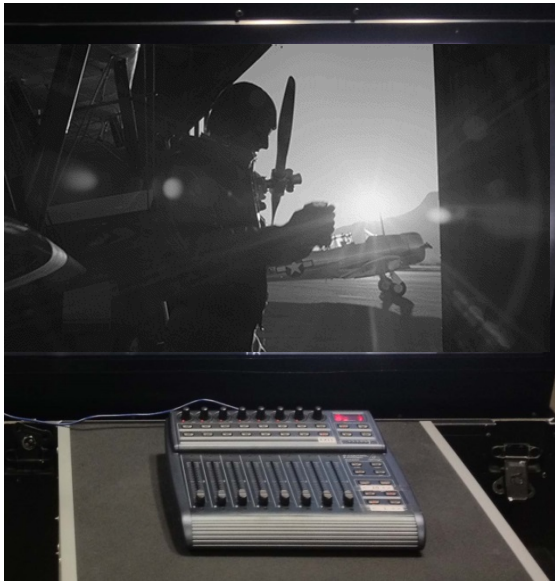
  (Spatio-temporal worst case scenario)



Results and Verification: Verification Study

- *Surround:* dark room (minimum veiling glare)
- *Display Resolution:* 1920 x 1080 pixel
- *Viewing distance:* 1 picture height
- *Study participants:* 8 expert viewers who perform image evaluation tasks every day.
  - All Participants had 20/20 vision, 3 without eyesight correction, 4 with glasses, 1 with contact lenses
- *Study Task:* Method of adjustment exploiting the pop-out effect of motion when phase-shifting quantization.

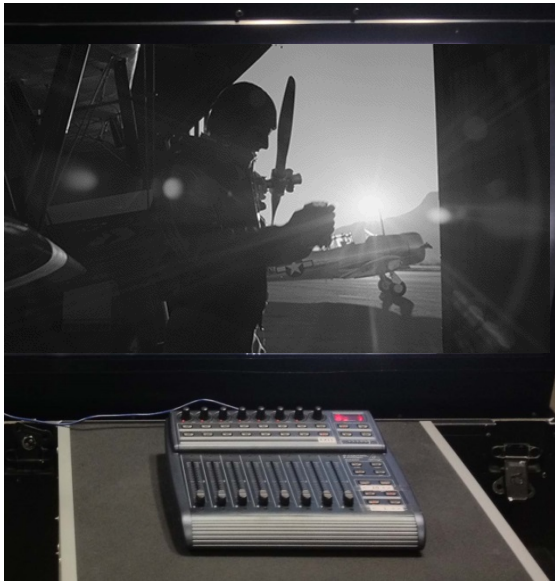
  (Spatio-temporal worst case scenario)



Results and Verification: Verification Study

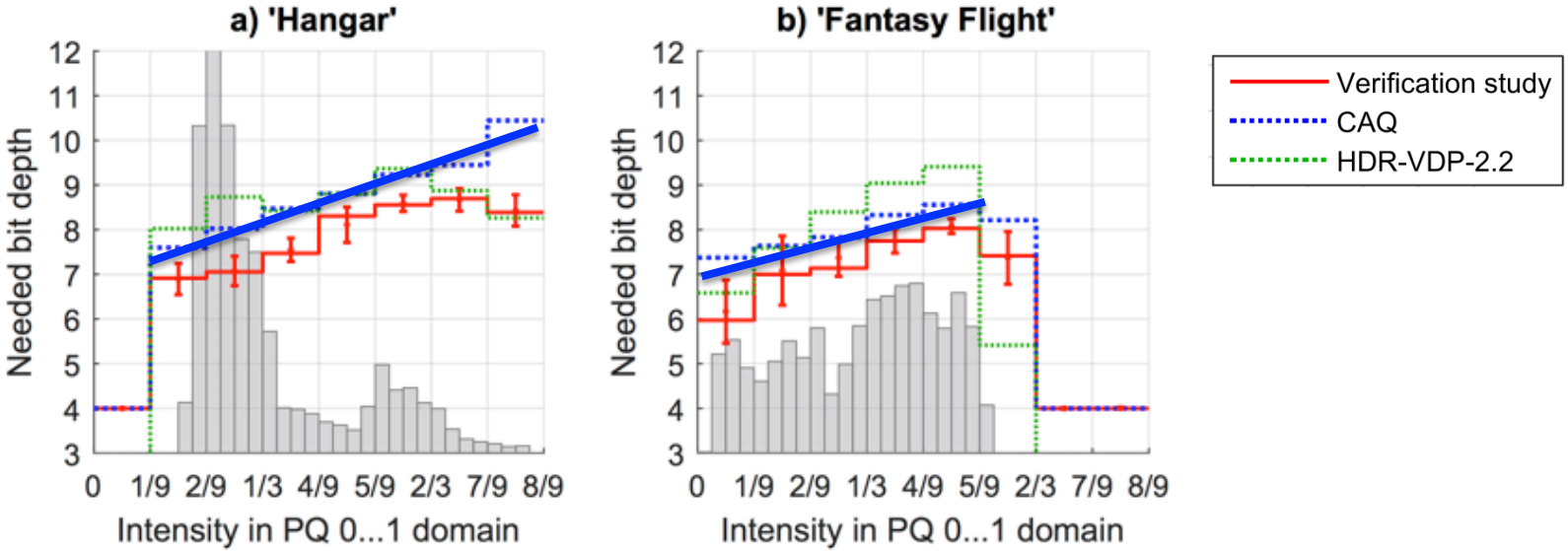
- *Surround:* dark room (minimum veiling glare)
- *Display Resolution:* 1920 x 1080 pixel
- *Viewing distance:* 1 picture height
- *Study participants:* 8 expert viewers who perform image evaluation tasks every day.
  - All Participants had 20/20 vision, 3 without eyesight correction, 4 with glasses, 1 with contact lenses
- *Study Task:* Method of adjustment exploiting the pop-out effect of motion when phase-shifting quantization.

  (Spatio-temporal worst case scenario)



Results:

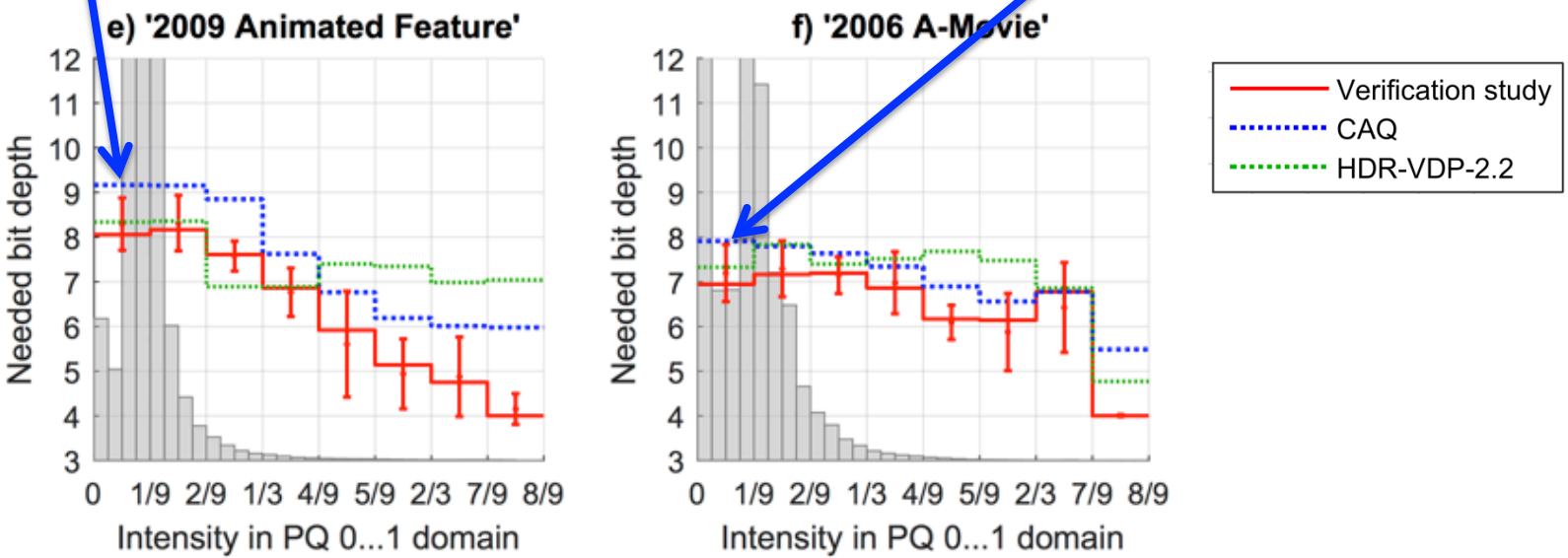
- Needed quantization for images captured by digital cameras is typically limited by photon shot noise: (relatively more in dark areas)



Results:

HOCHSCHULE DER MEDIEN

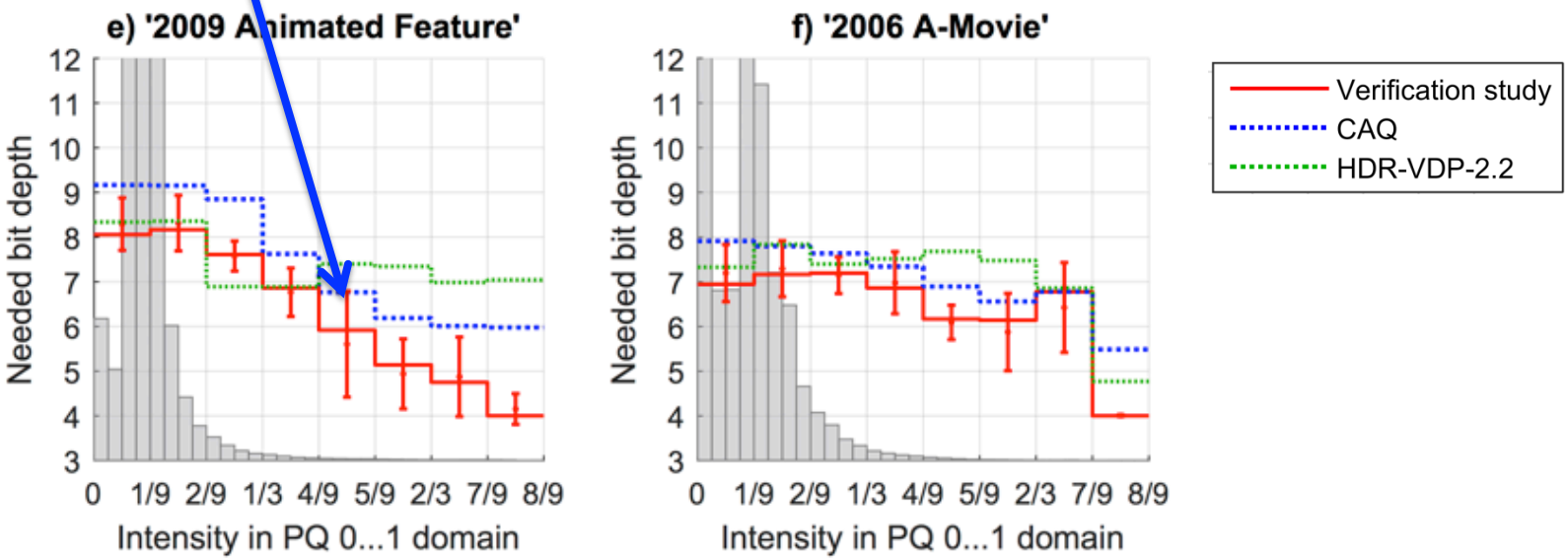
- Animated Content (e) typically needs much higher bit depths compared to content originated on analog film (f)



Results:

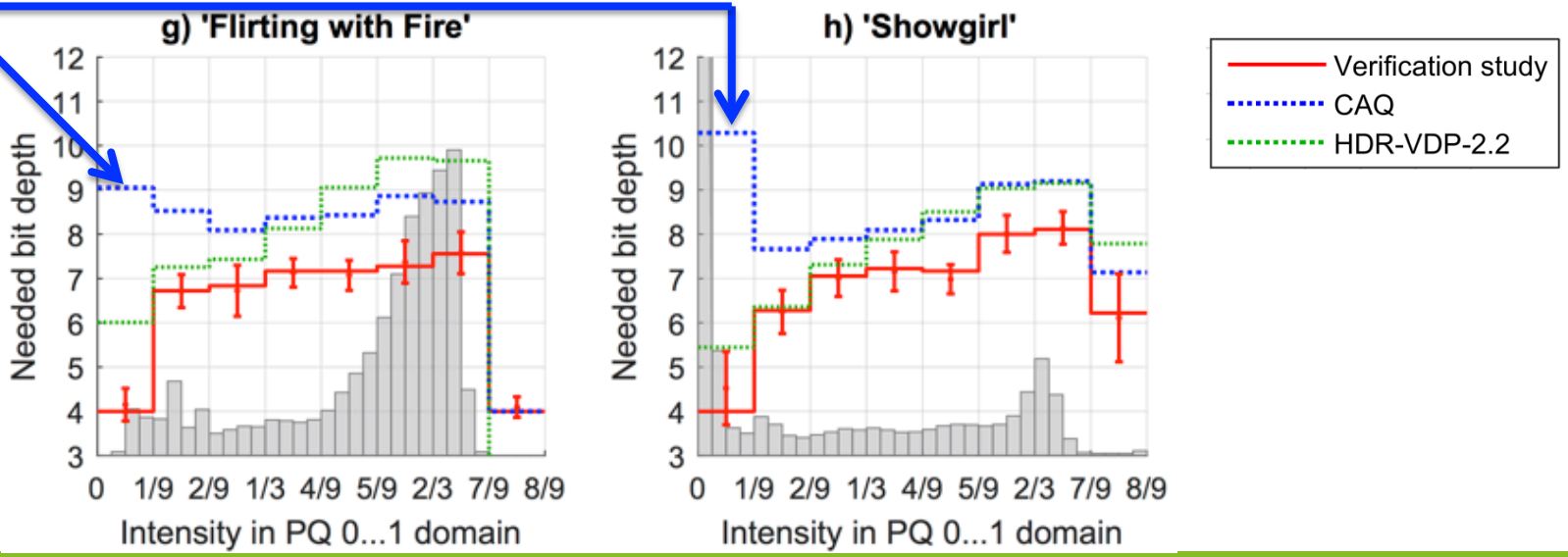
HOCHSCHULE DER MEDIEN

- Small objects can also be quantized coarser:



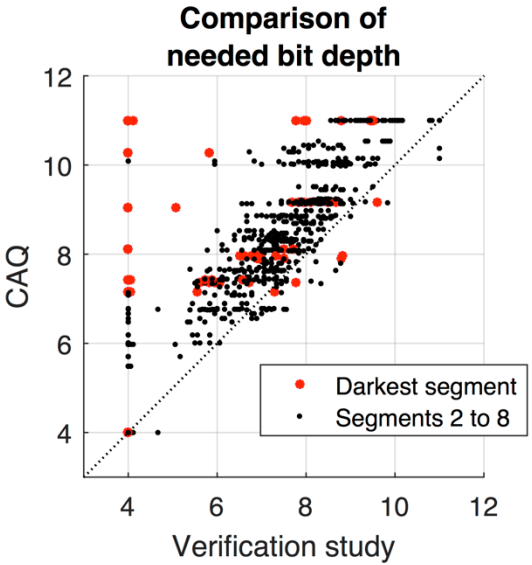
Limitations:

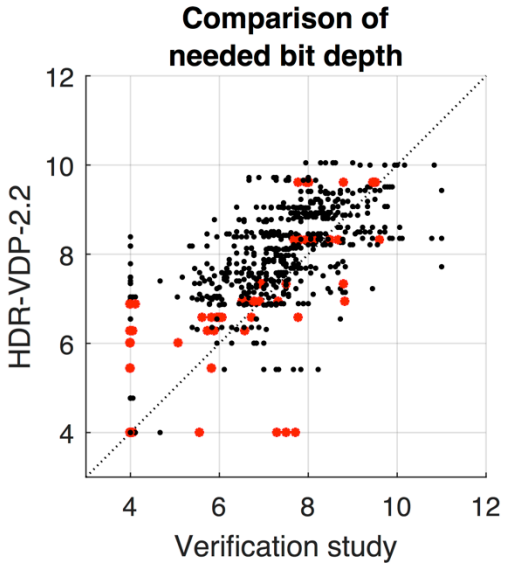
- Flare is not detected by CAQ (by design).
- HDR-VDP 2.2 *does* detects flare.
- For our application, reducing bit-depth based on flare is unwanted because possible sources of flare may be cropped before viewing.



Results: All Test Sequences

- For our specific application CAQ prediction has a higher correlation compared to using HDR-VDP 2.2 for predicting needed quantization.
- Spearman rank order correlation:
  - CAQ: 0.78
  - HDR-VDP 2.2: 0.74
- CAQ analysis runs nearly 10^3 times faster compared to HDR-VDP 2.2.





Conclusion

- We presented CAQ - a fast method for re-quantizing images.
- CAQ can quantize most HDR image sequences with less than 10 bits without introducing visible quantization artifacts by means of a one dimensional lookup table.
- CAQ provides a spatial map of needed bit depth per pixel. This can open up new applications like quantizing at even lower bit-depths or to apply dithering only locally where needed.