HDR Image Quality in the Evolution of JPEG XT

Thomas Richter
University of Stuttgart
IZUS/TIK

Motivation: Why JPEG XT

- 10918/T.81 is still the dominant standard for photographic images
- An entire toolchain exists to record, manipulate and display images encoded in this specification
- Market penetration of JPEG 2000 and JPEG XR is very low in the consumer market.
- JPEG XT is a backwards compatible extension of JPEG, addressing its major deficits:
    - Only lossy 8 bits/component
    - Other coding modes exist, but are not compatible to the 8 bit mode and never became popular.

Which images to consider?

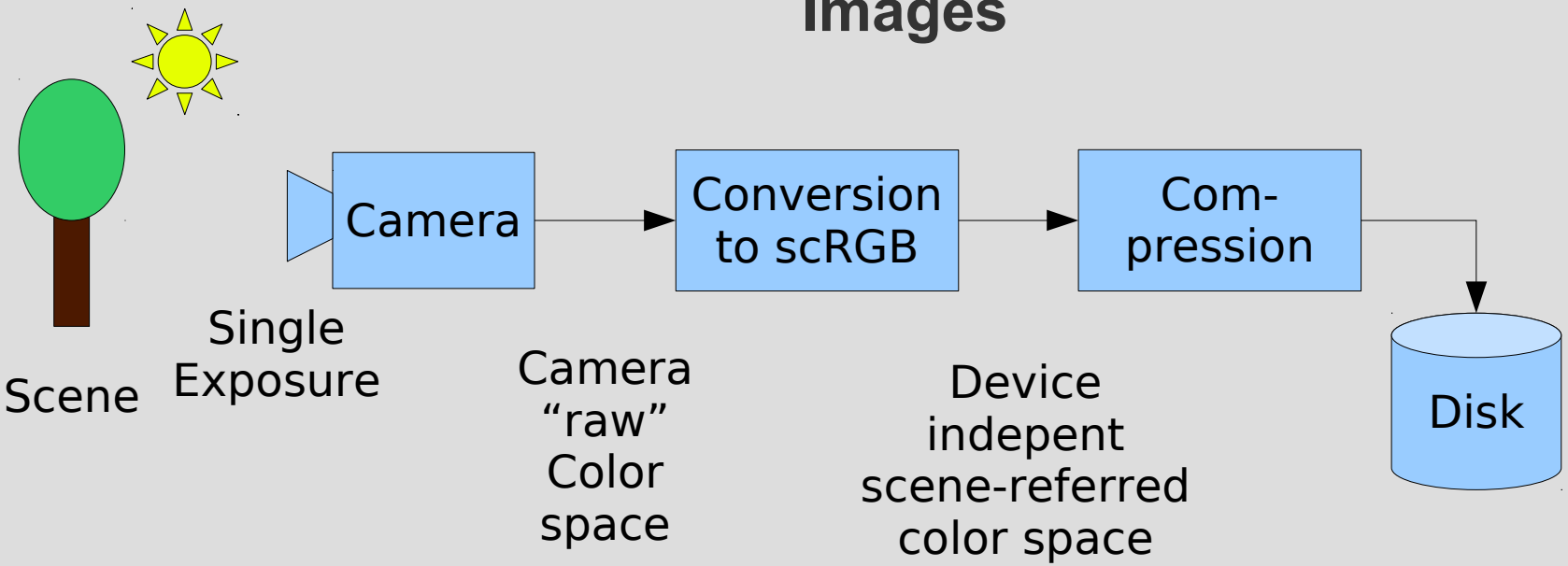
- Images requiring a higher bitdepth
  - More than 8 bits per sample.
- Images represented in floating point
  - Also typically 16 bits/sample, but in a half-logarithmic representation.
- Multi-Exposure images
  - Multiple images with varying exposure, combined into a single image?

Use Cases and Workflows

- Use “use cases” and “workflows” as definition rather than technical parameters
- Use cases can be roughly classified according to the capture and reproduction process
- Not an exhaustive list...

	Reproduction	
Accquisition		

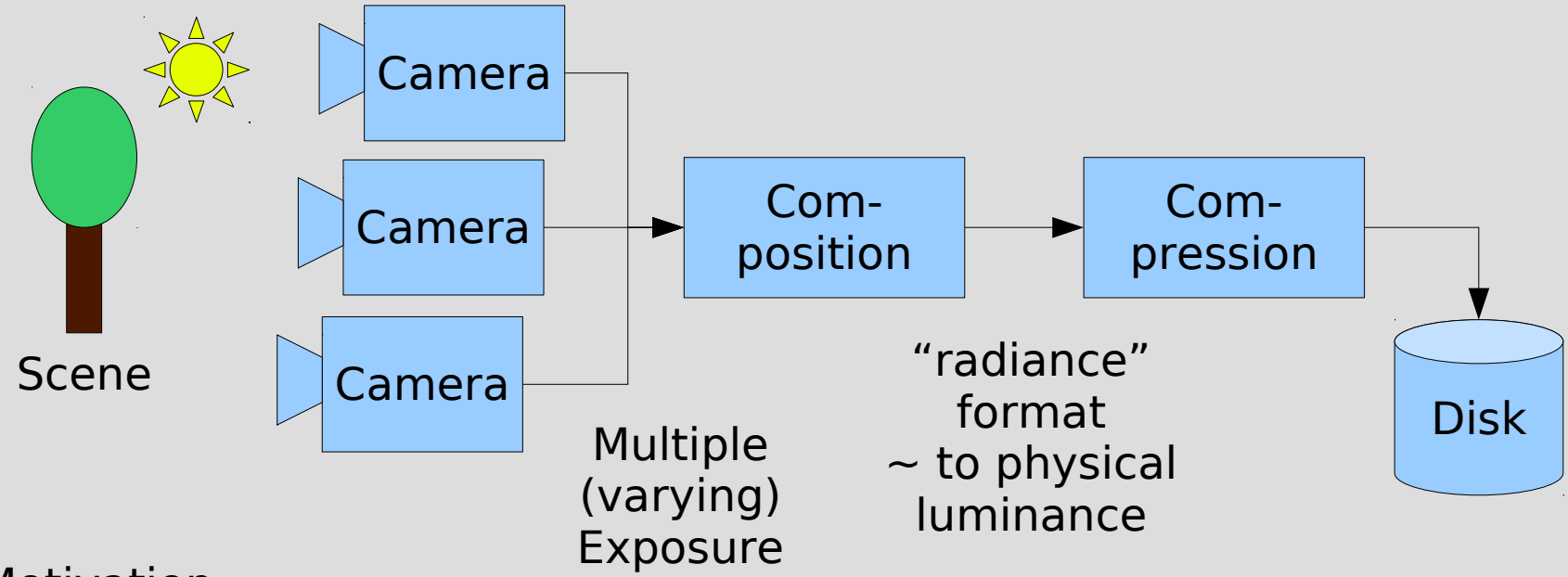
Capturing - Use Case 1: Archival of (Single Exposure) Images



Motivation:

- Archive images independent of the camera model and maker
- Camera color spaces often unknown
- Image decoding depends on proprietary tools of the manufacturer
  - Life time of such tools is limited by the life-cycle of the camera and/or the operating system they run on.
- **High bit depth required for high-precision**
  - Do not introduce loss by quantizing the colors to 8bpp as in traditional JPEG
- Images are processed ("developed") later
  - Tone mapping as final step.

Capturing - Use Case 2: Artistic Photography

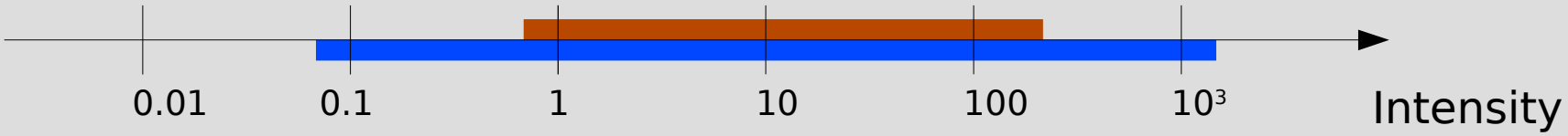


Motivation:

- Artistic “alienation” of natural scenes, artitistic expression
- Multiple captures with varying exposures, merged into a single image by proprietary tools
- Output is typically represented in floating point (e.g. exr or pfm)
- Pixel values might be calibrated to physical luminance (“radiance format”)
- **High bit depth required to represent high variation in luminance**
- Images are later tone mapped for the final application.

Dynamic Range vs. Precision

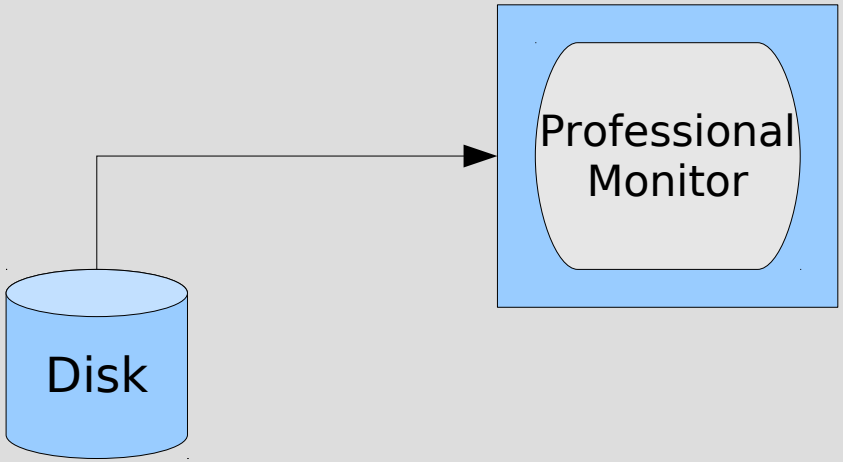
- Both single and multiple exposure require a higher bit depths than 8 bits/sample
  - e.g. 16 bits/sample



- But the dynamic range covered by them is different
- Single exposure case requires 16 bits for **higher precision**
- Multiple exposure case requires 16 bits for **larger range**
- **Test protocols must be different**

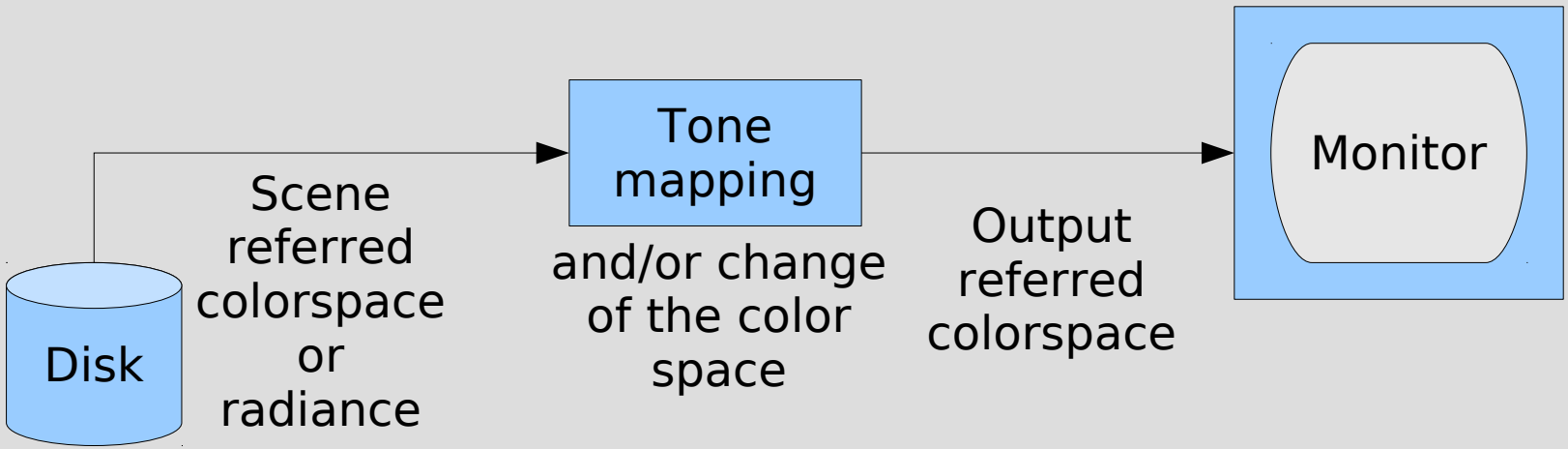
Don't confuse precision with dynamic range.

Direct Reproduction – Use Case 1



- The easy case: The output device supports the color space of the images
- Requires high quality, specialized output devices
- And a common colorspace of archived images and reproduction device
- (In-frame) dynamic range is higher than that of consumer devices, but does typically not cover the range covered by the **multi-exposure** case.
- Typically only applicable in the **high precision** case.

Reproduction on LDR Equipment – Use Case 2



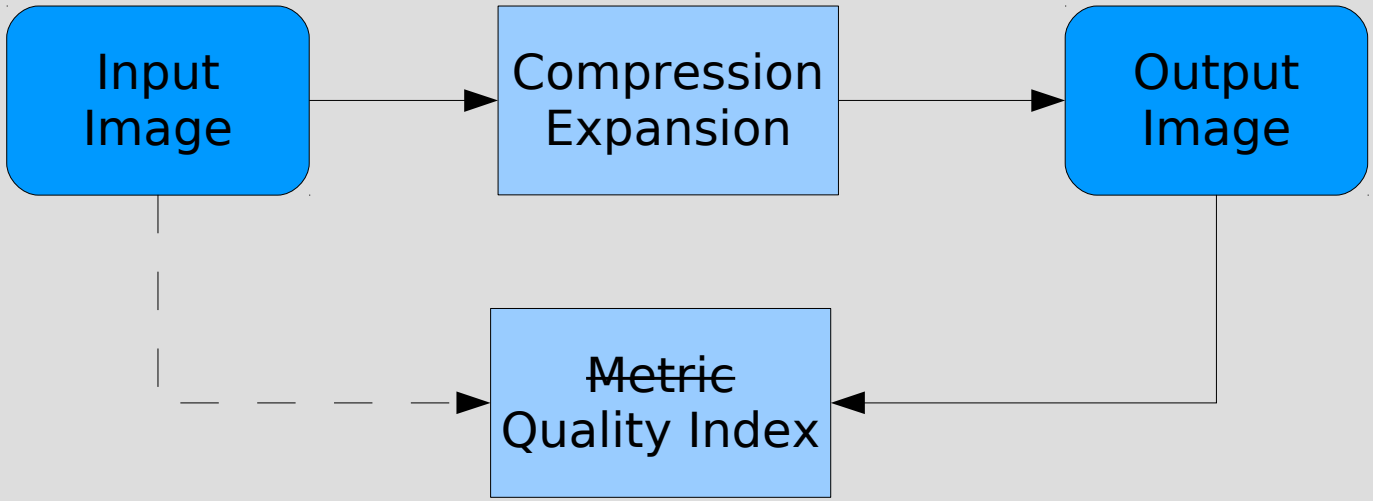
- Images are transformed from the scene referred colorspace to the output referred colorspace (e.g. sRGB).
- Reduction of the dynamic range by tone mapping.
- Output colorspace is typically 8 bit/sample.
- In medical applications, “windows” into the luminance range of the images are selected by the user.

Matrix of Use Cases

	Direct Reproduction	Rendering on Low Dynamic Range Equipment
High bit precision	Single exposure, professional equipment end to end	Single exposure, professional photography, one to many, e.g. web
Floating Point Input	Probably not meaningful?	Artistic photography, alienation of images

- Each entry in the table requires a separate test protocol (to be seen)
- There are probably more cases not covered here...

Metrics: Direct Measurements



- The known approach
- Feasible for top-left entry of the matrix
- Subjective test protocols are well studied and understood.
- Many objective ~~metrics~~ quality indices known
  - Full reference: PSNR, SSIM, VDP...
  - Reduced reference, no reference...
- Unfortunately, most quality indices are only suitable for LDR
  - SSIM and variations are only calibrated to 8 bits/component
- HDR-VDP [Mantiuk, Myszkowski, Seidel 2004] is a notable exception.
- PSNR can also be defined for multi-exposure/float, but...

Why is PSNR not feasible in the multi-exposure case
Problems of PSNR: Do you see the difference?

Original image in exr,
tone-mapped to 8bpp
Scaled to fit on screen

Reconstructed image in exr,
tone-mapped to 8bpp
Scaled to fit on screen

NOTE: Error is neither visible when checking luminance windows.

PSNR of the image in the **HDR domain** is 0.35dB.
(Problem here is a slight round-off error for some very light pixels)

PSNR is already working bad for LDR images.
It is a disaster for HDR images

Elementary Quality Indices for HDR

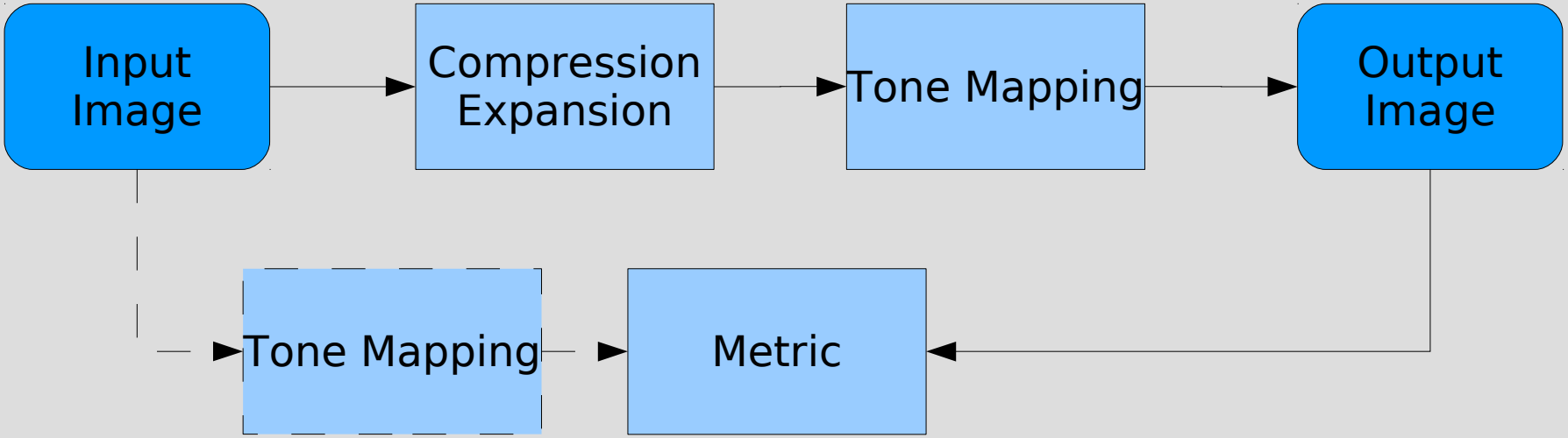
- Something as simple as PSNR for LDR is missing for HDR
  - PSNR is not suitable, as seen.
- Suitable for mathematical analysis, e.g. rate-distortion theory
- Suggestion: Mean relative error (and its log)

$$MRE(\vec{x},\vec{y}) := \sum_i \frac{|x_i - y_i|^2}{|x_i|^2 + |y_i|^2} \approx \sum_i \Delta \log x_i$$

- Approximately the coding error in the logarithmic domain
- Approximates the eye sensitivity (“Weber's Law”)
- One-point approximation of SSIM [QOMEX 2009]
- Mathematical analysis simple: MSE-optimal coding in the log domain is MRSE-optimal.
- Results from rate-distortion theory are applicable immediately.

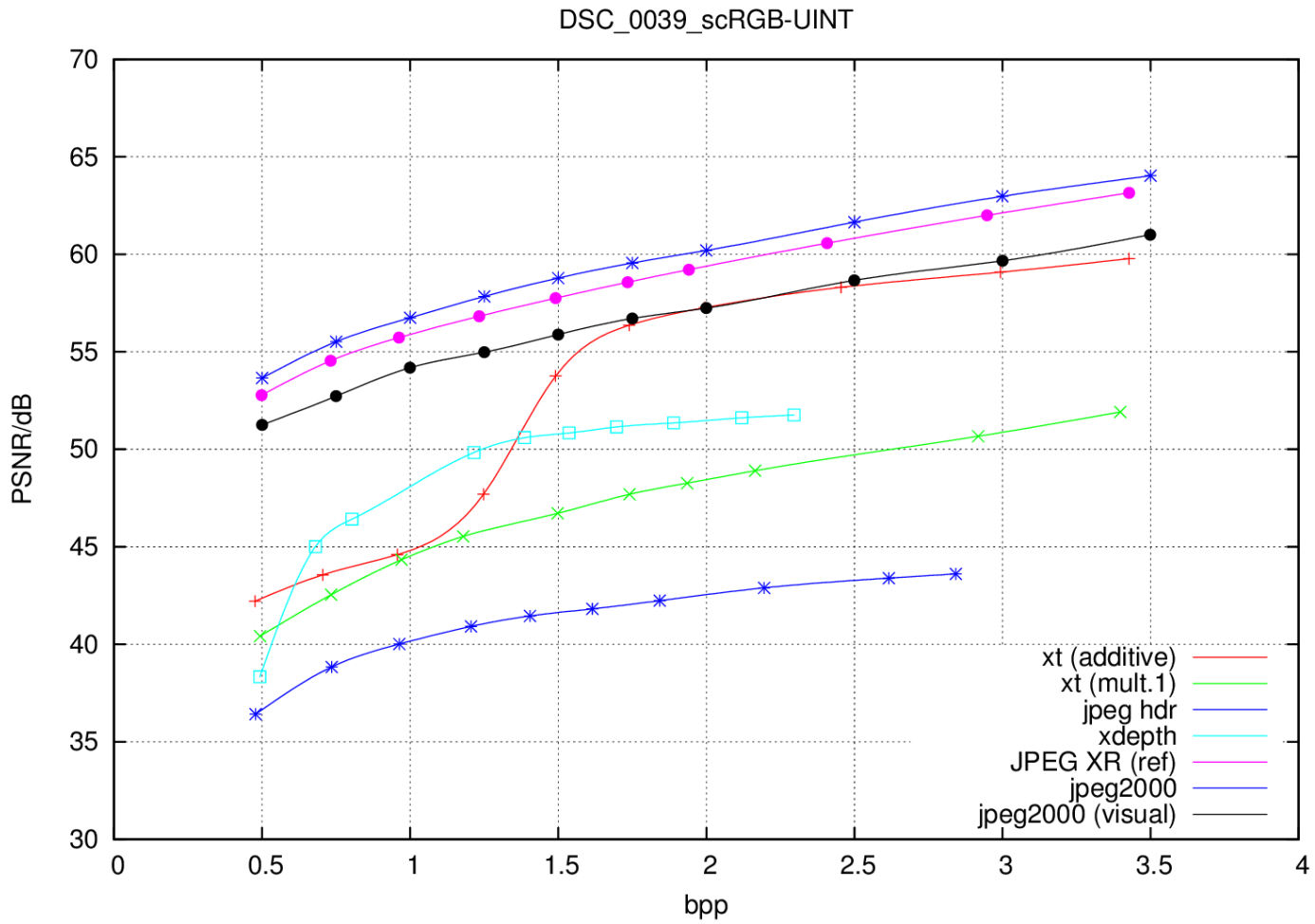
Measure Behind Tonemapping

(as already done on the previous slide)

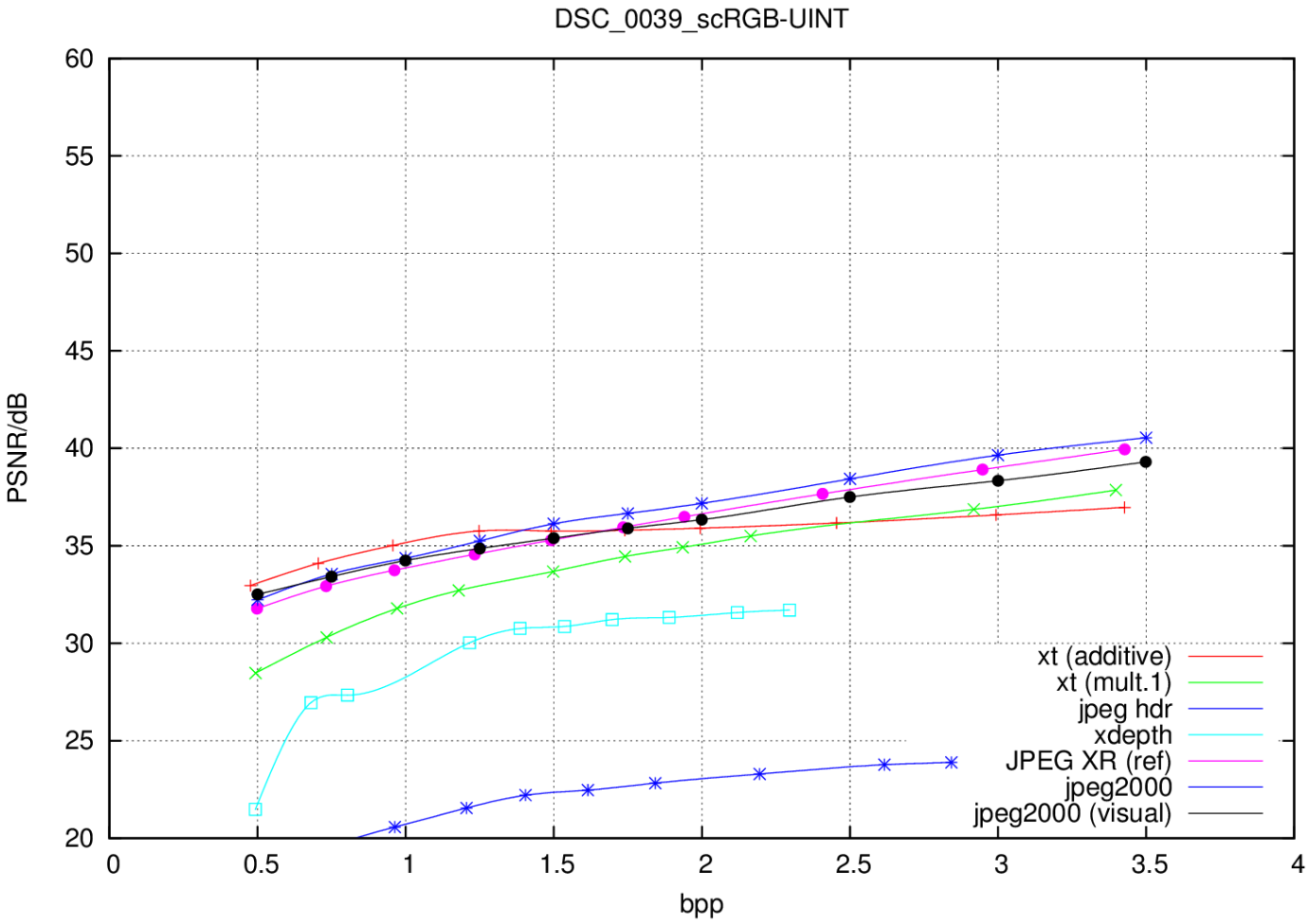


- Useful for the right column of the use cases table
- Subjective test protocols applicable, requiring standard equipment.
- LDR quality indices applicable
  - In all variations (full,reduced,no reference)
- Tone mapping is an **open variable** here
  - May become defined if input and output color space are defined
  - Otherwise left to the intention of the photographer and hard to specify
  - There is one possible trap/systematic error here...

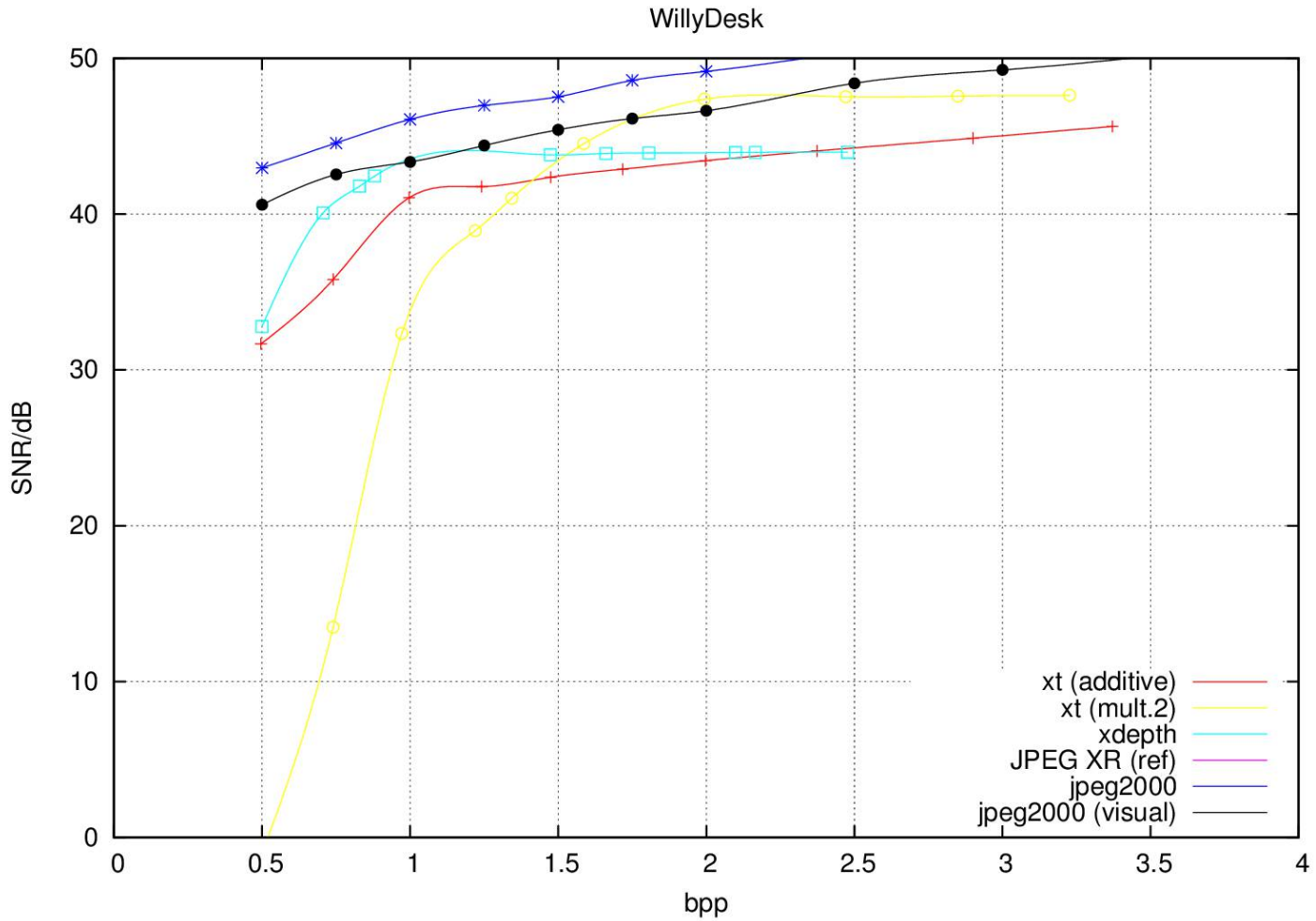
Not every method is good for every use case: 16 bit integer, direct measurement



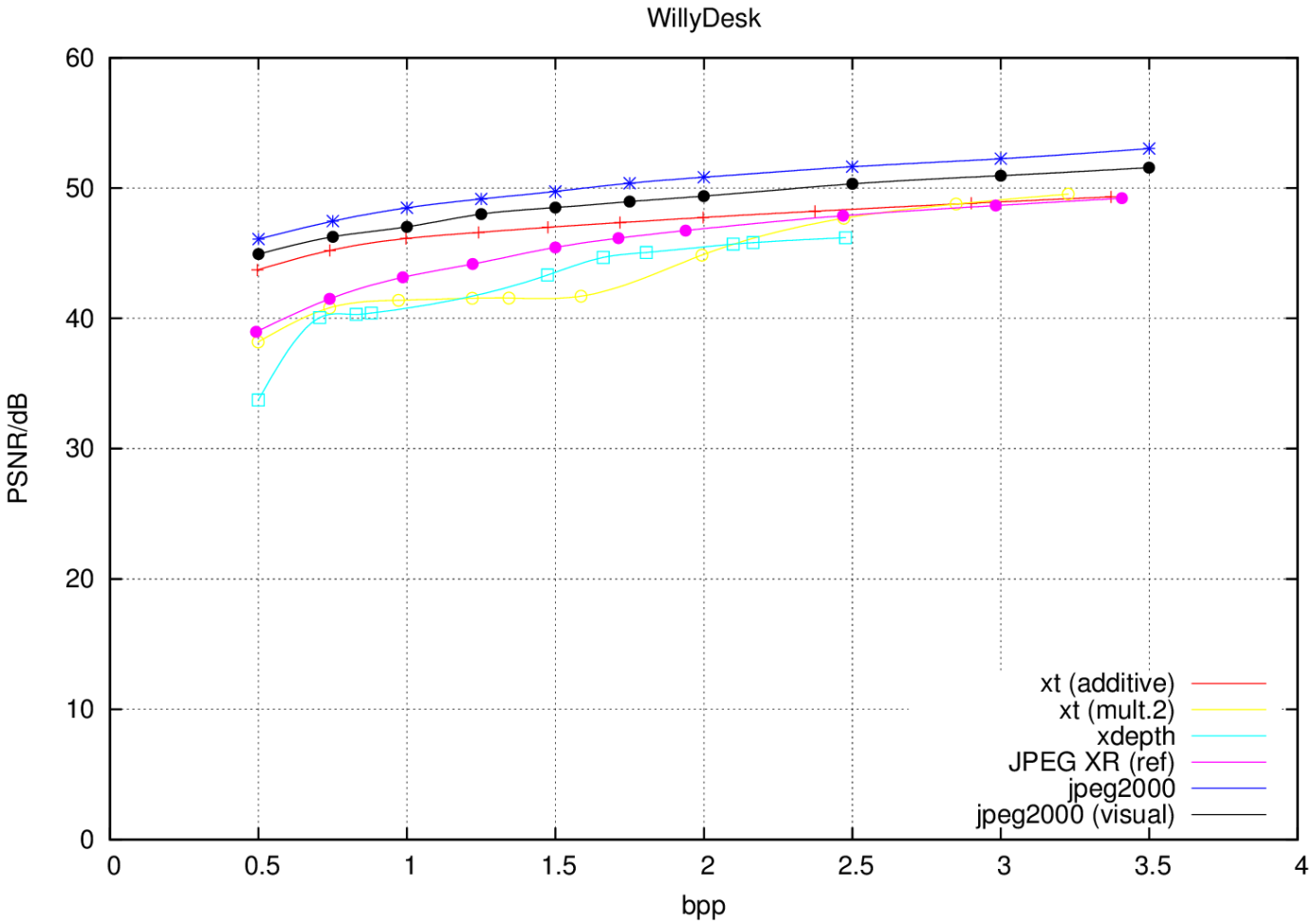
Not every method is good for every use case
16 bit integer – with tone mapping



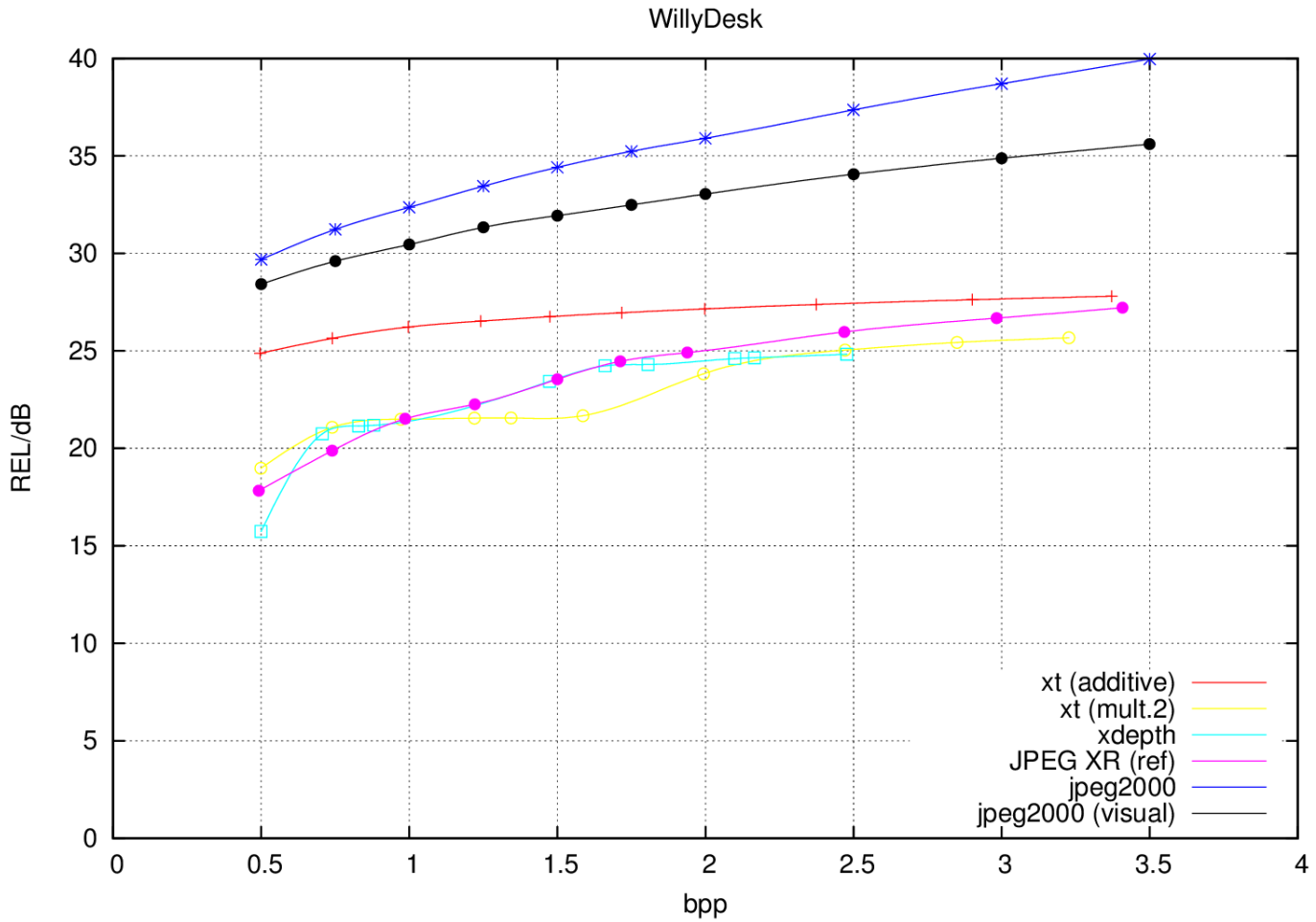
Not every method is good for every use case
float – direct measurement (questionable...)



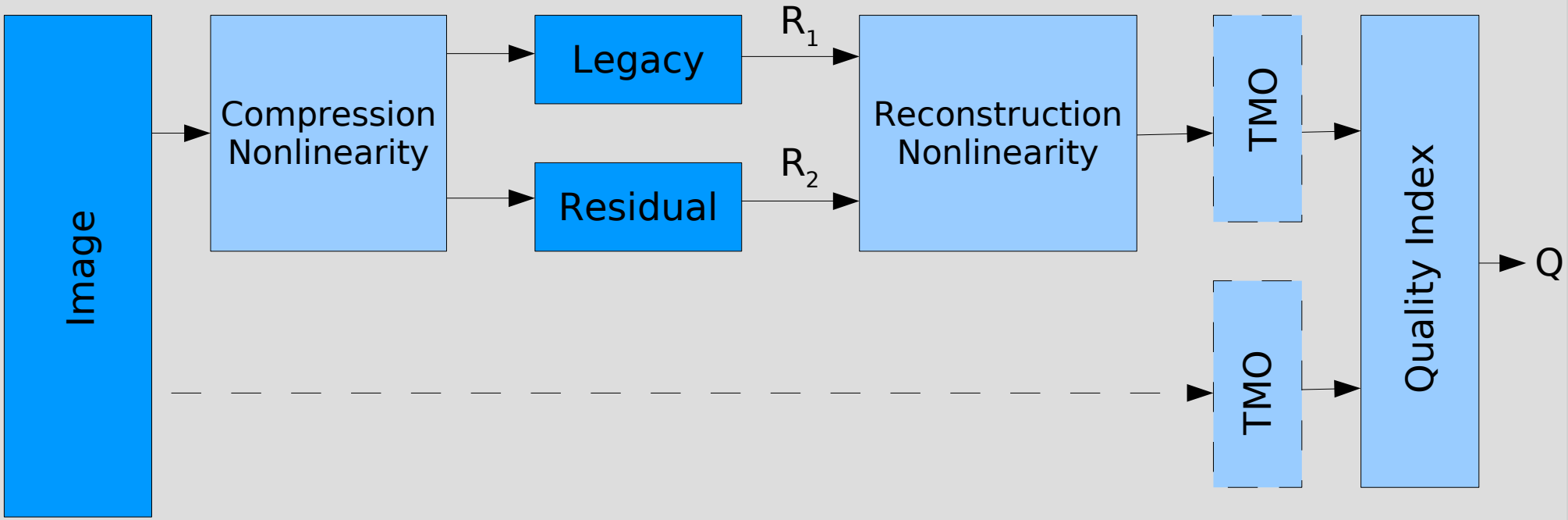
Not every method is good for every use case
float – with tone mapping



Defect of the QI or defect of the compressor replacing MSE with MRSE

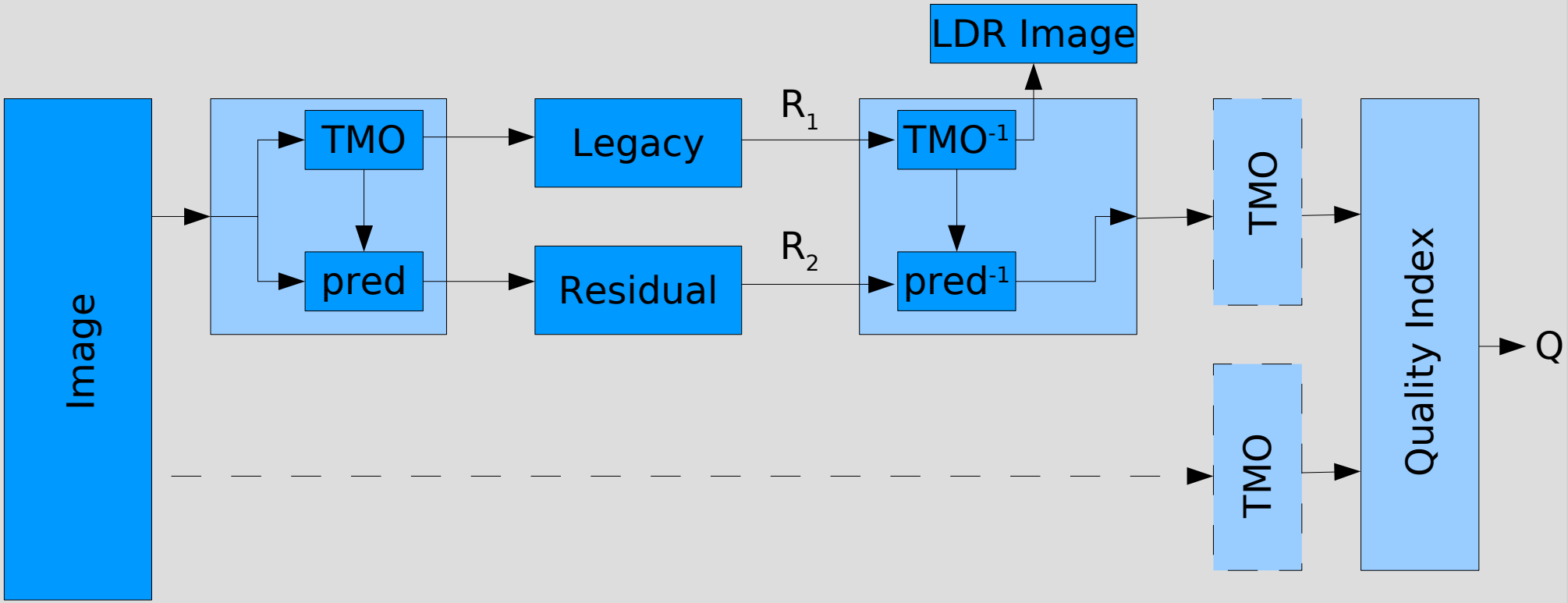


(Many) Open Questions



- Quality Index / Test protocol
  - Simple enough to allow mathematical analysis
  - Complex enough to capture important aspects of HVS
- Model of coding engine nonlinearities (most contributions are nonlinear)
- Model of the TMO(s)
- (Approximate) solution of the rate allocation problem

(Many) Open Questions



- Quality Index / Test protocol
  - Simple enough to allow mathematical analysis
  - Complex enough to capture important aspects of HVS
- Model of coding engine nonlinearities (most contributions are nonlinear)
- Model of the TMO(s)
- (Approximate) solution of the rate allocation problem

Conclusions

- Test protocols should be defined along the desired use cases
  - Both subjective and objective
- Traditional test protocols are typically not suitable **because** they do not take the full workflow into consideration
- LDR quality indices, even PSNR, cannot be blindly applied in the HDR regime
- New quality indices need to be developed that also operate in the HDR domain (Use case: Top left corner of the matrix)
- Such quality indices need not to be “perfect”, but should allow a mathematical analysis of the problem.

Thank You!

Test JPEG Image Quality Online:
http://jpegonline.rus.uni-stuttgart.de/index.py