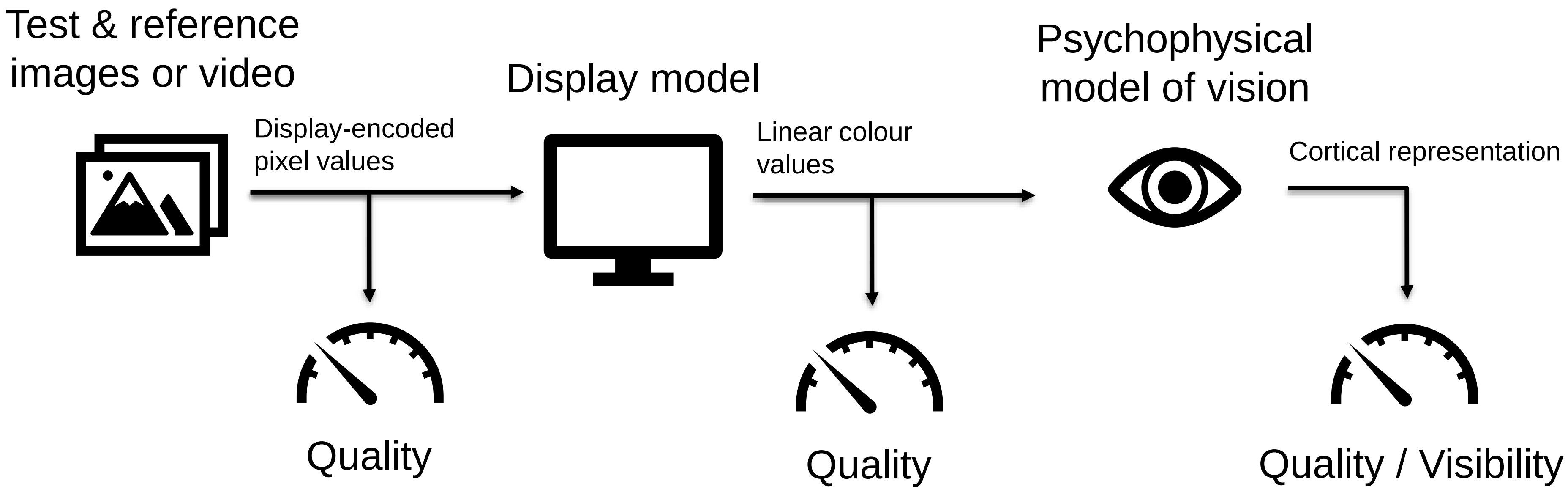
The family of VDP metrics for image and video quality predictions

Rafał Mantiuk

Modelling visual quality



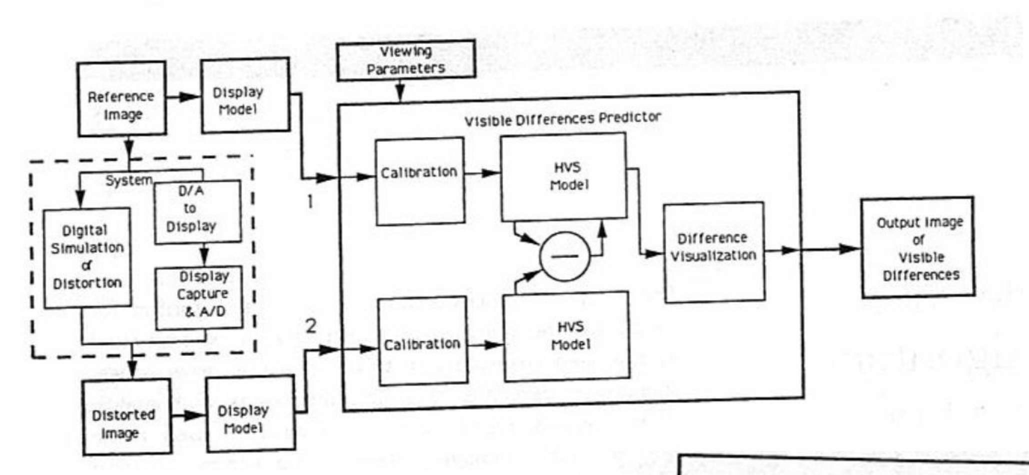
Most metric
PSNR, SSIM, LPIPS, …

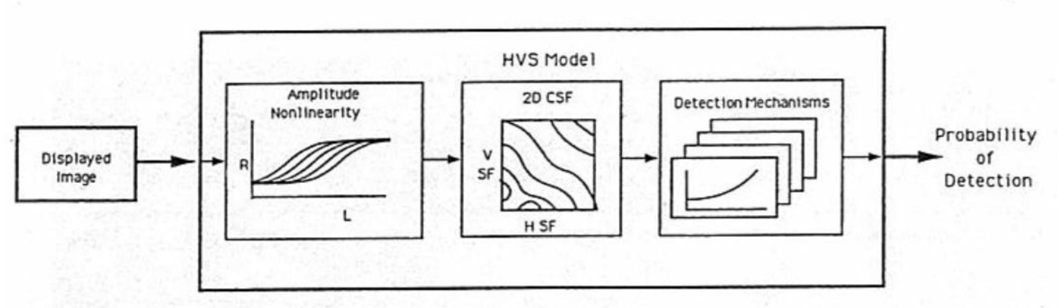
Few metrics
FLIP, PU-encoding,
SAST rescaling

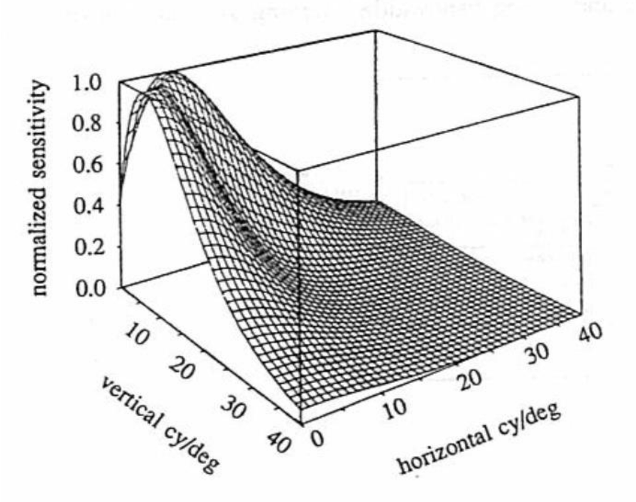
Visual
Difference
Predictors

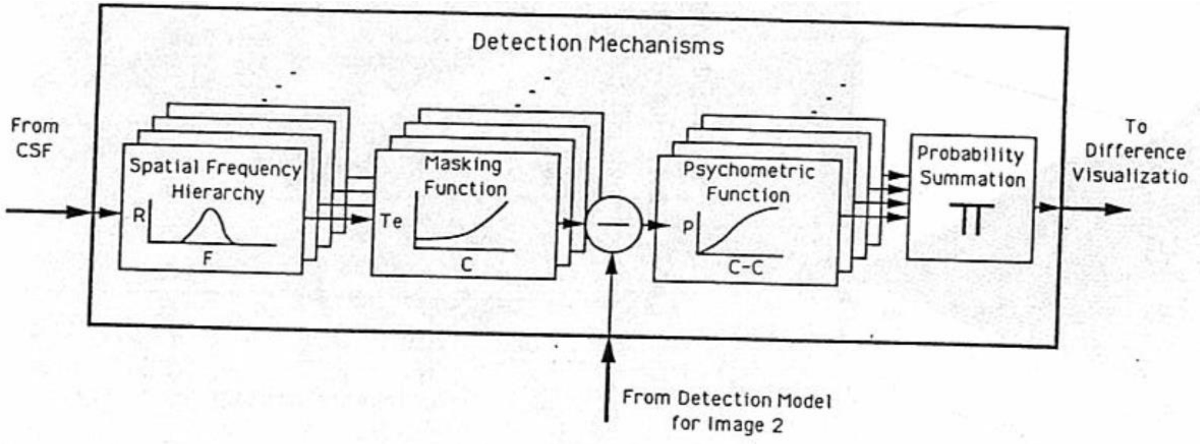
Correlation ⟵ **Goal** ⟶ Causation

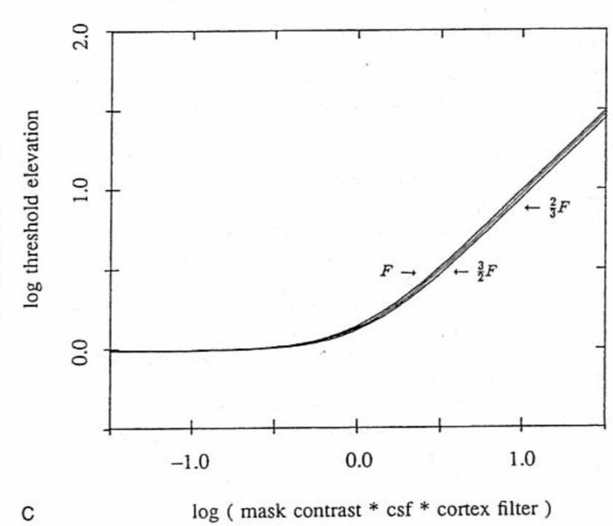
Visual Difference Predictor [Daly 1993]





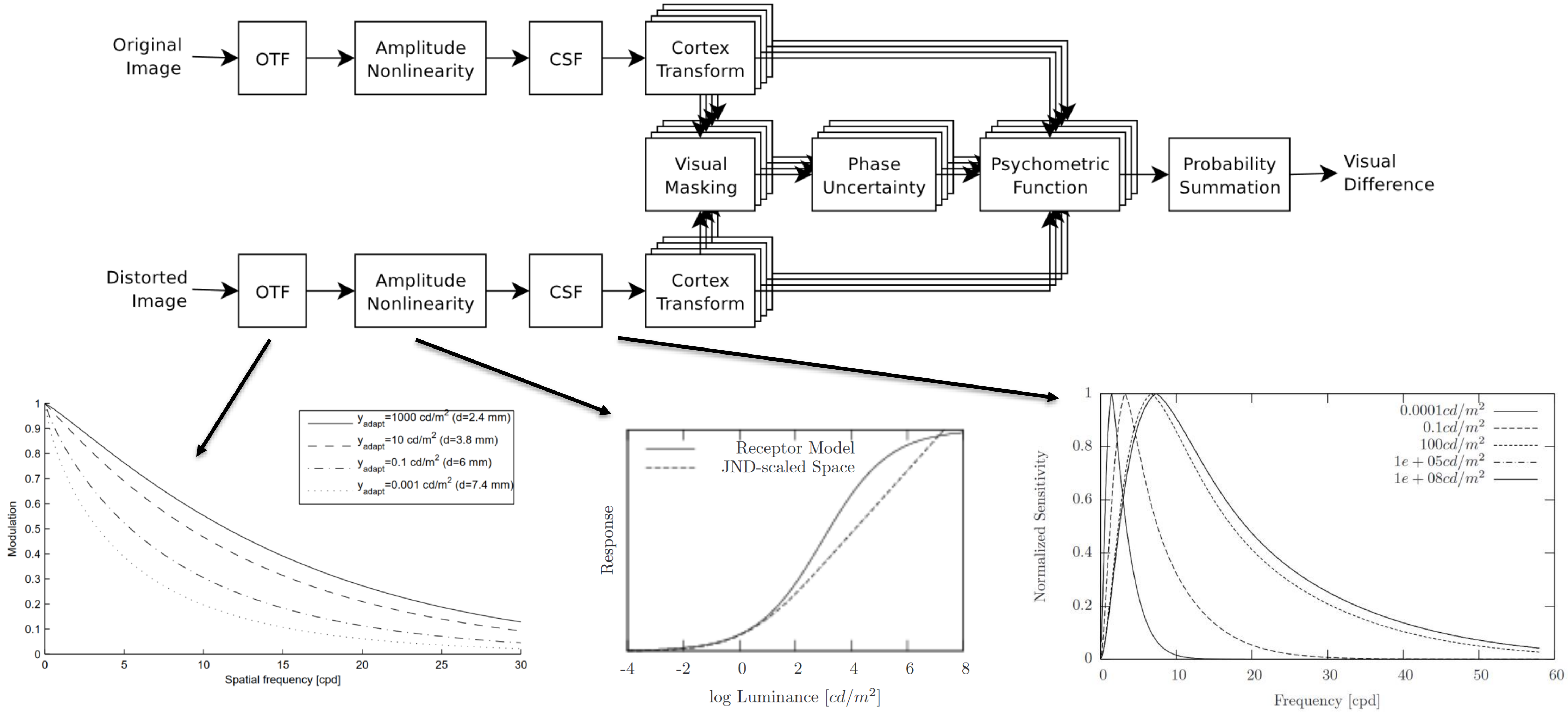






S. J. Daly, ‘Visible differences predictor: an algorithm for the assessment of image fidelity’, in *Digital Images and Human Vision*, vol. 1666, A. B. Watson, Ed., MIT Press, 1993, pp. 179–206. doi: 10.1117/12.135952.

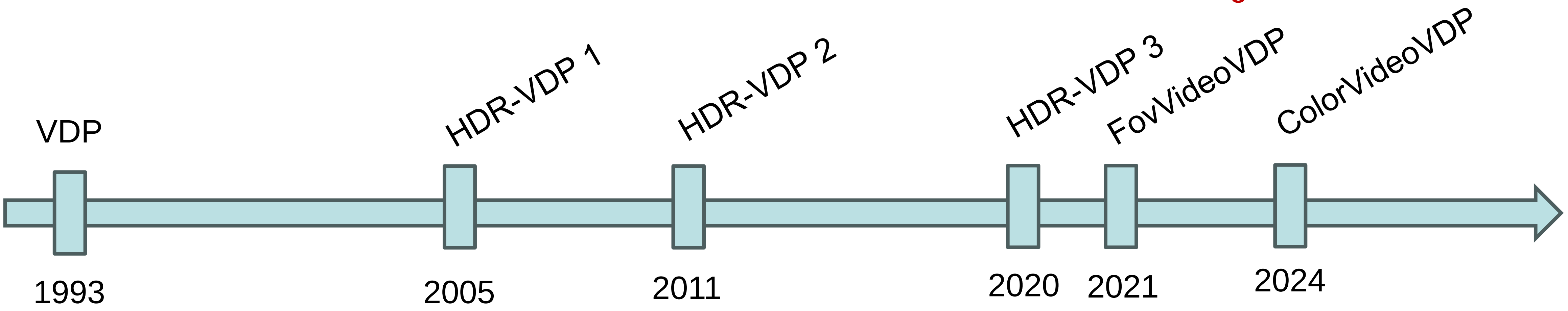
HDR-VDP 1.0 [Mantiuk, Daly, Myszkowski, Seidel, 2005]



Timeline

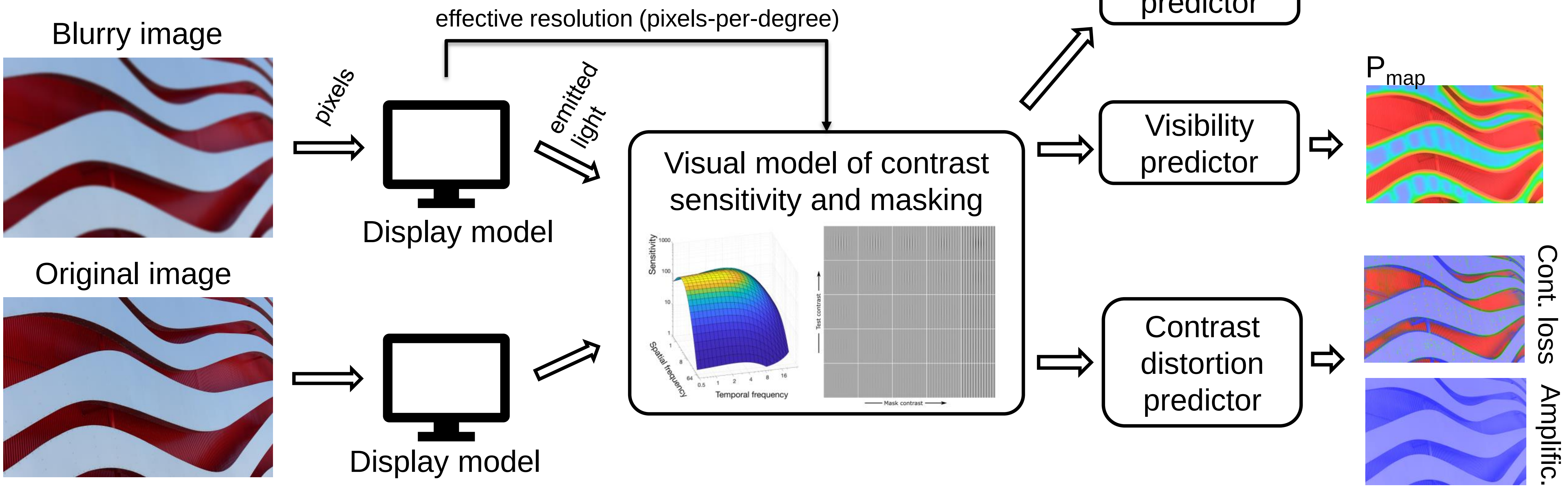
SIGGRAPH
Test of Time Award

1st place in
HDR Video Quality Measurement
Grand Challenge

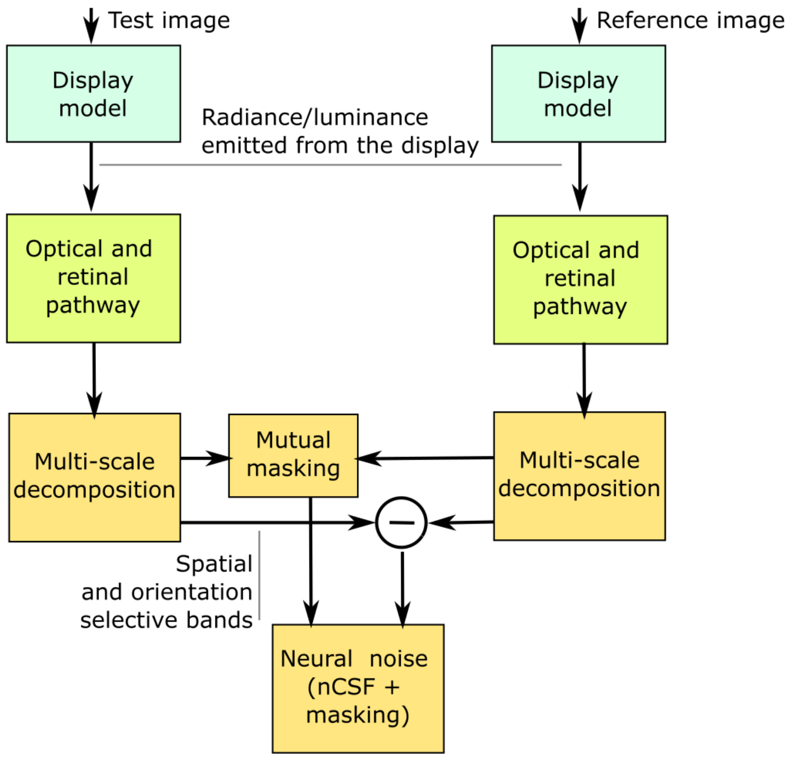


What is HDR-VDP-3?

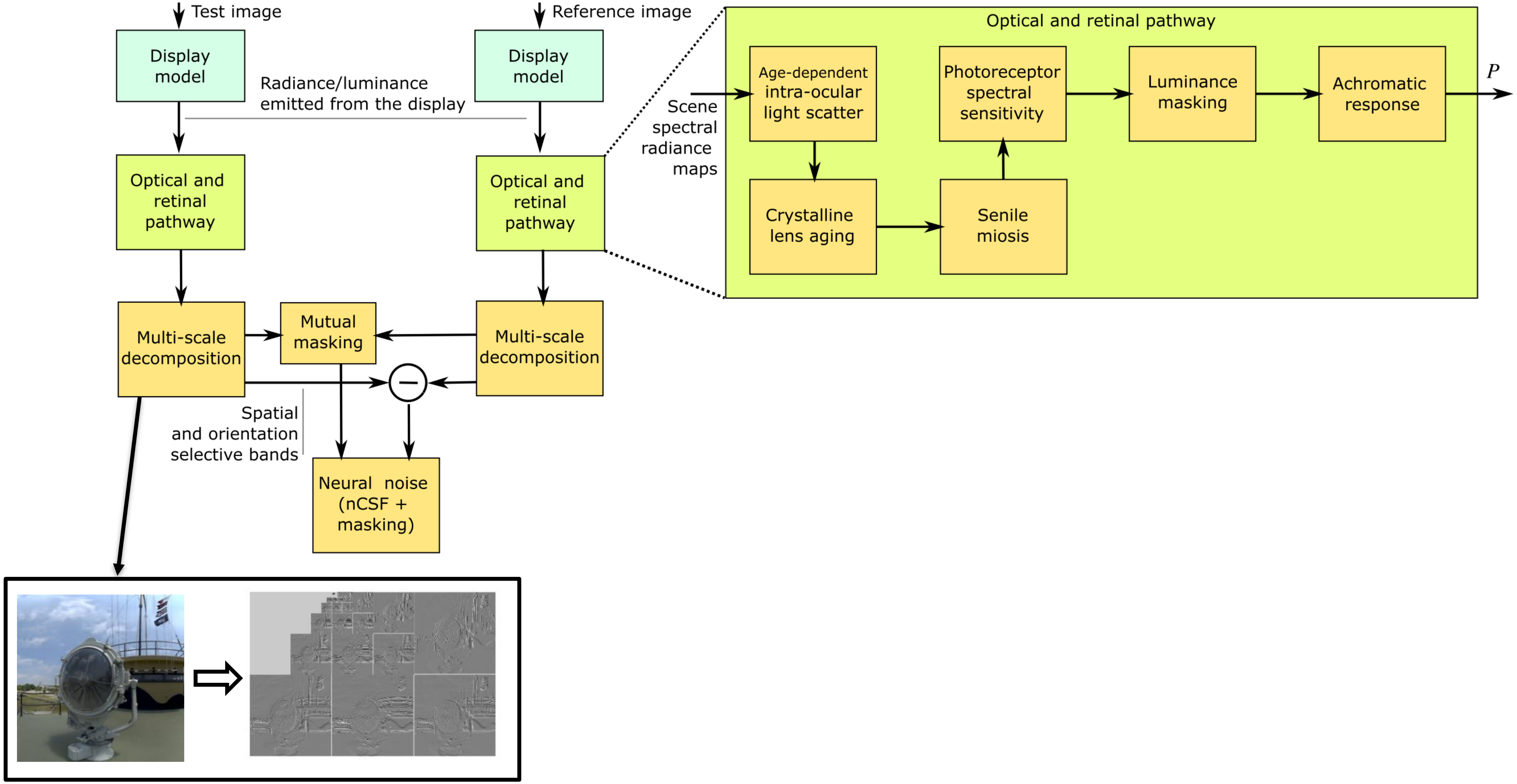
- A metric based on psychophysical models of contrast sensitivity and masking
- A metric that operates on physical radiance/luminance units
  - (rather than digital pixel values)



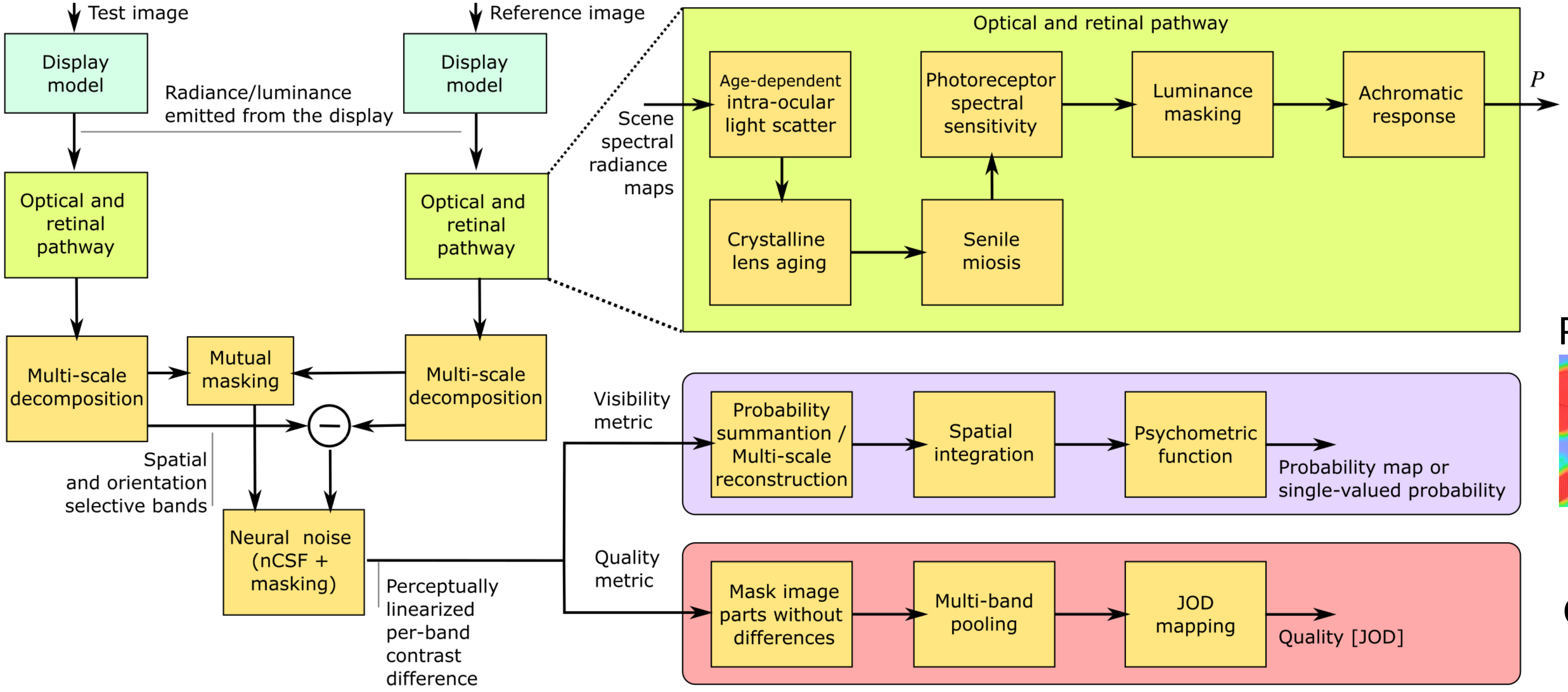
HDR-VDP-3 in more detail



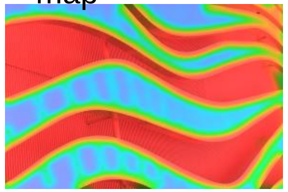
HDR-VDP-3 in more detail



HDR-VDP-3 in more detail

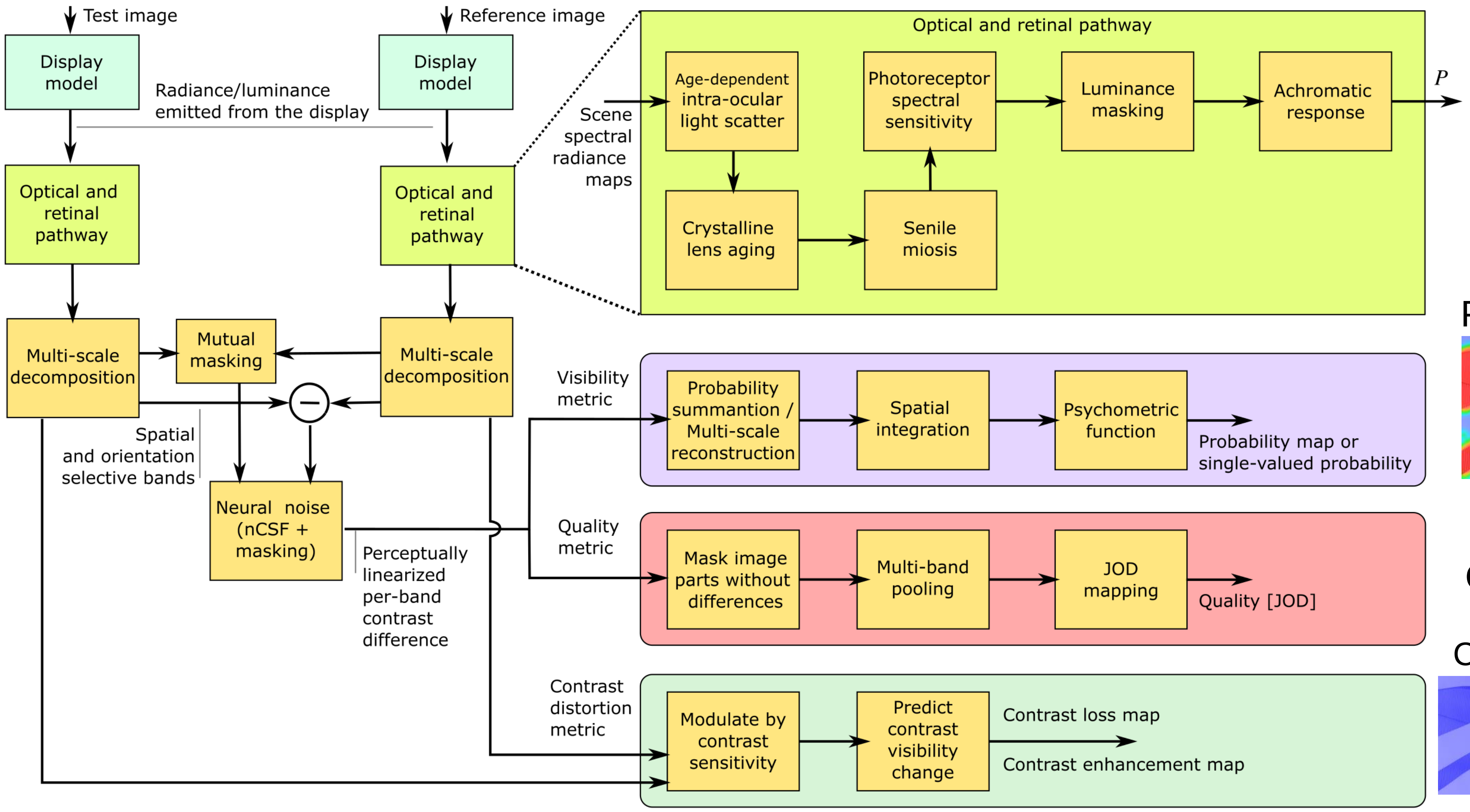


P_{map}

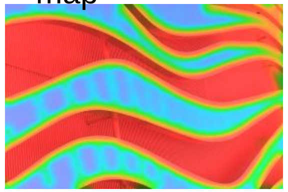


Q_{JOD}=5.69

HDR-VDP-3 in more detail

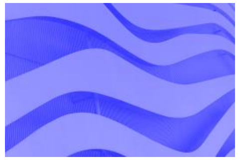


P_{map}

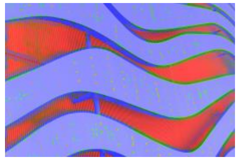


Q_{JOD}=5.69

Cont. loss



Amplification



SIGGRAPH 2024
DENVER+ 28 JUL — 1 AUG

ColorVideoVDP: A visual difference predictor for image, video and display distortions

Rafał K. Mantiuk,
Param Hanji,
Maliha Ashraf,
Yuta Asano,
and Alexandre Chapiro

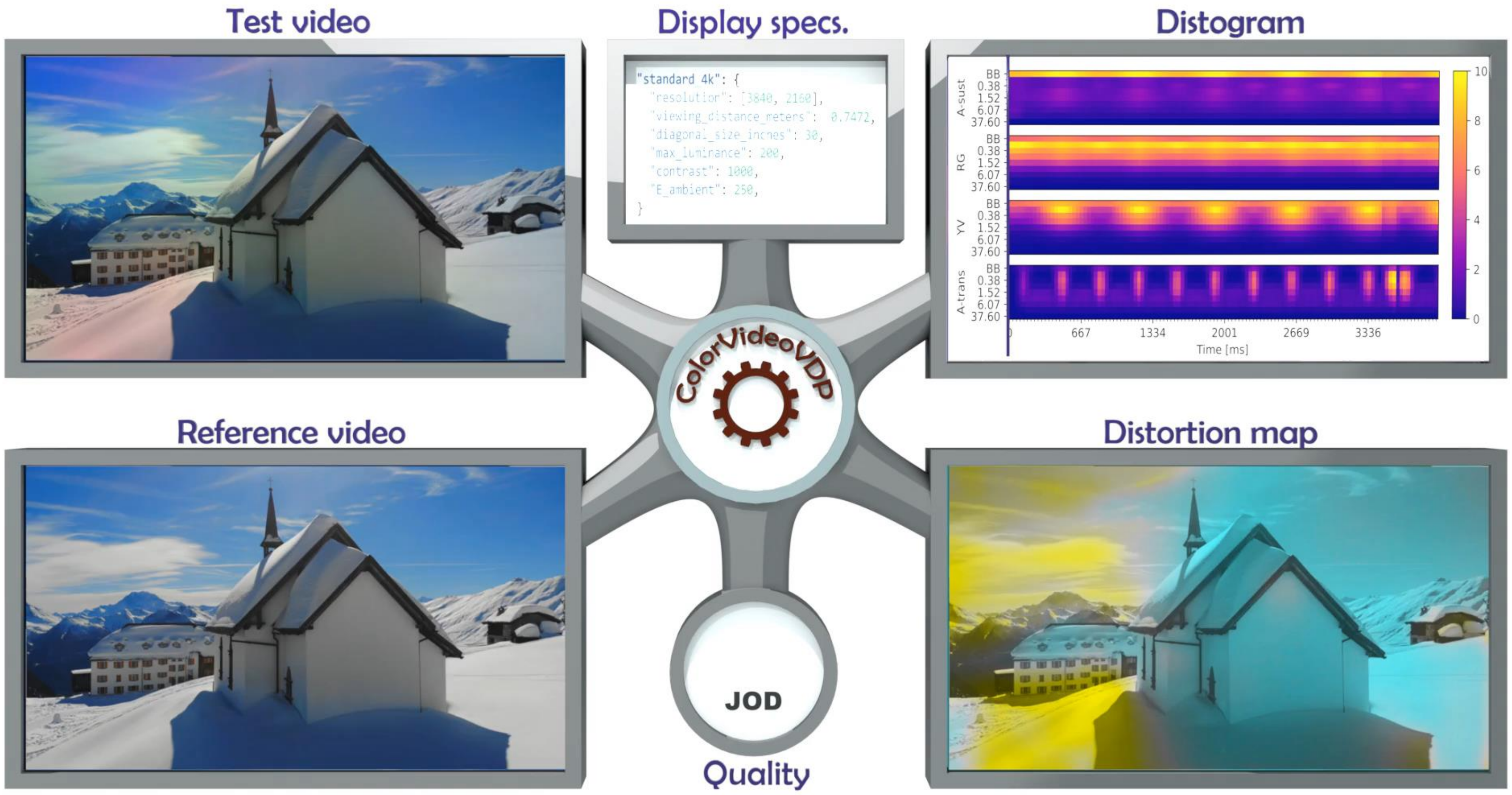
UNIVERSITY OF CAMBRIDGE

Meta

University of Cambridge

What is ColorVideoVDP?

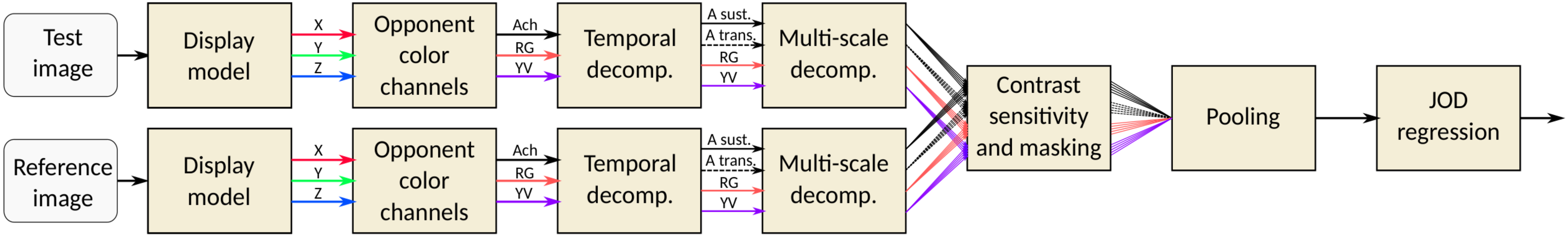
- A video quality metric
  - Models human colour and spatiotemporal vision
- Built on castleCSF
- Differentiable (PyTorch)
- Calibrated for
  - Video streaming
  - Display distortions



ColorVideoVDP – design goals

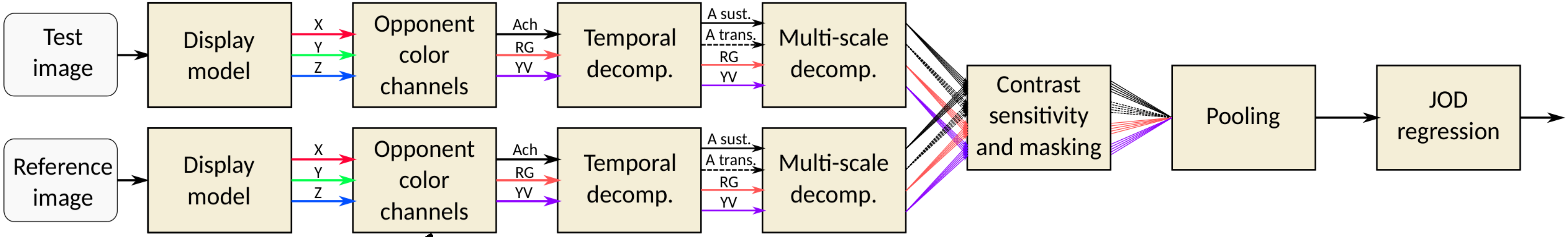
- Rely on psychophysical models of low-level vision as much as possible
  - Less chance for overfitting
  - More robust to corner cases
  - Can adapt to viewing conditions
- Make it explainable
  - No black boxes 
- Make it simple
  - A feature is not needed if the impact on the predictions is small
- Make it fast
  - ColorVideoVDP runs fast on a GPU
- Make it full differentiable
  - Important for optimization and training

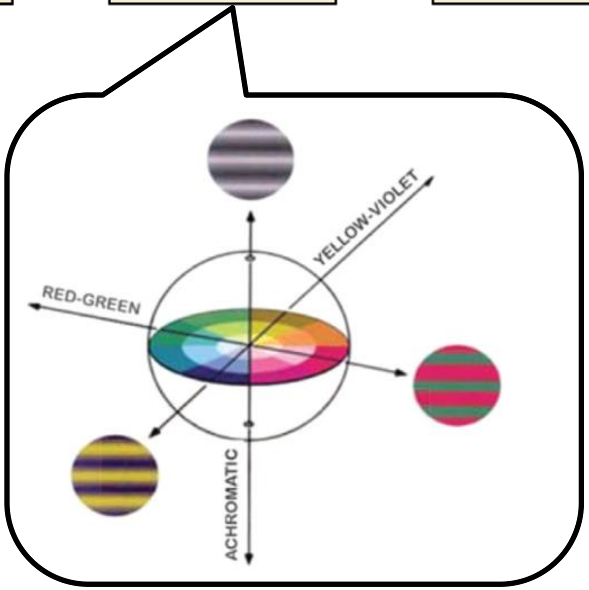
ColourVideoVDP




```
"standard_4k": {
  "resolution": [3840, 2160],
  "viewing_distance_meters":  0.7472,
  "diagonal_size_inches": 30,
  "max_luminance": 200,
  "contrast": 1000,
  "E_ambient": 250,
}
```

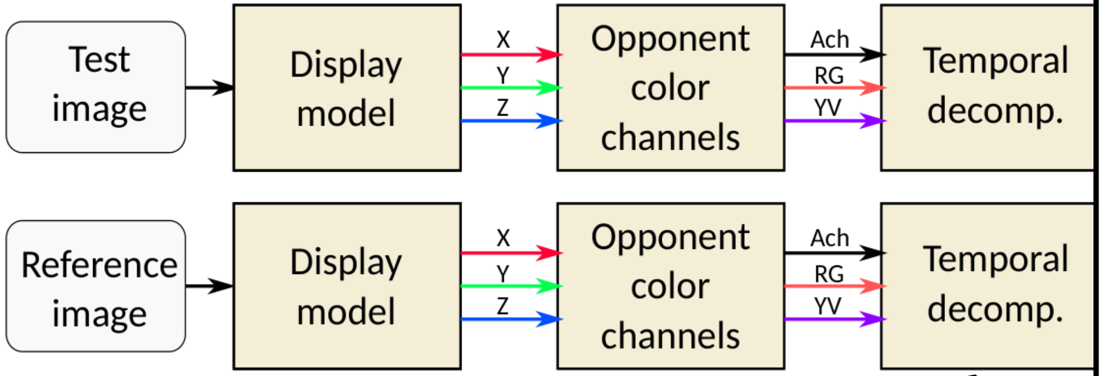
ColourVideoVDP

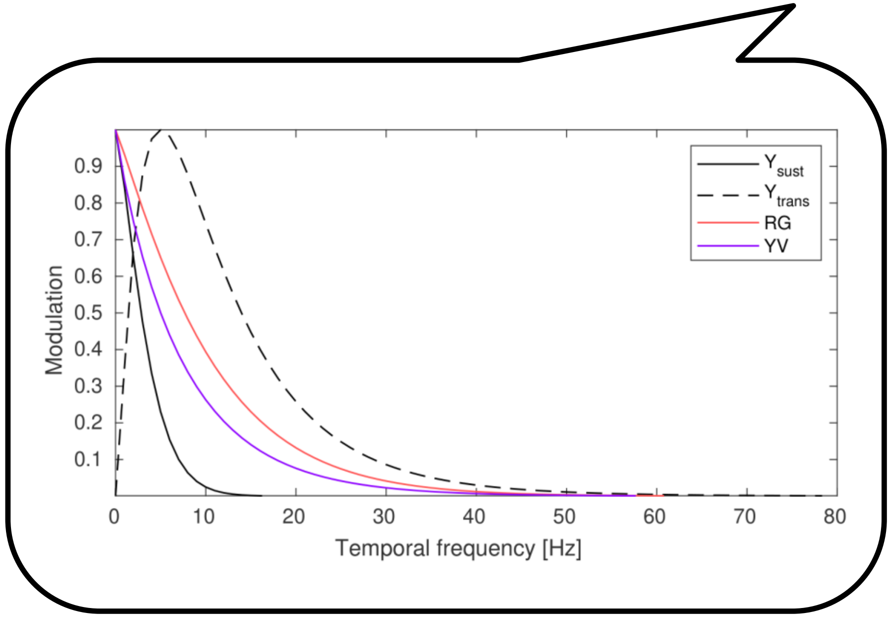


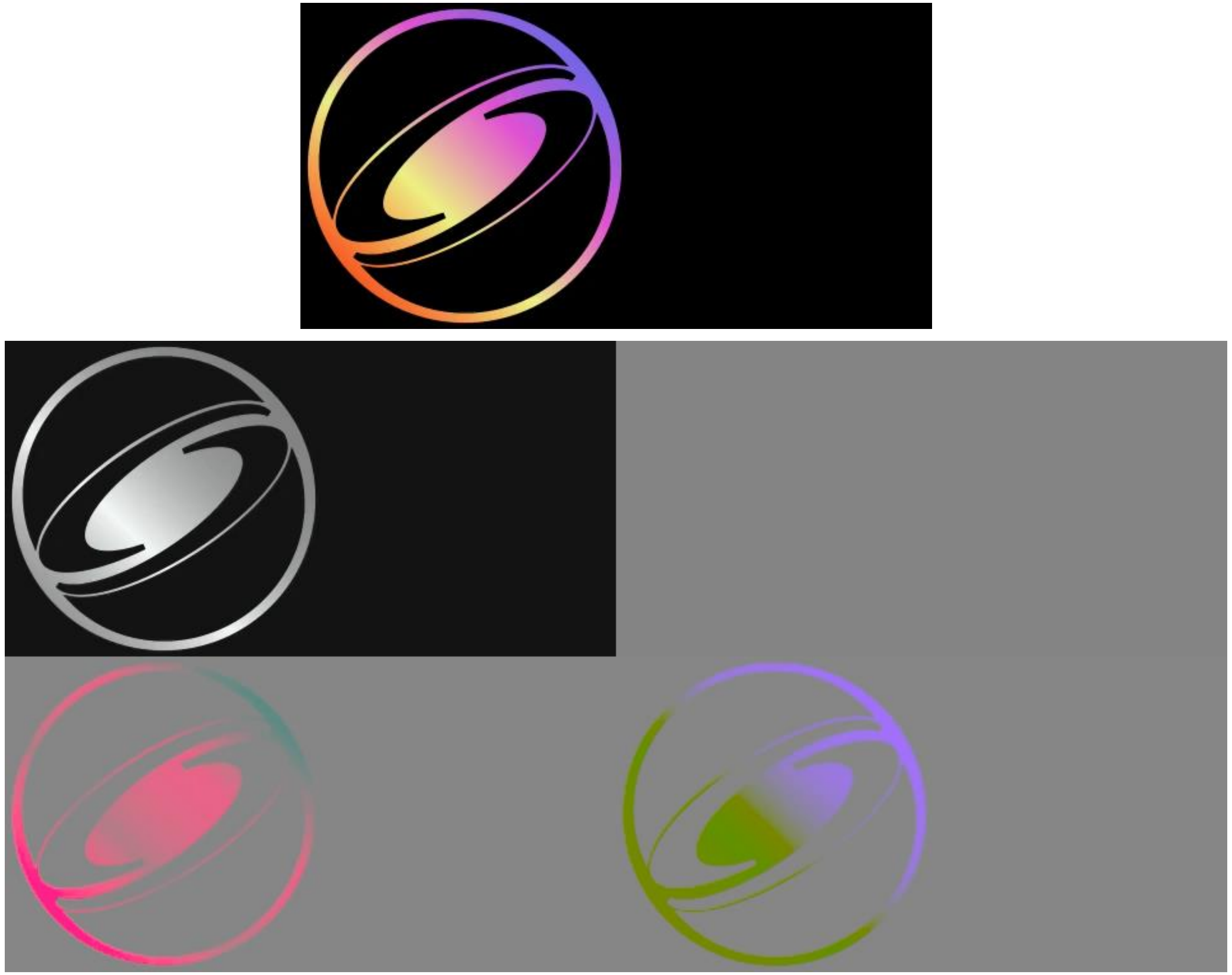


DKL color space
[Derrington, A. M. and Krauskopf, J. and Lennie, JoP, 1984]

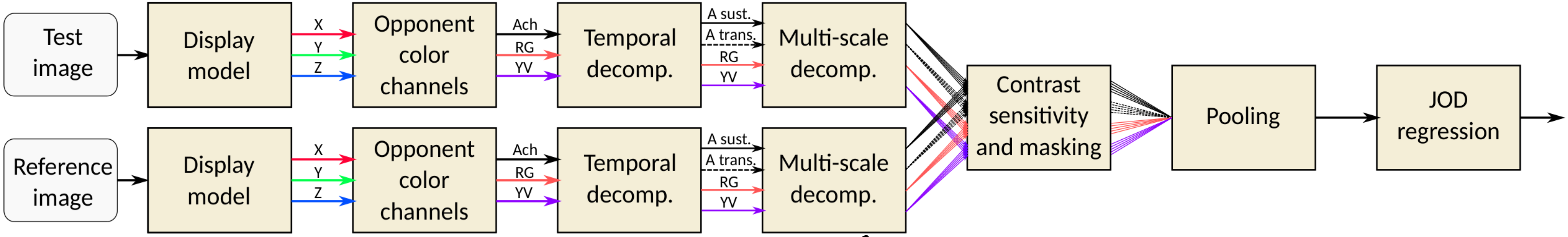
ColourVideoVDP







ColourVideoVDP



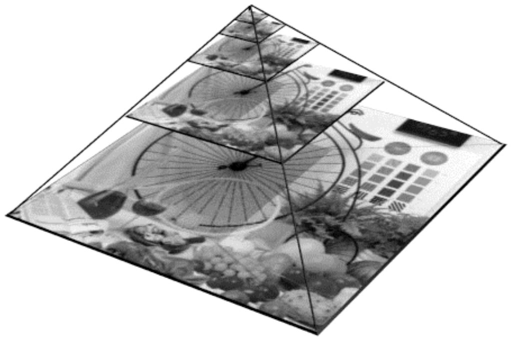
Contrast in complex images

DKL contrast

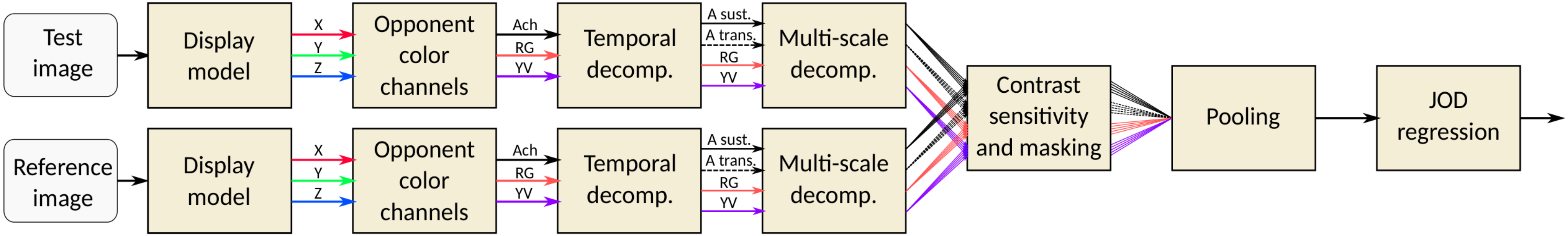
Laplacian pyramid coefficient

$$C_{b,c,f}(\boldsymbol{x}) = \frac{\mathcal{L}_{b,c,f}(\boldsymbol{x})}{\Uparrow \mathcal{G}_{b+1,S,f}(\boldsymbol{x})}$$

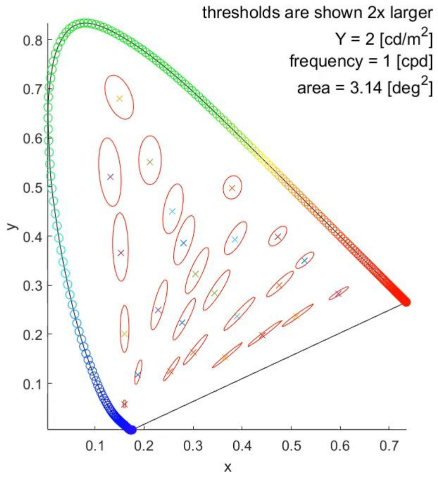
Upscaled Gaussian pyramid coefficient – sustained achromatic channel



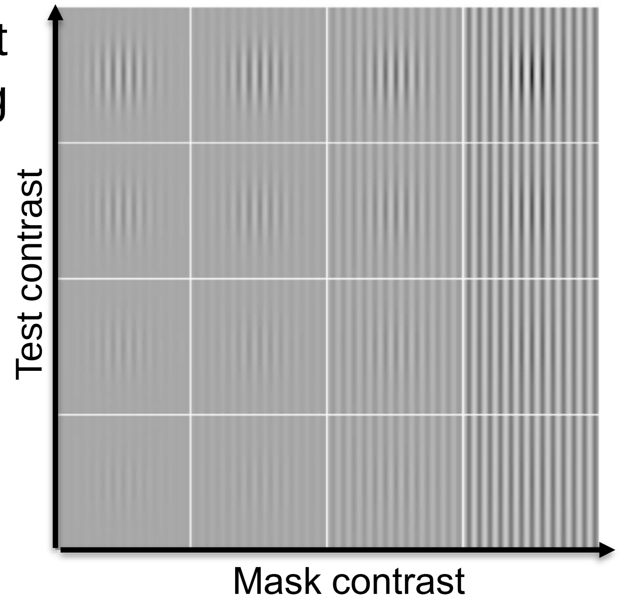
ColourVideoVDP



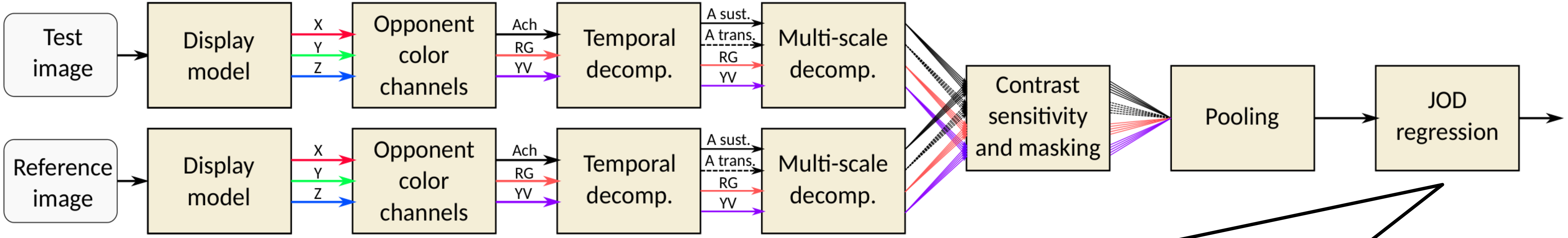
castleCSF minimum detectable contrast difference



Contrast masking

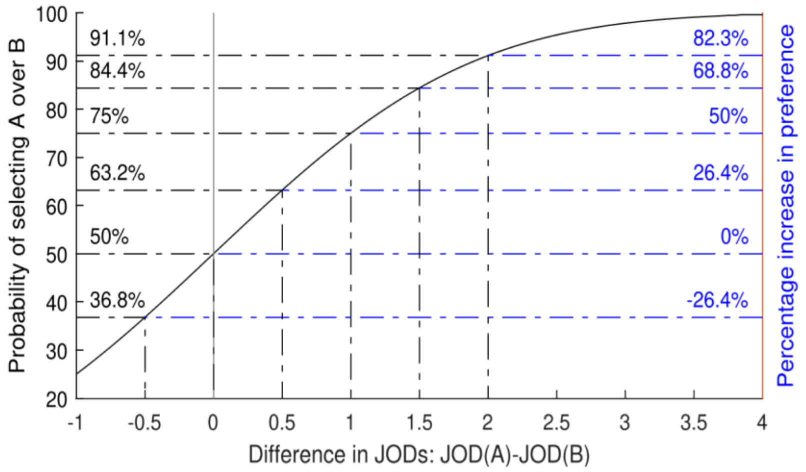


ColourVideoVDP



The quality is scaled in the units of
Just Objectionable Differences [JOD]
1 JOD difference ≈ 50% increase in preference

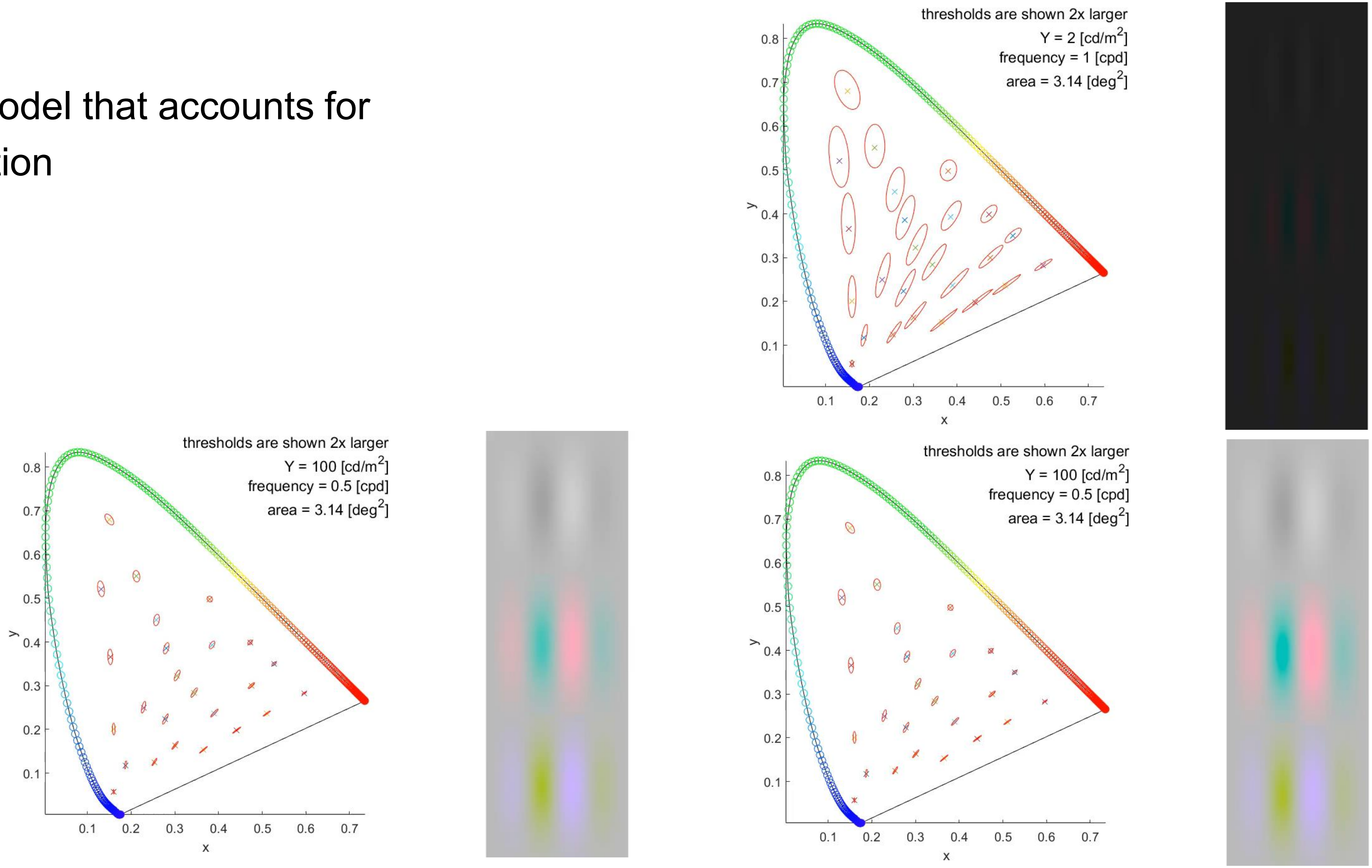
Can express supra-threshold (well-visible) differences



castleCSF

Contrast Sensitivity Model that accounts for

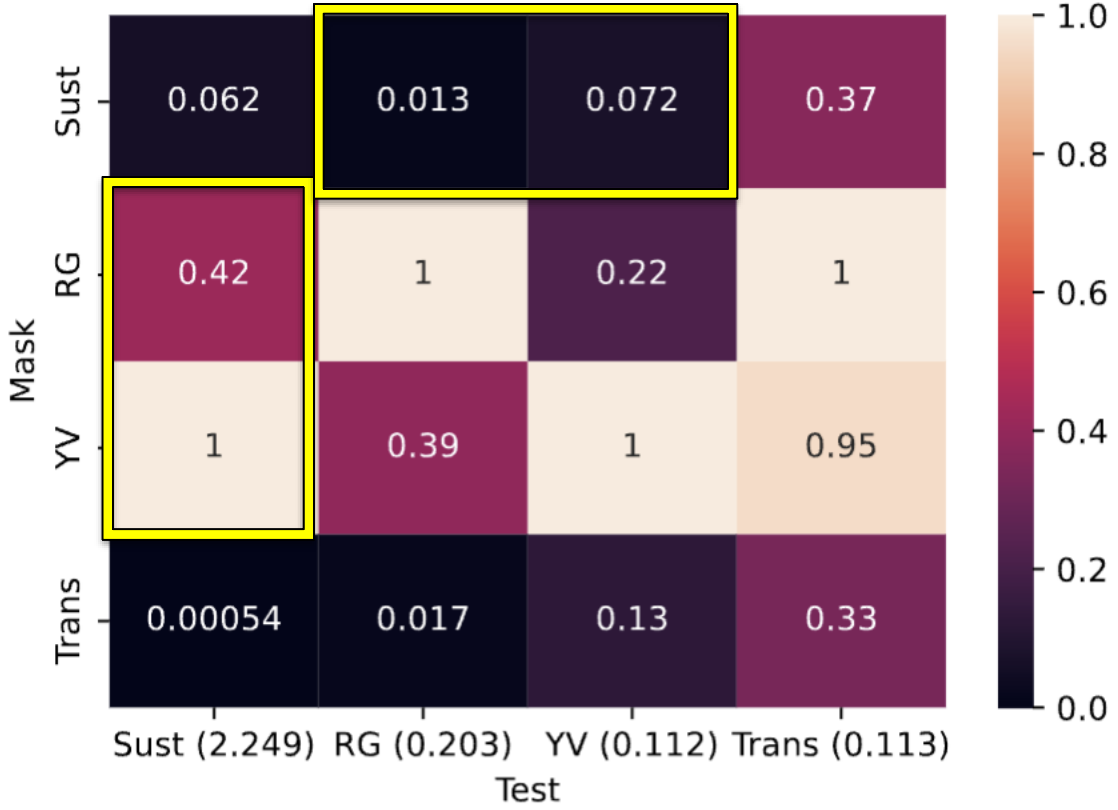
- Chromatic modulation
- Area
- Spatial freq.
- Temporal freq.
- Luminance
- Eccentricity



[Ashraf et al., JoV 2024]

Cross-channel masking

Trained on image/video quality datasets



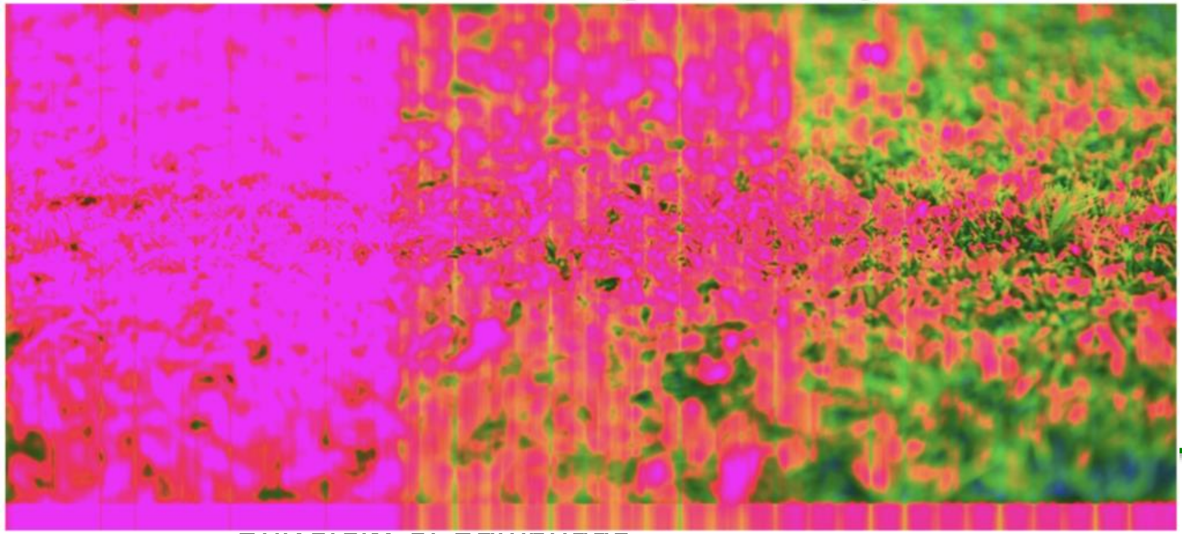
- Red-green and yellow-violet mask achromatic sustained channel
- But luminance does not mask the chromatic channels
- Consistent with the literature [Switkes et al. 1988]

Cross-channel contrast masking - example

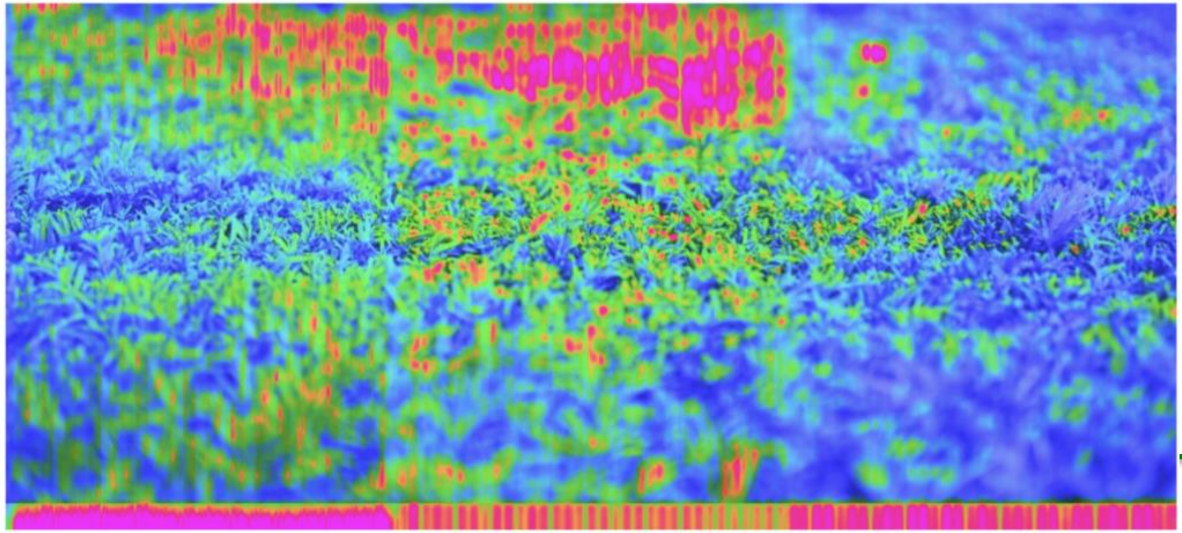
Reference image

Test image

Visual difference map — no masking

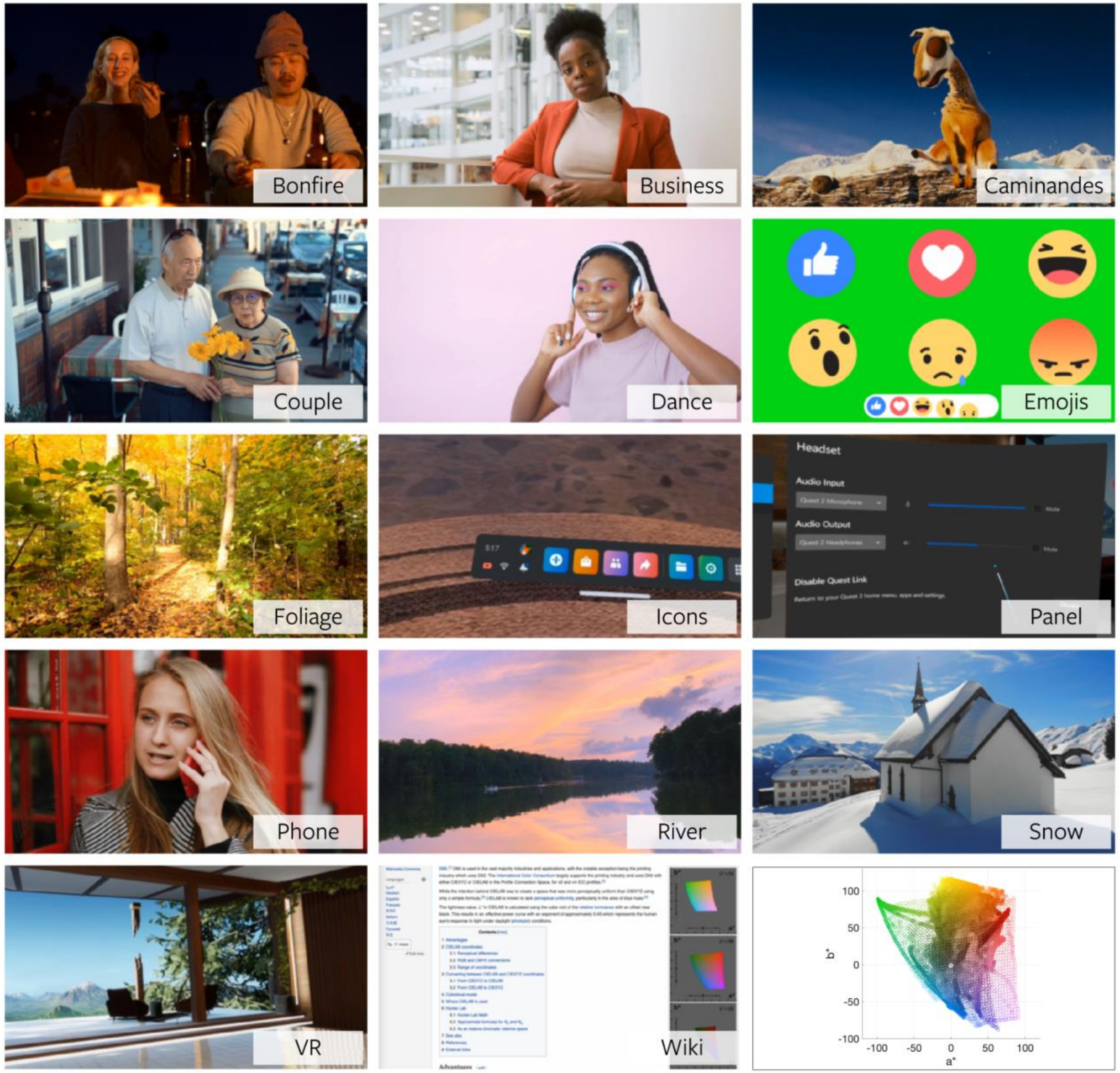


Visual difference map — with masking



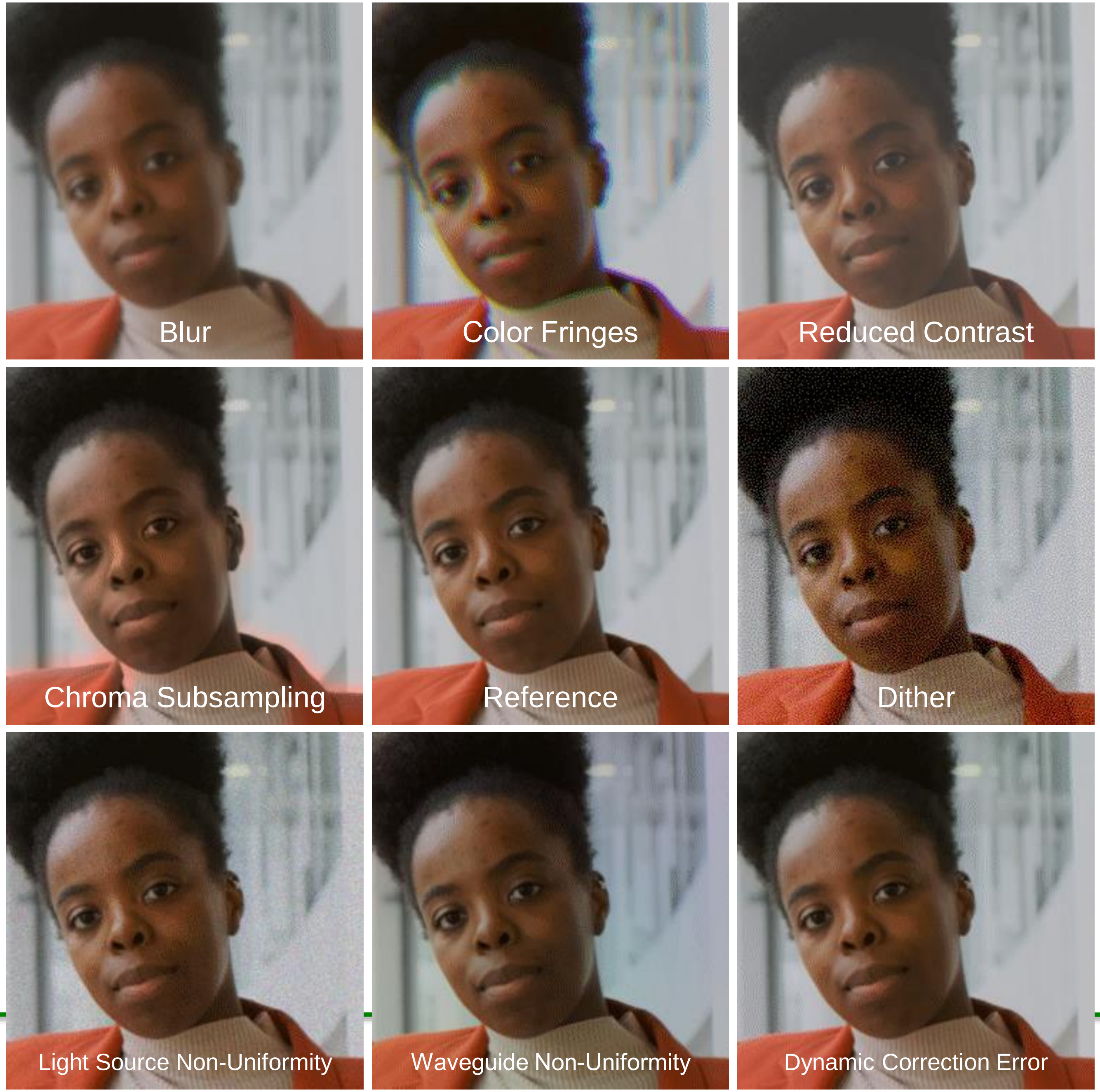
XR Display Artifact Video Dataset (XR-DAVID)

- 14 scenes including representative use-case scenarios



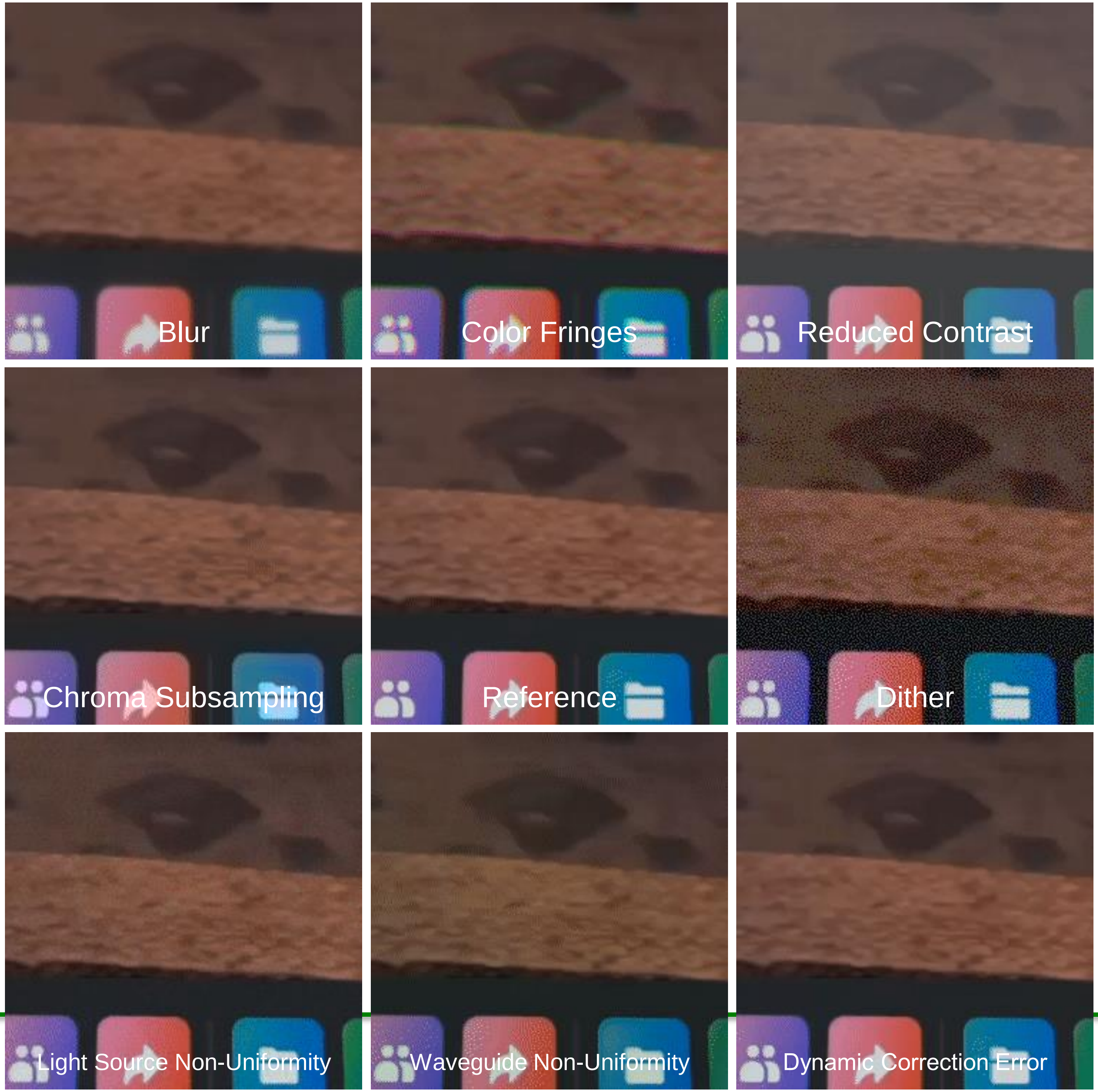
XR Display Artifact Video Dataset (XR-DAVID)

- 14 scenes including representative use-case scenarios
- 8 artifacts common in display and optics applications
  - 40 PPD, 300 nit display

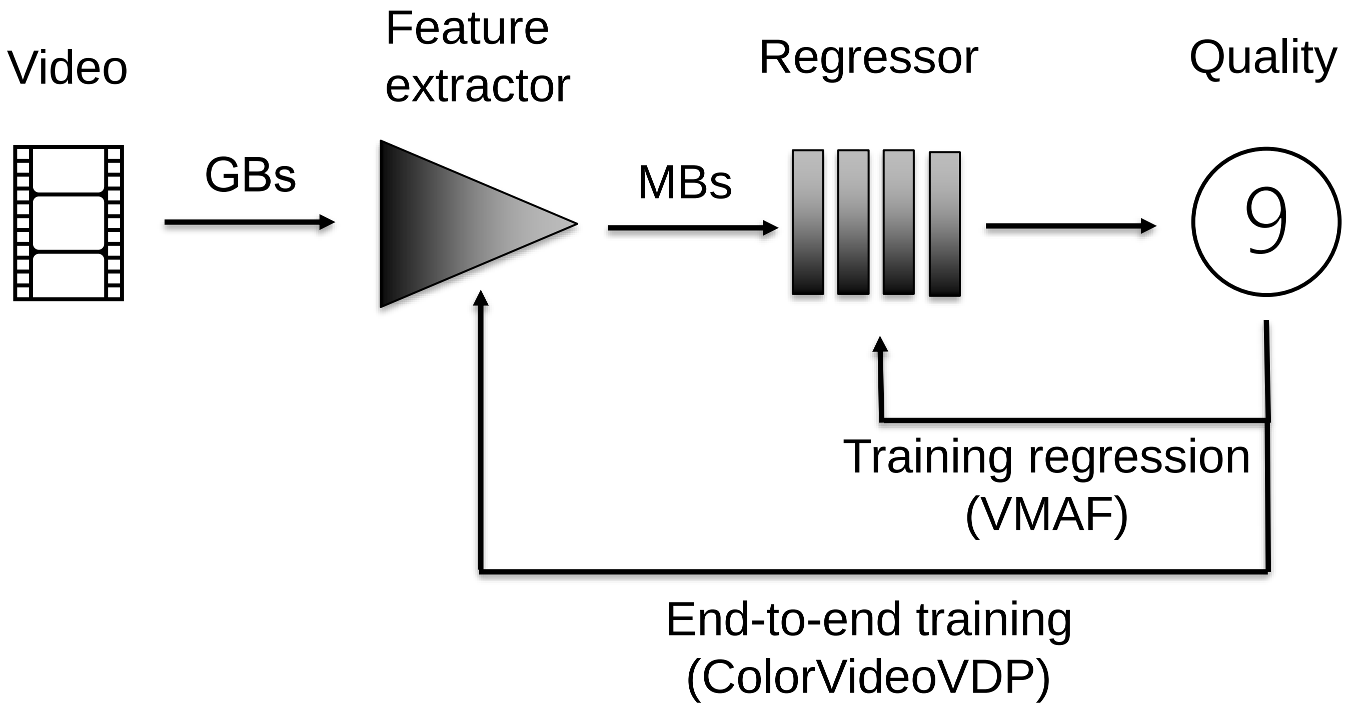


XR Display Artifact Video Dataset (XR-DAVID)

- 14 scenes including representative use-case scenarios
- 8 artifacts common in display and optics applications
  - 40 PPD, 300 nit display
- 77 participants
  - Pairwise comparisons with ASAP active sampling



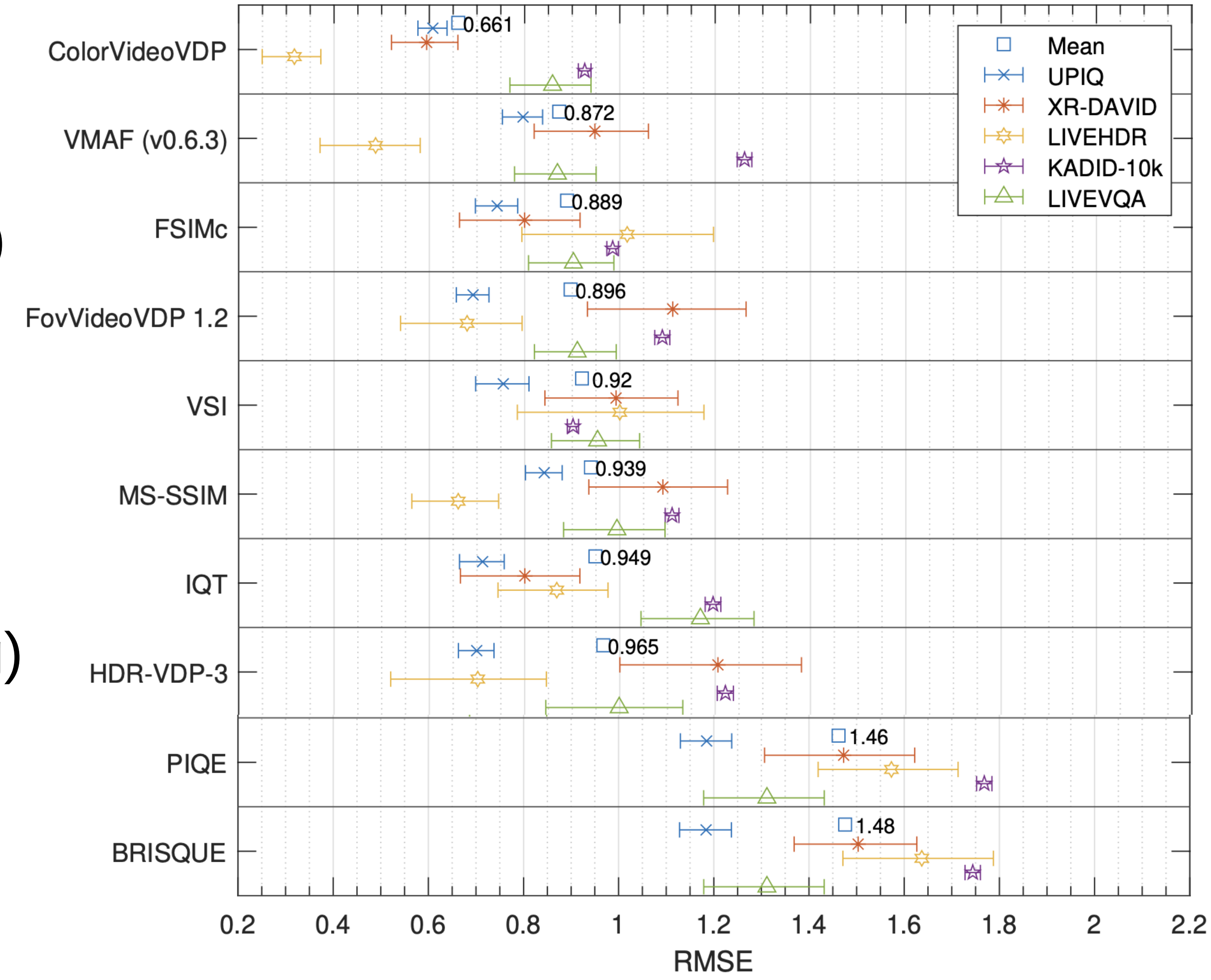
Training of ColorVideoVDP



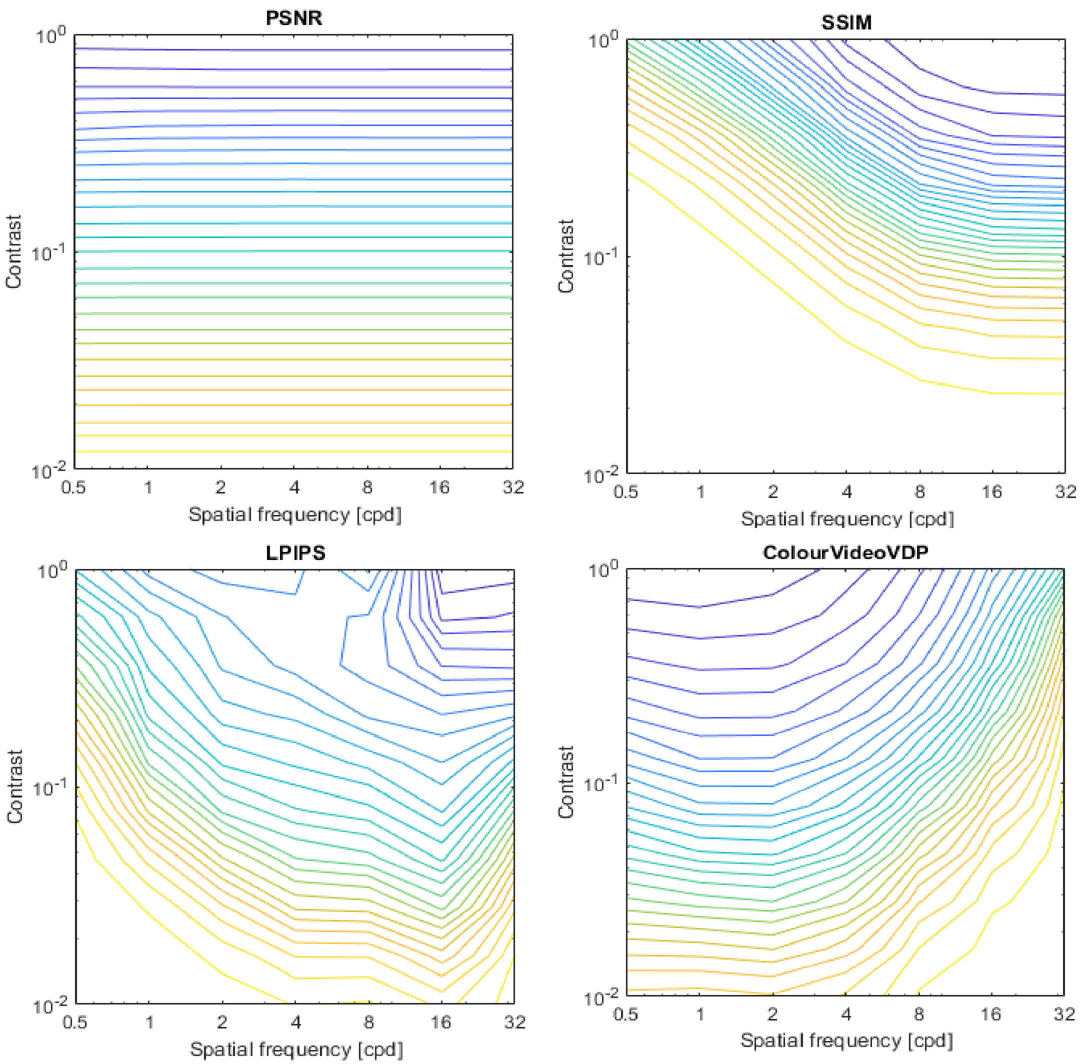
- 32 parameters in ColorVieoVDP
- ColorVideoVDP is fully differentiable – we can back-propagate gradients

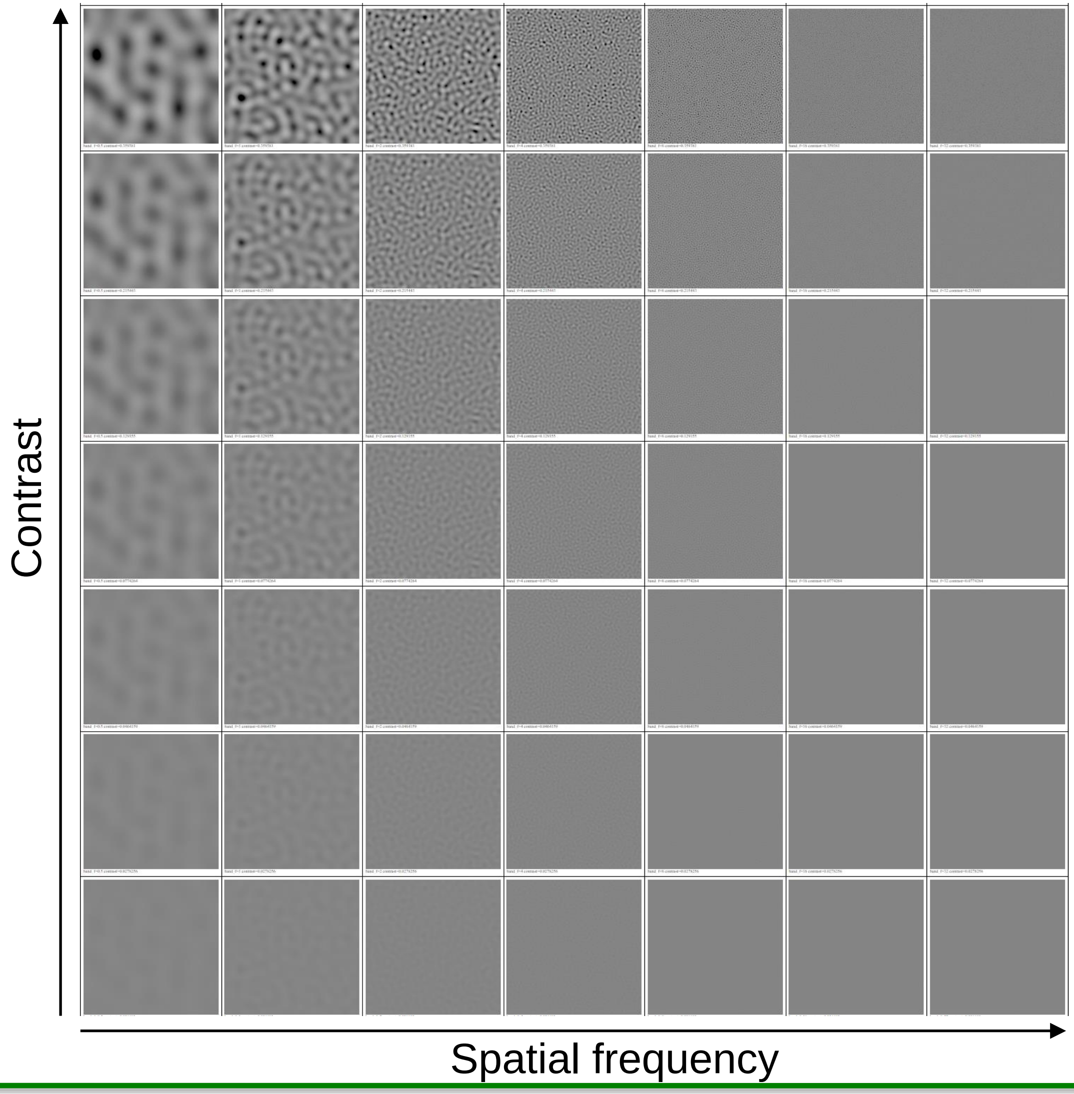
ColorVideoVDP predictions

- The metric was trained on
  - UPIQ (SDR and HDR images)
  - XR-DAVID (new, display distort.)
- And tested on
  - A test set of the datasets above
  - KADID-10k
  - LIVEVQA
  - LIVEHDR (HDR video streaming)
- Compared with 19 metrics



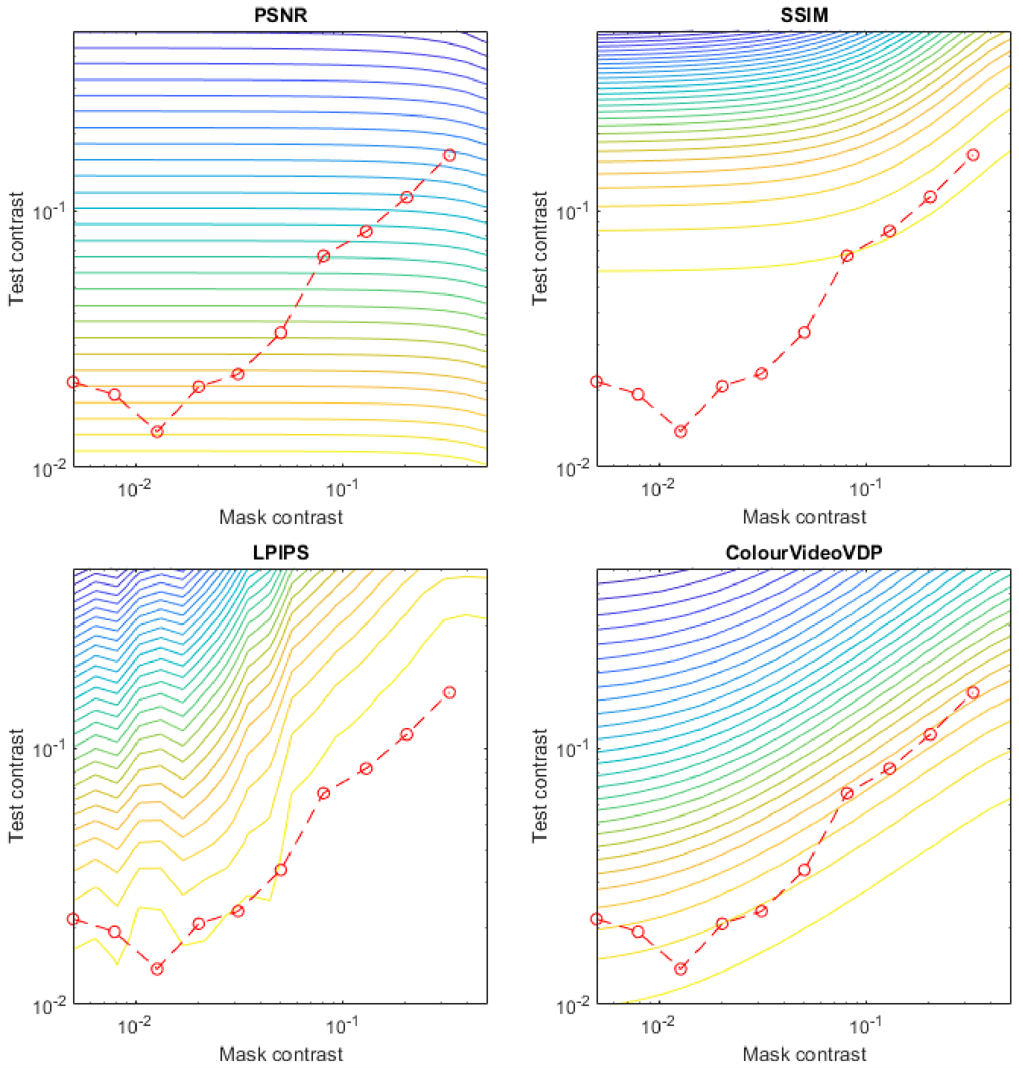
Metric performance on band-limited noise

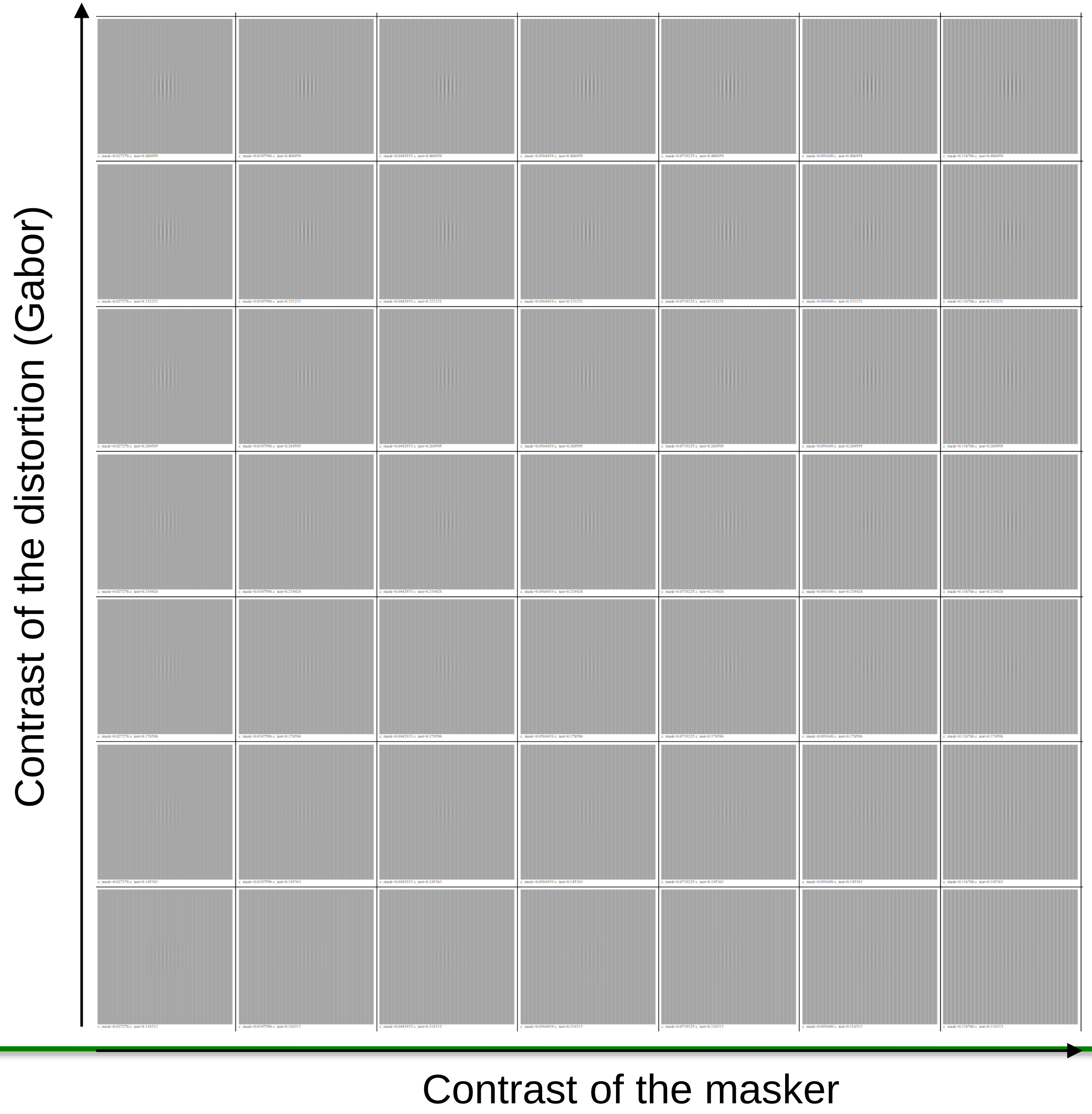




Violet – large difference; Orange – small difference

Metric performance on masking patterns

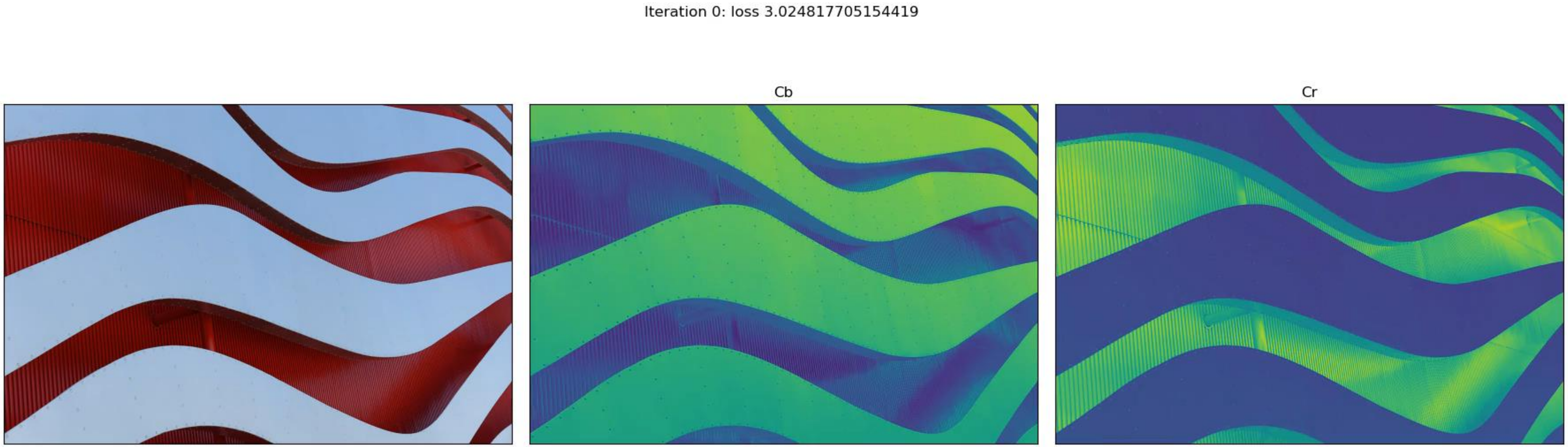




Violet – large difference; Orange – small difference

Example: ColorVideoVDP as a differentiable loss

- Adaptive chroma subsampling
- Removes invisible information to improve image/video coding



Summary

- Family of VDP visibility and quality metrics

Acknowledgements: Scott Daly, Karol Myszkowski, Kil Joong Kim, Allan G. Rempel and Wolfgang Heidrich, Manish Narwaria, Patrick Le Callet, Dounia Hammou, Param Hanji, Gyorgy Denes, Alexandre Chapiro, Anjul Patney, Maliha Ashraf

Metric	HDR	Color	Video	Visibility	Quality	Foveation	Glare	Aging	Differentiable
HDR-VDP 3.0 https://hdrvdp.sf.net/	✓			✓	✓		✓	✓	
FovVideoVDP https://github.com/gfxdisp/FovVideoVDP	✓		✓		✓	✓			
ColorVideoVDP https://github.com/gfxdisp/ColorVideoVDP	✓	✓	✓		✓				✓