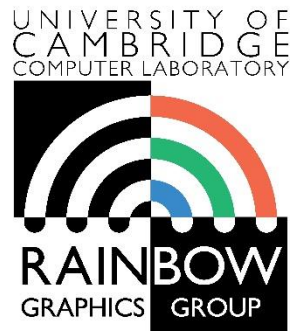
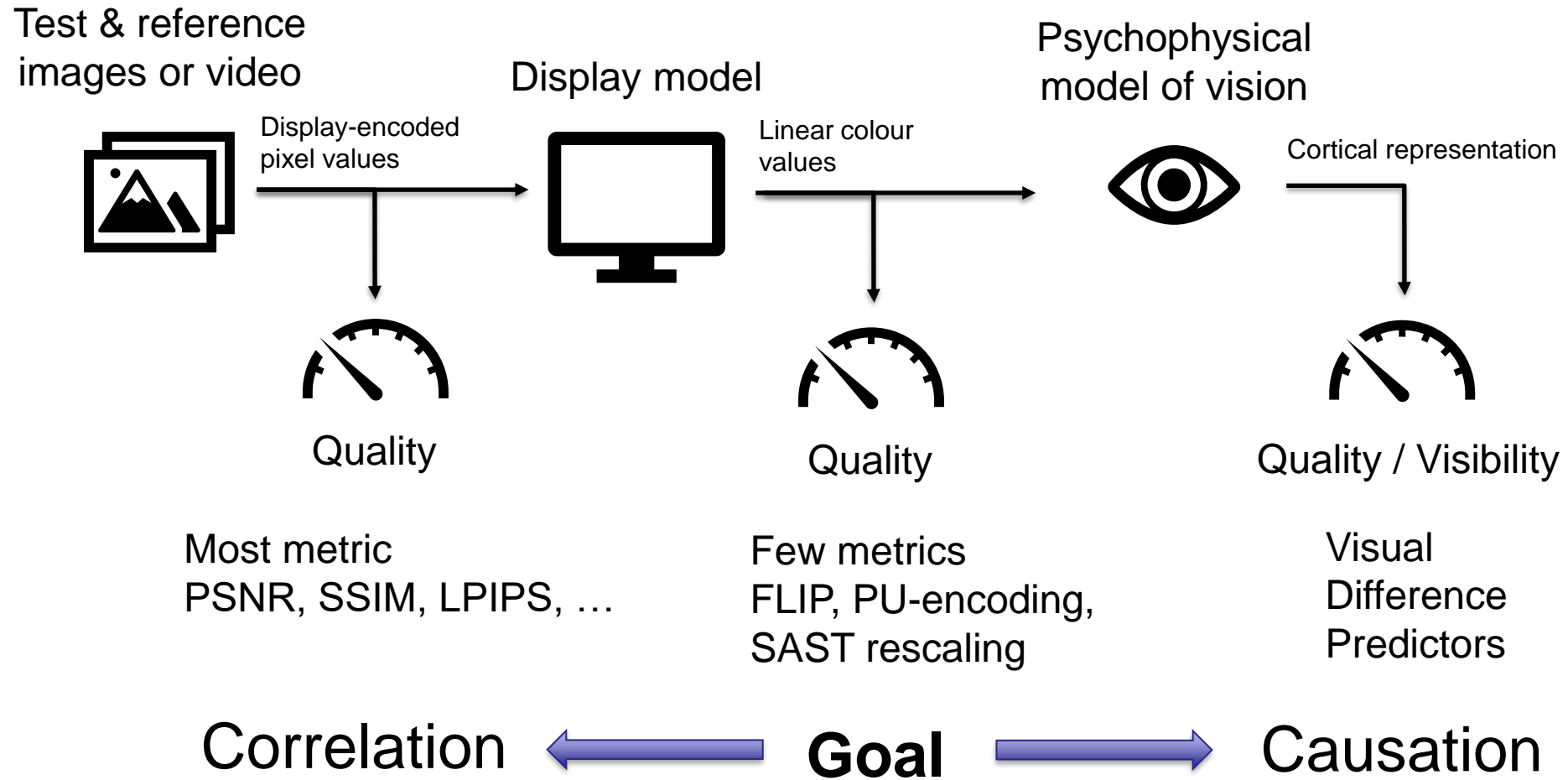


The family of VDP metrics for image and video quality predictions

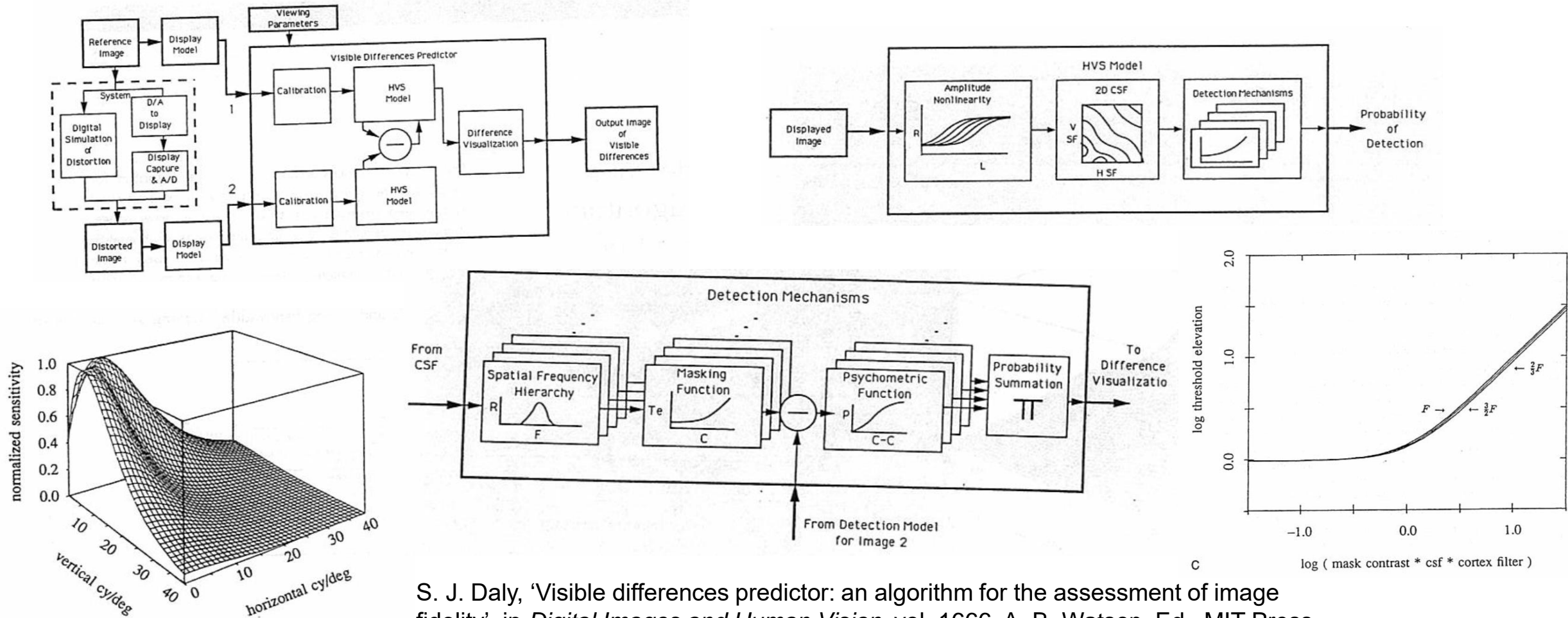
Rafał Mantiuk



Modelling visual quality

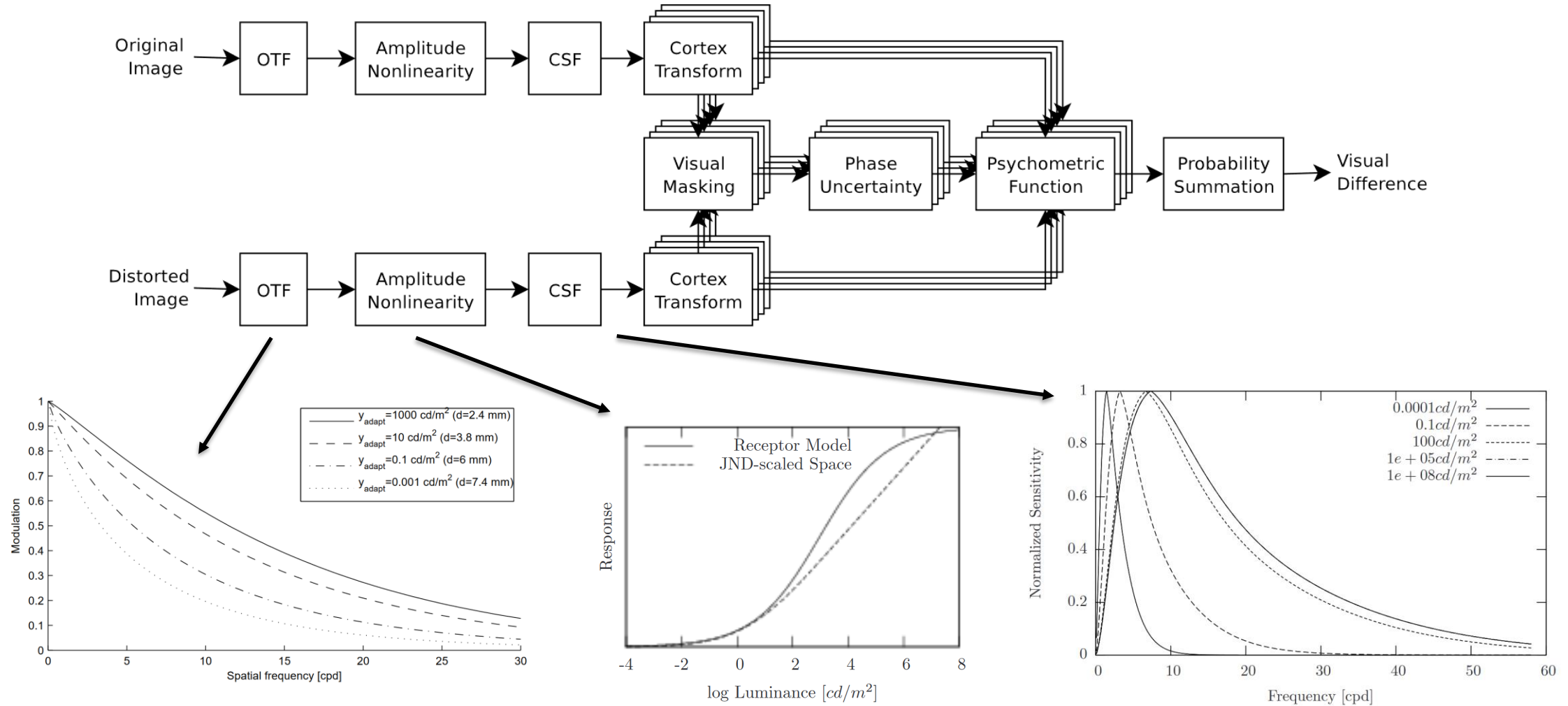


Visual Difference Predictor [Daly 1993]



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](https://doi.org/10.1117/12.135952).

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



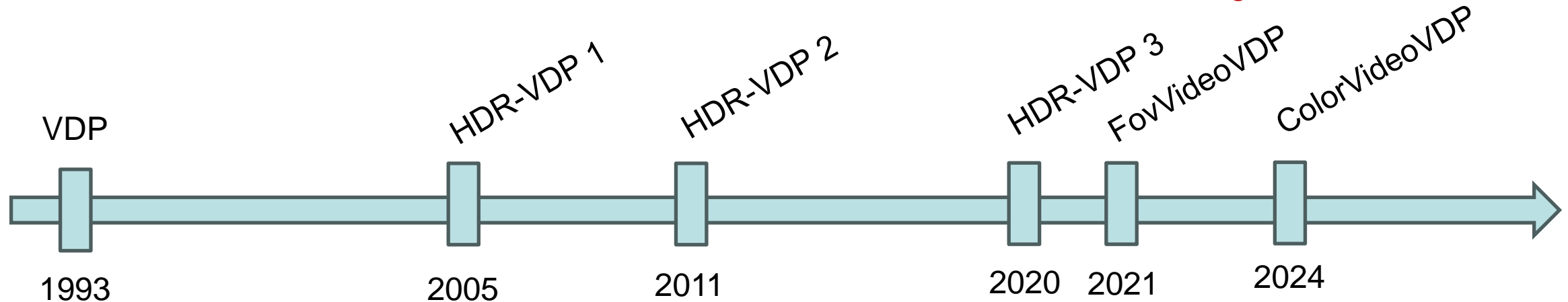
1993



SIGGRAPH

Test of Time Award

1st place in HDR Video Quality Measurement Grand Challenge



The *visual quality predictor* (VQP) algorithm for describing the human visual response is motivated by the need quantitatively to describe the visual consequences of decisions regarding the design and quality control of imaging products, intended to be used in the development of a new generation of imaging system hardware, and imaging media, it is a design tool that can find wide areas of application. The differences due to the imaging systems may begin as mathematical differences, i.e., incorrect colour, incorrect contrast, incorrect focus, etc., but they may also be due to differences in the image displayed, i.e., incorrect hue, incorrect brightness and chrominance) once the image is displayed. The goal of the VQP is to determine the degree to which these physical differences determine visible differences. Commonly used methods (e.g. 1, 2, 3, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100) as the system's MTF and noise power spectra and calculate a single number describing image quality. Although these techniques perform reasonably well for many aspects of testing media, they do not take into account the nonlinear aspects of human visual quality, the effects of adaptive algorithms, or the nonlinear aspects of analog media. The problems with these techniques lie in their lack of physical information in the complexity, their inability to deal with the effects of the nonlinear aspects of human visual quality, and their lack of simplicity relative to the complexity of the visual system.

To solve these problems the VDP uses a digital image-processing approach. Using actual images, rather than just parameters of the imaging system, enables the preservation of phase information. This information is necessary to predict visual distortion because of the masking properties of the visual system, in which the location of the image error is as important as the magnitude. Further, nonlinearities in the media or algorithms pose problems for the current approaches that use power spectra and MTF, because of their implications of linearity. An image-processing approach can easily incorporate such system parameters as MTF and noise power spectra through simulation, yet it also allows for more exact simulation of

Abstract Non-invasive and real-time monitoring of tissue oxygenation, intravascular and extracellular pH, and tissue temperature is the focus of high-resolution optical imaging. This paper reports on the development of a new generation of high-resolution optical imaging (HROI) images and videos. HROI images contain actual observations of physical targets, which cover an area 14 times of magnification, instead of 4 times magnification, found in standard images. The additional pixels provide more details of the targets, which are not visible in standard images. The HROI images are generated by a new generation of HROI cameras, which are capable of displaying contrast of 500:1, as compared to the contrast of 700:1 in LCD displays. With the development of high-resolution optical imaging techniques, a need for an intravascular fluid quality assessment of the circulating blood is expected.

In this paper we propose several modifications to the intravascular fluid quality (VFP) test. The modifications include the use of a probe with a larger diameter, the use of a larger diameter catheter, the use of a larger diameter catheter corresponding to real vessel observation. The proposed network aims to accept the spectra of light scattered from the vessel wall, the scattering of the light in the optics (FOV), nonlinear response to laser for the full range of contrast, and the use of a larger diameter catheter. The proposed network is capable of enhanced HROI display, capable of displaying the range of brightness that is close to that found in real tissue.

Keywords: High-resolution optical imaging; high-contrast range; HROI; projection; VFP; contrast sensitivity; VFP test.

Some imaging and rendering systems commonly use physically accurate lighting information in the form of Radiance Transfer (RT) images, textures, or maps. These data are useful in order to accurately represent scene appearances. Unlike their long-standing counterparts, HDR images, RT images contain the entire radiance range and full range of instances that is visible to a human observer. RT image data can be acquired via on-camera measurement campaigns, using multi-exposure techniques,¹ which involve taking several pictures at different exposures and then combining them together to create a single RT image. Another way to obtain RT images is reading out the RT values from a scene's physically-based rendering engine, which is often provided by the user. However, RT images can not be directly displayed on conventional LCDs or CRT monitors due to their limited dynamic range and color. Therefore, methods of instance compression (low mapping) and gamut mapping are required.^{2–5} Since traditional monitors cannot accurately display HDR data, new displays of extended contrast and numerous other features have been developed to better represent the information contained for HDR images. Several methods for HDR images^{6–10} and video^{11–13} have been proposed.

When designing an image interface or processing application, it is desirable to measure the visual quality of the resulting images. To avoid tedious subjective tests, where a group of people has to assess the quality, degradation, objective visual quality metrics can be used. The most successful objective metrics are based on models of the Human Visual System (HVS) and can predict such effects as a nonlinear response to luminance, limited sensitivity to spatial and temporal frequencies, and visual masking.¹⁴

R.M.: E-mail: marcel@nasa.gov; Telephone: +49 691 5105-427

[illegible]

Keywords: visual search, image quality, visual search, high-density, visual processing

Links: [DOI](#), [PMID](#), [PubMed](#)

1 Introduction

Violating results in computer graphics and imaging is a challenge.

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd



Keywords: Image Matrix; High-Dynamic Range

ness, which produces the probability of detecting differences in each part of the image. In other cases, we may want to use image processing algorithms that modify an image but should not introduce disturbing artifacts. An example is tone mapping, where an image tone mapped for a low-dynamic range display must be different from the original high-dynamic range image, but it should preserve the general visibility of the contrast. For such applications, we need to use a contrast detection metric. HDR-VDP-3 addresses all these applications using the same core visual model.

The short paper is not meant to be a complete description of the metrics, but a concise high-level overview with references to relevant papers, which [WACV 2019](#) HEUR-V2D 1 has the same processing pipeline as HEUR-V2D which is explained in detail in [Guentzel et al., 2011](#). In this short paper, we first explain HEUR-V2D with respect to the metrics [Guentzel et al., 2011](#) and [Guentzel et al., 2012](#). Then, we explain the differences with respect to HEUR-V2D 2 (Section 2) and finally explain how the metric was adapted to assess quality assessment for the WACV HEUR-V2D Quality Measurement Grand Challenge (Section 3).

HNR-VDP-3 is the third major iteration of the motif, which was originally inspired by seminal works on stability detection metrics by Daly [2005](#), Lubin [2006](#), and Watson [2009](#). The original HNR-VDP-1 [\[Mannan et al., 2008\]](#) is an extension of the VDP by Daly [2005](#). The extension incorporated changes allowing to compare high dynamic range images, for example of glass, photoreceptor response [\[Mannan et al., 2008\]](#), (presence of the PQ-function later

the 1998 cooling), and export sensitivity, which adapts to local abundance. Relative to the VEP, this metric, based



FoViduoVR is a video-differences system that models the spatial, temporal, and perceptual aspects of perception. While many other systems are available, no work provides the first practical treatment of these three central aspects of visual consciousness. The complex interaction between spatial

[illegible][illegible][illegible]

© 2006 Blackwell Publishing Ltd, *Journal of Internal Medicine* 260: 495–503



Figure 5 shows the ColorVideoDiff outputs for a pair of test reference videos. It consists of four panels: 'Reference video' (a person in a blue shirt), 'Output video' (a person in a blue shirt), 'Distance map' (a heatmap showing distance), and 'Diff again' (a heatmap showing difference).

NEW Reference Sources
Baker, H., Huxford, P., and Hogg, M. *Malika Adoni, Yoko Asano, and Hiroshi Chugan*. 2014. *Colloidal Gold: A novel reference product for image, radio, and stellar analysis*. In *2014 IEEE 11th Technical Days (TECHNICAL DAYS 2014)*.

[illegible]

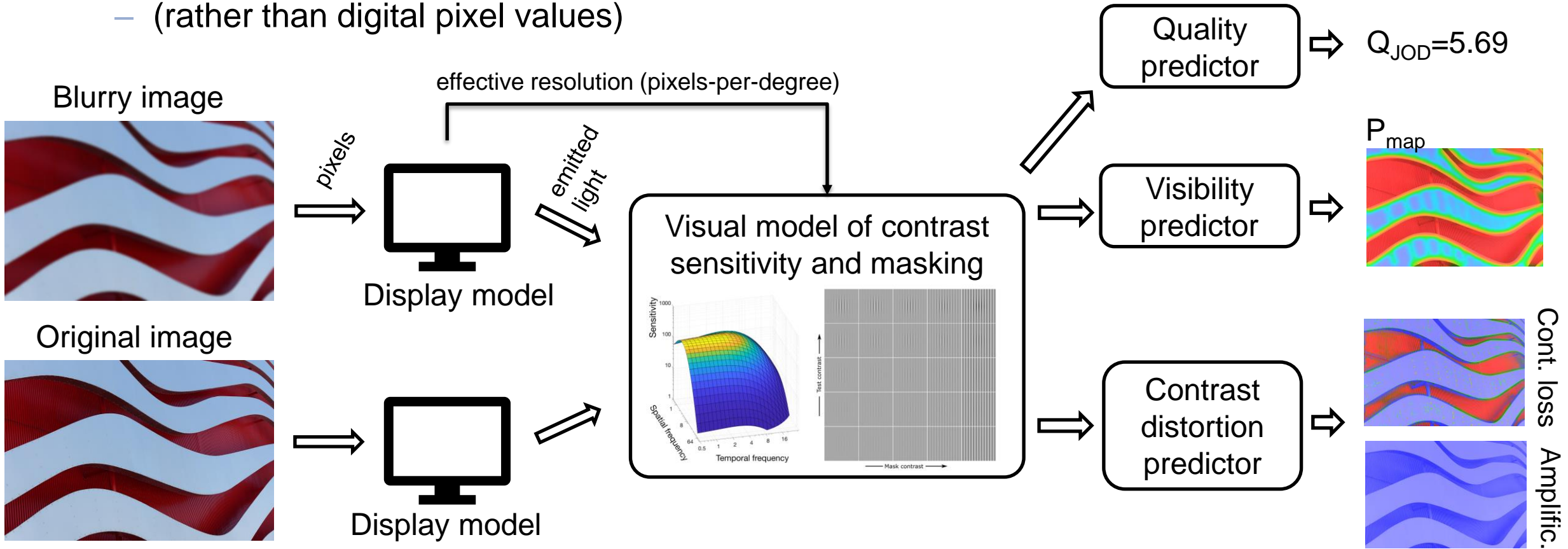
© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd, *Journal of Internal Medicine* 260: 395–404

For more information on this article, please contact the author at marco@cs.cmu.edu.
 Manuscript received 12/15/04; revised manuscript received 10/10/05; accepted manuscript received 10/10/05.
 Copyright © 2006 ACM 0734-4015/06/0000-0000\$5.00
 DOI: 10.1145/1135590.1135591

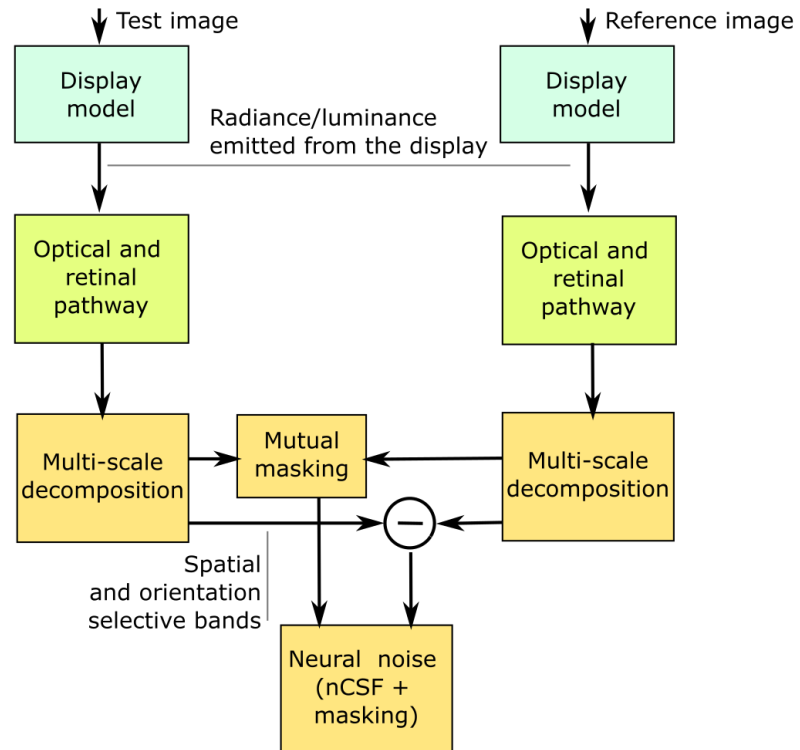
©Copyright 20 Technical Papers July 30-Aug 5, 2014, Boston, MA

What is HDR-VDP-3?

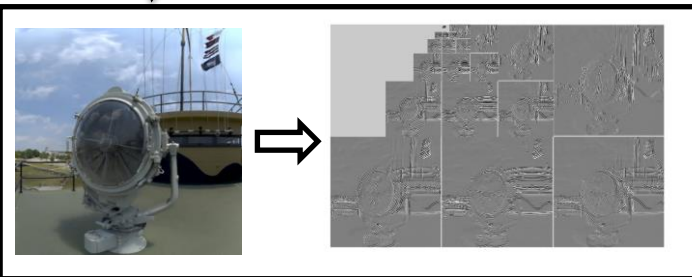
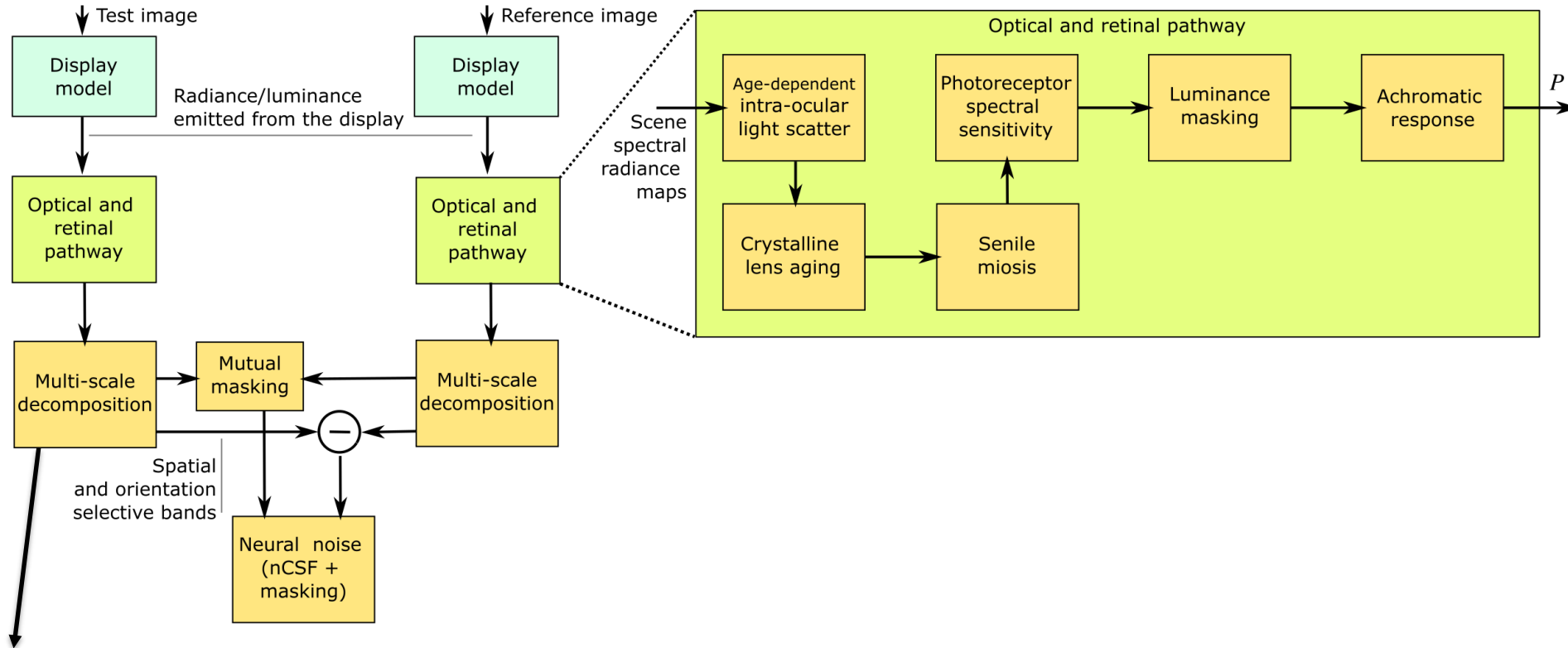
- A metric based on psychophysical models of **contrast sensitivity** and **masking**
- A metric that operates on **physical radiance/luminance units**
 - (rather than digital pixel values)



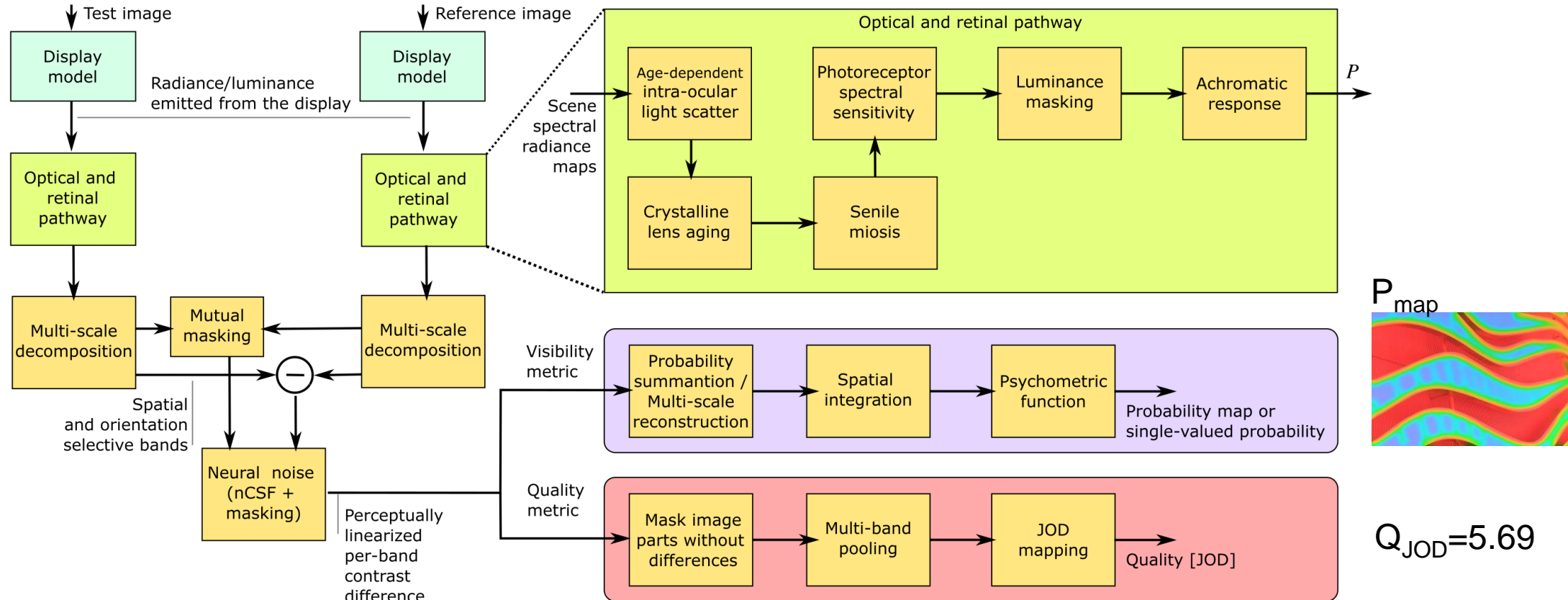
HDR-VDP-3 in more detail



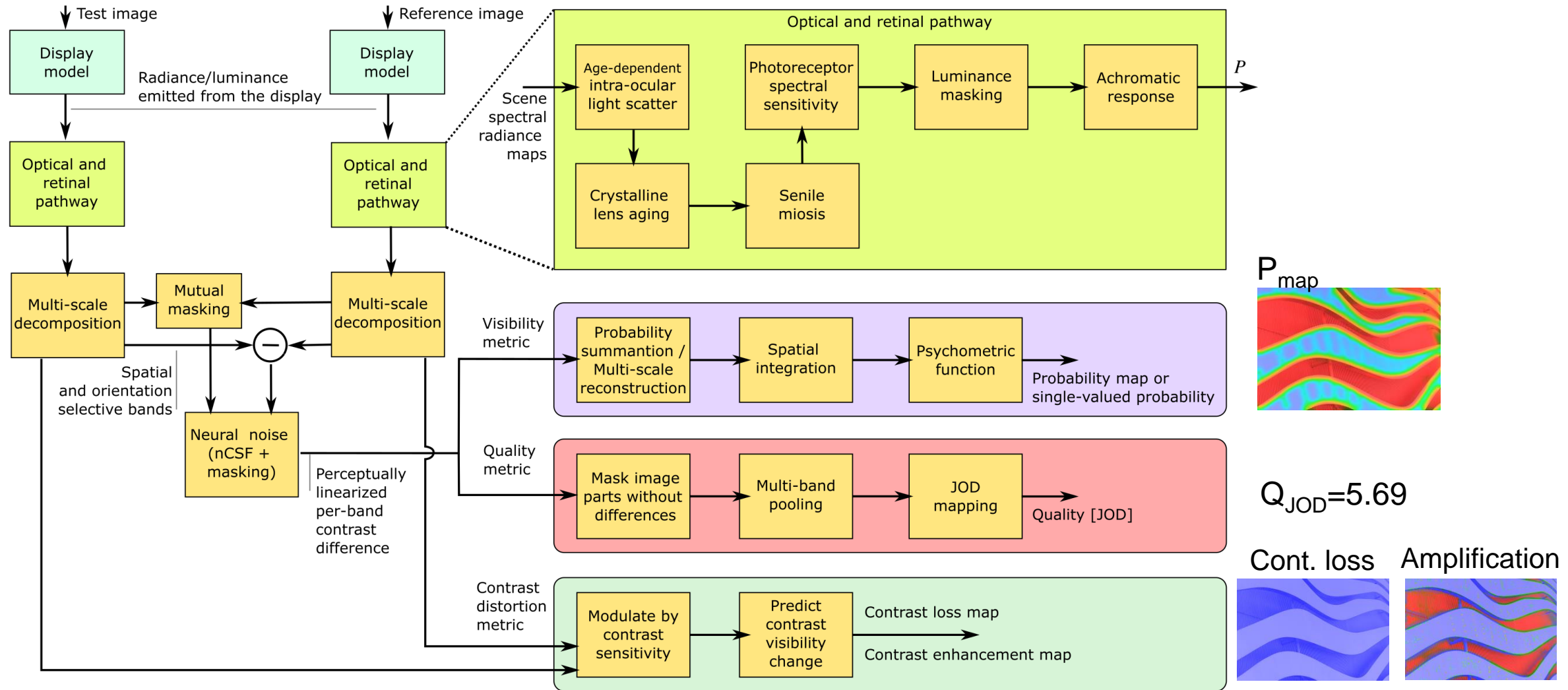
HDR-VDP-3 in more detail



HDR-VDP-3 in more detail



HDR-VDP-3 in more detail





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

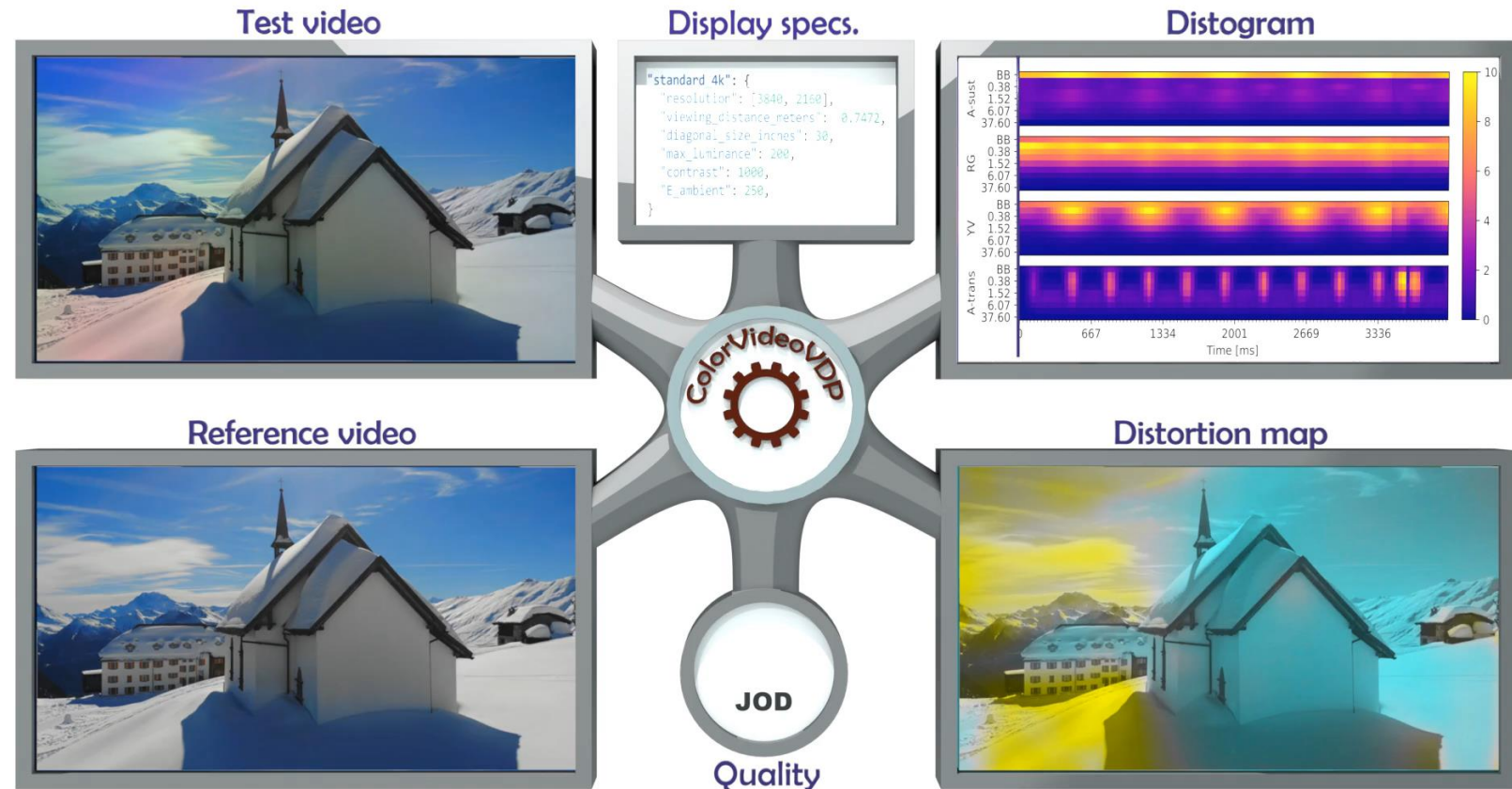
Rafał K. Mantiuk,
Param Hanji,
Maliha Ashraf,
Yuta Asano,
and Alexandre Chapiro




UNIVERSITY OF
CAMBRIDGE

What is ColorVideoVDP?

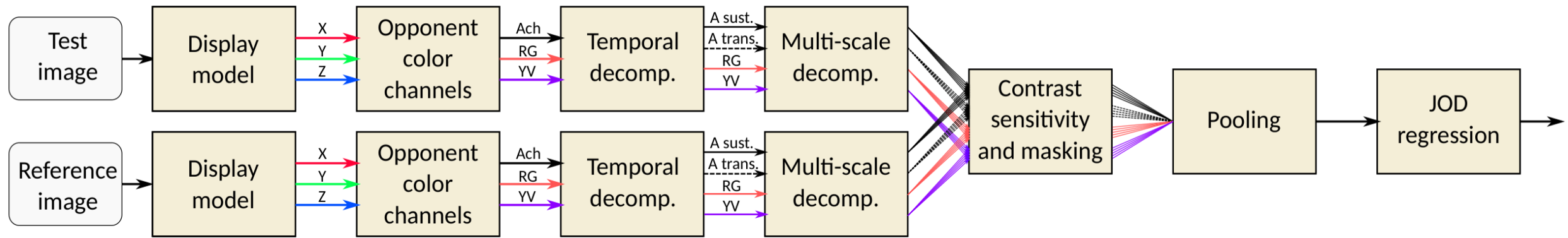
- A video quality metric
 - Models human colour and spatiotemporal vision
- Built on **castleCSF**
- Differentiable (PyTorch)
- Calibrated for
 - Video streaming
 - Display distortions



ColorVideoVDP – design goals

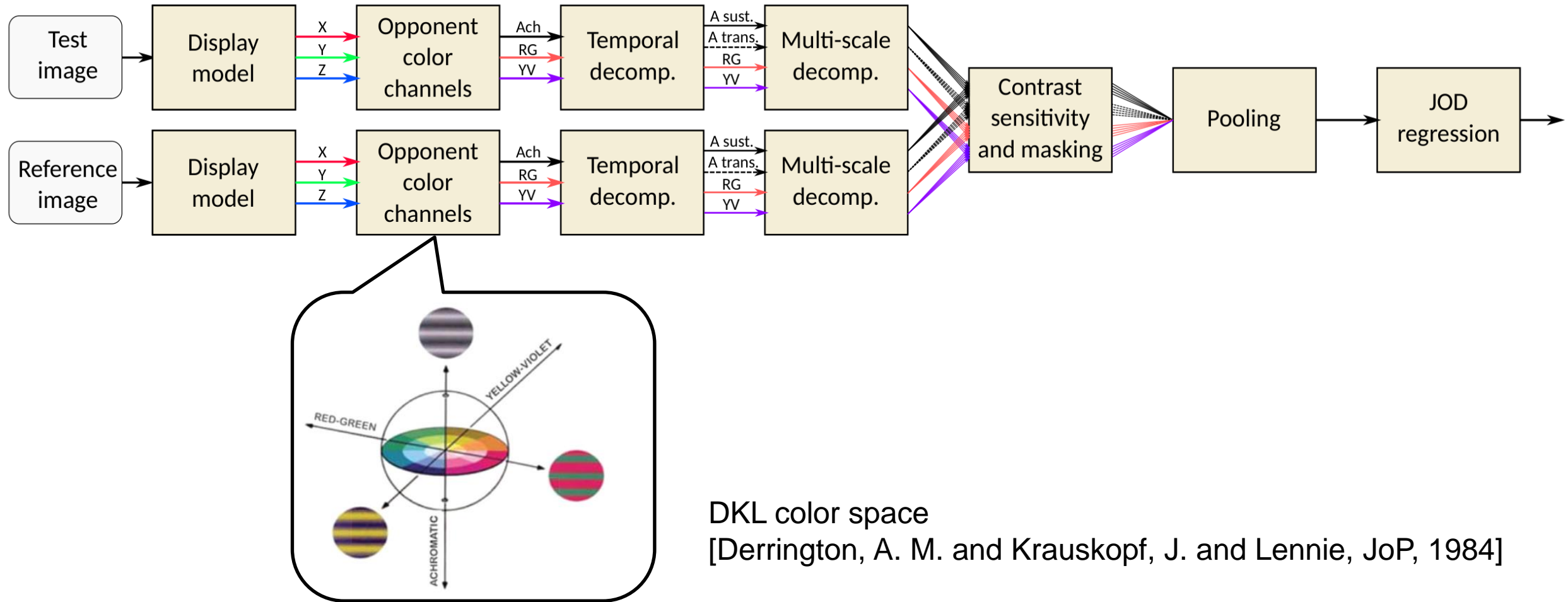
- Rely on [psychophysical models of low-level vision](#) as much as possible
 - Less chance for overfitting
 - More robust to corner cases
 - Can adapt to viewing conditions
- Make it [explainable](#)
 - No black boxes 
- Make it [simple](#)
 - A feature is not needed if the impact on the predictions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

ColourVideoVDP

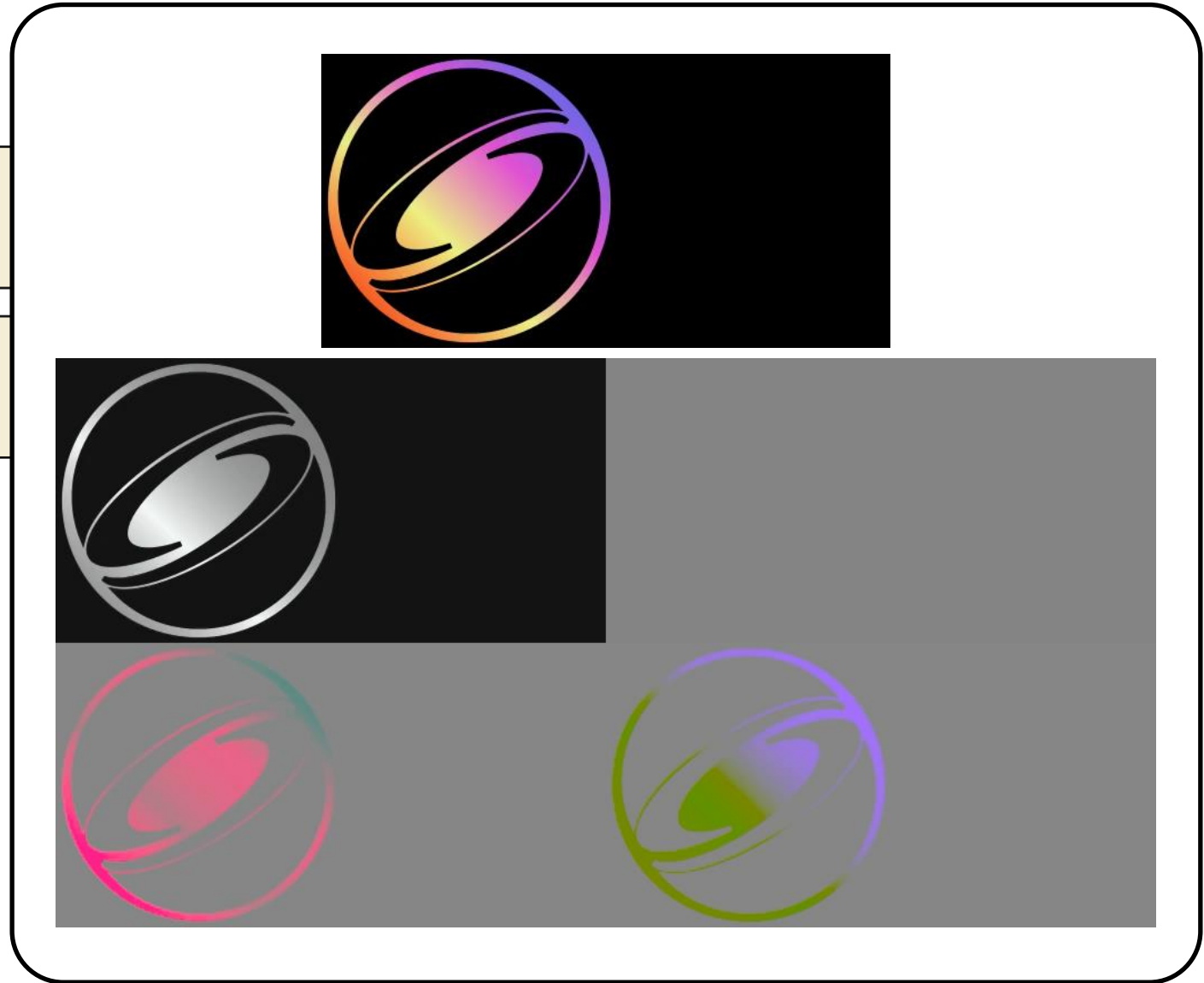
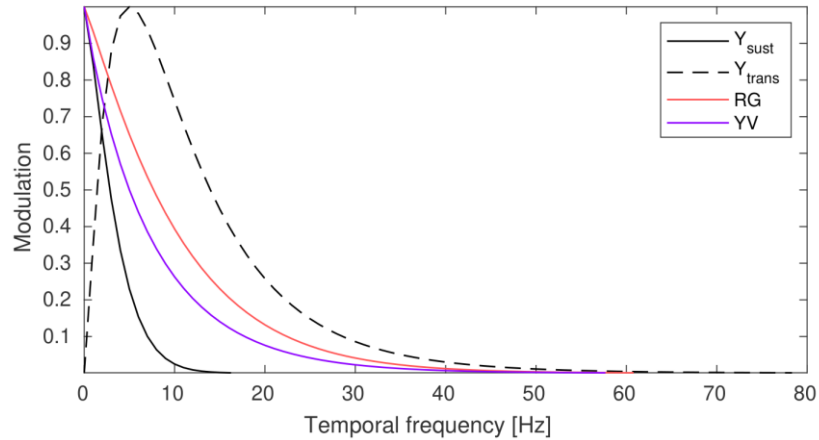
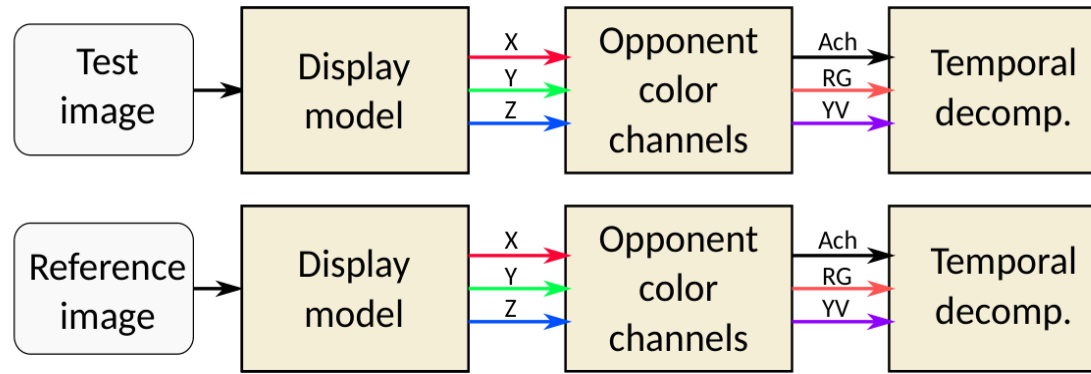


```
"standard_4k": {  
  "resolution": [3840, 2160],  
  "viewing_distance_meters": 0.7472,  
  "diagonal_size_inches": 30,  
  "max_luminance": 200,  
  "contrast": 1000,  
  "E_ambient": 250,  
}
```

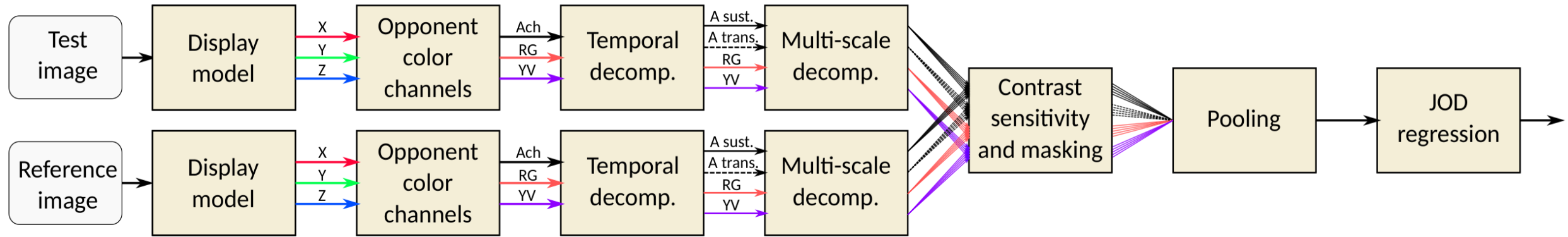
ColourVideoVDP



ColourVideoVDP



ColourVideoVDP



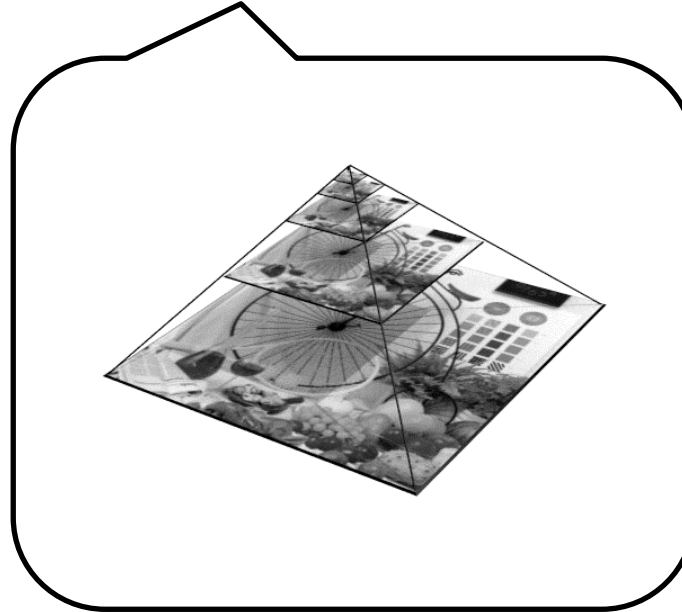
Contrast in complex images

DKL contrast

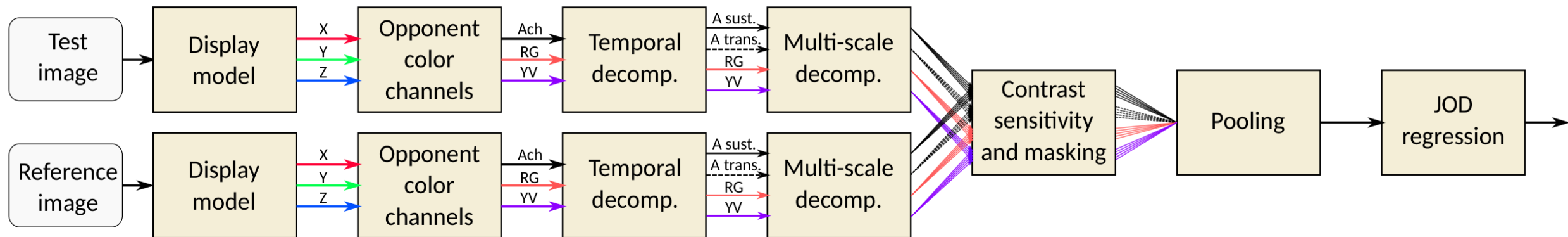
Laplacian pyramid coefficient

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

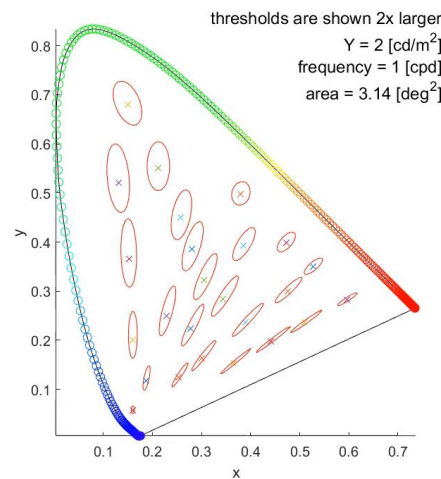
Upscaled Gaussian pyramid coefficient – sustained achromatic channel



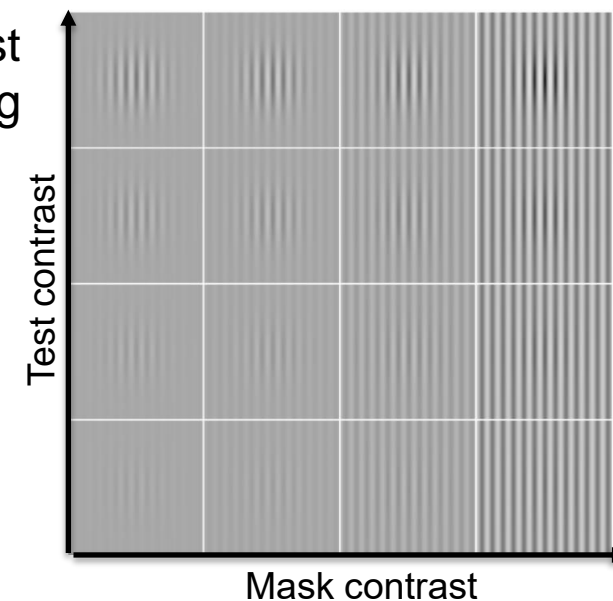
ColourVideoVDP



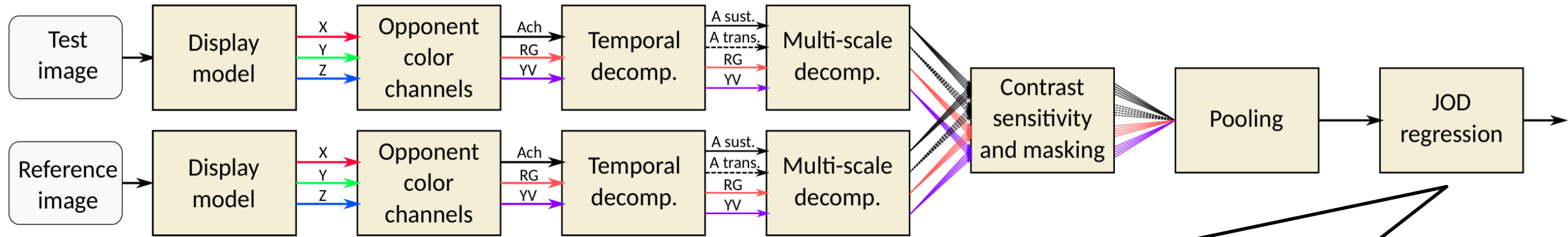
castleCSF
minimum
detectable
contrast
difference



Contrast
masking

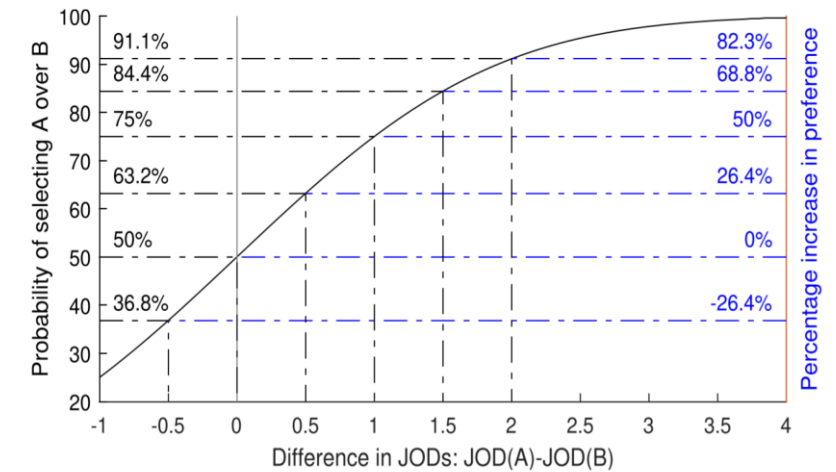


ColourVideoVDP



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

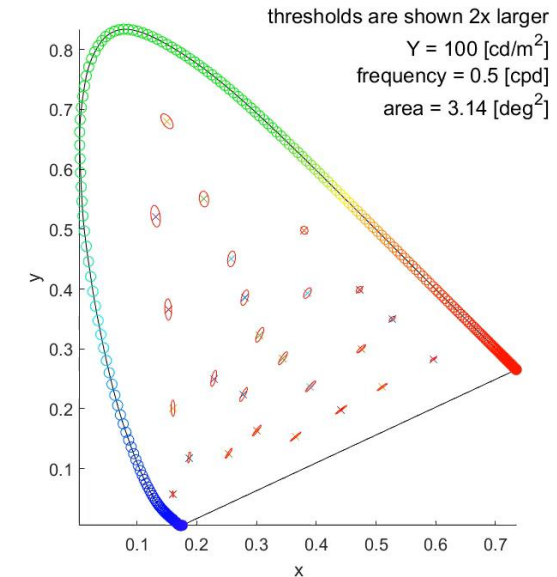
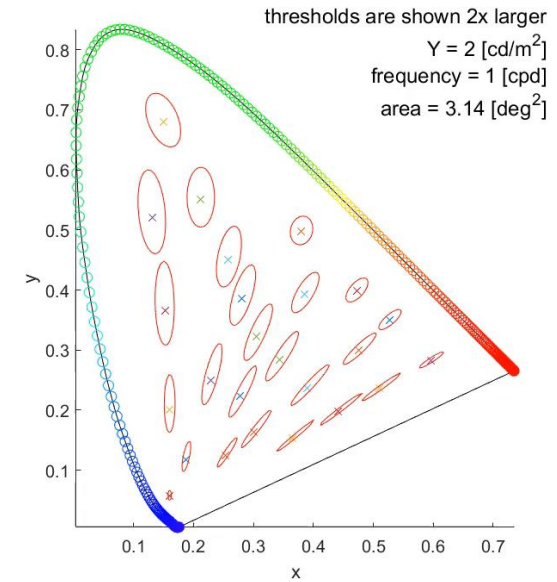
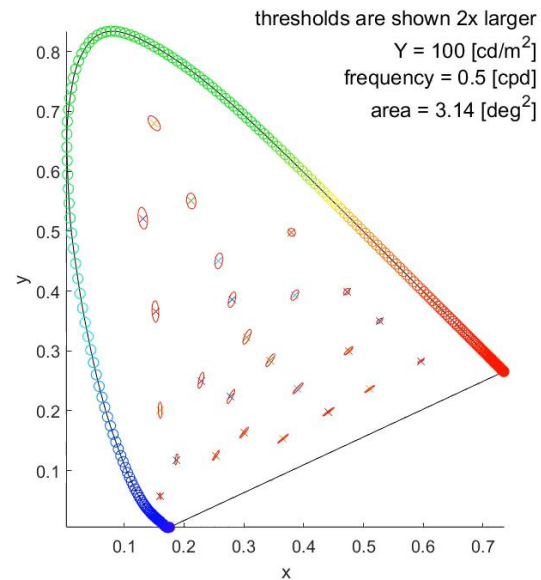
Can express supra-threshold (well-visible)
differences



castleCSF

Contrast Sensitivity Model that accounts for

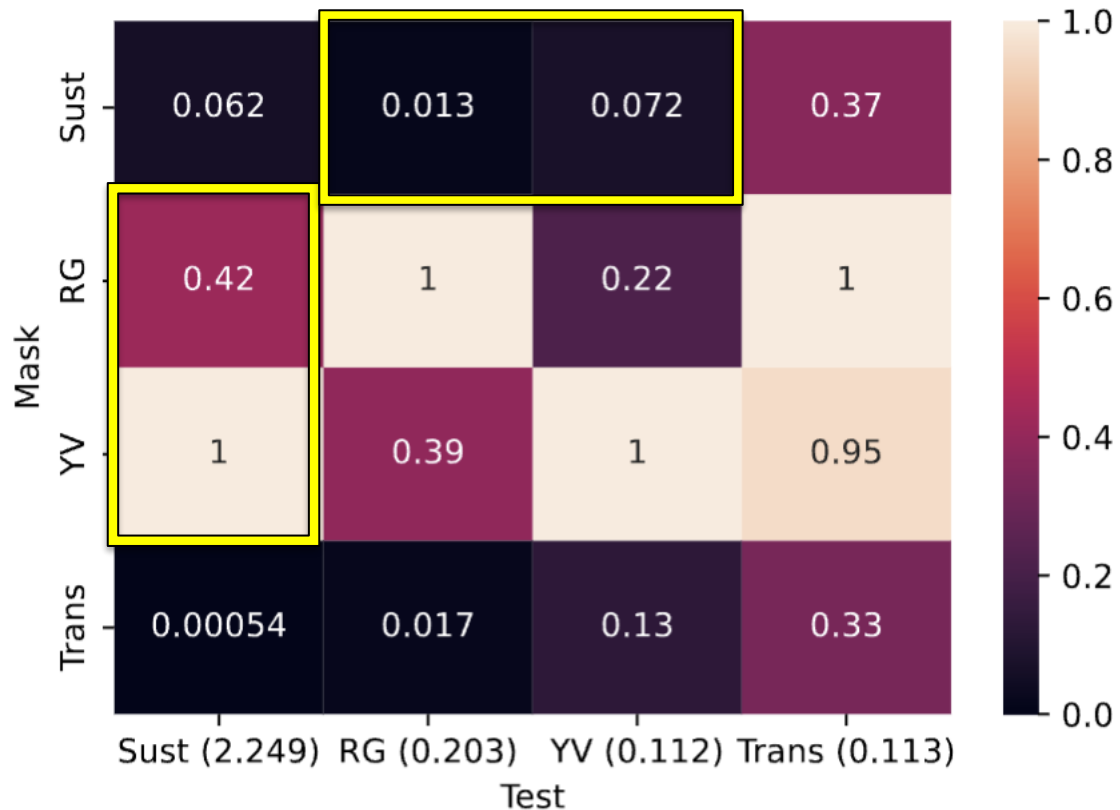
- Chromatic modulation
- Area
- Spatial freq.
- Temporal freq.
- Luminance
- Eccentricity



[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

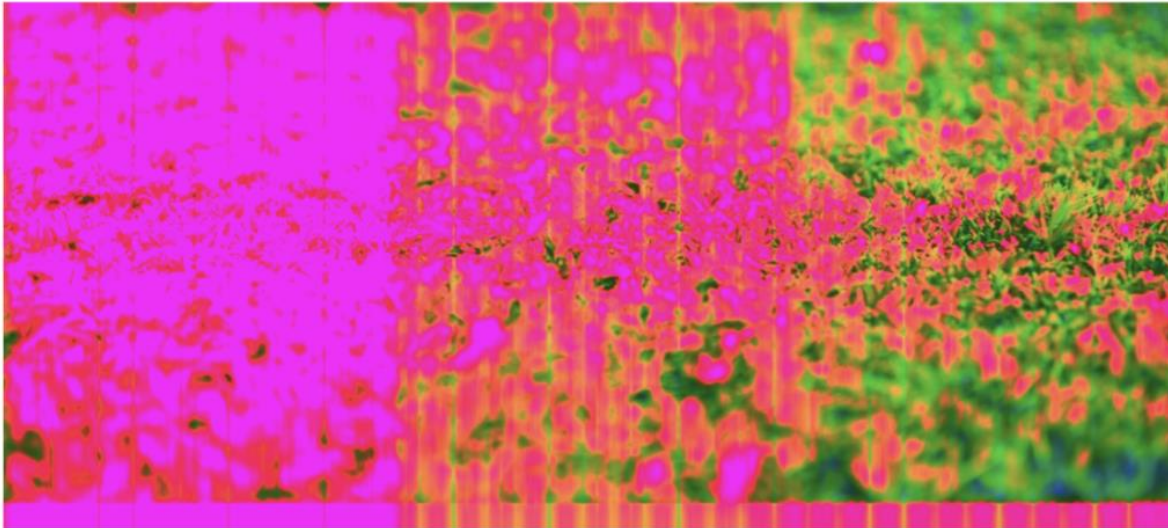
Reference image



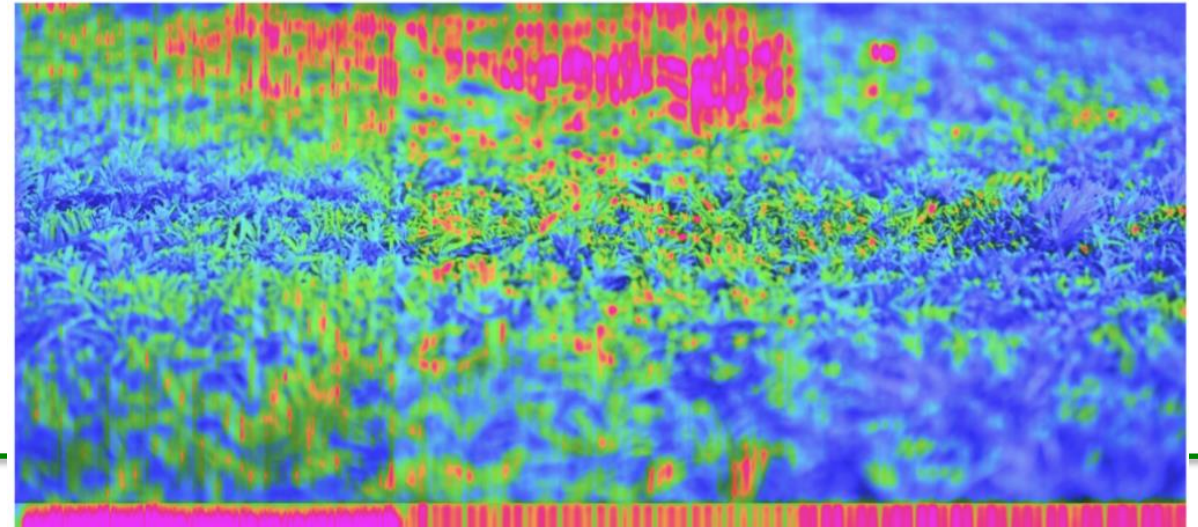
Test image



Visual difference map — no masking

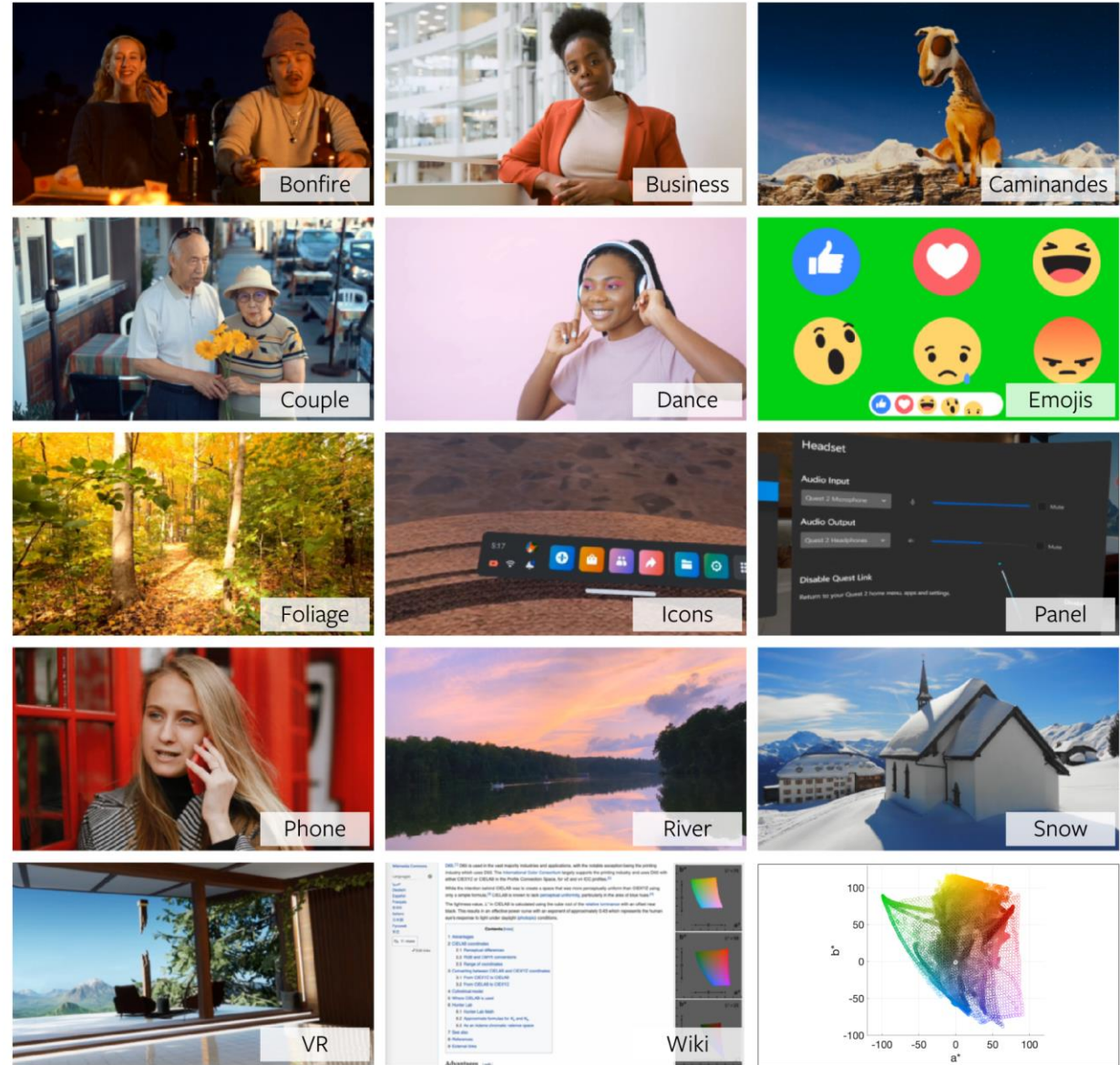


Visual difference map — with masking



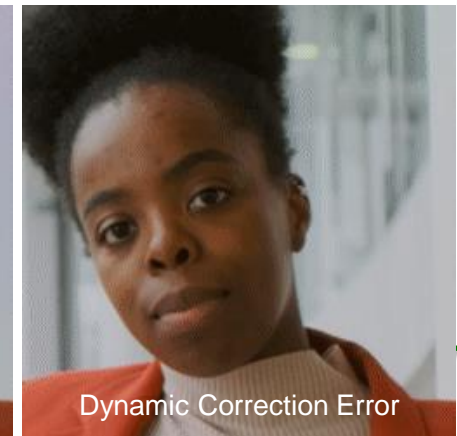
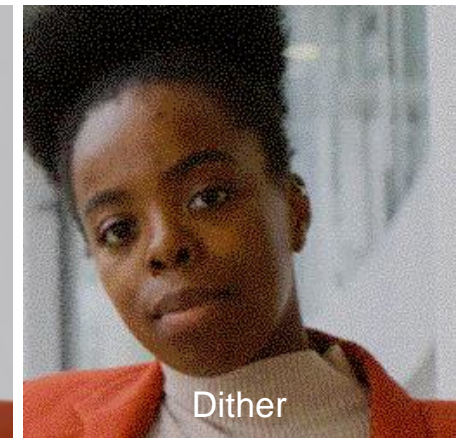
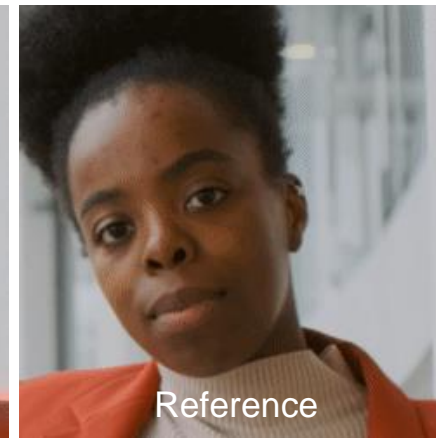
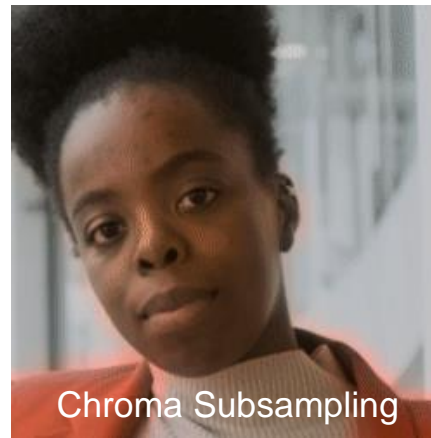
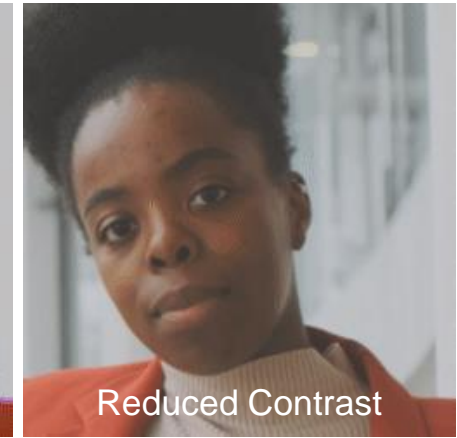
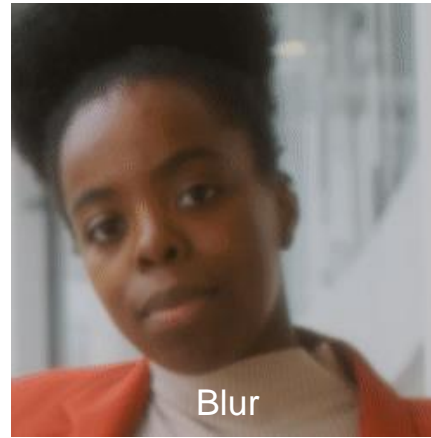
XR Display Artifact Video Dataset (XR-DAVID)

- 14 scenes including representative use-case scenarios



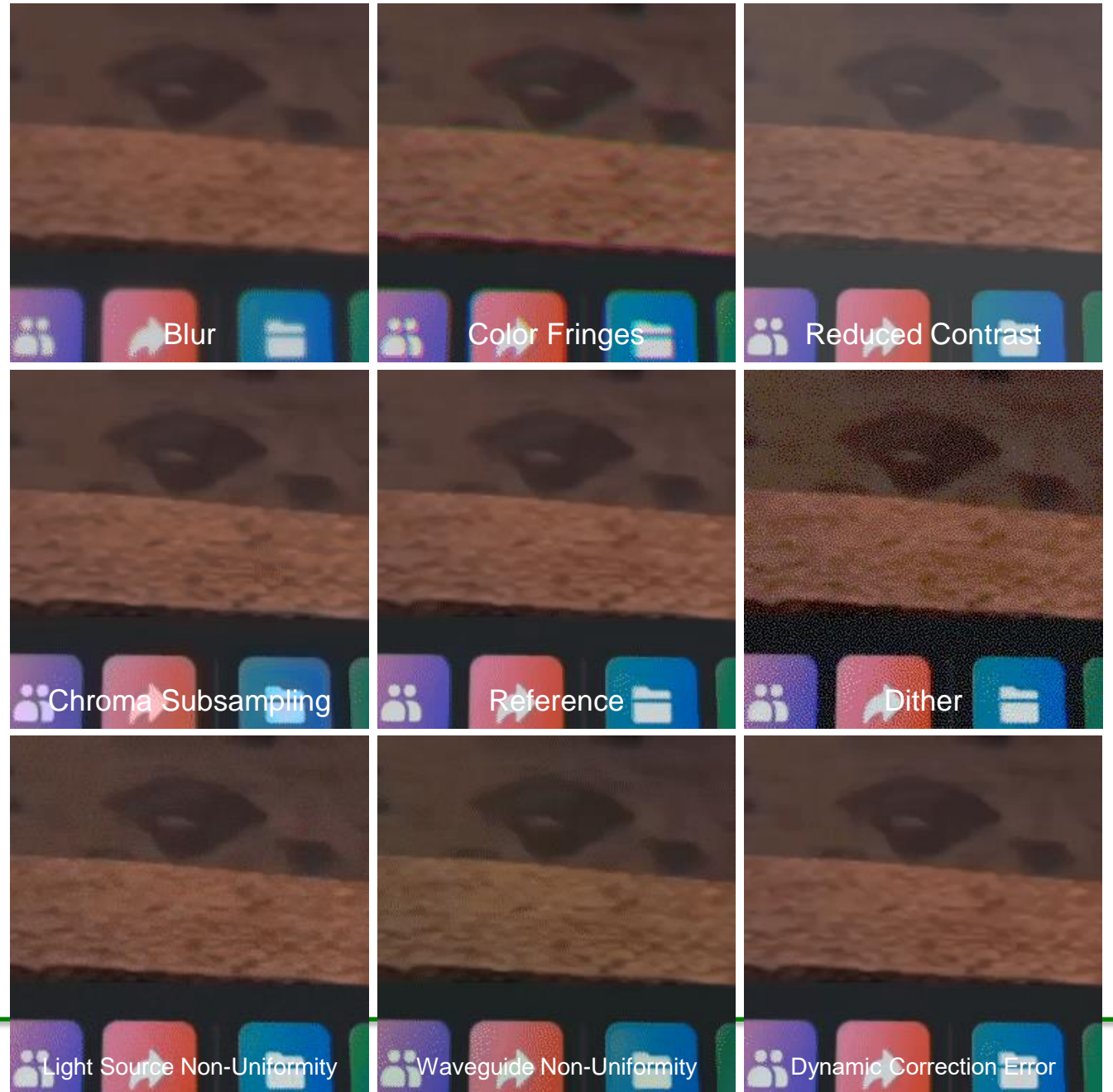
XR Display Artifact Video Dataset (XR-DAVID)

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

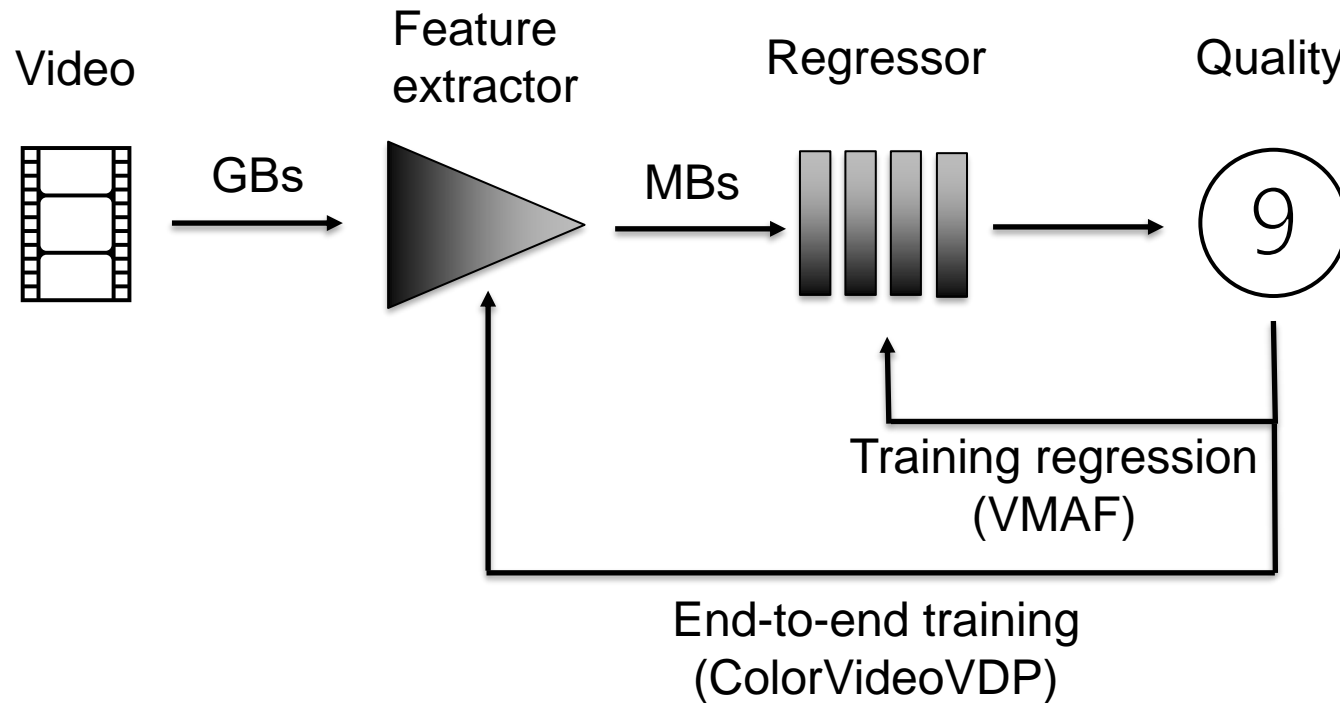


XR Display Artifact Video Dataset (XR-DAVID)

- 14 scenes including representative use-case scenarios
- ⑩ 8 artifacts common in display and optics applications
 - 40 PPD, 300 nit display
- ⑩ 77 participants
 - ⑩ Pairwise comparisons with ASAP active sampling



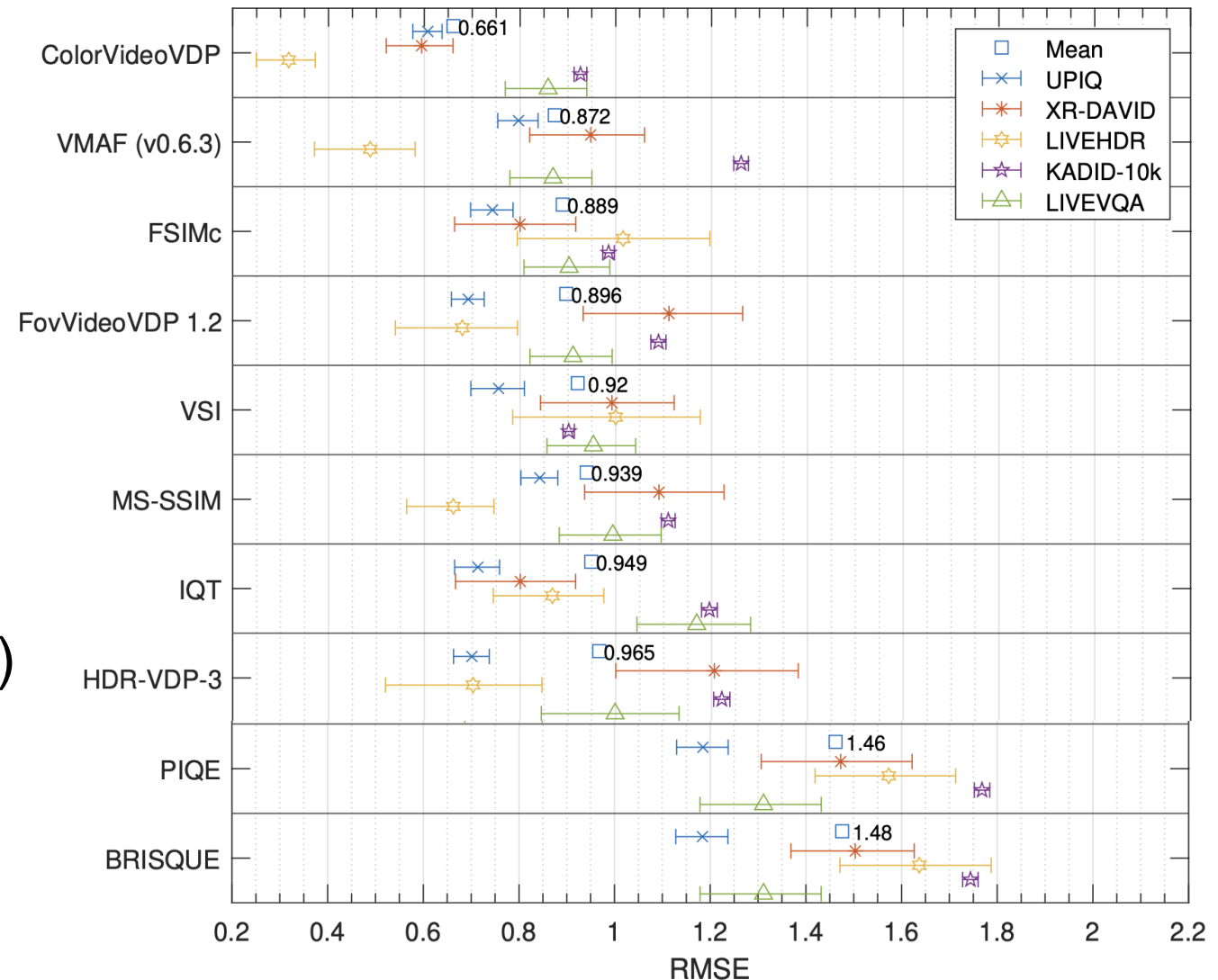
Training of ColorVideoVDP



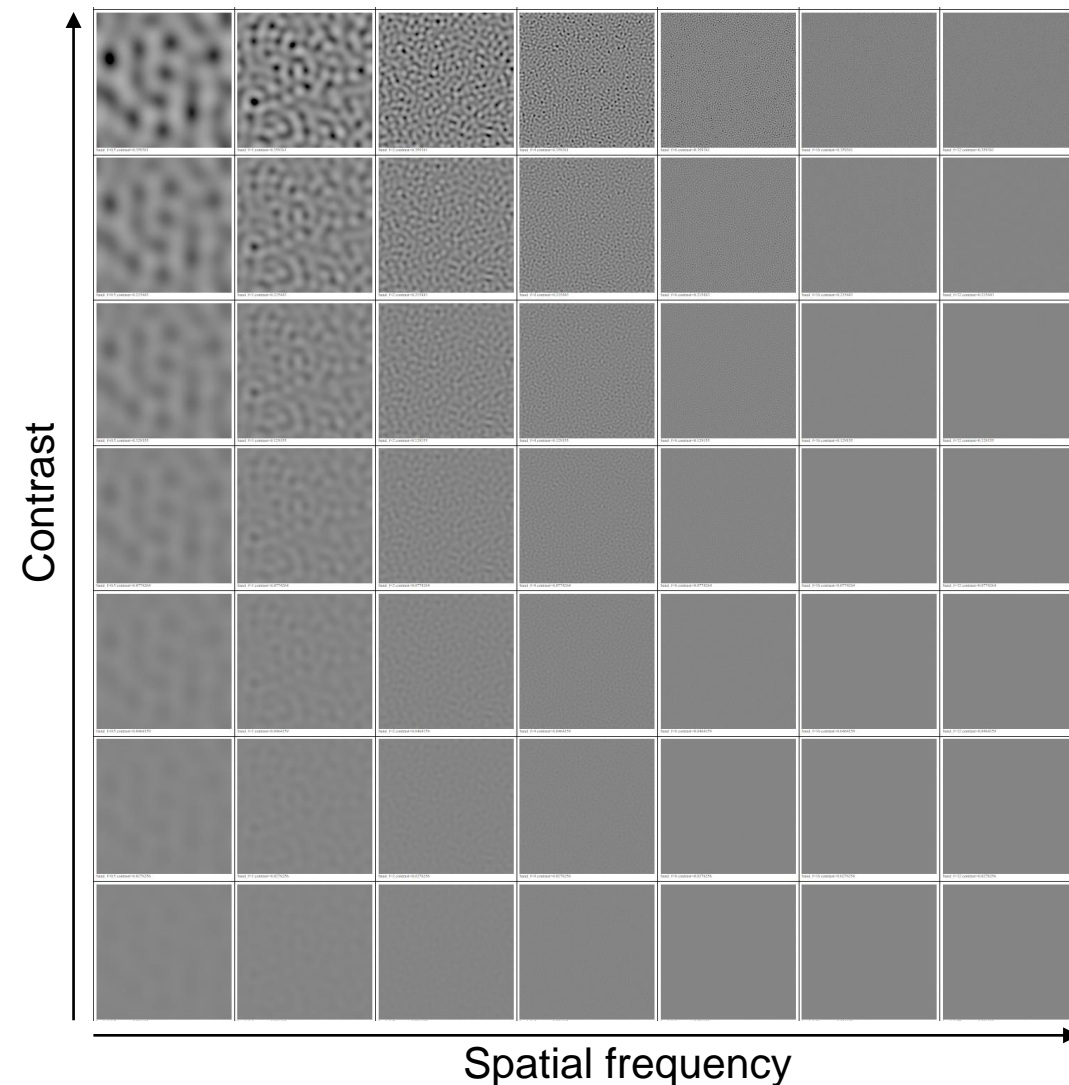
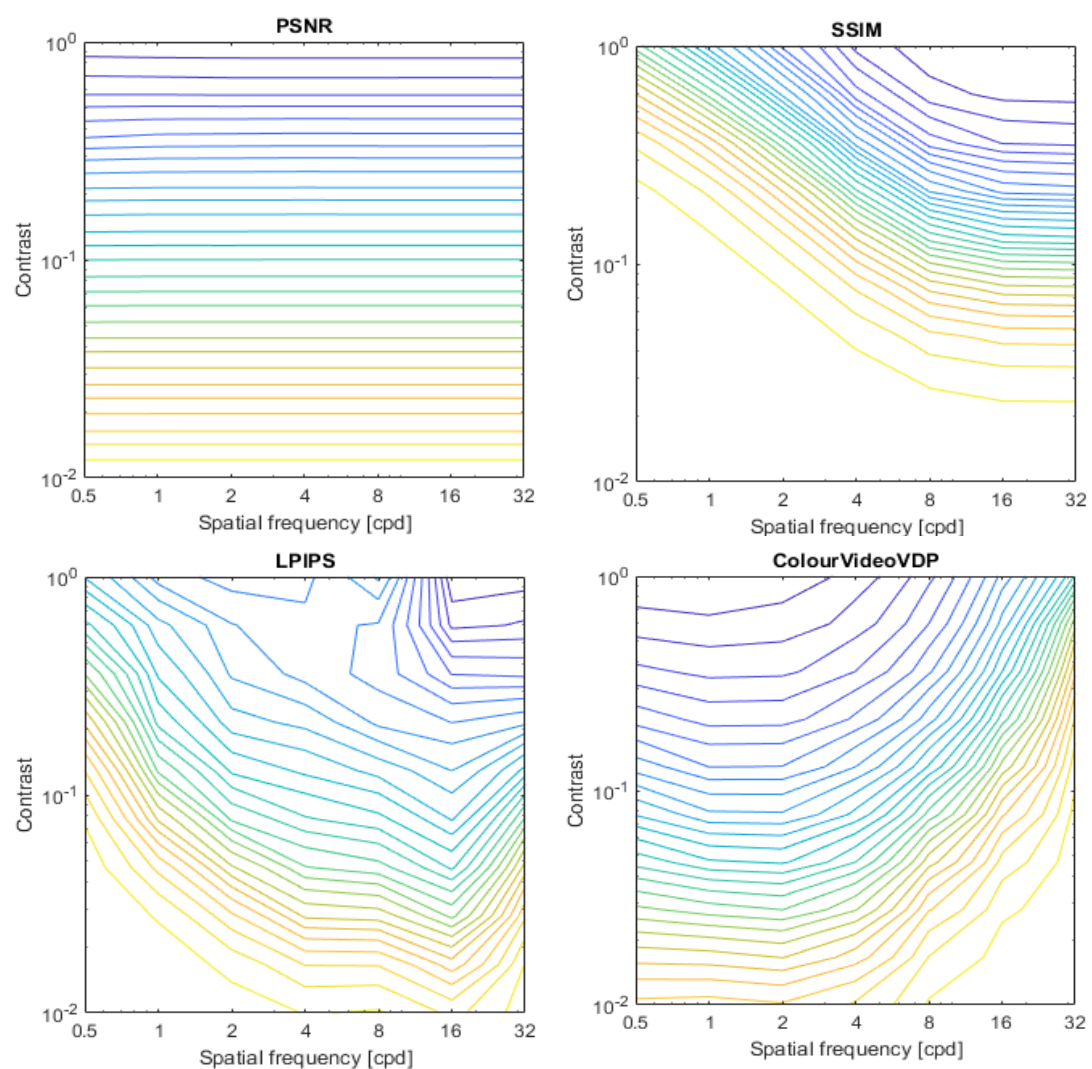
- 32 parameters in ColorVieoVDP
- ColorVideoVDP is fully differentiable – we can back-propagate gradients

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

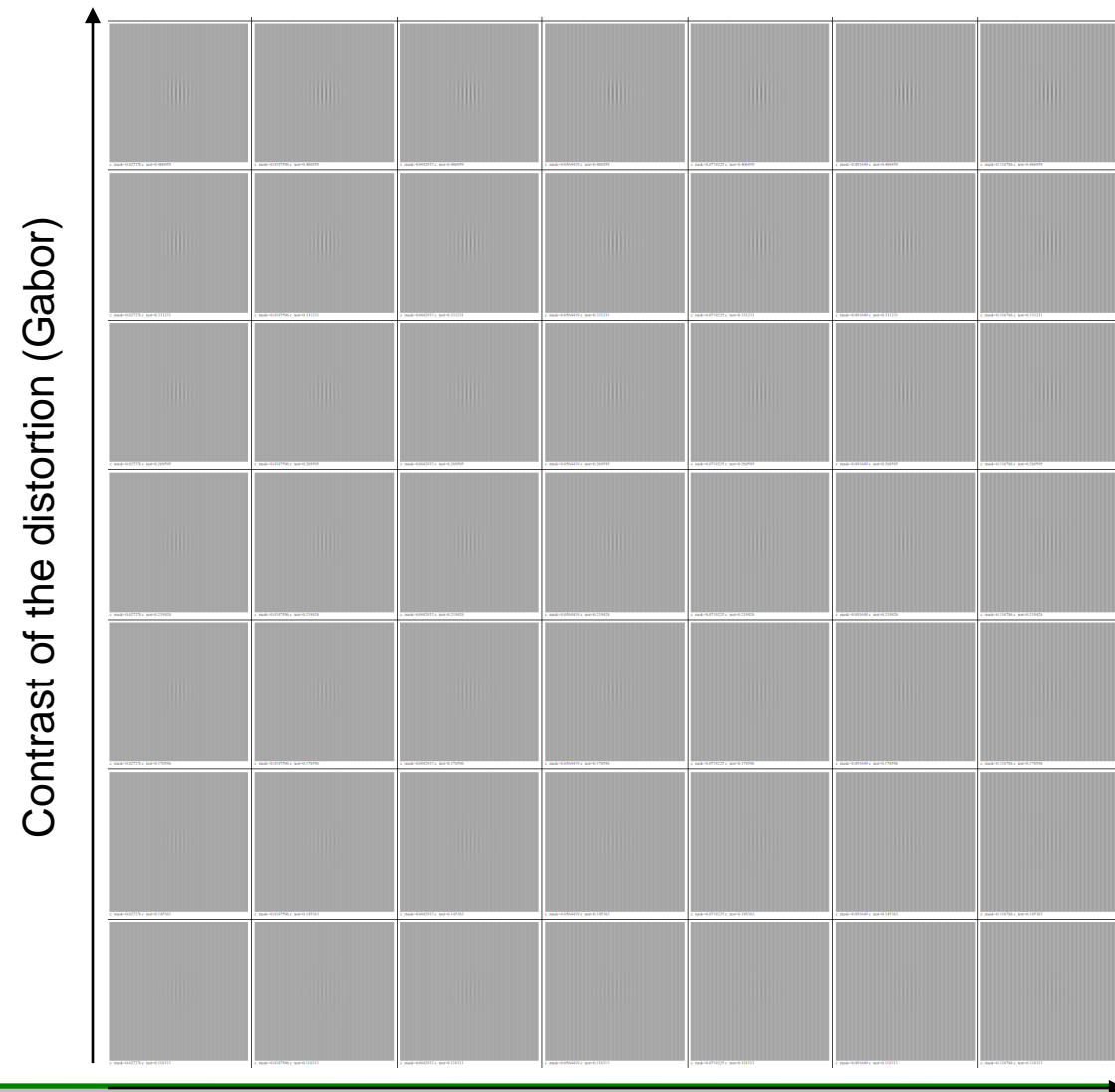
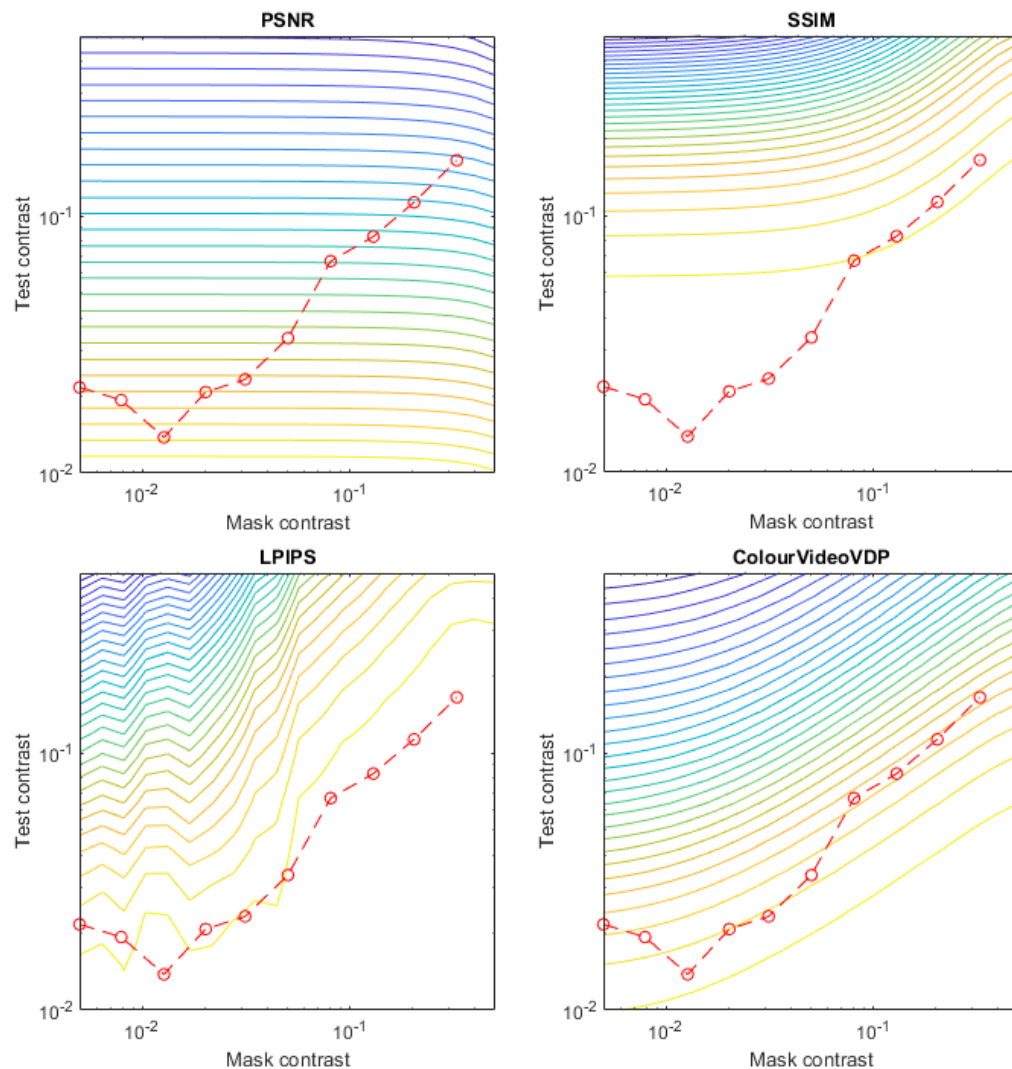


Metric performance on band-limited noise



Violet – large difference; Orange – small difference

Metric performance on masking patterns



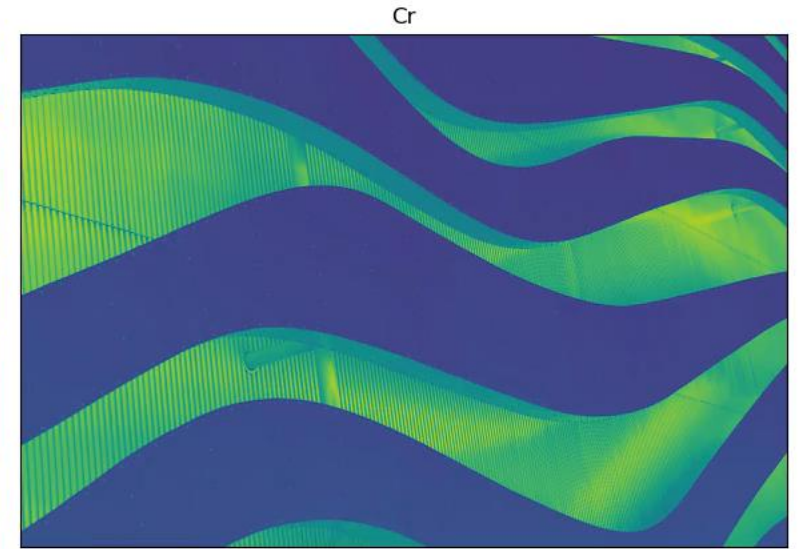
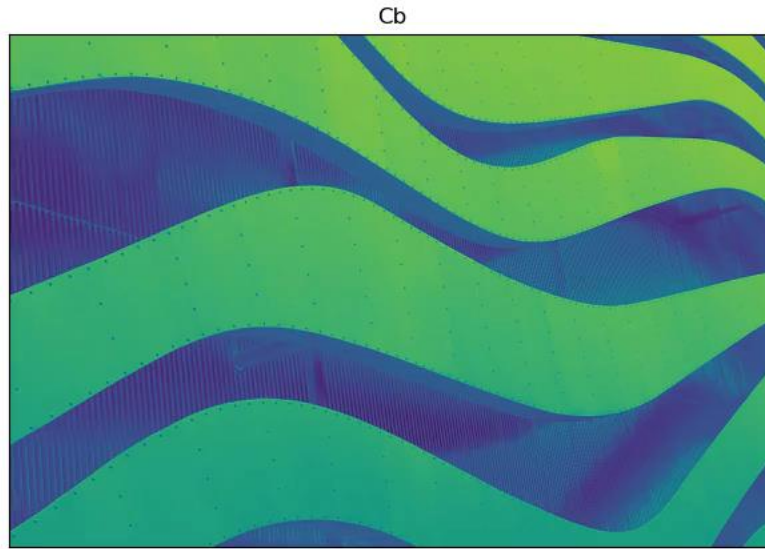
Violet – large difference; Orange – small difference

Contrast of the masker

Example: ColorVideoVDP as a differentiable loss

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

Iteration 0: loss 3.024817705154419



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	✓	✓	✓		✓				✓