JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Learning a Practical SDR-to-HDRTV Up-conversion using New Dataset and Degradation Models

C. Guo[1,2],
L. Fan[1,3],
Z. Xue[2,4],
X. Jiang[1,2]

Peng Cheng Laboratory (PCL), Shenzhen, China **1**

2 *State Key Laboratory of Media Convergence and Communication (MCC), Communication University of China (CUC), Beijing, China*

Shenzhen Graduate School, Peking University, Shenzhen, China **3**

4 *Academy of Broadcasting Science (ABS), National Radio and Television Administration (NRTA), Beijing, China*

THU-PM-155

Scan the QR code for more information

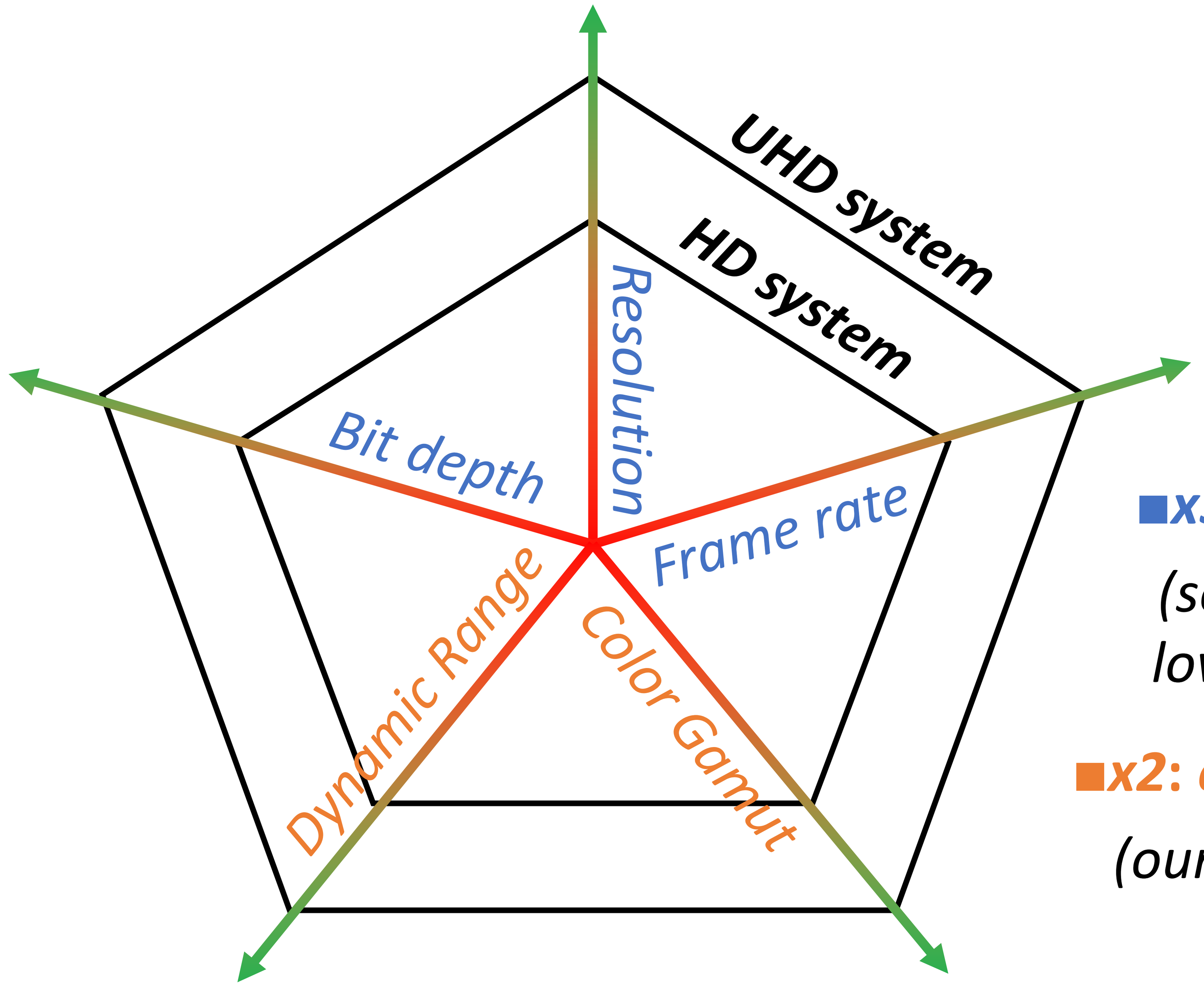
guocheng@cuc.edu.cn

THU-PM-155: *Guo et al. Learning a Practical SDR-to-HDRTV Up-conversion using New Dataset and Degradation Models*

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Our Scope/Task:

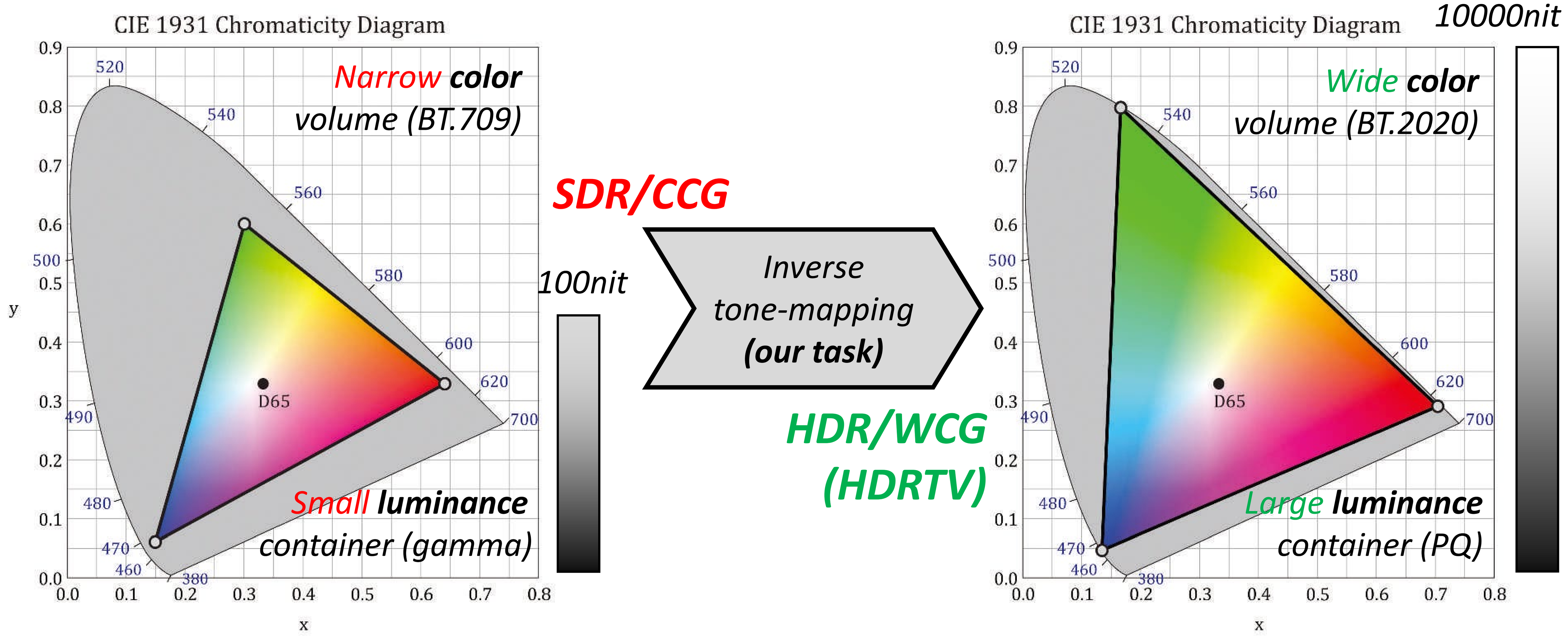
*Computer Vision **(CV)*** → *Image-to-image **(I2I)*** → *Low-level Vision* → *Image Enhancement* → *Inverse tone-mapping (a.k.a. up-conversion)*



■**x3:** *on "Resolution"*
(solved by other low-level vision tasks)

■**x2:** *on "Expressiveness"*
(our concern →

Converting** SDR/CCG (standard dynamic range/conventional color gamut) **to** HDR/WCG (high dynamic range/wide color gamut) **content



Our Concerns

Our Responses

1. *(new respective)* ***Training data*** *(of learning-based methods)* ⇨ *Proposing new **HDRTV4K dataset***

2. *(conventional respective)* ***Model design*** *(of neural network)* ⇨ *Designing new **Luminance Segmented Network***

3. *(from IQA perspective)* ***Assessment criteria*** *(of our task)* ⇨ *Using HDR/WCG **tailored metrics** & **subj. exp.***

THU-PM-155: *Guo et al. Learning a Practical SDR-to-HDRTV Up-conversion using New Dataset and Degradation Models*

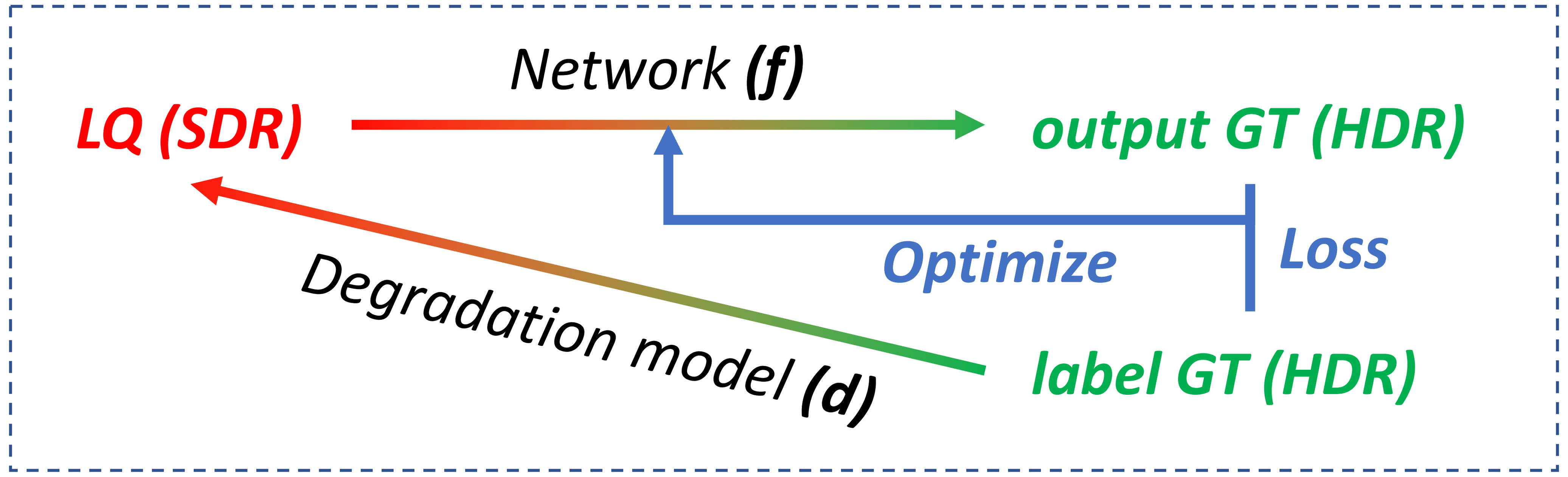
JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

1. New HDRTV4K dataset

Motivation

Workshop on vision dataset understanding (VDU, in both CVPR22 & 23):

"Data is the fuel of CV, yet its impact has long been underestimated."



Training (paired, supervised) paradigm of restoration-like I2I task

There're 2 ingredients of training:

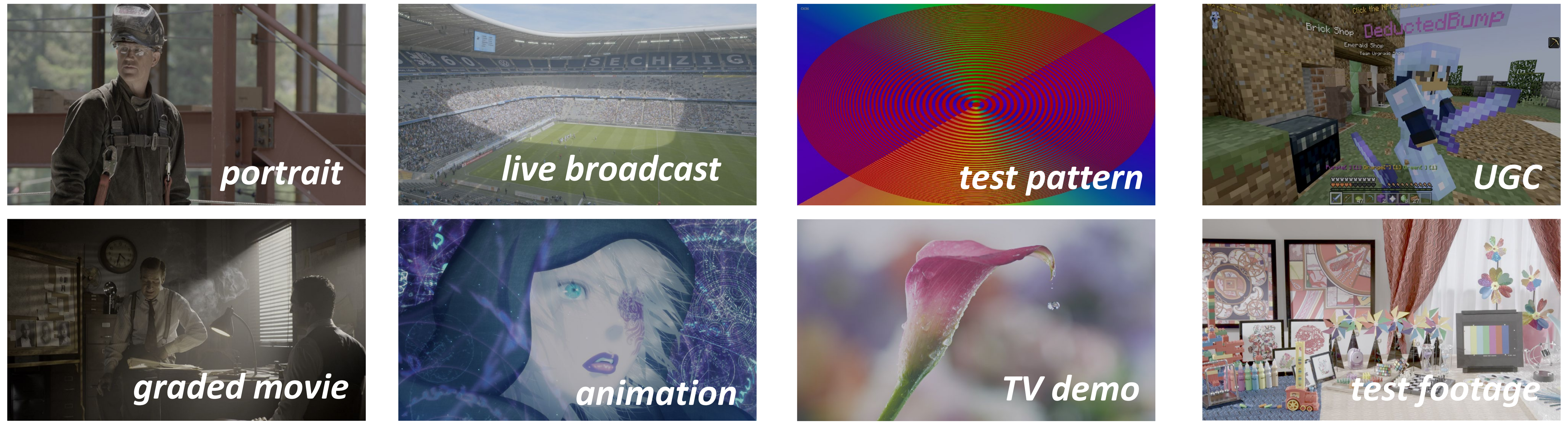
*1. **How** the target **GT** should be, and 2. **What degradation** network can learn*

Specific Concerns

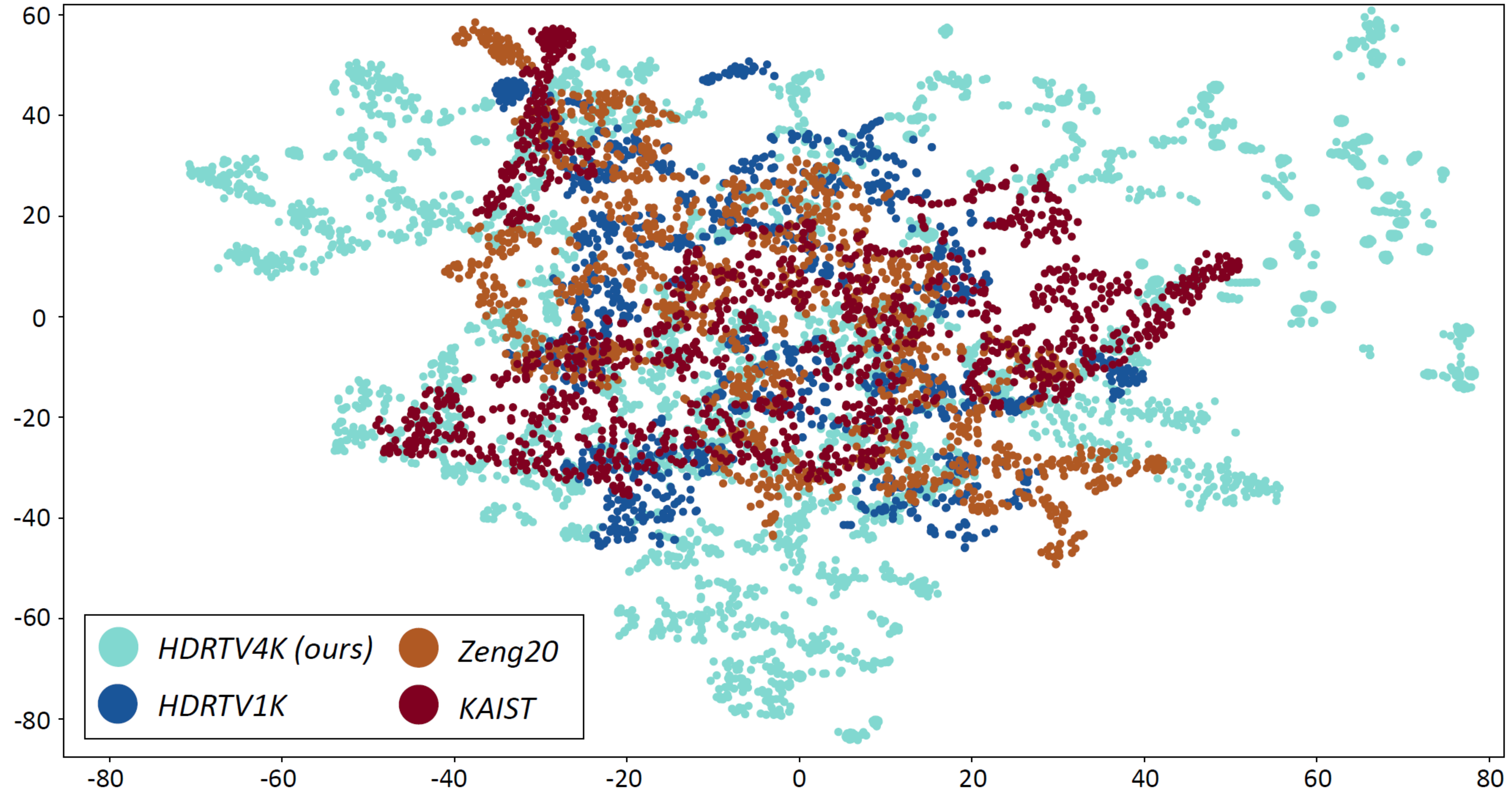
Training set			Model's benefit
***1.1.** HDR/WCG frames (as label)*	*diversity*	paired training	*generalization ability*
	quality (HDR/WCG volume)		*HDR/WCG expansion*
↓ Degradation model (DM) ↓			
***1.2.** SDR/CCG Frames (as input)*	*extent of degradation*		*recovery capability*
	style or aesthetic		*aesthetic performance*

1.1. HDRTV4K label HDR with more diversity

*We collect **3848** individual HDR/WCG **frames** from more categories*



Our representation in t-SNE latent space is more disperse than others



THU-PM-155: *Guo et al. Learning a Practical SDR-to-HDRTV Up-conversion using New Dataset and Degradation Models*

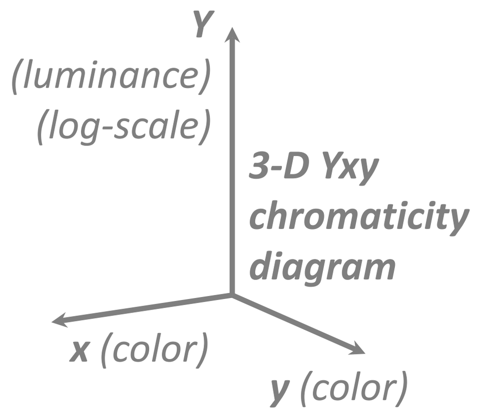
JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

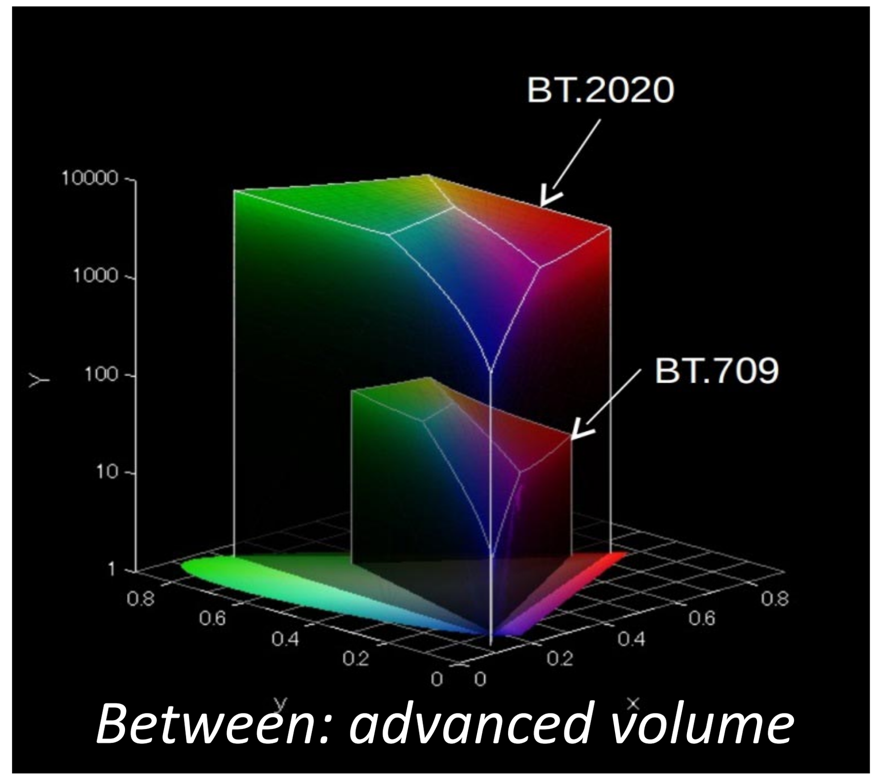
1. New HDRTV4K dataset

1.1. HDRTV4K label HDR with better quality

Usually quality is manifested in compression artifact, noise etc.
Yet, due to HDR's discrepancy in luminance and color container,
*its **quality** is also **advanced HDR/WCG volume**.*

Measured from 2 aspects:
1. Spatial fraction (F__P)
2. Numerical extent (E__)
*__ is for **HL(HighLight)** or **WG(Wide-Gamut)***





Metrics on ↓, from aspect →	***Spatial** fraction*	***Numerical** energy*
*Advanced **luminance** volume (>100nit)*	*FHLP*	*EHL*
*Advanced **color** volume (outside BT.709)*	*FWGP*	*EWG*

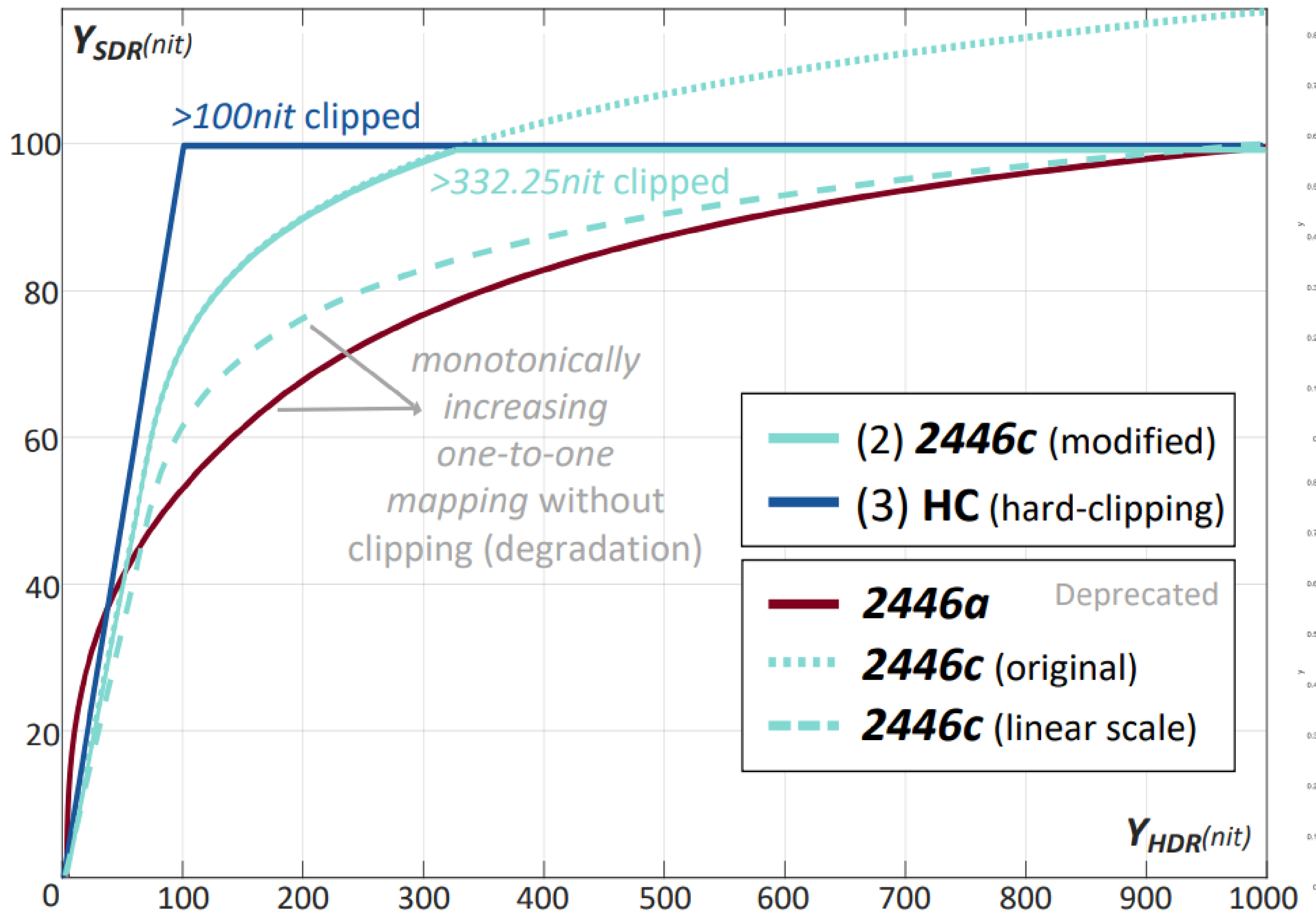
We use these metrics to compare label HDR in different dataset

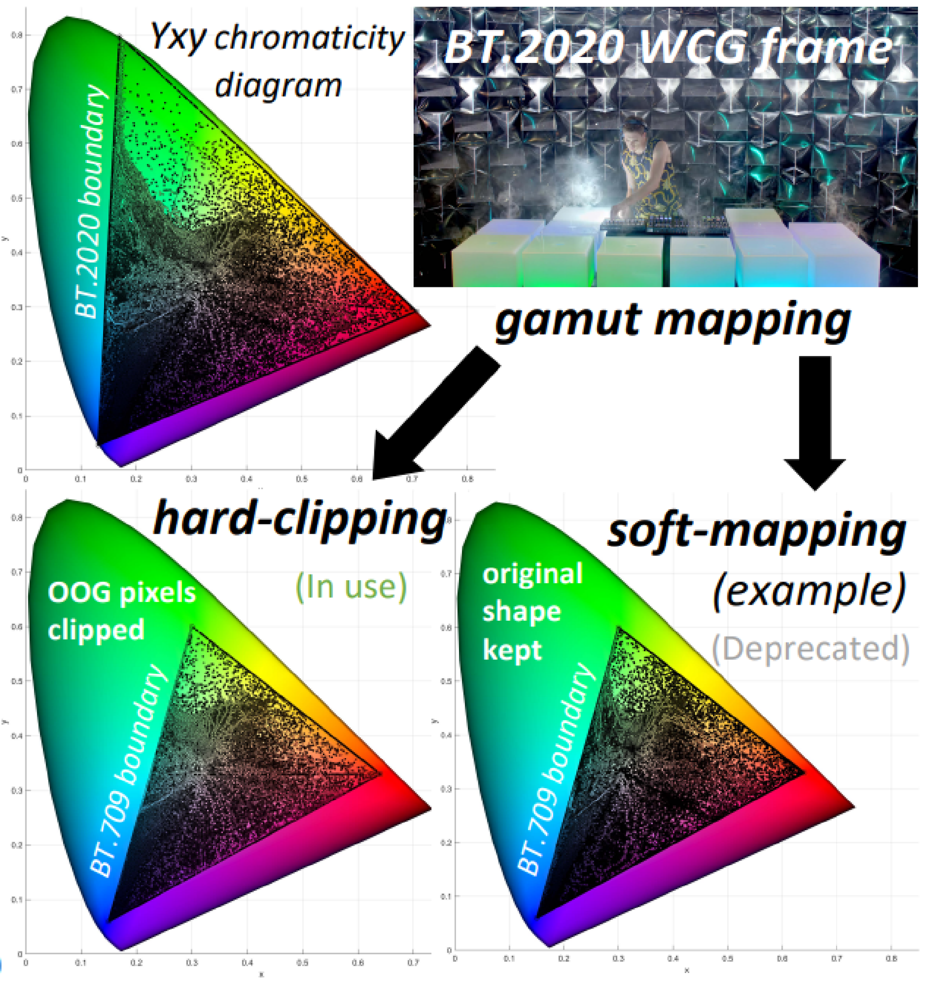
Metrics→		***Extent of HDR (luminance)***		***Extent of WCG***	
Dataset↓		*FHLP*	*EHL*	*FWGP*	*EWG*
Current open datasets	*KAIST*	*1.5250*	*0.2025*	*5.4771*	*0.1104*
	Zeng20	*0.0197*	*0.0012*	*0.4792*	*0.0034*
	HDRTV1K	*1.2843*	*0.1971*	*2.9089*	*0.1633*
*HDRTV4K **(ours)***		*5.3083*	*0.9595*	*2.6369*	*0.5123*

***Our label HDR** contains **more advanced luminance and color volume**,*
hence network will have more chance to produce them.

1.2. New degradation models (DMs) w. explicitly defined degradation

These're DMs we use: **1. OCIO2** **2. 2446c+GM** **3. HC+GM**





Hard-clipping** on both luminance and color to produce **more truncation

1.2. New degradation models (DMs) w. better style consistency

HDR
5.308
21.200
9.827

degrade

SDR from DM →	ours 2.	ours 3.	ours 1.	current DM
Over-exposed pixels (%)	*1.739*	*4.252*	*1.580*	*5.439*
Luminance level	*11.669*	*14.602*	*18.887*	*28.219 (bad)*
Saturation level	*10.183*	*10.377*	*9.977*	*14.641 (bad)*

DMs' will not excessively enhance style/aesthetic during HDR-to-SDR, so
network will not learn a vise-versa SDR-to-HDR style deterioration

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

2. Luminance Segmented Network (LSN)

Problem formulation

Different degradation occurs in different luminance range of input SDR

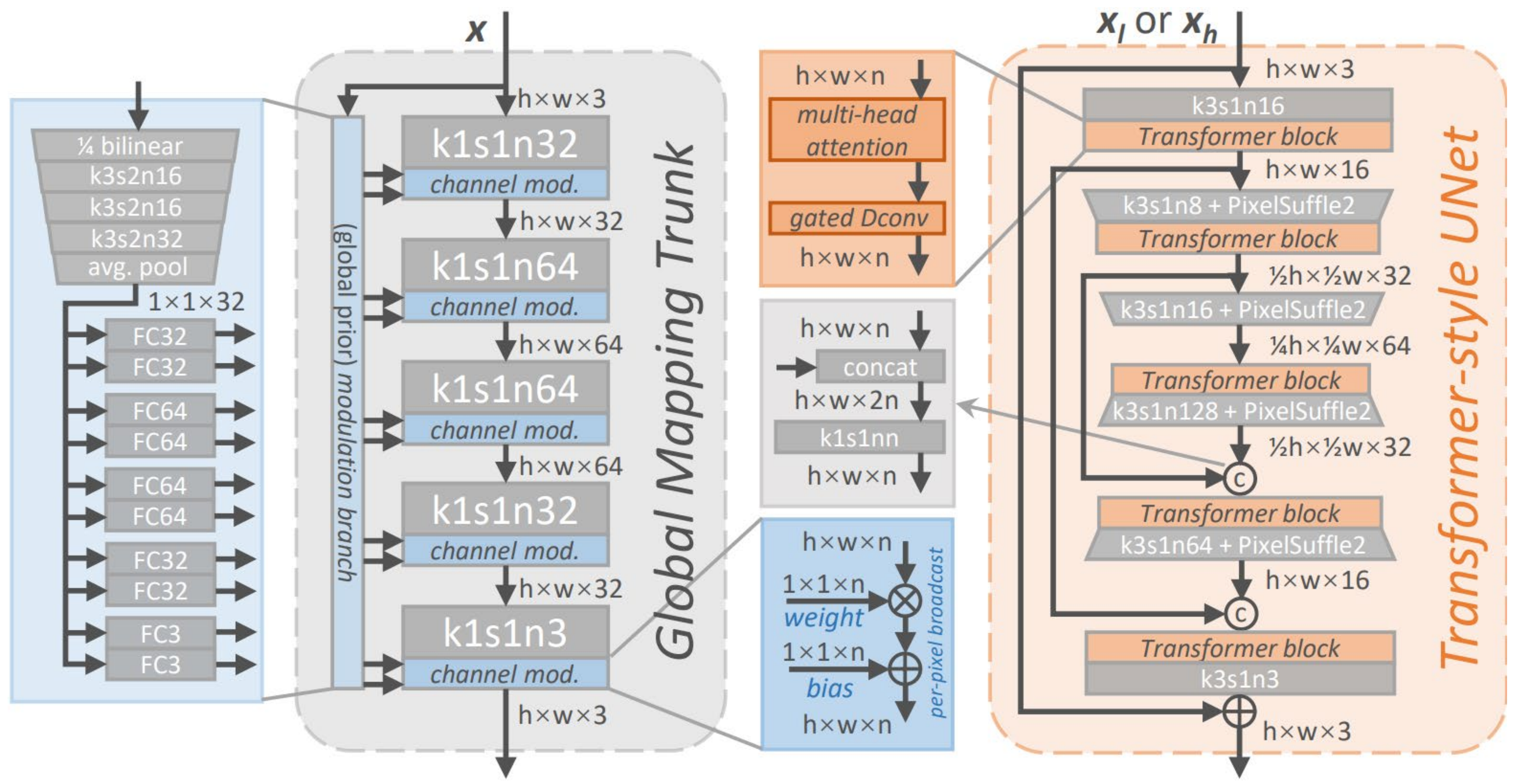
Input SDR (x) | *Bright area* | *Mid-tone* | *Dark area*

Degradation type: *Over-exposure* *Banding/quantization etc.*

Extent of degradation: **Greater** **Smaller** **Greater**

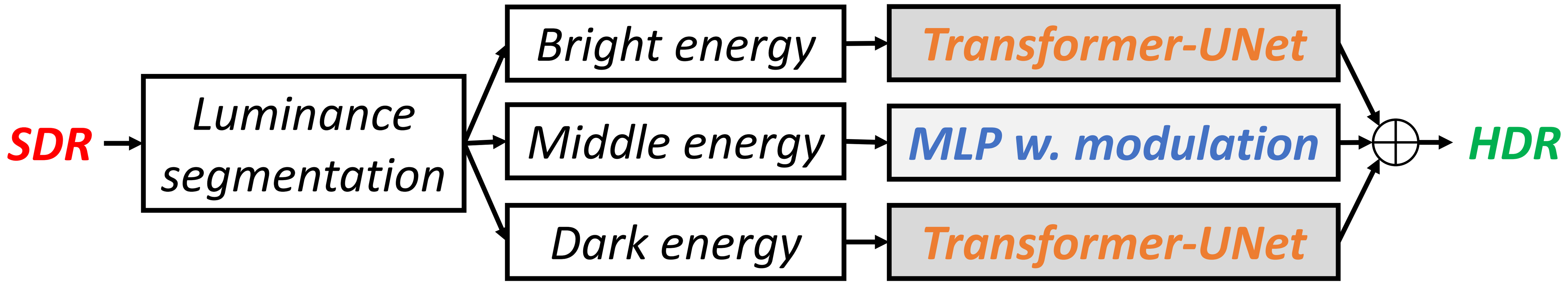
Requirement on network's

Recover capability: **more** **less** **more**



Hence, we use ***MLP with modulation*** *(less capability) on mid-tone, and* ***Transformer-UNet*** *(stronger capability) on bright & dark area*

Network Design



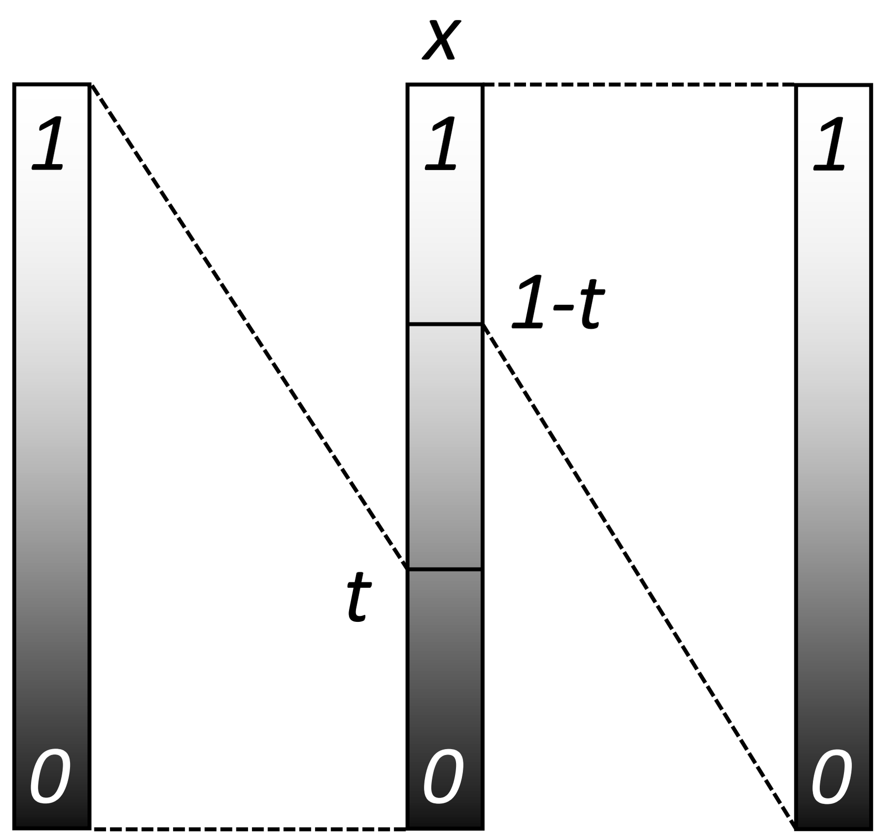
Luminance segmentation

1. *Treat the whole SDR as mid-tone for MLP (multi-layer perceptron) to process, and let Transformer-UNet to predict the residual (the part beyond MLP's capability):*

   Middle energy = x

2. *Map SDR's bright & dark range to more significant value ([0, 1]) which is easier for Transformer-UNet to process:*

Dark energy = max(0, (t-x)/t)



Bright energy = max(0, (x-1)/t + 1)

THU-PM-155: *Guo et al. Learning a Practical SDR-to-HDRTV Up-conversion using New Dataset and Degradation Models*

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

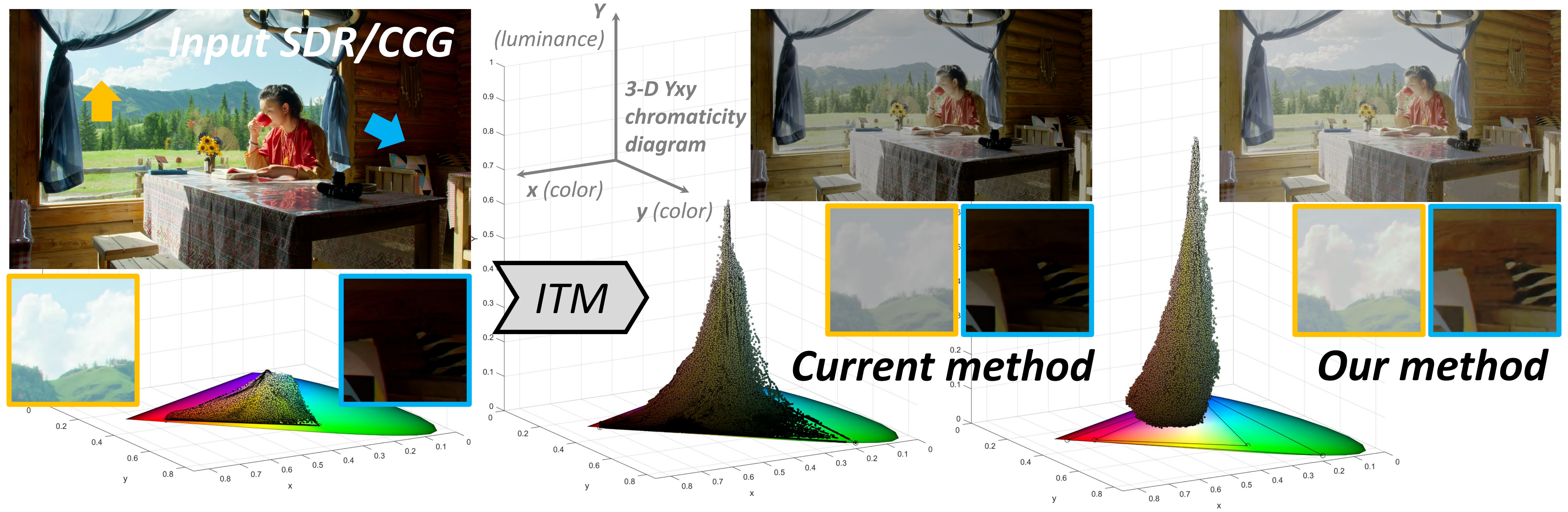
3. Experiments and its criteria

Assessment criteria

Common I2I restoration low-level tasks (e.g. super-resolution, denoising, frame-interpolation) only need to assess if output is close to GT/clean version, yet our task is different due to aesthetic factors and HDR container discrepancy etc.

Type	Criteria	Detailed fig.	Tailored metrics	Subj. exp.
Aesthetic	Brightness appearance	×	ALL (avg. luminance)	√
	Saturation appearance	×	ASL (avg. saturation)	√
Enhancement-related	Bright & dark area Recover/enhancement	√	×	√
DNN-related	Less artifact	√	×	√
Unique to HDR/WCG	Expansion of advance color & luminance	√ (Yxy diagram)	FHLP, EHL, EWGP, EWG	×

Visual results



Subjective experiment

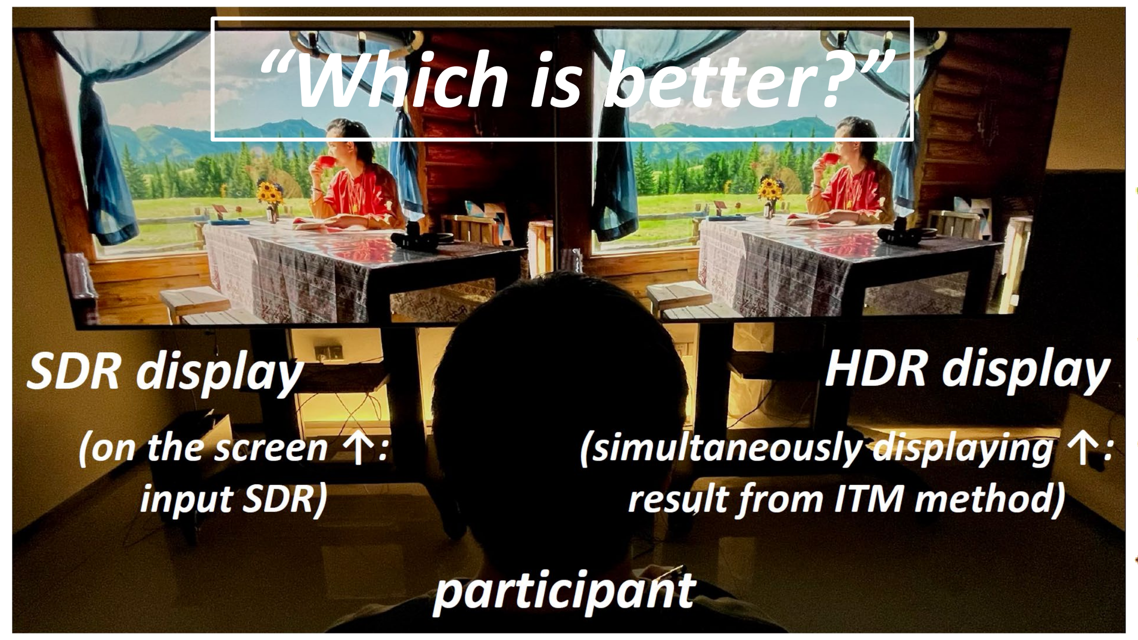
***SDR** input & **HDR** output by different methods are **displayed** "side-by-side"*

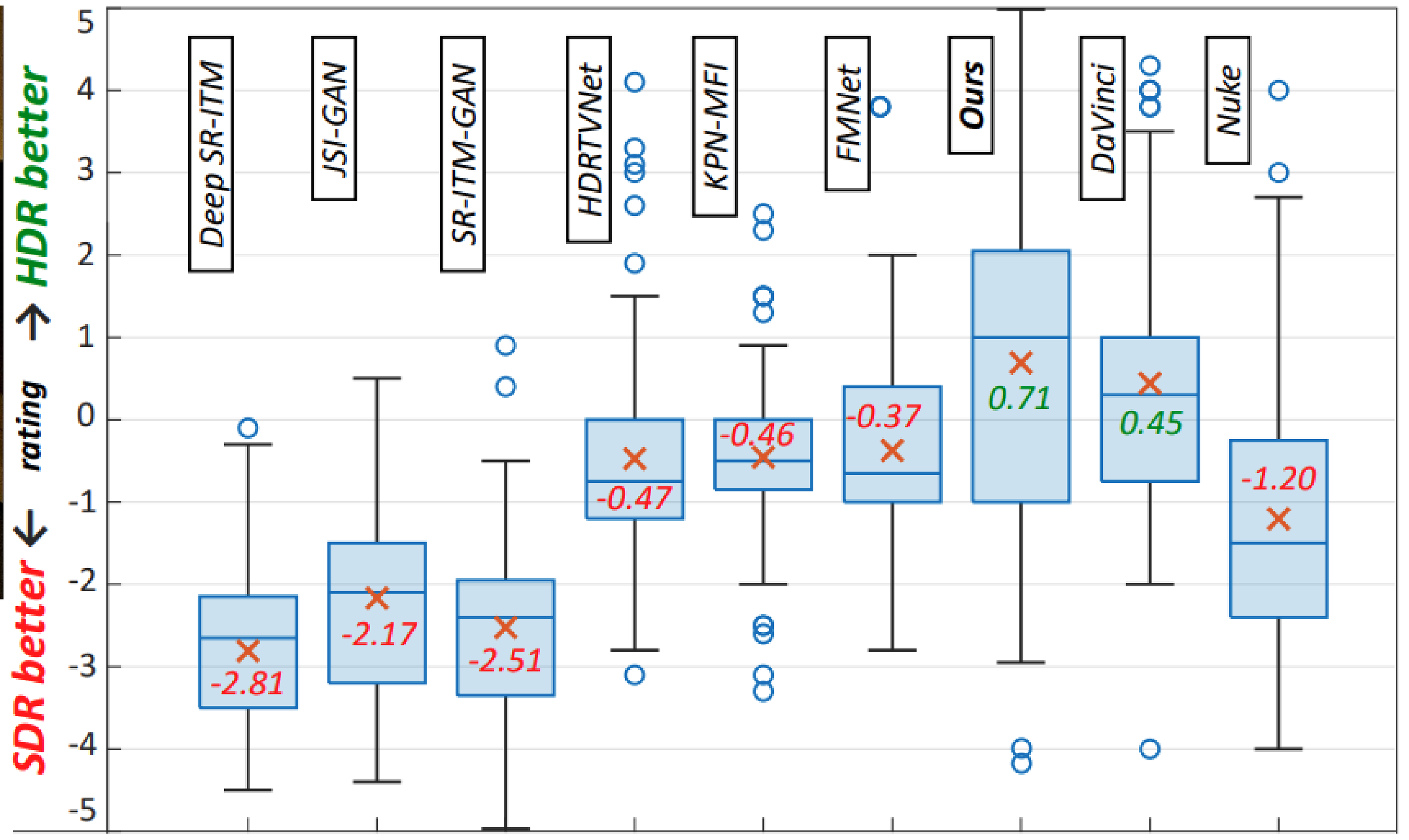
*participants are asked to **score** from -5 to 5 & select at least 1 **attribution***

SDR better — same — HDR better

-5 — 0 — 5

brightness | color | info. recover | artifact





Our result is the few which is Recognized better than SDR

HDR/WCG tailored metrics

Metrics		D. SR-ITM	JSI-GAN	SR-ITM-GAN	HDRTVNet	FMNet	Ours
Expansion of advanced color & luminance	FHLP	0.232	0.204	0.133	0.308	0.226	4.251
	EHL	0.372	0.136	0.550	0.625	0.474	2.599
	FWGP	1.094	1.333	0.000	2.433	2.459	0.687
	EWG	0.172	0.212	0.000	0.069	0.220	0.361
Style or aesthetic	ALL	9.580	9.659	7.853	9.759	9.758	20.42
	ASL	5.485	5.741	6.400	5.817	5.770	7.522

Thanks!

Please refer to our paper for more demonstration.