Learning a Deep Color Difference Metric for Photographic Images

Haoyu Chen[1] Zhihua Wang[1,2] Yang Yang[3] Qilin Sun[4]
Kede Ma[1]

[1]City University of Hong Kong

[2]Shenzhen MSU-BIT University

[3]Shenzhen Transsion Holdings Co., Ltd.

[4]The Chinese University of Hong Kong (Shenzhen)

Quick Preview

We proposed CD-Flow, a normalizing flow-based CD metric for photographic images.

- Step 1: Utilize a multi-scale autoregressive normalizing flow to learn a coordinate transform,
- Step 2: Computing the Euclidean distance in the transformed feature space.

- Properties of the learned feature transform:
  - Consistent with the working mechanism of human color perception.
  - Proper as a mathematical metric.
  - Accurate to explain human data of perceptual CDs.
  - Robust to slight geometric distortions.

Introduction

Modular and segregated view of cortical color processing:

- Visual perception of colorrelated quantities is separate from the perception of form, motion direction, and depth order in natural scenes.
- Investigate color perception under minimal conditions on form (e.g., uniformly colored patches).

Representative methods:

- JPC79 [1], CMC(l:c) [2], BFD(l:c) [3], CIELAB [4], CIE94 [5], and CIEDE2000 [6].

Naïve application of these metrics to photographic images:

- Compute the mean of the CDs between co-located pixels.
- Empirically shown to correlate poorly to human perception of CDs [7].

Problem Definition

- Denote RGB image space as $\mathcal{X}$ with an unknown distribution $p_{\mathcal{X}}$ and the transformed representation space as $\mathcal{Z}$ with a latent distribution $p_{\mathcal{Z}}$.
- Given a training dataset $\mathcal{D} = \{(x^{(i)}, y^{(i)}), \Delta V^{(i)}\}_{i=1}^{M}$:
  - $x^{(i)}, y^{(i)} \in \mathcal{X}$ form the i-th image pair of the same visual content but different color appearances.
  - $\Delta V^{(i)}$ represents the corresponding human perceptual CD collected from a subjective experiment.
  - M is the number of training pairs.
- Our goal:
  - learn a flow-based invertible transform f.
  - f maps RGB images to latent representations with Gaussian conditionals for CD assessment.

Feature Transform

- K scales of flow processing: $f = f_1 \circ f_2 \circ \cdots \circ f_K$ for multi-scale color and form interaction and abstraction.
- $z_{2(k-1)}$ is processed and split into z_{2k-1} and z_{2k}.
- z_{2k} further undergoes the $(k+1)$-th scale of processing and splitting.
- At the final K-th scale, we only process $z_{2(K-1)}$ to z_{2K-1} without splitting.

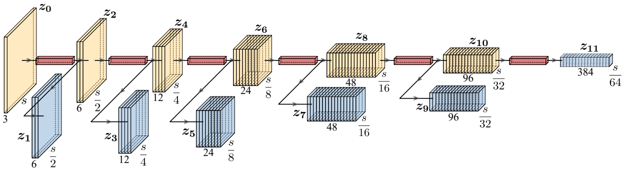

Figure: Feature transform of the proposed CD-Flow.

Feature Transform

- The probability density of the latent representation $z = \{z_1, z_3, \ldots, z_{2K-1}\}$ can then be conditionally factorized as

$$p(z) = \prod_{k=1}^{K-1} p\left(z_{2k-1} \middle| \left\{z_{\geq(2k+1)}\right\}\right) p(z_{2K-1}) \tag{1}$$

- Due to the bijectivity of the normalizing flow:

$$p(z) = \prod_{k=1}^{K-1} p(z_{2k-1}|z_{2k})p(z_{2K-1}). \tag{2}$$

- $p(z_{2k-1}|z_{2k})$, for $k \in \{1, 2, \cdots, K-1\}$ is modeled as conditionally independent Gaussians.
- $p(z_{2K-1})$ is modeled as (unconditionally) independent Gaussians.

CD Distance & Loss Function

CD distance:

- CD distance between two input images x and y is defined as Euclidean distance between two latent color representations $f(x)$ and $f(y)$:

$$\Delta E(x,y) = \sqrt{\frac{(f(x) - f(y))^T (f(x) - f(y))}{D}}. \quad (3)$$

Loss function:

- Measure the ℓ_p-norm induced distance between the predicted CD computed by Eq. (3) and the perceptual CD of the given image pair (x,y):

$$\ell(x,y) = \|\Delta E(x,y) - \Delta V(x,y)\|_p. \quad (4)$$

Loss Function

- Introduce a multi-scale version of Eq. (4) to put more emphasis on coarser-scale latent representations:

$$\ell_{\mathrm{ms}}(x,y) = \sum_{k=1}^{K} \|\Delta E_k(x,y) - \Delta V(x,y)\|_p, \tag{5}$$

where

$$\Delta E_k(x,y) = \sqrt{\frac{(f_{k:}(x) - f_{k:}(y))^T (f_{k:}(x) - f_{k:}(y))}{D_k}}. \tag{6}$$

Loss Function

- Incorporate the commonly used maximum likelihood objective in normalizing flow [8]:

$$\begin{aligned}\ell_{\mathrm{nl}}(x) &= -\log p_{\mathcal{X}}(x) \\ &= -\log p_{\mathcal{Z}}(f(x)) - \log \left| \det\left(\frac{\partial f(x)}{\partial x}\right)\right| .\end{aligned} \tag{7}$$

- During training, randomly sample a mini-batch $\mathcal{B}$ from the training dataset $\mathcal{D}$ in each iteration, and optimize the model parameters:

$$\ell(\mathcal{B}) = \frac{1}{|\mathcal{B}|}\sum_{(x,y)\in\mathcal{B}}\Big(\ell_{\mathrm{ms}}(x,y) + \lambda\big(\ell_{\mathrm{nl}}(x) + \ell_{\mathrm{nl}}(y)\big)\Big), \tag{8}$$

where λ is the trade-off to balance the magnitudes of different loss terms.

Main Results

- Compare the proposed CD-Flow with existing CD measures:

Method	Perfectly aligned pairs			Non-perfectly aligned pairs			All		
	STRESS↓	PLCC↑	SRCC↑	STRESS↓	PLCC↑	SRCC↑	STRESS↓	PLCC↑	SRCC↑
CIELAB	31.244	0.793	0.775	29.639	0.690	0.579	31.872	0.716	0.666
CIE94	34.721	0.790	0.772	29.916	0.693	0.572	34.326	0.710	0.654
CIEDE2000	29.975	0.825	0.821	30.347	0.667	0.563	31.439	0.726	0.686
S-CIELAB	30.094	0.822	0.819	31.804	0.631	0.522	32.780	0.700	0.657
Hong06	60.557	0.794	0.810	57.070	0.543	0.461	61.227	0.645	0.632
Ouni08	29.977	0.826	0.821	30.355	0.668	0.563	31.444	0.726	0.685
CD-Net	**20.891**	**0.867**	**0.870**	**22.543**	**0.818**	**0.776**	**21.431**	**0.846**	**0.842**
CD-Flow	**16.613**	**0.896**	**0.904**	**21.374**	**0.856**	**0.794**	**18.473**	**0.871**	**0.865**

- Robustness of CD-Flow to mild geometric distortions (including translation, rotation, and dilation).

Method	Translation			Rotation			Dilation		
	STRESS↓	PLCC↑	SRCC↑	STRESS↓	PLCC↑	SRCC↑	STRESS↓	PLCC↑	SRCC↑
CIELAB[9]	29.414	0.620	0.577	32.633	0.529	0.495	31.511	0.519	0.467
CIE94[5]	29.141	0.645	0.596	31.943	0.566	0.519	30.323	0.567	0.505
CIEDE2000[6]	28.035	0.654	0.613	31.255	0.566	0.527	29.928	0.566	0.512
CD-Net[10]	**19.825**	**0.845**	**0.842**	**22.463**	**0.784**	**0.772**	**21.704**	**0.787**	**0.773**
CD-Flow	**19.311**	**0.852**	**0.856**	**20.139**	**0.837**	**0.816**	**21.352**	**0.827**	**0.797**

Main Results

- Generalizability of CD-Flow on COM dataset [6]:

Method	BFD-P [3]		Leeds [11]		Witt [12]		RIT-DuPont [13]		COM dataset [6]	
	STRESS↓	PLCC↑	STRESS↓	PLCC↑	STRESS↓	PLCC↑	STRESS↓	PLCC↑	STRESS↓	PLCC↑
CIELAB [9]	45.054	0.749	40.093	0.295	51.689	0.565	30.348	—	45.202	0.693
CIE94 [5]	35.798	0.830	**30.494**	**0.584**	**31.857**	0.793	**20.982**	—	**33.235**	**0.814**
CIEDE2000 [6]	**31.935**	**0.861**	**19.247**	**0.772**	**30.358**	**0.825**	**20.239**	—	**28.979**	**0.862**
CD-Net	39.312	0.791	38.558	0.449	33.640	**0.828**	42.999	—	38.872	0.786
CD-Flow	**34.661**	**0.833**	34.275	0.476	31.965	0.820	36.504	—	35.061	0.801

- Generalizability of CD-Flow on TID2013 subset [14]:

Method	STRESS↓	PLCC↑	SRCC↑
CIEDE2000 [6]	18.203	0.730	0.751
PieAPP [15]	20.918	0.620	0.653
LPIPS [16]	15.420	0.816	0.804
DISTS [17]	**15.235**	**0.821**	0.805
CD-Net [10]	15.962	0.801	**0.826**
CD-Flow	**14.110**	**0.837**	**0.832**

References

[1] Roderick McDonald. Industrial pass/fail colour matching. *Journal of the Society of Dyers and Colourists*, 96(7):372–376, 1980.

[2] Frank J. J. Clarke, Roderick McDonald, and Bryan Rigg. Modification to the JPC79 colour–difference formula. *Journal of the Society of Dyers and Colourists*, 100(4):128–132, 1984.

[3] Ming Ronnier Luo and Bryan Rigg. BFD(l:c) colour-difference formula. *Journal of the Society of Dyers and Colourists*, 103(2):86–94, 1987.

[4] Marc Mahy, Luc Van Eycken, and André Oosterlinck. Evaluation of uniform color spaces developed after the adoption of CIELAB and CIELUV. *Color Research & Application*, 19(2):105–121, 1994.

[5] Roderick McDonald and Kenneth J. Smith. CIE94-A new colour-difference formula. *Journal of the Society of Dyers and Colourists*, 111(12):376–379, 1995.

[6] Ming Ronnier Luo, Guihua Cui, and Bryan Rigg. The development of the CIE 2000 colour-difference formula: CIEDE2000. *Color Research & Application*, 26(5):340–350, 2001.

[7] Benhur Ortiz-Jaramillo, Asli Kumcu, Ljiljana Platisa, and Wilfried Philips. Evaluation of color differences in natural scene color images. *Signal Processing: Image Communication*, 71:128–137, 2019.

[8] George Papamakarios, Eric T. Nalisnick, Danilo Jimenez Rezende, Shakir Mohamed, and Balaji Lakshminarayanan. Normalizing flows for probabilistic modeling and inference. *Journal of Machine Learning Research*, 22(57):1–64, 2021.

[9] Alan R. Robertson. The CIE 1976 color-difference formulae. *Color Research & Application*, 2(1):7–11, 1977.

[10] Zhihua Wang, Keshuo Xu, Yang Yang, Jianlei Dong, Shuhang Gu, Lihao Xu, Yuming Fang, and Kede Ma. Measuring perceptual color differences of smartphone photography. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.

[11] Dong-Ho Kim. New weighting functions for the modified CIELAB colour-difference formulae. *Textile Coloration and Finishing*, 9(6):51–57, 1997.

[12] Klaus Witt. Geometric relations between scales of small colour differences. *Color Research & Application*, 24(2):78–92, 1999.

[13] Roy S. Berns, David H. Alman, Lisa Reniff, Gregory D. Snyder, and Mitchell R. Balonon-Rosen. Visual determination of suprathreshold color-difference tolerances using probit analysis. *Color Research & Application*, 16(5):297–316, 1991.

[14] Nikolay Ponomarenko, Lina Jin, Oleg Ieremeiev, Vladimir Lukin, Karen Egiazarian, Jaakko Astola, Benoit Vozel, Kacem Chehdi, Marco Carli, Federica Battisti, and Chung-Chieh Jay Kuo. Image database TID2013: Peculiarities, results and perspectives. *Signal Processing: Image Communication*, 30:57–77, 2015.

[15] Ekta Prashnani, Hong Cai, Yasamin Mostofi, and Pradeep Sen. PieAPP: Perceptual image-error assessment through pairwise preference. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1808–1817, 2018.

[16] Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 586–595, 2018.

[17] Keyan Ding, Kede Ma, Shiqi Wang, and Eero P. Simoncelli. Image quality assessment: Unifying structure and texture similarity. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(5):2567–2581, 2020.