Work Sharing
&
Paper Reading

Author: Hu Shiyu

2019-09-20

Topic 1

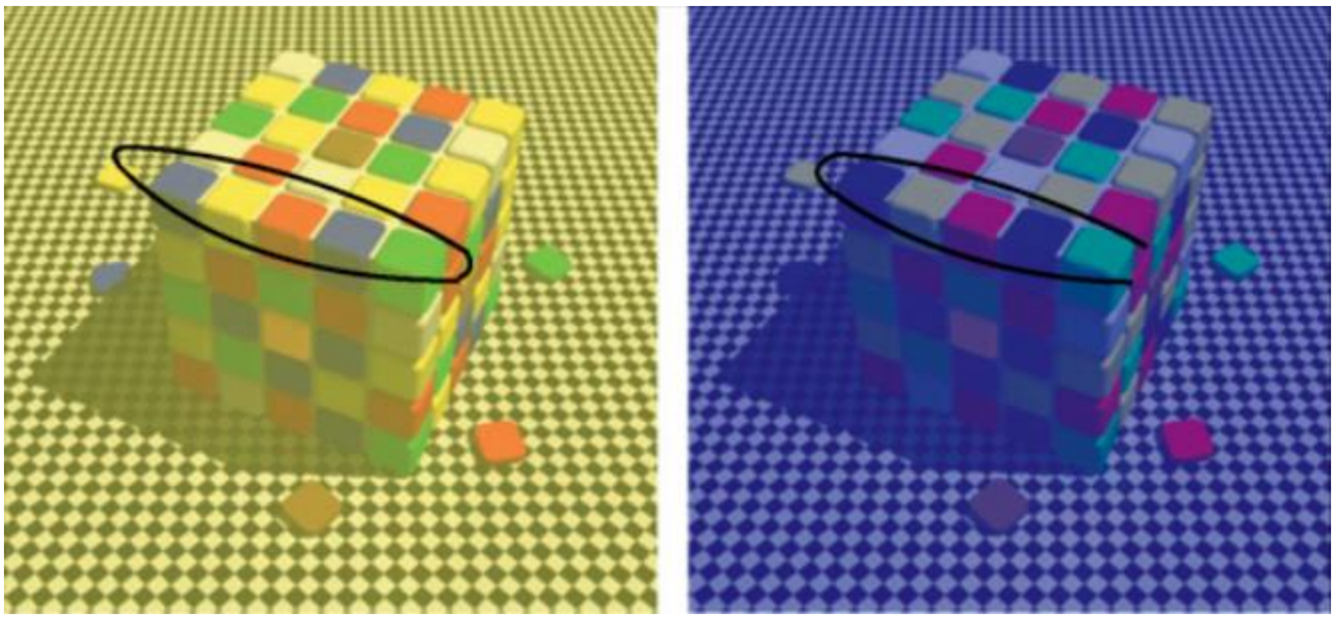
- 颜色恒常性研究

Part 1. 颜色恒常性研究

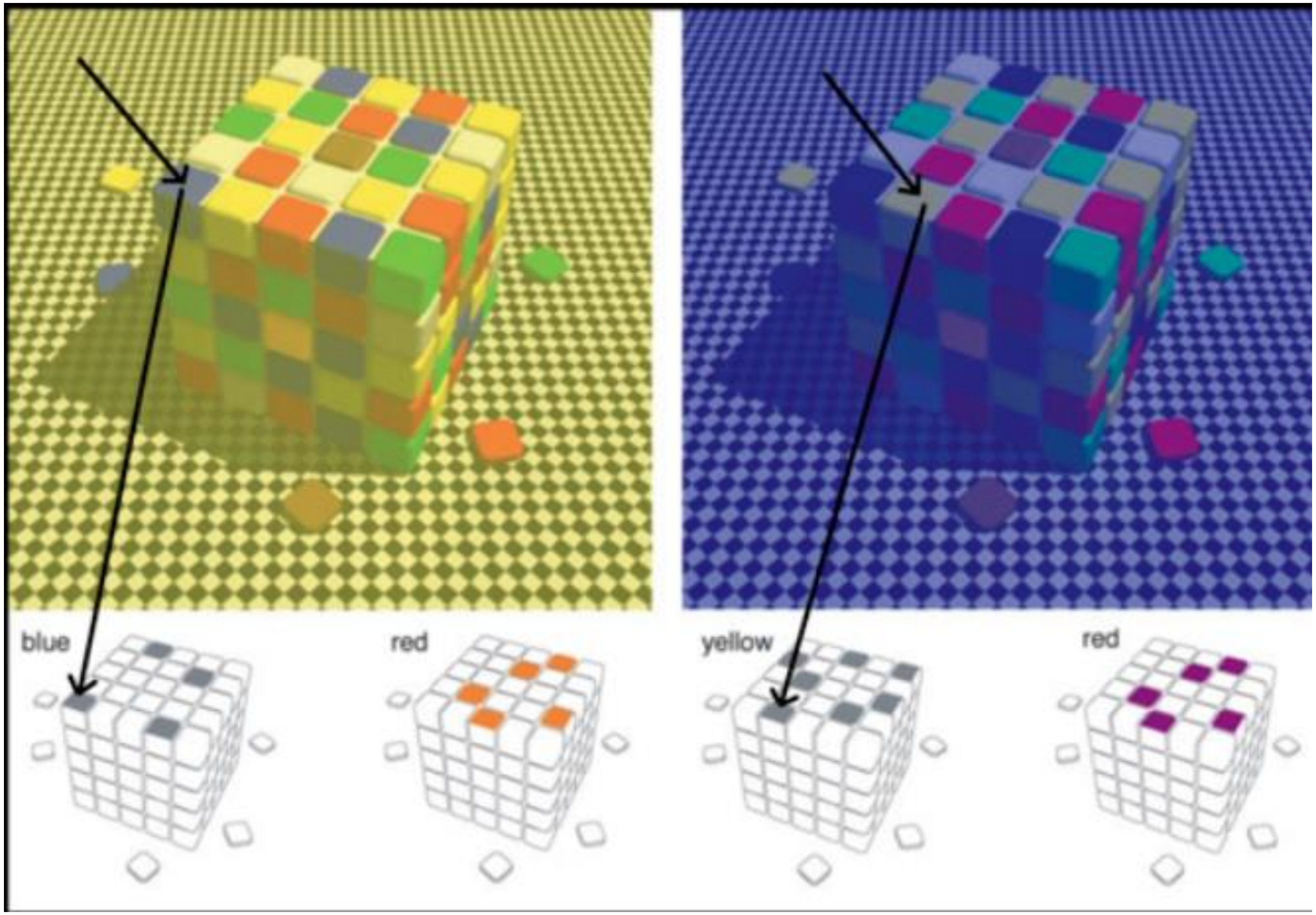
Section 1. 基础知识介绍

1.1 定义

方块的颜色：蓝黄红蓝绿



为什么人眼和计算机的结果完全不同?



颜色恒常性(Color Constancy)

颜色并不是一个真实存在的实例，而是**大脑和视网膜处理的结果**。
色彩的恒常性，即**在外界光源变化**的情况下，人类所**感知的颜色依然能够保留物体原始的色彩**。

白平衡(White Balance)

白平衡是**摄像器材**对不同光照条件下的图片进行**校正**，使校正后的图片颜色接近于**标准光源下拍摄的图片颜色**，从而实现**近似于人类颜色恒常性**的功能。

Fig. 1. Illustration of the influence of differently colored light sources on the measured image values. These images are adapted from [7] and show the same scene, rendered under four different light sources.

1.1 定义

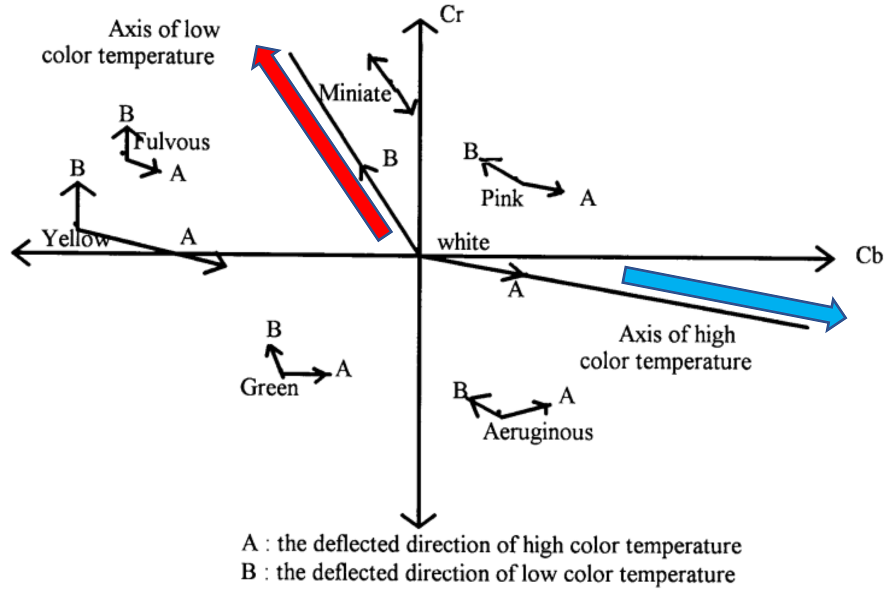


Fig.5 the experimental data of colors under different light sources.

当光源的色温较**高**时，光源呈现**蓝**色；
当光源的色温较**低**时，光源呈现**红**色。

物体在不同的光源下会反射不同的光，从而呈现出不同的颜色。

颜色恒常性在**消除光源差异、恢复物体真实颜色**中十分重要。

颜色恒常性的研究方法可以分成两大类：

•基于假设的方法Assumption-based Method
•基于学习的方法Learning-based Method

基于假设的方法Assumption-based Method

•**双反射模型Dichromatic Reflection Model**
- Specular highlights method, Constrained Dichromatic Reflection, Log-relative Chromaticity constraint

•**朗伯体模型Lambertian Model**
- Gray World Method, White Patch Method, Shade of Gray, Grey Edge Method

优点：模型轻量化并且易于理解
缺点：一旦实际图片不满足模型的假设条件，模型的性能将大大降低

基于学习的方法Learning-based Method

- **Low-level**
  - Color-by-Correlation, Gamut Mapping, Bayesian color constancy

缺点：在生成low-level的颜色特征时，丢弃了空间信息和纹理信息，因此可能会产生模糊估计

- **High-level**
  - CNN模型

优点：随着深度学习的发展，这一类方法在benckmark上取得SOTA的效果
缺点：大部分是针对例如图像分类等任务设计的，模型结构对于颜色恒常性问题过于复杂，直接使用的话效果一般

1.3 理论基础 朗伯体反射模型

物体表面反射光谱

$$\boxed{E(\lambda,\mathbf{x})}=\boxed{m^{b}(\mathbf{x})}\,\boxed{s(\lambda,\mathbf{x})}\,\boxed{e(\lambda)}$$

场景照明

反射系数的几何依赖项
(与光源照射方向和
物体表面方向有关)

光谱反射率

颜色的度量值
(反射光谱与响应函数
在可见光谱上的积分)

相机的光谱响应函数

$$\boxed{f^{c}(\mathbf{x})}=\int_{\omega}E(\lambda,\mathbf{x})\rho^{c}(\lambda)d\lambda=m^{b}(\mathbf{x})\int_{\omega}s(\lambda,\mathbf{x})e(\lambda)\boxed{\rho^{c}(\lambda)}d\lambda$$

$$\mathbf{c}^{b}(\mathbf{x})=\int_{\omega}s(\lambda,\mathbf{x})e(\lambda)\rho^{c}(\lambda)d\lambda$$

漫反射系数

$$\mathbf{f}(\mathbf{x})=\mathrm{m}^{b}(\mathbf{x})\boxed{\mathbf{c}^{b}(\mathbf{x})}$$

$$f^c(\mathbf{x}) = \int_\omega E(\lambda,\mathbf{x})\rho^c(\lambda)d\lambda = m^b(\mathbf{x})\int_\omega s(\lambda,\mathbf{x})e(\lambda)\boxed{\rho^c(\lambda)}d\lambda$$

$\Big\downarrow$ 假设相机敏感度函数为狄拉克函数

$$\boxed{\rho^c(\lambda) = \delta(\lambda-\lambda_c)}$$

$$f^c(\mathbf{x}) = m^b(\mathbf{x})s(\lambda_c,\mathbf{x})e(\lambda_c)$$

$$\text{光源下拍摄的图像} \longleftarrow \boxed{\frac{f_1^c(\mathbf{x})}{f_2^c(\mathbf{x})}} = \frac{m^b(\mathbf{x})s(\lambda_c,\mathbf{x})e^1(\lambda_c)}{m^b(\mathbf{x})s(\lambda_c,\mathbf{x})e^2(\lambda_c)} = \boxed{\frac{e^1(\lambda_c)}{e^2(\lambda_c)}} \longrightarrow \text{光源}$$

$$f_1 = D_{1,2}f_2$$

两个光源下的图像可以通过对角变换得到

1.3 理论基础 Von Kries模型

标准光源下拍摄的图像

未知光源下拍摄的图像

$$\begin{pmatrix} R_c \\ G_c \\ B_c \end{pmatrix} = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} R_u \\ G_u \\ B_u \end{pmatrix} \longrightarrow \boxed{\mathbf{f}_c} = \boxed{\mathbf{D}_{u,c}}\boxed{\mathbf{f}_{\mathrm{u}}}$$

对角矩阵
(将未知光源u 下的图像变换
到经典光源c (通常为 D65
光源) 下的图像)

可以通过对角变换调整输入图像的3个颜色通道增益，将图像
颜色校正到标准白光下的颜色

颜色恒常问题可以转化为未知光源估计问题

自然界物体颜色:

- 消色物体（Achromatic object）
- 有色物体（Chromatic object）

消色(Achromatic)

- 从白到黑的一系列中性灰色，只有亮度差别，没有色调和饱和度特征
- 当消色物体对可见光谱的光波的反射率在80%以上时则呈现白色；在4%以下时，则呈现黑色；在这两者之间的，是具有各种亮度的灰色

消色物体

- 黑、白、灰色物体
- 对照明光线具有**非选择性吸收的特性**，即光线照射到消色物体上时，被吸收的入射光中的各种波长的色光是等量的；被**反射或透射的光线其光谱成分也与入射光的光谱成分相同**

有色物体

- 呈现其他颜色的物体
- 对光源进行**有选择的吸收和反射**，从而呈现出不同的颜色

消色物体对于光源无差别吸收的特性可以用于估计环境光照
此时光源估计的问题转化为确定消色物体的问题

角度误差(Angular Error)

标定光源 $\mathbf{e}_e = \left[e_e^R, e_e^G, e_e^B\right]$

估计光源 $\mathbf{e}_u = \left[e_u^R, e_u^G, e_u^B\right]$

角度误差
(越小表示对图像的
光源估计越准确)

$$\boxed{d_{ang}(\mathbf{e}_e, \mathbf{e}_u)} = \cos^{-1}\left(\frac{\boxed{\mathbf{e}_e} \cdot \boxed{\mathbf{e}_u}}{\|\mathbf{e}_e\| \cdot \|\mathbf{e}_u\|}\right)$$

- 角度误差值**越小**，则表示算法对于图像光源估计**越准确**
- 常用的统计指标有角度误差的**均值，中位数和worst-25%**
  - 角度误差分布是偏向一侧的**非对称分布**，有时使用**角度误差中值比使用角度误差平均值更能描述光源估计算法的性能**
  - worst-25%则是所有角度误差中前 25%的最大值的平均值,用于描述光源估计算法对实验效果**最差的情况的鲁棒性**

1.5 数据集 Color Checker Dataset

Color Checker

- Color Checker一般是标准的24色色板，不同的颜色会在环境光照下反射不同光
- 根据**最后一排黑白灰的色块信息**，可以对环境光照进行计算，得到相应的ground truth

1.5 数据集 Color Checker Dataset

- Color Checker Dataset是利用Color Checker上的信息来进行光源估计的数据集。
- 数据集包含**568张图片**（室内246张，室外322张），采用**Canon 5D和Canon 1D**两款相机拍摄。

1.5 数据集　Color Checker Dataset

Color Checker Dataset是**使用时间最长和使用范围最广**的光源估计数据集之一，目前有3个版本：

•Gehler version (2008)

- 所有图像是Canon RAW格式
- 作者提供利用Canon Digital Photo Professional program的 automatic mode处理后的tiff格式
- **数据集经过一定的处理，包含一些非线性信息，如gamma校正**

•Shi's re-processing version (2011, widely used)

- 对RAW data进行再一次处理，生成12-bit的PNG格式
- 在RGB空间是线性的图片(gamma=1)
- **但是在使用过程中产生3套不同的ground truth，并不适合对算法进行评测**（很多只在Shi's re-processing version上评测的算法可能会有问题，因为很多作者并没有指出自己的模型是在哪一套ground truth上进行的测试）

•New RECommended version (2018)

- 作者在发现ground truth不统一后，重新修正了数据集的ground truth
- **如果要用Color Checker Dataset进行光源估计，这是最合适的版本**

1.5 数据集　NUS-8 Dataset

original camera dependent RAW images

original camera default output JPG images

using DCRAW extract the original linear image

color checker information including mask positions and colors of all patches

1736 images from 8 different cameras

CAMERA	RAW	JPEG	PNG	MASK	GROUNDTRUTH
Canon EOS-1Ds Mark III	ZIP1 ZIP2 ZIP3 ZIP4 ZIP5	ZIP1 ZIP2	ZIP1 ZIP2 ZIP3 ZIP4	ZIP	MAT
Canon EOS 600D	ZIP1 ZIP2 ZIP3 ZIP4 ZIP5	ZIP1 ZIP2	ZIP1 ZIP2 ZIP3 ZIP4	ZIP	MAT
Fujifilm X-M1	ZIP1 ZIP2 ZIP3 ZIP4 ZIP5	ZIP1 ZIP2	ZIP1 ZIP2 ZIP3 ZIP4	ZIP	MAT
Nikon D5200	ZIP1 ZIP2 ZIP3 ZIP4 ZIP5	ZIP1 ZIP2	ZIP1 ZIP2 ZIP3 ZIP4	ZIP	MAT
Olympus E-PL6	ZIP1 ZIP2 ZIP3 ZIP4 ZIP5	ZIP1 ZIP2	ZIP1 ZIP2 ZIP3 ZIP4	ZIP	MAT
Panasonic Lumix DMC-GX1	ZIP1 ZIP2 ZIP3 ZIP4 ZIP5	ZIP1 ZIP2	ZIP1 ZIP2 ZIP3 ZIP4	ZIP	MAT
Samsung NX2000	ZIP1 ZIP2 ZIP3 ZIP4 ZIP5	ZIP1 ZIP2	ZIP1 ZIP2 ZIP3 ZIP4	ZIP	MAT
Sony SLT-A57	ZIP1 ZIP2 ZIP3 ZIP4 ZIP5	ZIP1 ZIP2	ZIP1 ZIP2 ZIP3 ZIP4	ZIP	MAT

groudtruth illuminants extracted form color-checkers

- NUS-8数据集也是**使用Color Checker作为参照物**进行光源估计
- 和Color Checker Dataset相比，NUS-8包含的**图片数目和相机型号更丰富**（1736张图，来源于8种不同型号的相机），且提供了**不同格式**的图片

Cube Dataset

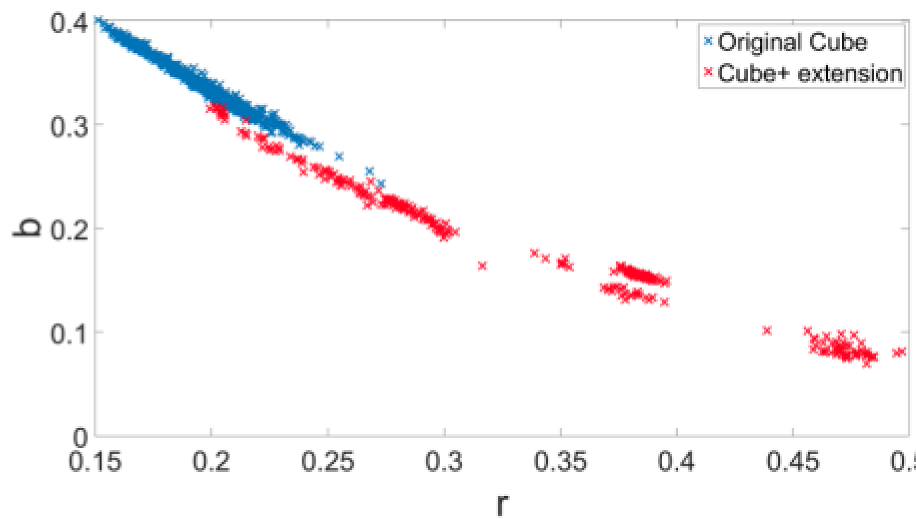
Spyder Cube

- Spyder Cube是一个包含**金属球、白色面、黑色面以及吸光孔**的元件
- 和Color Checker类似，Spyder Cube通过**不同的消色表面对环境光的反射信息**，对环境光照进行计算，得到相应的ground truth

1.5 数据集 Cube Dataset

- Cube数据集就是利用**Canon 550D**拍摄的**1365张室外**照片
- 在每张照片的右下角均放置spyder cube用于后期ground truth的计算

1.5 数据集 Cube+ Dataset



- Cube+数据集是在Cube数据集的基础
  照片
- 在颜色分布上，Cube+数据集也拥有

Part 1. 颜色恒常性研究

Section 2. 研究方法&Paper Reading

Color Constancy

Research Website on Illuminant Estimation

Search

Color Constancy

Home | Color Constancy | Evaluation | Source-code | Contact

FEATURED

About this webpage

This website presents work about Color Constancy. It contains an explanation of the topic, as well as an overview of various approaches to obtain it. Further, results of several methods applied to some of the publicly available data sets are provided, as well as reproduction errors (as proposed by Finlayson and Zakizadeh at BMVC 2014). Finally, links to various publicly available data sets and source-codes are provided. If you have any questions, comments, remarks or additions to this website, please write an e-mail to the contact person of this website. Feel free to refer to this website in your publications if you use any of the information that you recovered from this place (like the pre-computed results of several algorithms on various data sets). Drop me a line if you would like to have your publication mentioned on this page.

Posted in **Uncategorized**

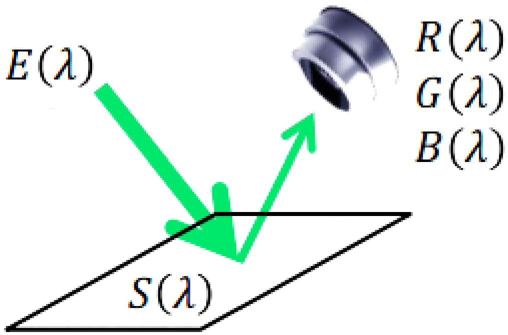
SOURCE-CODE

- **(2002) Barnard's Software**
- **(2004) C.C. Through I.I.C.**
- **(2007) Grey-Edge**
- **(2007) Top-down**
- **(2007, 2011) C.C. Using N.I.S.**
- **(2008) Bayesian Framework**
- **(2008) Spatial Correlations**
- **(2010) Generalized Gamut**
- **(2011) Thin-plate Spline**
- **(2012) Weighted Grey-Edge**
- **(2013) Edge-based Spatio-spectral**

http://colorconstancy.com/index.html

Shades of Gray and Color Constancy (CIC 2004)

$E(\lambda)$: Spectral distribution
$S(\lambda)$: *Lambertian* surface
$C(\lambda)$: Color signal



$$C(\lambda) = E(\lambda)S(\lambda)$$

$$R = \int_{\omega} E(\lambda)S(\lambda)R(\lambda)d\lambda$$

$$G = \int_{\omega} E(\lambda)S(\lambda)G(\lambda)d\lambda$$

$$B = \int_{\omega} E(\lambda)S(\lambda)B(\lambda)d\lambda$$

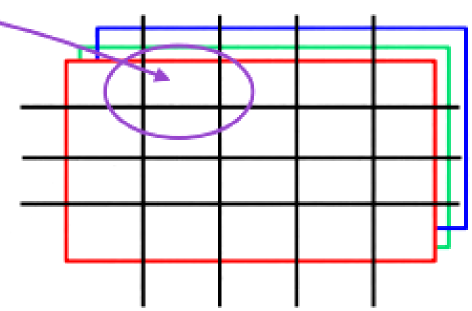
Sensor response curve
Or
Sensitivity function

An image represented by three *N*-dimensional vector

- Given image *I*

$$\underline{R} = [R_1, R_2, \cdots, R_N]^T$$
$$\underline{G} = [G_1, G_2, \cdots, G_N]^T$$
$$\underline{B} = [B_1, B_2, \cdots, B_N]^T$$

I:image

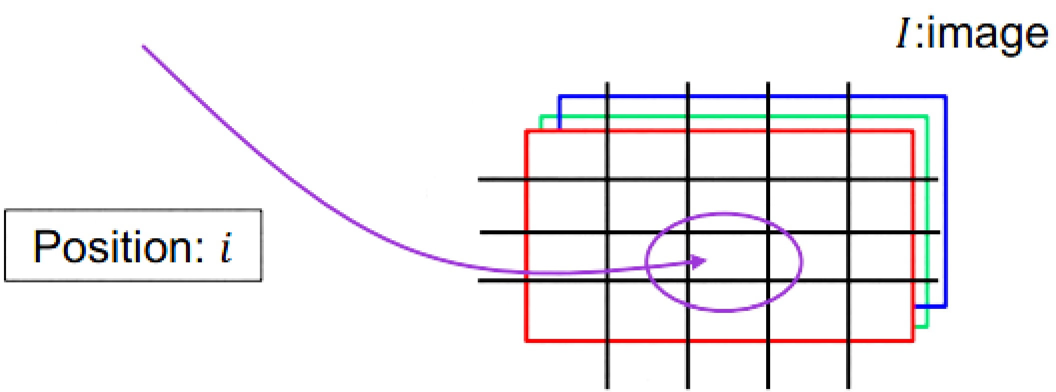


One pixel intensity over the image

$$R_i = \int_{\omega} E(\lambda) S_i(\lambda) R(\lambda) d\lambda$$

$$G_i = \int_{\omega} E(\lambda) S_i(\lambda) G(\lambda) d\lambda$$

$$B_i = \int_{\omega} E(\lambda) S_i(\lambda) B(\lambda) d\lambda$$

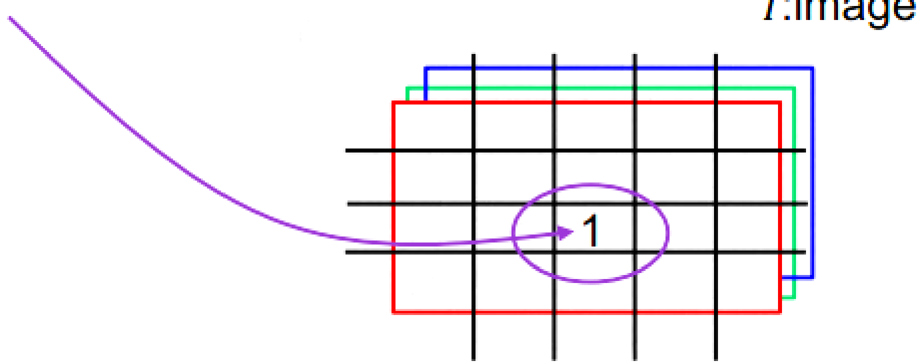


Assuming that at least a white patch exist in an image

$$\max_{i\in\{1,2,\cdots,N\}} R_i = \int_\omega E(\lambda)R(\lambda)d\lambda = R_e$$

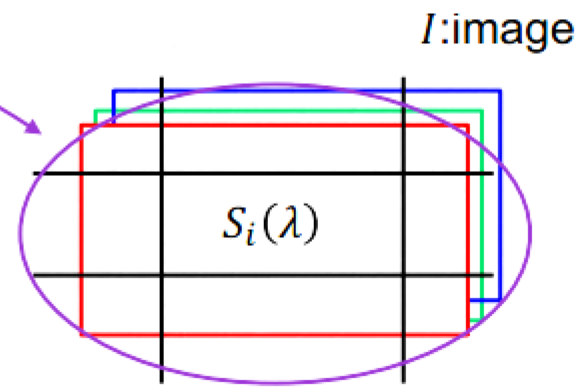
$$\max_{i\in\{1,2,\cdots,N\}} G_i = \int_\omega E(\lambda)G(\lambda)d\lambda = G_e$$

$$\max_{i\in\{1,2,\cdots,N\}} B_i = \int_\omega E(\lambda)B(\lambda)d\lambda = B_e$$



Assuming that a scene average is grey

$$\mu(S(\lambda)) = \sum_{i=1}^{N} \frac{S_i(\lambda)}{N} = k$$



$$\mu(\underline{R}) = \int_{\omega} E(\lambda)\left(\sum_{i=1}^{N} \frac{S_i(\lambda)}{N}\right) R(\lambda)d\lambda = kR_e$$

$$\mu(\underline{G}) = \int_{\omega} E(\lambda)\left(\sum_{i=1}^{N} \frac{S_i(\lambda)}{N}\right) G(\lambda)d\lambda = kG_e$$

$$\mu(\underline{B}) = \int_{\omega} E(\lambda)\left(\sum_{i=1}^{N} \frac{S_i(\lambda)}{N}\right) B(\lambda)d\lambda = kB_e$$

Minkowski family norm

Definition of p-norm for $\underline{X} = \{X_1, X_2, \cdots, X_N\}^T$

$$\|\underline{X}\|_p = \left\{\sum_{i=1}^{N} |X_i|^p\right\}^{1/p}$$

Example of 2 norm

- Equal to Euclidean distance

$$\|\underline{X}\|_2 = \left\{\sum_{i=1}^{N} |X_i|^2\right\}^{1/2} = \sqrt{X_1^2 + X_2^2 + \cdots + X_N^2}$$

2.1 Shades of Gray — Minkowski family norm

Mean of p-norm for $\underline{X} = \{X_1, X_2, \cdots, X_N\}^T$

$$\mu_p(\underline{X}) = \frac{\|\underline{X}\|_p}{N^{1/p}} = \sqrt[p]{\frac{X_1^p + X_2^p + \cdots + X_N^p}{N}}$$

Monotonically increasing sequence

$$\frac{\|\underline{X}\|_p}{N^{1/p}} \le \frac{\|\underline{X}\|_q}{N^{\frac{1}{q}}}, \qquad if \;\; p \le q$$

Infinity norm

$$\|\underline{X}\|_\infty = \max_{0 \le i \le N} |X_i|$$

Gray-World

$$\begin{bmatrix} R_e \\ G_e \\ B_e \end{bmatrix} = \begin{bmatrix} \mu_1(\underline{R}) \\ \mu_1(\underline{G}) \\ \mu_1(\underline{B}) \end{bmatrix}$$

Max-RGB

$$\begin{bmatrix} R_e \\ G_e \\ B_e \end{bmatrix} = \begin{bmatrix} \mu_\infty(\underline{R}) \\ \mu_\infty(\underline{G}) \\ \mu_\infty(\underline{B}) \end{bmatrix}$$

2.1 Shades of Gray

Gray-World

$$\begin{bmatrix} R_e \\ G_e \\ B_e \end{bmatrix} = \begin{bmatrix} \mu_1(\underline{R}) \\ \mu_1(\underline{G}) \\ \mu_1(\underline{B}) \end{bmatrix}$$

$$\begin{array}{c} \boxed{\begin{array}{c} \mu_1(\underline{R}) \\ \mu_1(\underline{G}) \\ \mu_1(\underline{B}) \end{array}} \begin{array}{c} \le \\ \le \\ \le \end{array} \boxed{\begin{array}{c} \mu_2(\underline{R}) \le \cdots \\ \mu_2(\underline{G}) \le \cdots \\ \mu_2(\underline{B}) \le \cdots \end{array}} \begin{array}{c} \le \\ \le \\ \le \end{array} \boxed{\begin{array}{c} \mu_\infty(\underline{R}) \\ \mu_\infty(\underline{G}) \\ \mu_\infty(\underline{B}) \end{array}} \end{array}$$

Max-RGB

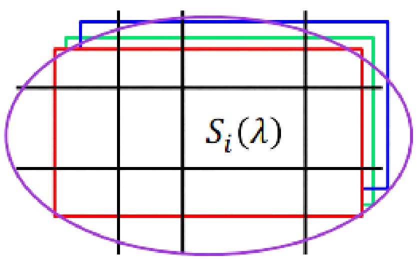
$$\begin{bmatrix} R_e \\ G_e \\ B_e \end{bmatrix} = \begin{bmatrix} \mu_\infty(\underline{R}) \\ \mu_\infty(\underline{G}) \\ \mu_\infty(\underline{B}) \end{bmatrix}$$

$$R_{p,i} = \left[\int_\omega E(\lambda)S_i(\lambda)R(\lambda)d\lambda\right]^p = \int_\omega E_p(\lambda)\sigma_i(\lambda)R(\lambda)d\lambda = R_i^p$$

$$G_{p,i} = \left[\int_\omega E(\lambda)S_i(\lambda)G(\lambda)d\lambda\right]^p = \int_\omega E_p(\lambda)\sigma_i(\lambda)G(\lambda)d\lambda = G_i^p$$

$$B_{p,i} = \left[\int_\omega E(\lambda)S_i(\lambda)B(\lambda)d\lambda\right]^p = \int_\omega E_p(\lambda)\sigma_i(\lambda)B(\lambda)d\lambda = B_i^p$$

I:image



$$\mu_p(S(\lambda)) = \left[\sum_{i=1}^{N} \frac{\{S_i(\lambda)\}^p}{N}\right]^{1/p} = k_p$$

Assuming that the average of pixels raised to the power of p is gray

2.1 Shades of Gray

$$\mu_p(\underline{R_p}) = \left[\int_\omega E_p(\lambda)\left(\sum_{i=1}^{N}\frac{\{S_i(\lambda)\}^p}{N}\right)R(\lambda)d\lambda\right]^{1/p} = k_p R_e$$

$$\mu_p(\underline{G_p}) = \left[\int_\omega E_p(\lambda)\left(\sum_{i=1}^{N}\frac{\{S_i(\lambda)\}^p}{N}\right)G(\lambda)d\lambda\right]^{1/p} = k_p G_e$$

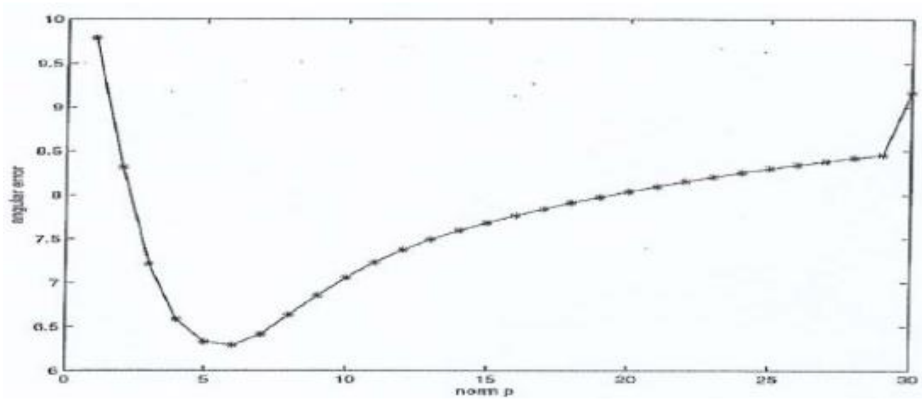
$$\mu_p(\underline{B_p}) = \left[\int_\omega E_p(\lambda)\left(\sum_{i=1}^{N}\frac{\{S_i(\lambda)\}^p}{N}\right)B(\lambda)d\lambda\right]^{1/p} = k_p B_e$$

where $\underline{R_p} = \{R_1^p, R_2^p, \cdots, R_N^p\}^T$

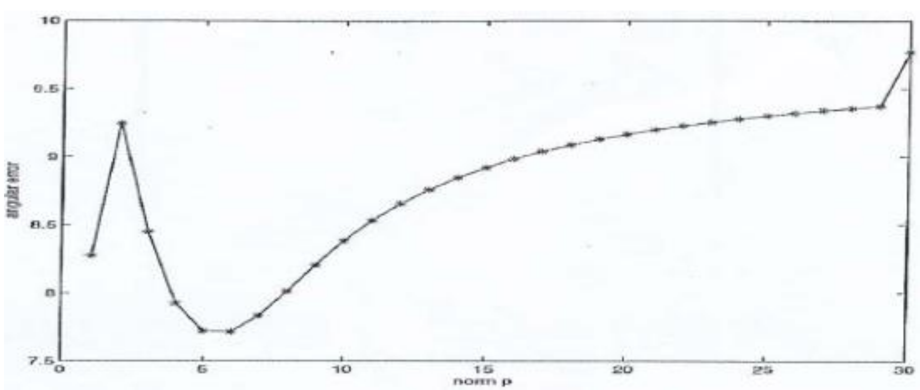
, $\underline{G_p} = \{G_1^p, G_2^p, \cdots, G_N^p\}^T$

, $\underline{B_p} = \{B_1^p, B_2^p, \cdots, B_N^p\}^T$

2.1 Shades of Gray p值选择



The figure shows the angular error of the group A images for 30 values of p

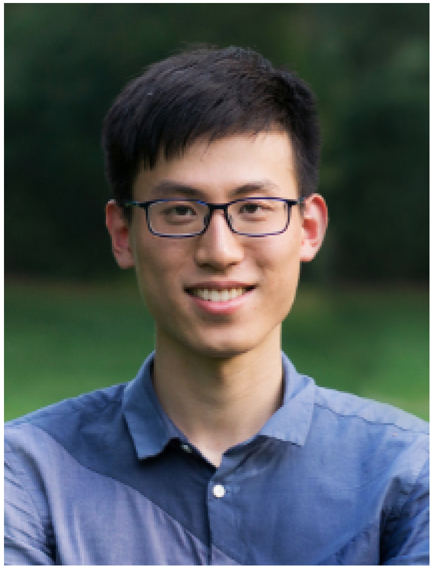


The figure shows the angular error of the group B images for 30 values of p

norm	ang_err		dist_err	
p	group A	group B	group A	group B
1	9.78	8.27	0.0788	0.0624
2	8.32	9.24	0.0640	0.0668
3	7.22	8.45	0.0535	0.0593
4	5.59	7.93	0.0476	0.0549
5	6.33	7.72	0.0448	0.0530
6	6.29	7.71	0.0440	0.0527
7	6.42	7.83	0.0445	0.0533
8	6.64	8.01	0.0456	0.0543
9	6.86	8.21	0.0469	0.0555
10	7.06	8.38	0.0481	0.0566
11	7.23	8.53	0.0492	0.0576
12	7.37	8.66	0.0502	0.0584
13	7.49	8.76	0.0510	0.0590
14	7.59	8.85	0.0517	0.0596
15	7.68	8.92	0.0523	0.0601
16	7.76	8.99	0.0523	0.0605
17	7.84	9.04	0.0534	0.0609
18	7.91	9.09	0.0539	0.0612
19	7.97	9.13	0.0544	0.0615
20	8.04	9.17	0.0549	0.0618
21	8.10	9.20	0.0553	0.0620
22	8.15	9.23	0.0557	0.0622
23	8.20	9.26	0.0561	0.0624
24	8.25	9.28	0.0565	0.0626
25	8.30	9.30	0.0568	0.0627
26	8.34	9.32	0.0571	0.0629
27	8.38	9.34	0.0574	0.0630
28	8.42	9.36	0.0577	0.0631
29	8.46	9.37	0.0580	0.0632
∞	9.16	9.77	0.0630	0.0659

Fully Convolutional Color Constancy with Confidence-weighted Pooling (CVPR 2017)

- *代码:*
  - *https://github.com/yuanming-hu/fc4*
- *论文:*
  - *http://openaccess.thecvf.com/content_cvpr_2017/papers/Hu_FC4_Fully_Convolutional_CVPR_2017_paper.pdf*
- *测试环境:*
  - *TensorFlow-GPU + python2.7*



Yuanming Hu 胡渊鸣

I am a second-year Ph.D. student at MIT CSAIL, advised by Frédo Durand and Bill Freeman.

I graduated with honor from Tsinghua University (Yao class) in July 2017. I worked on deep learning and computer vision, during my internship with Stephen Lin at Microsoft Research Asia. My undergraduate thesis is on automatic photo post-processing using reinforcement learning and adversarial learning (presented at SIGGRAPH 2018). I completed my master thesis (The ChainQueen Differentiable Physical Simulator) with Wojciech Matusik in November 2018. My work has been partly supported by an Edwin Webster fellowship and a Snap Research fellowship.

2.2 FC4 网络结构

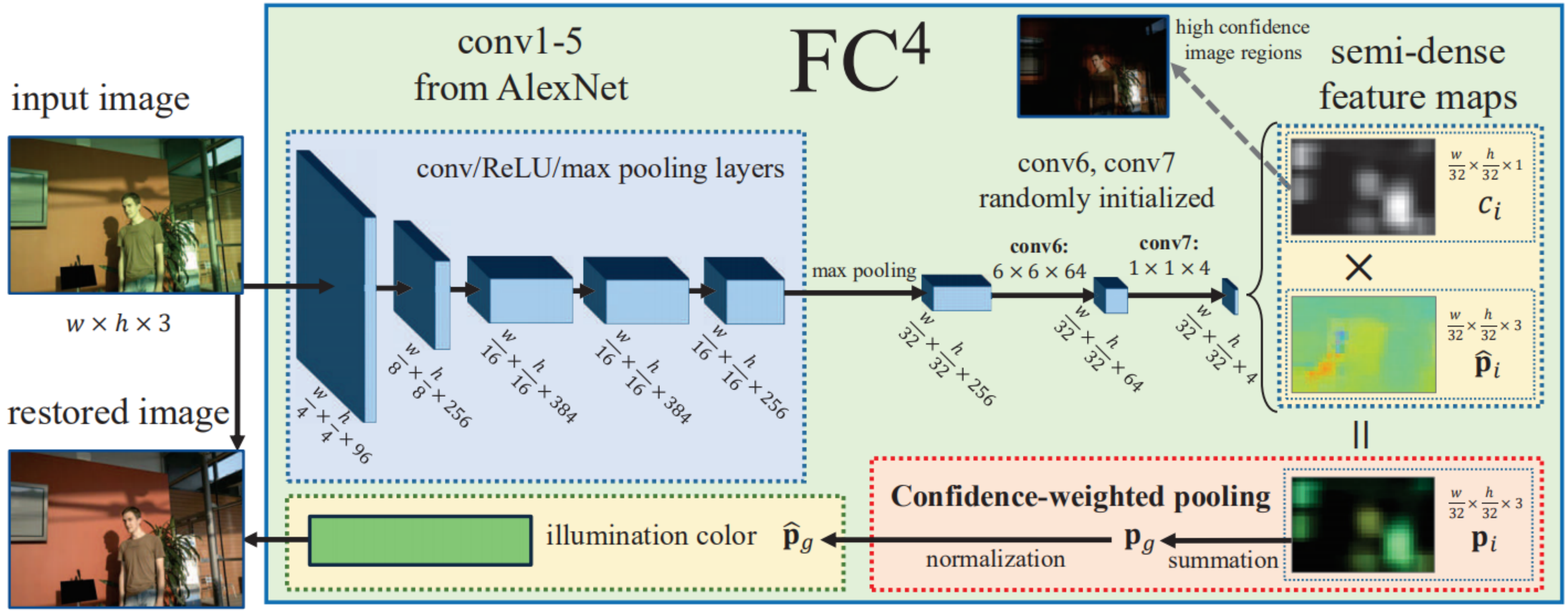


Figure 2: The architecture of AlexNet-FC^4. Replacing AlexNet (conv1-conv5) with SqueezeNet v1.1 (conv1-fire8 plus an extra 2 × 2 pooling) yields SqueezeNet-FC^4.

- 论文使用**全卷积网络可以接受任意图像尺寸的输入**
- 提出SqueezeNet和AlexNet两个版本，相较于AlexNet网络，**SqueezeNet更轻量级，在保持同等分类精准率的前提下模型参数缩小了50倍**

2.2 FC4 网络结构

在ImageNet上预训练的AlexNet或者SqueezeNet的前5层

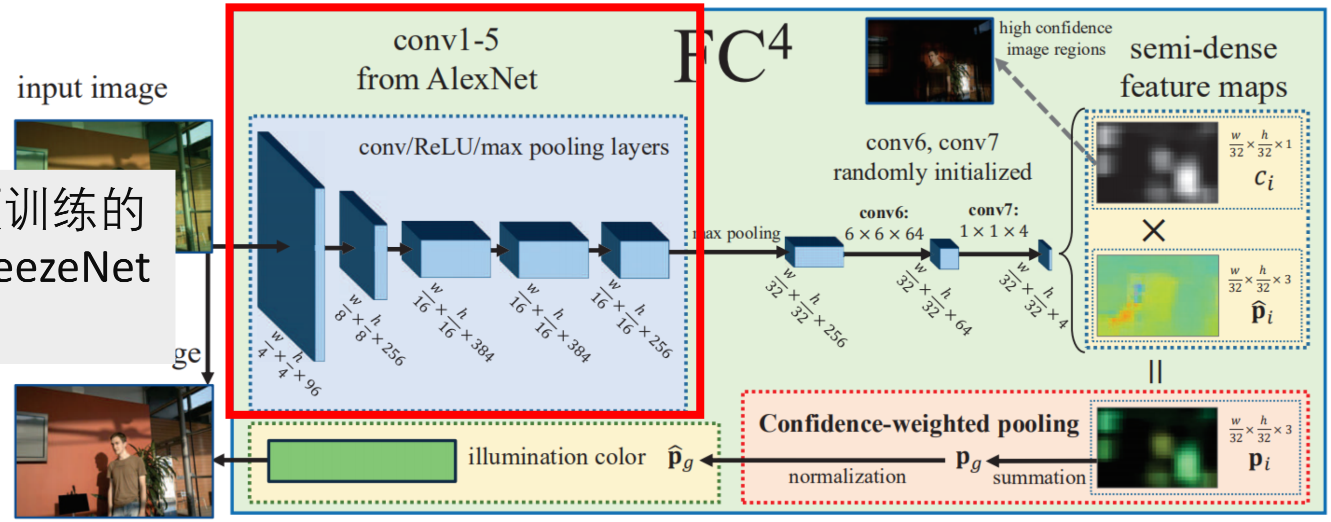


Figure 2: The architecture of AlexNet-FC^4. Replacing AlexNet (conv1-conv5) with SqueezeNet v1.1 (conv1-fire8 plus an extra 2×2 pooling) yields SqueezeNet-FC^4.

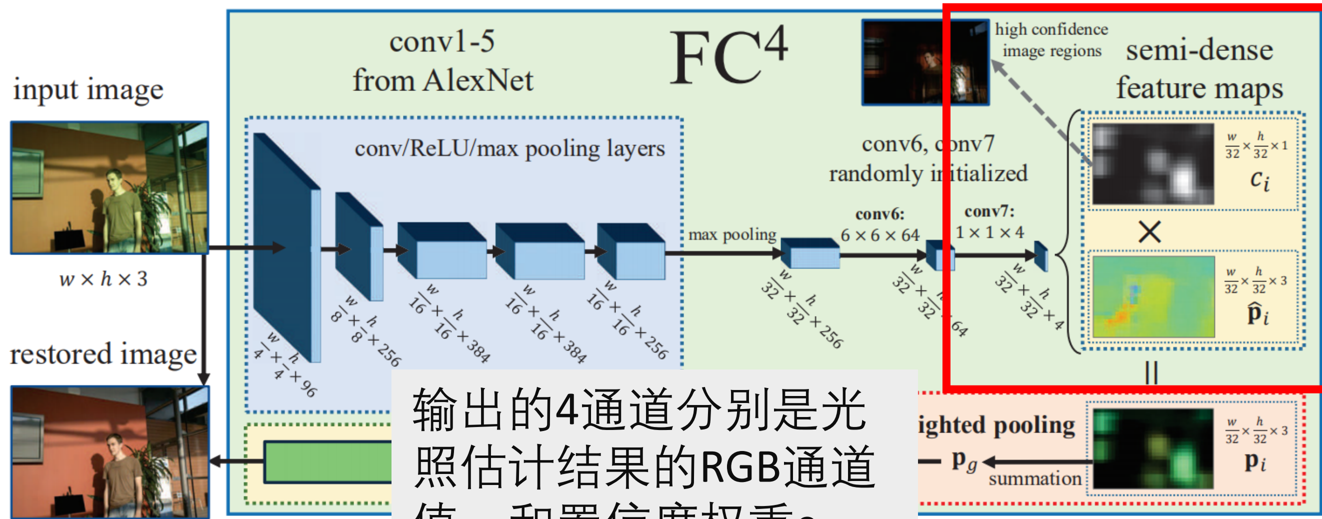
选择网络结构原则：

(1) the network should be able to extract sufficient semantic features to discriminate ambiguous patches (such as textureless walls) for illumination estimation

(2) the network should not be illumination invariant, but rather it should be sensitive to different lighting colors

选取的网络应该能够提取足够的语义特征来选择合适的patchs用于照明估计，并对不同的光源颜色敏感

2.2 FC4 网络结构



输出的4通道分别是光照估计结果的RGB通道值，和置信度权重c

Figure 2: The architecture of AlexNet-FC[4]. Replacing AlexNet (conv1-conv5) with SqueezeNet v1.1 (conv1-fire8 plus an extra 2×2 pooling) yields SqueezeNet-FC[4].

语义信息选择机制：

- 选择图像中**语义信息丰富的色块**用于估计，避免语义不明确的色块影响照度估计
- 网络中采用**更大的带有更多的语义的512*512patch**（以往论文中大多是32*32的patch）
- 利用置信度权重，可以**将监督信号仅派发给训练期间具有语义的区域**

$\hat{P}_g^*$：真实光照值(ground truth)

$\hat{P}_g$：光照估计值

$c(R_i)$： R_i 区域的权重值

$g(R_i)$： R_i 区域的照度估计

光源估计值计算：

$$\hat{P}_g = \text{normalize}\left(\sum_{i \in R} c(R_i) g(R_i)\right)$$

损失函数计算（AE值）：

$$L(\hat{P}_g) = \frac{180}{\pi} \arccos(\hat{P}_g \cdot \hat{P}_g^*)$$

2.2 FC4 结果

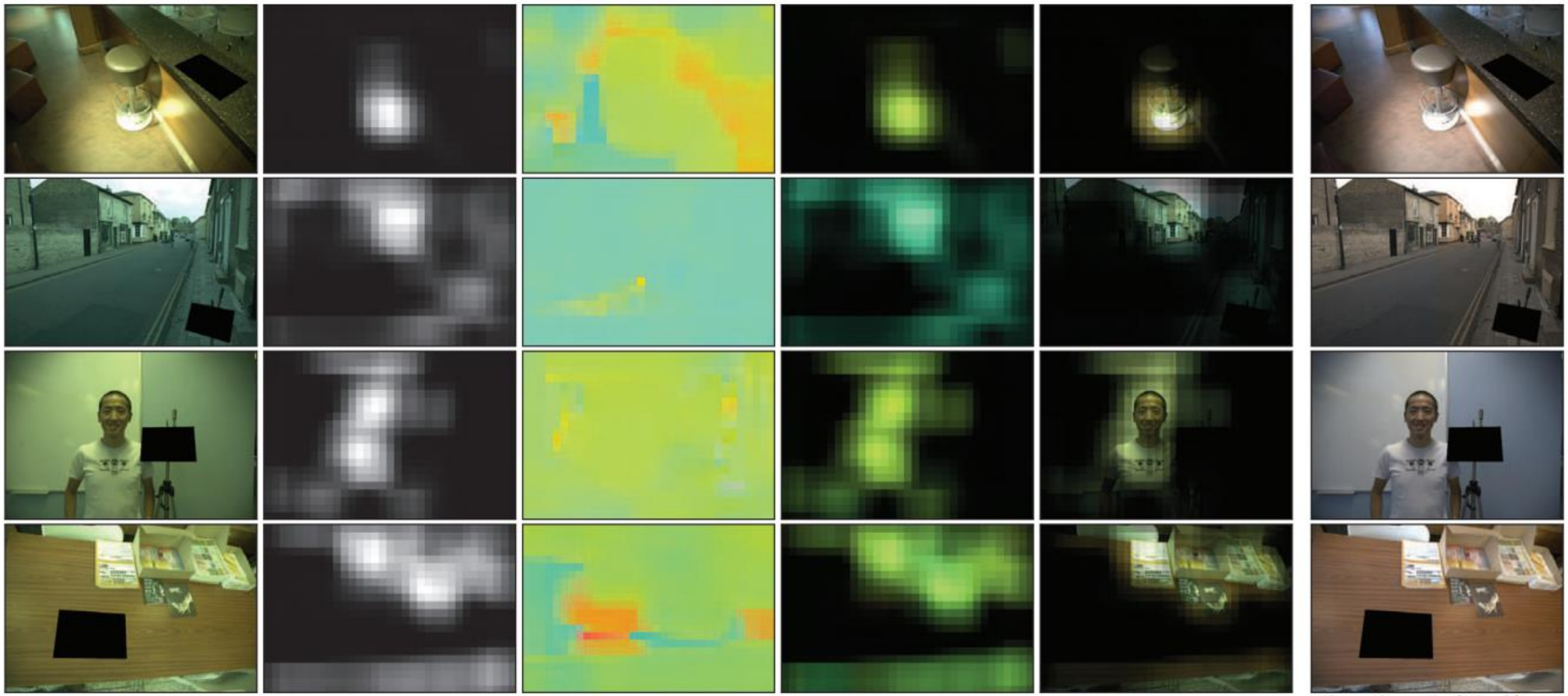


Figure 3: Examples of outputs by our network. Note that noisy estimates in regions of little semantic value are masked by the confidence map, resulting in more robust estimation. The angular errors are 0.54, 4.63, 1.78 and 4.76 degrees, respectively.

因为FC4网络会学习到图片中的颜色信息作为语义来影响照度估计，所以**需要将图片中颜色丰富的color checker去掉**以免干扰结果。
作者的做法是根据color checker的坐标信息，将其**位置处像素置黑**，去除干扰。

2.2 FC4 结果

Method	Mean	Med.	Tri.	Best 25%	Worst 25%	95% Quant.
White-Patch [10]	7.55	5.68	6.35	1.45	16.12	-
Edge-based Gamut [3]	6.52	5.04	5.43	1.90	13.58	-
Gray-World [11]	6.36	6.28	6.28	2.33	10.58	11.3
1st-order Gray-Edge [42]	5.33	4.52	4.73	1.86	10.03	11.0
2nd-order Gray-Edge [42]	5.13	4.44	4.62	2.11	9.26	-
Shades-of-Gray [19]	4.93	4.01	4.23	1.14	10.20	11.9
Bayesian [21]	4.82	3.46	3.88	1.26	10.49	-
General Gray-World [4]	4.66	3.48	3.81	1.00	10.09	-
Intersection-based Gamut [3]	4.20	2.39	2.93	0.51	10.70	-
Pixel-based Gamut [3]	4.20	2.33	2.91	0.50	10.72	14.1
Natural Image Statistics [22]	4.19	3.13	3.45	1.00	9.22	11.7
Bright Pixels [27]	3.98	2.61	-	-	-	-
Spatio-spectral (GenPrior) [13]	3.59	2.96	3.10	0.95	7.61	-
Cheng et al. 2014 [14]	3.52	2.14	2.47	0.50	8.74	-
Corrected-Moment (19 Color) [17]	3.50	2.60	-	-	-	8.60
Exemplar-based [26]	3.10	2.30	-	-	-	-
Corrected-Moment (19 Color)* [17]	2.96	2.15	2.37	0.64	6.69	8.23
Corrected-Moment (19 Edge) [17]	2.80	2.00	-	-	-	6.90
Corrected-Moment (19 Edge)* [17]	3.12	2.38	2.59	0.90	6.46	7.80
Regression Tree [15]	2.42	1.65	1.75	0.38	5.87	-
CNN [7]	2.36	1.98	-	-	-	-
CCC (dist+ext) [5]	1.95	1.22	1.38	0.35	4.76	5.85
DS-Net (HypNet+SelNet) [38]	1.90	1.12	1.33	0.31	4.84	5.99
AlexNet-FC^4	1.77	1.11	1.29	0.34	4.29	5.44
SqueezeNet-FC^4	1.65	1.18	1.27	0.38	3.78	4.73

Quasi-unsupervised Color Constancy (CVPR 2019)

- *代码:*
    - *https://github.com/claudio-unipv/quasi-unsupervised-cc*
- *论文:*
    - *http://openaccess.thecvf.com/content_CVPR_2019/papers/Bianco_Quasi-Unsupervised_Color_Constancy_CVPR_2019_paper.pdf*
- *测试环境:*
    - *pytorch-gpu + python3.7*

Simone Bianco Assistant Professor

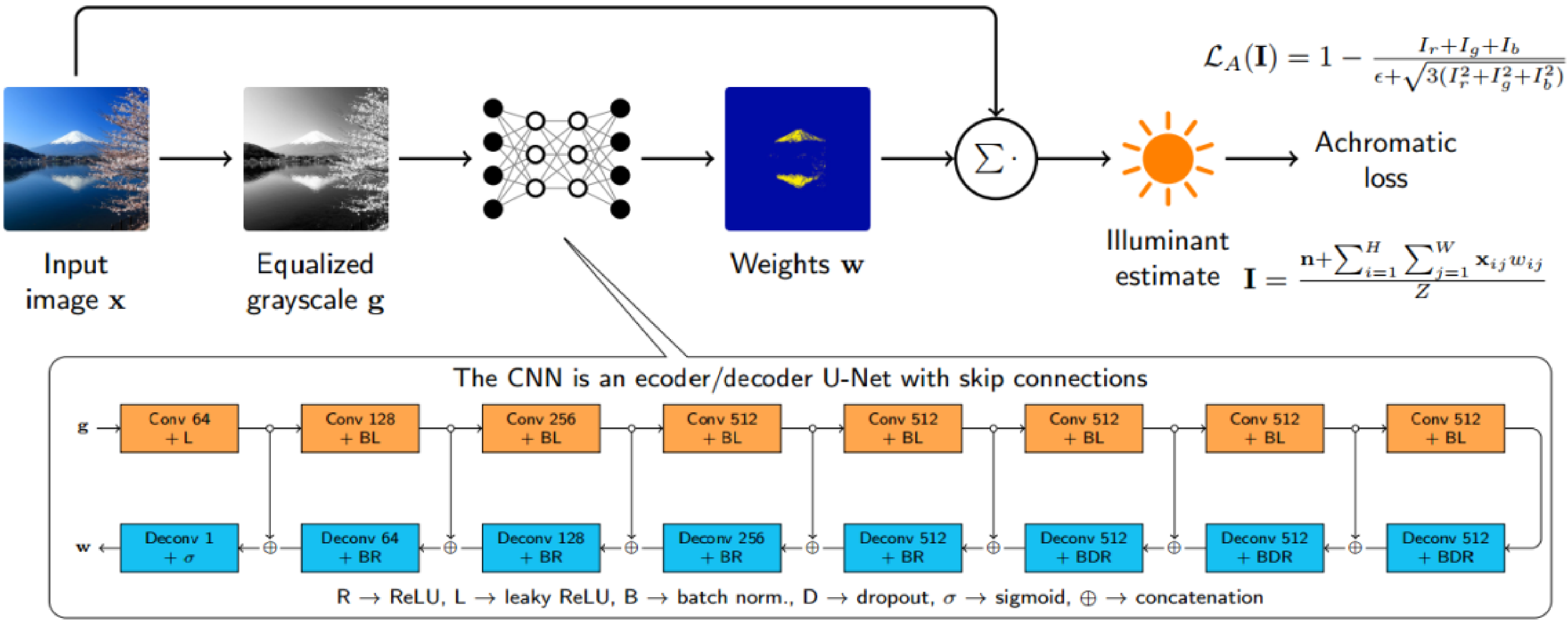
e-mail: bianco@disco.unimib.it

DISCo (Department of Informatics, Systems and Communication)
University of Milan-Bicocca
Viale Sarca 336, Building U14
Room 1011, tel: +390264487827

Biography

Simone Bianco obtained the PhD in Computer Science at DISCo (Dipartimento di Informatica, Sistemistica e Comunicazione) of the University of Milano-Bicocca, Italy, in 2010. He obtained the BSc and the MSc degree in Mathematics from the University of Milano-Bicocca, Italy, respectively in 2003 and 2006. He is currently an assistant professor and his research interests include computer vision, machine learning, optimization algorithms, and color imaging.

2.3 Quasi-unsupervised CC 网络结构

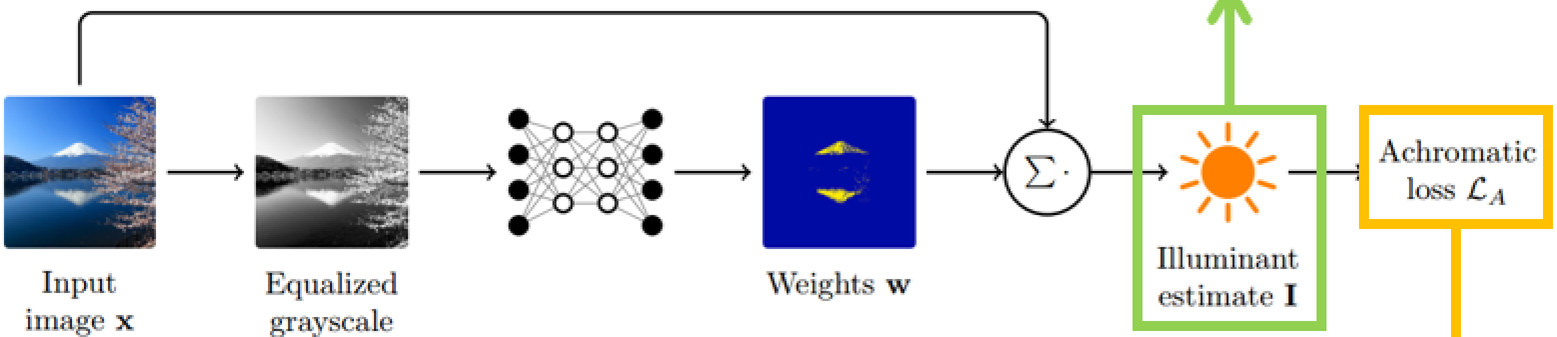


Quasi-unsupervised：

No ground truth about the color of the illuminant is needed. Instead, the method exploits the assumption **that images have been approximately balanced** either manually or by unspecified automatic processing pipelines.

2.3 Quasi-unsupervised CC 网络结构

$$\mathbf{I} = \frac{\mathbf{n} + \sum_{i=1}^{H} \sum_{j=1}^{W} \mathbf{x}_{ij} w_{ij}}{Z}$$



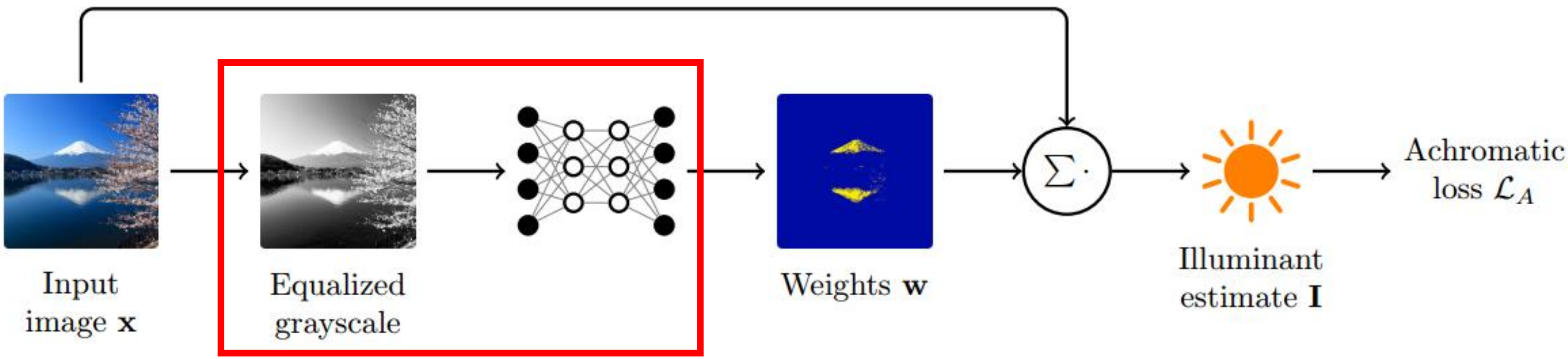
$$\boxed{\mathcal{L}_A(\mathbf{I})} = 1 - \frac{I_r + I_g + I_b}{\boxed{\epsilon} + \sqrt{3(I_r^2 + I_g^2 + I_b^2)}}$$

$I_r \simeq I_g \simeq I_b$

achromatic loss
close to zero

$\epsilon = 10^{-4}$

small quantity used
to stabilize the ratio



网络输入：

- 第一种：灰度图（Grayscale）
- 第二种：6通道图像（Directions）
  - 用Sobel算子计算水平和垂直空间的导数
  - 将导数归一化为单位向量
  - 两个导数乘以三通道得到一个六通道的图像

2.3 Quasi-unsupervised CC 网络结构

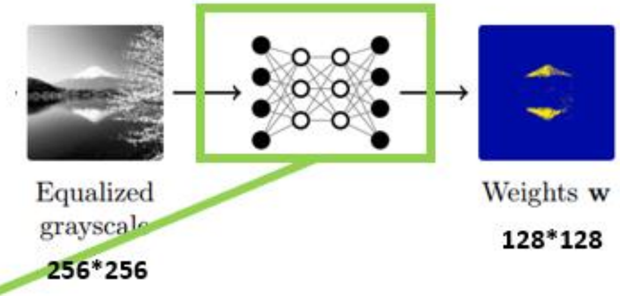
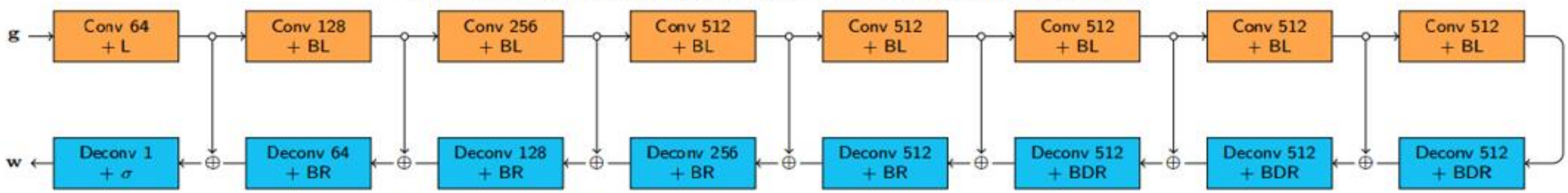


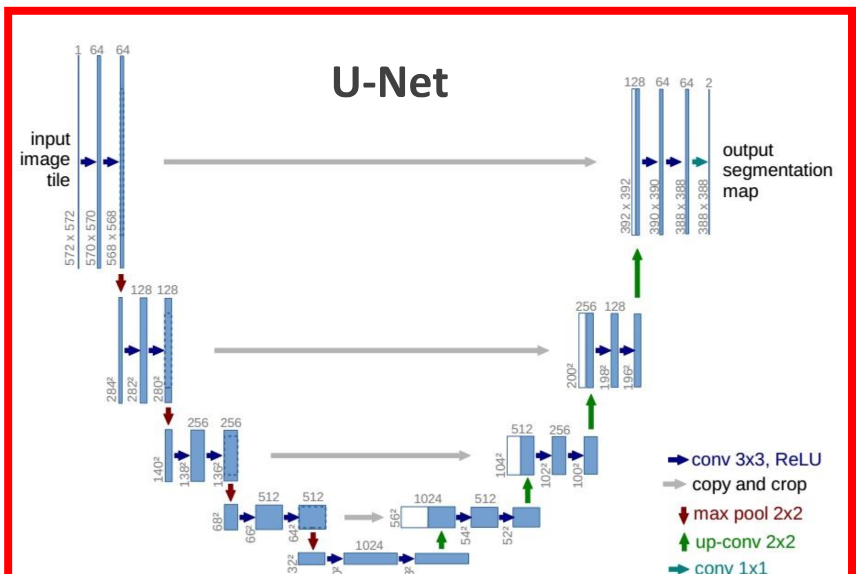
image-to-image translation:

- Pix2pix模型用在image-to-image任务上，此处grayscale-to-weight有相似之处
- Pix2pix模型在着色问题上有效，说明该模型结构对于图像颜色（尤其是消色区域的颜色）具有识别能力，符合颜色恒常性任务

The CNN is an ecoder/decoder U-Net with skip connections



R → ReLU, L → leaky ReLU, B → batch norm., D → dropout, σ → sigmoid, ⊕ → concatenation



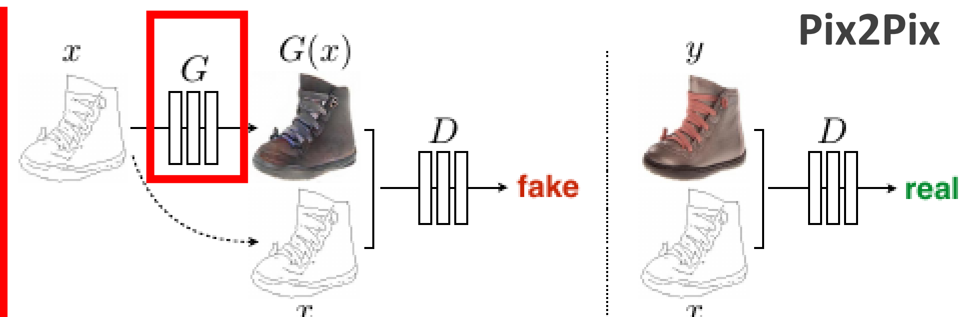


Figure 2: Training a conditional GAN to map edges→photo. The discriminator, D, learns to classify between fake (synthesized by the generator) and real {edge, photo} tuples. The generator, G, learns to fool the discriminator. Unlike an unconditional GAN, both the generator and discriminator observe the input edge map.

使用一般数据集的Quasi-unsupervised Learning：

- **Ilsvrc12**：
  - 用于互联网大规模视觉识别挑战赛；由来自1000个不同类别的大约120万个样本组成
- **Places365**：
  - 作为场景识别系统的基准，包括代表365种不同类别场景的约180万张图像
- **Flickr100k**：
  - 评估图像检索算法，包含通过搜索146个最受欢迎的标签从Flickr照片共享服务收集的100071张图像

$$\underbrace{\boxed{\mathcal{L}_A(\mathbf{I})}}_{\text{achromatic loss}} = 1 - \frac{\overbrace{\boxed{I_r + I_g + I_b}}^{\text{network estimate}}}{\epsilon + \sqrt{3(I_r^2 + I_g^2 + I_b^2)}}$$

使用颜色恒常性数据集的Quasi-unsupervised Learning & Fine tuning：

使用一般数据集对无监督模型进行训练，然后使用经典的颜色恒常性数据集对模型进行进一步微调

- Color Checker dataset (Shi’s re-processing version)
- NUS-8 Camera Dataset

$$\text{chromatic loss} \longleftarrow \boxed{\mathcal{L}_C(\mathbf{I},\hat{\mathbf{I}})} = 1 - \frac{I_r \cdot \hat{I}_r + I_g \cdot \hat{I}_g + I_b \cdot \hat{I}_b}{\epsilon + \sqrt{\underbrace{\boxed{(I_r^2 + I_g^2 + I_b^2)}}_{\text{network estimate}}\underbrace{\boxed{(\hat{I}_r^2 + \hat{I}_g^2 + \hat{I}_b^2)}}_{\text{ground truth}}}}$$

模型包含多个变体：

- **训练数据集的不同**
  - Ilsvrc12
  - Places365
  - Flickr100k

对于训练数据集的差异：**网络的输出与所选择的训练集关系不大**。在大多数情况下，在Ilsvrc12和Places365上训练的网络得到的权重非常相似

- **输入的数据形式不同**
  - Equalized grayscale
  - Gradient directions
  - Both

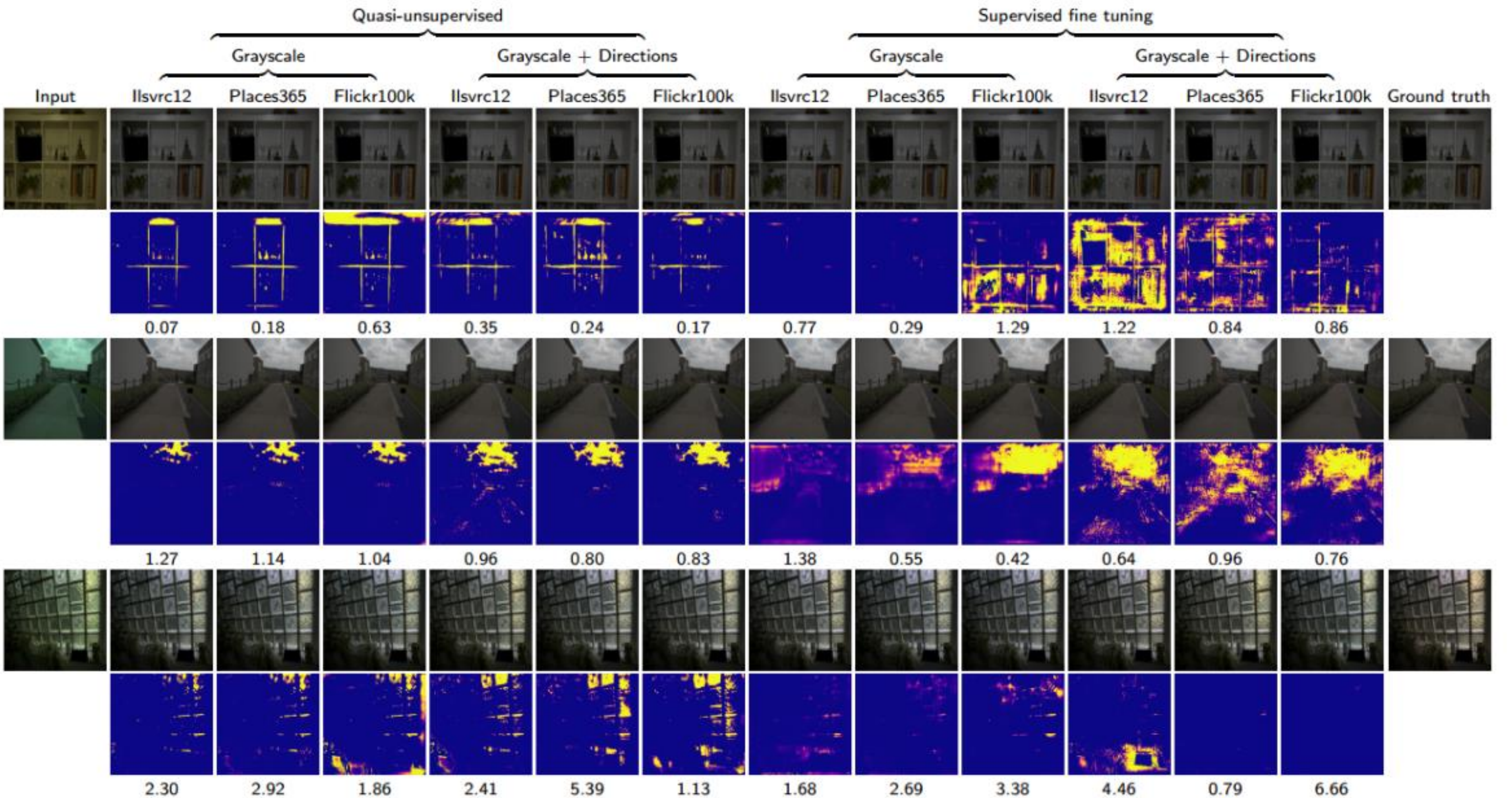
对于输入数据的差异：相对于使用梯度方向的网络，**仅处理灰度图像的网络会选择更容易识别区域中的像素**

- **训练的方式不同**
  - Quasi-unsupervised
  - Quasi-unsupervised + fine tuning

对于训练方式的差异：**fine-tuning网络倾向于选择更少的像素**。可能的原因是在fine-tuning时移除了一些可能会输出大权重的噪声项

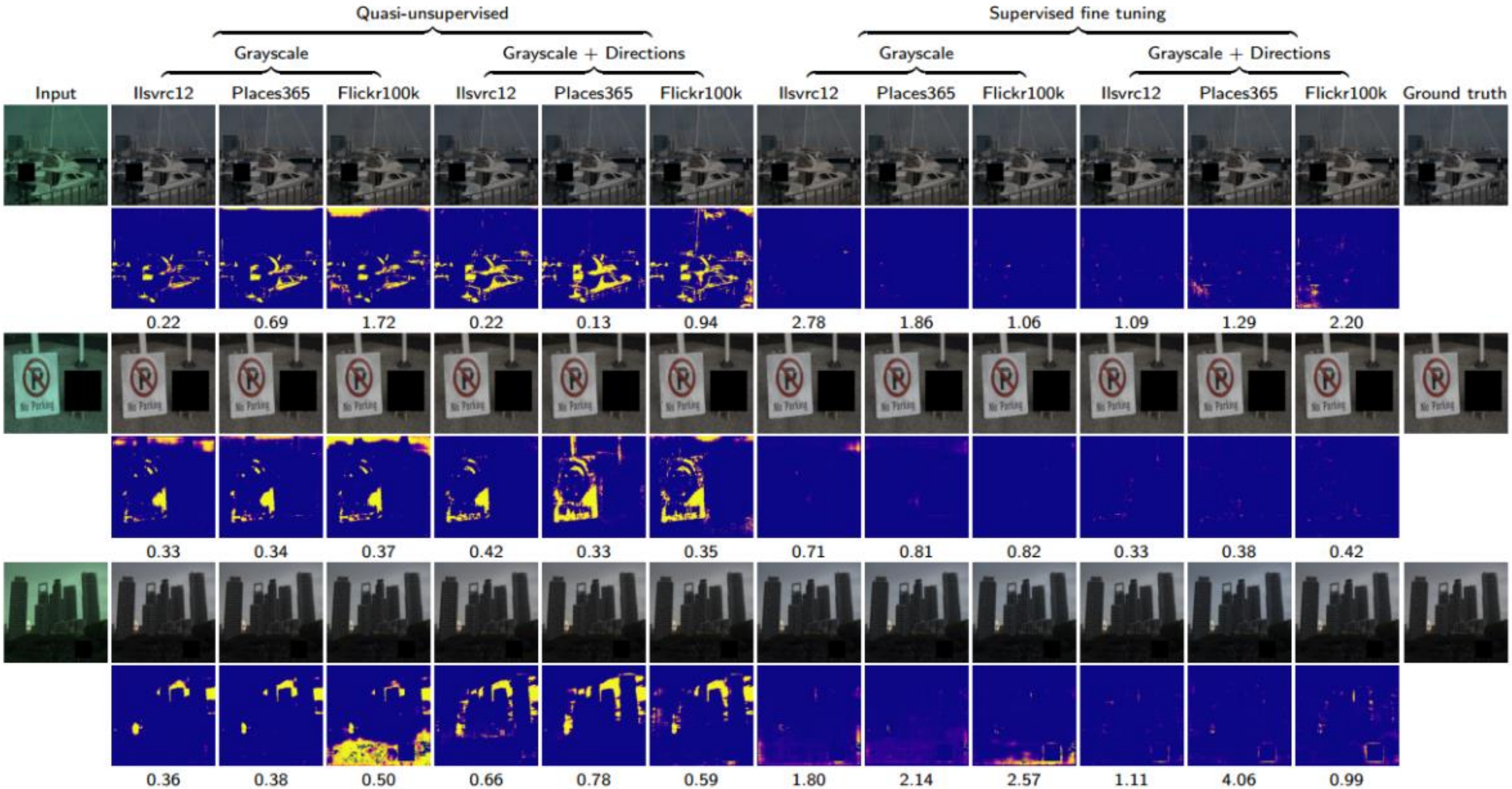
2.3 Quasi-unsupervised CC 可视化结果

Color Checker images processed by variants of the method



2.3 Quasi-unsupervised CC 可视化结果

NUS-8 images processed by variants of the method



2.3 Quasi-unsupervised CC 测试结果

Quasi-unsupervised Learning

Training set	Test set	Input	Mean	Median	Max
Ilsvrc12	CC	Grayscale	4.04	2.67	27.88
	CC	Directions	3.67	2.53	**17.62**
	CC	Both	3.46	**2.23**	21.17
Places365	CC	Grayscale	4.01	2.60	27.72
	CC	Directions	**3.43**	2.38	18.31
	CC	Both	3.60	2.45	21.47
Flickr100k	CC	Grayscale	4.09	2.67	27.09
	CC	Directions	3.70	2.48	20.86
	CC	Both	3.59	2.25	20.04
Ilsvrc12	NUS	Grayscale	3.14	2.24	22.39
	NUS	Directions	2.97	2.15	15.89
	NUS	Both	3.00	2.27	19.16
Places365	NUS	Grayscale	3.24	2.32	22.66
	NUS	Directions	**2.91**	2.24	16.05
	NUS	Both	3.07	2.20	17.12
Flickr100k	NUS	Grayscale	3.27	2.38	21.28
	NUS	Directions	2.95	**2.12**	16.40
	NUS	Both	2.98	2.16	**15.86**

Table 1. Statistics of angular errors (in degrees) obtained by variations of the proposed method on the CC and NUS datasets. Training has been performed on three datasets with different inputs: equalized grayscale, gradient directions, and their combination.

使用Sobel计算的梯度信息作为输入会比只用灰度图作为输入要效果好

Quasi-unsupervised Learning & Fine tuning

Dataset	Mean	Median	Max
CC	2.91 (-0.55)	1.98 (-0.25)	19.9 (-1.2)
NUS	1.97 (-1.03)	1.41 (-0.86)	20.5 (+1.6)

Table 2. Statistics of angular errors (in degrees) obtained by the network trained on Ilsvrc12 and fine tuned on the two test datasets. The values in brackets report the difference with respect to those obtained in the quasi-unsupervised setting.

mean和median的AE值在使用Color Checker和NUS数据集fine tuning后均降低

2.3 Quasi-unsupervised CC 测试结果

- **in dataset(in-db)**：模型的训练和测试是在同一个颜色恒常性数据集上（数据集划分为训练/验证/测试）
- **cross dataset(cross-db)**：模型训练和微调是在A颜色恒常性数据集，测试是在B颜色恒常性数据集
- **no dataset(no-db)**：模型没有在任何颜色恒常性数据集上进行训练，测试是在颜色恒常性数据集上进行

	Method	Color Checker Mean	Color Checker Med.	Color Checker Max	NUS Mean	NUS Med.										
	SoG [19] with GSA [3]	4.05	2.54	21.87	3.31	2.58	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	gGW [4] with GSA [3]	4.05	2.58	21.15	3.45	2.68	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Unsupervised (in-db)	GE1[46] with GSA [3]	**4.03**	**3.08**	18.89	3.18	2.48	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	GE2[46] with GSA [3]	4.13	3.34	**17.78**	3.41	2.52	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	Banic and Loncaric [2]				**2.96**	**1.70**	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	WP [34]	5.97	3.74	45.00	3.57	2.49	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	GW [9]	4.76	3.59	24.92	4.17	3.17	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Unsupervised (no-db)	Buzzelli et al. (gl. norm)[10]	4.84	4.12	20.80	4.88	4.17	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	Buzzelli et al. (ch. norm)[10]	5.48	4.81	**19.88**	4.32	3.37	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	Proposed	**3.46**	**2.23**	21.17	**3.00**	**2.25**	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	SoG [19]	3.85	2.43	**20.89**	3.42	2.45	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	gGW [4]	4.12	2.52	22.51	3.37	2.49	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	GE1 [46]	4.06	2.67	23.05	3.18	2.18	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Parametric (in-db)	GE2 [46]	4.18	2.68	24.05	3.19	2.18	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	BP[32]	3.98	2.61			2.48		2.45	2.48	2.67	2.30	2.18	2.15	2.49	2.62	3.13
	Cheng et al.[15]	**3.52**	**2.14**	28.35	**3.02**	**2.12**	23.28	**2.01**	1.89	2.15	**2.08**	**1.87**	2.02	**2.03**	2.33	2.72
	Grey Pixel (edge) [48]	4.60	3.10		3.15	2.20										
	SoG [19]	6.08	3.85	37.24	3.44	2.59	**18.40**	2.73	2.43	2.70	2.58	2.47	2.41	2.41	2.68	2.94
	gGW [4]	4.66	2.84	31.59	3.53	2.71	19.87	2.71	2.53	2.81	2.67	2.62	2.56	2.48	2.75	3.14
Parametric (cross-db)	GE1 [46]	4.06	2.67	**23.05**	3.18	2.18	21.81	2.37	**2.00**	2.06	2.23	**1.98**	**2.02**	**2.13**	2.34	**2.70**
	GE2 [46]	4.26	2.82	23.45	3.53	2.62	23.00	2.92	2.62	2.58	2.62	2.63	2.33	2.34	2.66	3.20
	Qian et al. [40]	**3.65**	**2.38**	26.12	**3.16**	**2.15**	21.93	**2.22**	2.07	**1.91**	**2.18**	2.06	2.11	2.14	**2.14**	3.05
	Bayesian [23]	4.70	3.44			2.81		2.80	2.35	3.20	3.10	2.81	2.41	3.00	2.36	3.53
	Spatio-Spectral (ML) [13]	3.55	2.93			2.54		2.80	2.32	2.70	2.43	2.24	2.28	2.51	2.70	2.99
	Spatio-Spectral (GP) [13]	3.47	2.90			2.39		2.67	2.03	2.45	2.26	2.21	2.22	2.29	2.58	2.89
	Natural Image Statistics [24]	4.09	3.13			2.69		3.04	2.46	2.95	2.40	2.17	2.28	2.77	2.88	3.51
	Exemplar-based [31]	2.89	2.27													
	Chakrabarti (Empirical)[12]	2.89	1.89													
	Chakrabarti (End-to-end)[12]	2.56	1.67													
	Cheng et al. [16]	2.42	1.65			1.58		1.57	1.62	1.58	1.65	1.41	1.61	1.78	1.48	
Supervised (in-db)	Color Dog [1]		1.49			1.76		1.72	1.85	1.81	1.94	1.46	1.69	1.89	1.77	
	Bianco et al. [8]	2.36	1.44	16.98		1.77		1.71	1.85	1.75	1.88	1.65	1.59	1.88	1.63	2.00
	FFCC [6]	1.78	**0.96**	**16.25**	1.99	**1.34**	**19.80**	**1.34**	1.33	1.35	**1.45**	**1.16**	**1.28**	1.47	**1.35**	**1.36**
	Oh and Kim [38]	2.16	1.47		2.41	2.15		2.18	1.75	2.75	2.00	2.22	1.53	1.65	3.11	2.68
	CCC (dist+ext) [5]	1.95	1.22		2.38	1.48										
	FC^4(AlexNet) [28]	**1.77**	1.11		2.12	1.53										
	DS-Net (HypNet+SelNet) [44]	1.90	1.12		2.24	1.46										
	Proposed + Fine Tuning	2.91	1.98	19.9	**1.97**	1.41	20.50	1.59	**1.26**	**1.34**	1.52	1.35	1.29	**1.30**	1.52	1.84
	Bayesian [23]	4.75	3.11		3.65	3.08										
	Exemplar-based [31]	6.50	5.10													
Supervised (cross-db)	Chakrabarti (Empirical) [12]	**3.87**	3.25		3.49	2.87		2.39	3.03	3.72	**3.07**	4.30	**2.00**	3.15	**3.92**	
	Chakrabarti (End-to-end) [12]	3.89	**3.10**		3.52	2.71		**2.18**	**2.42**	**3.01**	3.17	**3.29**	2.33	**3.13**	4.32	
	Cheng et al. [16]	5.52	4.52		4.86	4.40										
	FFCC [6]	3.91	3.15		**3.19**	**2.33**										

2.3 Quasi-unsupervised CC 测试结果

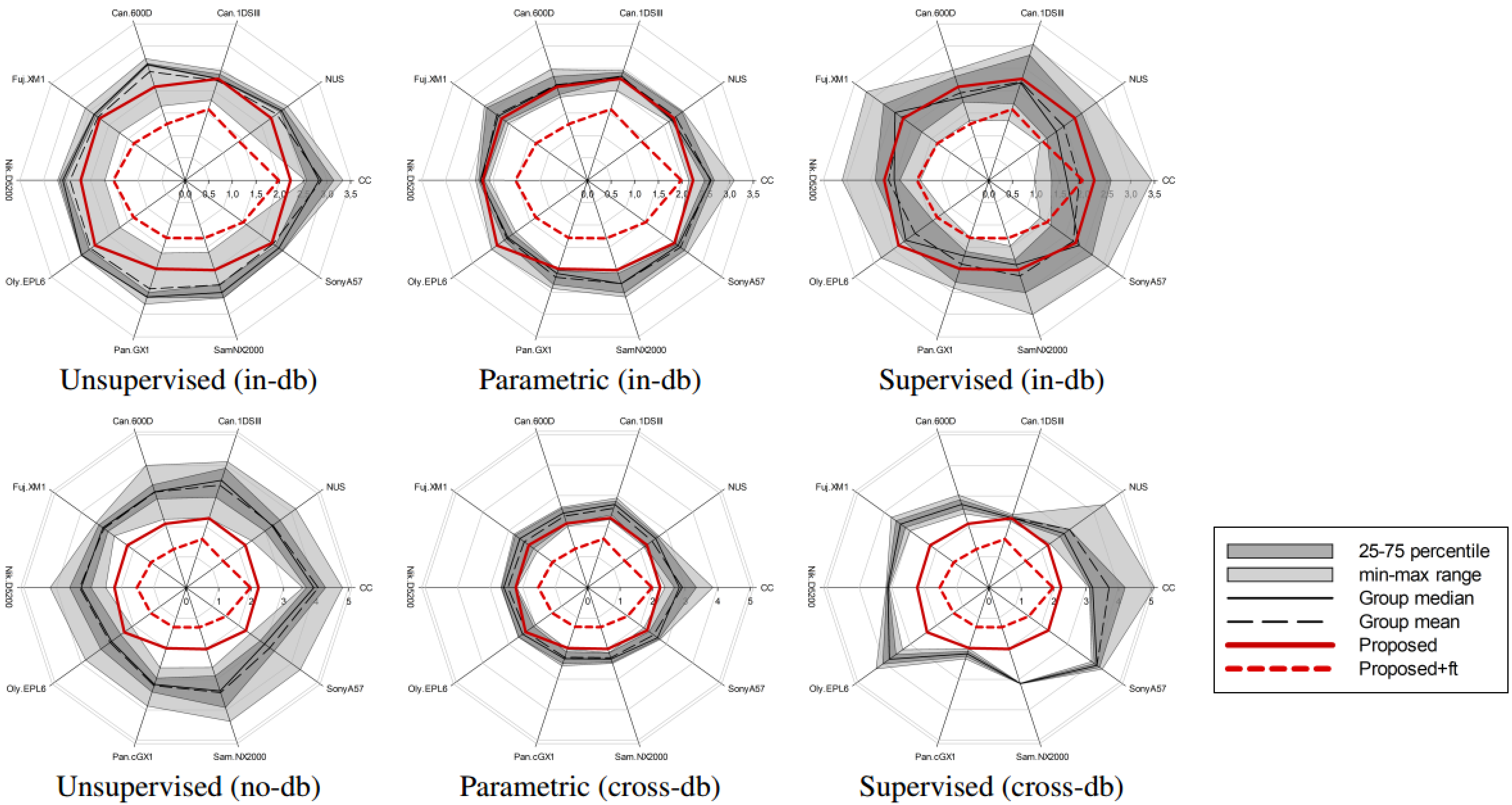


Figure S1. Visual summary of the median errors of the proposed method, with and without fine tuning, on CC, NUS and individual NUS cameras. The method is compared with six groups of algorithms. For each group are drawn the best and worst median error, the interquartile range, the median ad the mean.

Color Constancy by Rewighting Image Feature Maps (2018 Dec)

- *代码：*
  - *https://github.com/QiuJueqin/Reweight-CC*
- *论文：*
  - *https://arxiv.org/pdf/1806.09248.pdf*
- *测试环境：*
  - *keras + TensorFlow-GPU + python3.7*

Qiu Jueqin

COLOR SCIENCE · COMPUTATIONAL PHOTOGRAPHY · COMPUTER VISION

Zheda Rd., Xihu Distr., Hangzhou, 310027

(+86) 155-5802-2825 | jqx1991@gmail.com | QiuJueqin

Summary

I am currently pursuing my Ph.D degree in Color & Imaging Lab of Zhejiang University, under the supervision of Prof. Hai-song Xu. My research interests include color constancy, imaging simulation, and human-vision-system-based color reproduction. I have been cooperating with corporations as the technical lead during my postgraduate period, and are highly familiar with the digital image signal processing (ISP) pipeline in mobile platforms.

Education

Sep 2014 – Jun 2019	PhD Candidate, College of Optic. Sci. and Engr.	*Zhejiang University, China*
Sep 2010 – Jun 2014	B.S., School of Instr. Sci. and Opto-Electr. Engr.	*Beihang University, China*

2.4 Reweight-CC 数据集

- **Single datasets**
  - Color Checker RECommended dataset
  - NUS-8 camera dataset
- **MultiCam dataset**
  - Contain **Color Checker RECommended dataset**, **NUS-8 camera dataset** and **Cube dataset**
  - Over 3,000 full-resolution images from 11 camera models:
    - **Canon 1D**
    - **Canon 1Ds MkIII**
    - **Canon 5D**
    - **Canon 550D**
    - **Canon 600D**
    - **Fujifilm XM1**
    - **Nikon D5200**
    - **Olympus EPL6**
    - **Panasonic GX1**
    - **Samsung NX2000**
    - **Sony A57**

蒸馏得到所有的消色像素点：

$$x := \begin{cases} achromatic & \text{if } |(\boxed{u_x} - u_0) - (\boxed{v_x} - v_0)| \le T \text{ and } |(u_x - \boxed{u_0}) + (v_x - \boxed{v_0})| \le T \\ chromatic & \text{otherwise}, \end{cases}$$

$$[u_x, v_x] = [r_x/(r_x + g_x + b_x), g_x/(r_x + g_x + b_x)]$$

待确定像素

消色点

$$x := \begin{cases} achromatic & \text{if } \boxed{\mathcal{W}(u_x, v_x)} > 0, \\ chromatic & \text{otherwise}, \end{cases}$$

$$\mathcal{W}(u_x, v_x) = ReLU\,\boxed{(min\,(\mathbf{A}[u_x, v_x]^\intercal + \mathbf{b}))}$$

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & 1 \\ -1 & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} T - u_0 + v_0 \\ T + u_0 - v_0 \\ T - u_0 - v_0 \\ T + u_0 + v_0 \end{bmatrix}$$

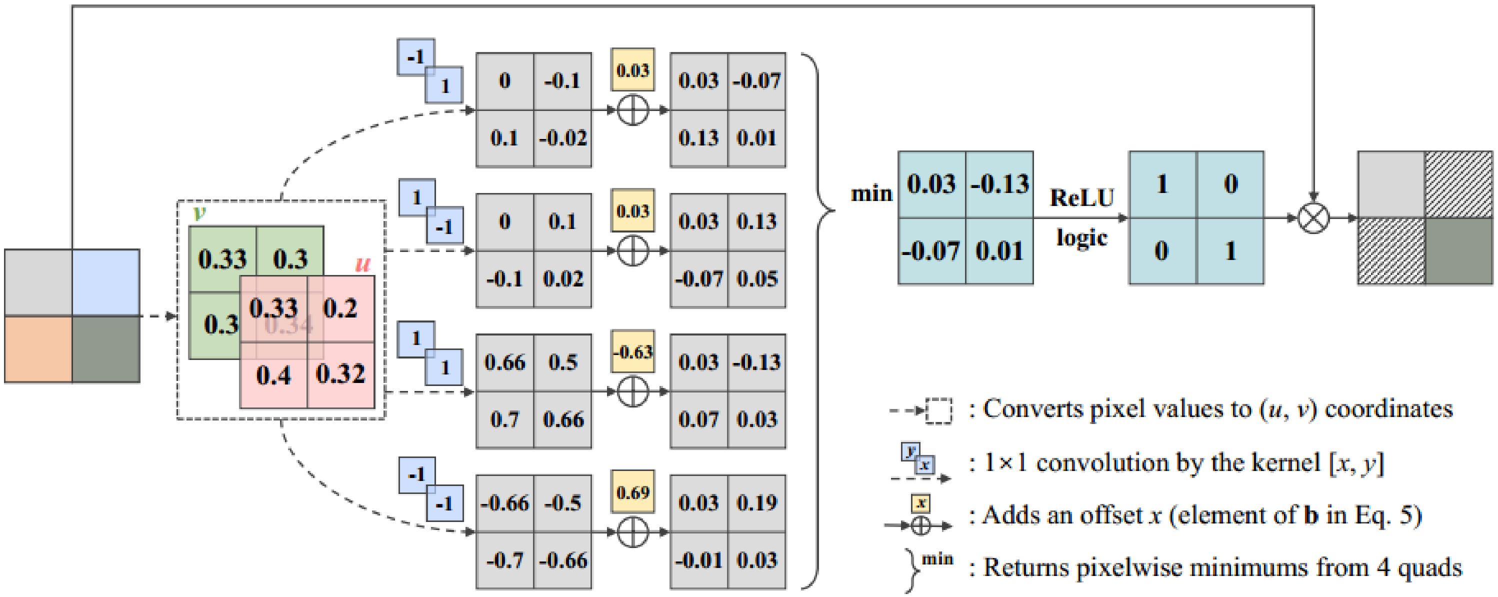
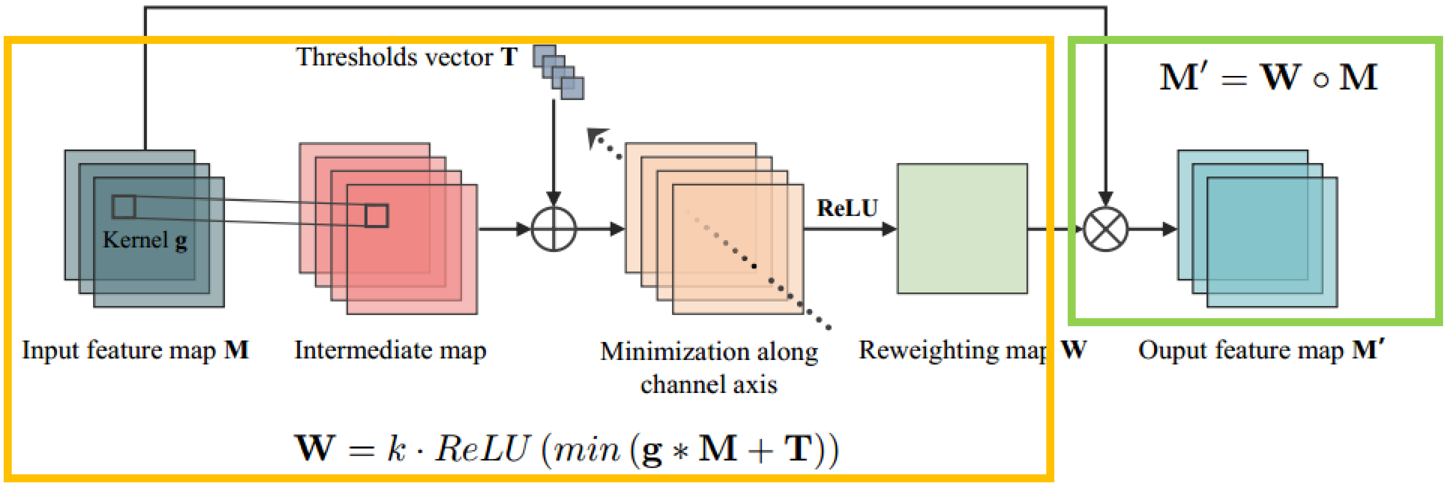


Figure 1: Visualization of near-achromatic pixels distillation by the function $\mathcal{W}$ in Eq. (6). Pixels with chromaticity coordinates (u, v) located within a small region (controlled by the threshold T) centering at the neutral point (u_0, v_0) will be considered near-achromatic. Without loss of generality, in this example we choose $(u_0, v_0) = (0.33, 0.33)$ and $T = 0.03$.

ReWU

- 学习如何在feature maps上加上“constraints” 和“thresholds”
- 加强有用的像素，并抑制对光照估计无用的像素

2.4 Reweight-CC Feature Map Reweight Unit (ReWU)



Here $\mathbf{g}$ is a $1 \times 1 \times C \times K$ tensor, where C is the number of channels in the input feature map, K is the number of kernels, which can also be interpreted as *the number of constraints* imposed on the feature map. $\mathbf{T}$ is a $K \times 1$ thresholds vector determining the tolerance for the pixels being away from the "hotspot" in the C-dimensional space. Symbol $*$ denotes the spatial convolution, $min(\cdot)$ herein is the minimization operation *along the channel axis*, and k is a trainable scaling parameter that allows to adjust the activation of the reweighting map if needed. The structure of the ReWU is illustrated in Fig. 2.

2.4 Reweight-CC 网络结构

Table 1: Architecture details of incarnate networks with 1, 2, and 3 hierarchical levels. The confidence estimation branches are not included but they share the same architectures as the illuminant estimation branches, only with the final 3-neuron layer replaced by a 1-neuron one.

	1-Hierarchy			2-Hierarchy			3-Hierarchy		
layer name	output size	Conv / FC	ReWU	output size	Conv / FC	ReWU	output size	Conv / FC	ReWU
input (hrchy_0)	224×224	–	1×1, 16	224×224	–	1×1, 16	224×224	–	1×1, 16
hrchy_1	112×112	3×3, 32 stride 2	1×1, 32	112×112	3×3, 32 stride 2	1×1, 32	112×112	3×3, 32 stride 2	1×1, 32
hrchy_2	–	–	–	112×112	3×3, 32	1×1, 32	112×112	3×3, 32	1×1, 32
hrchy_3				–	–	–	112×112	3×3, 64	1×1, 64
concat	35×1	–		67×1	–		131×1	–	
fc_1	64×1	1, 64	–	128×1	1, 128	–	256×1	1, 256	–
fc_2	32×1	1, 32		64×1	1, 64		128×1	1, 128	
fc_3	16×1	1, 16		32×1	1, 32		64×1	1, 64	
estimate	3×1	1, 3		3×1	1, 3		3×1	1, 3	

2.4 Reweight-CC 网络结构

Table 1: Architecture details of incarnate networks with 1, 2, and 3 hierarchical levels. The confidence estimation branches are not included but they share the same architectures as the illuminant estimation branches, only with the final 3-neuron layer replaced by a 1-neuron one.

	1-Hierarchy			2-Hierarchy			3-Hierarchy		
layer name	output size	Conv / FC	ReWU	output size	Conv / FC	ReWU	output size	Conv / FC	ReWU
input (hrchy_0)	224×224	–	1×1, 16	224×224	–	1×1, 16	224×224	–	1×1, 16
hrchy_1	112×112	3×3, 32 stride 2	1×1, 32	112×112	3×3, 32 stride 2	1×1, 32	112×112	3×3, 32 stride 2	1×1, 32
hrchy_2	–	–	–	112×112	3×3, 32	1×1, 32	112×112	3×3, 32	1×1, 32
hrchy_3				–	–	–	112×112	3×3, 64	1×1, 64
concat	35×1	–		67×1	–		131×1	–	
fc_1	64×1	1, 64	–	128×1	1, 128	–	256×1	1, 256	–
fc_2	32×1	1, 32		64×1	1, 64		128×1	1, 128	
fc_3	16×1	1, 16		32×1	1, 32		64×1	1, 64	
estimate	3×1	1, 3		3×1	1, 3		3×1	1, 3	

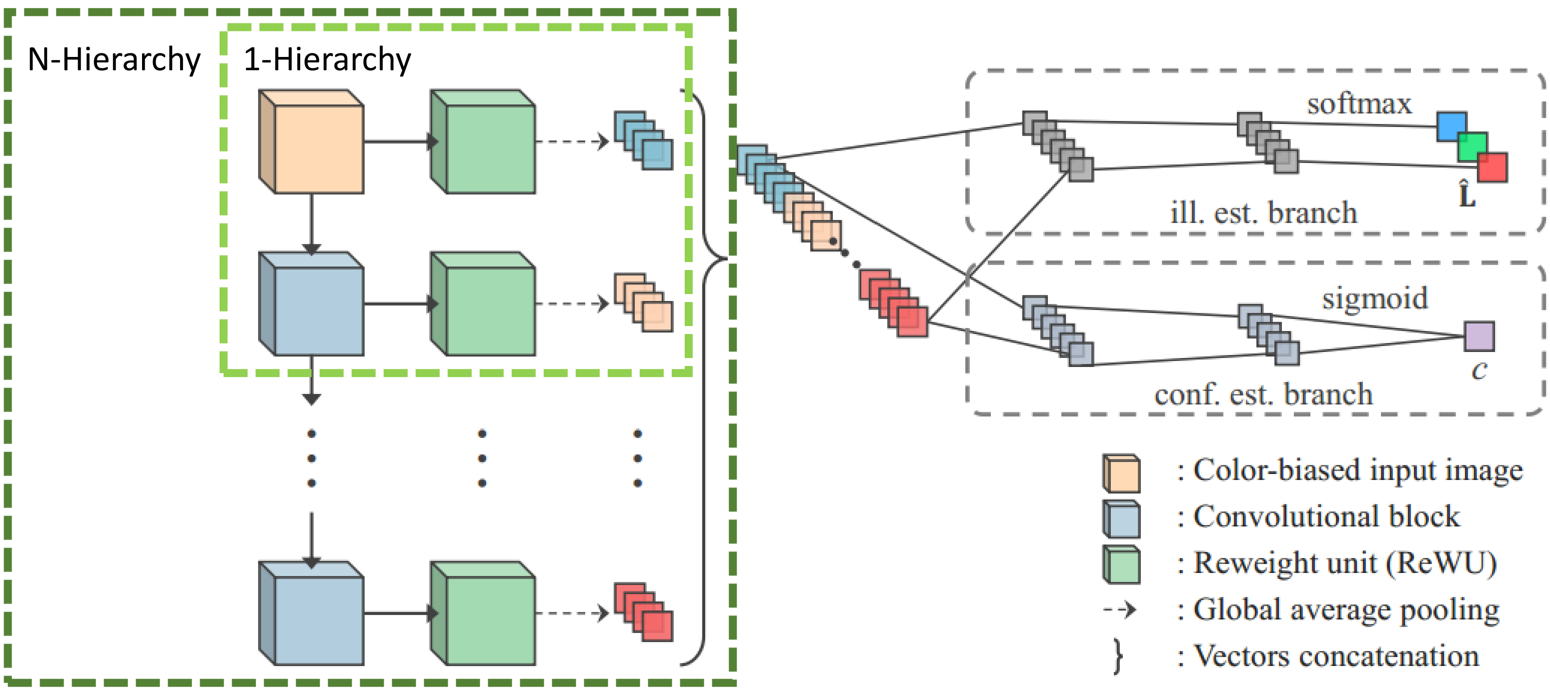
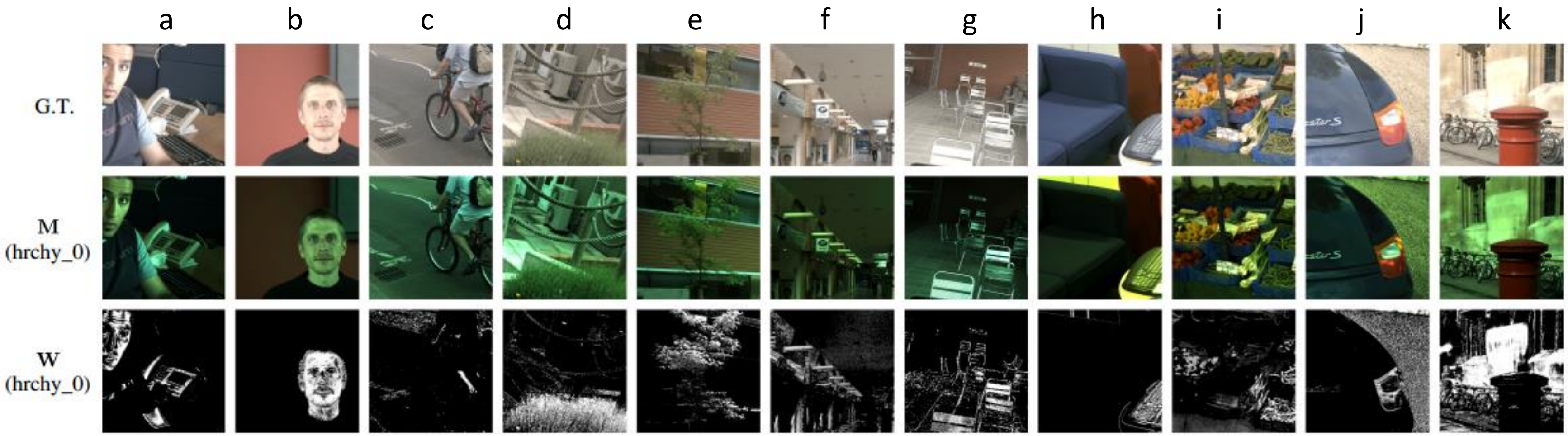


Figure 3: The architecture of the proposed regression network. The *illuminant estimation branch* produces a triplet that represents the estimated illuminant color, and the *confidence estimation branch* produces a single score that represents the confidence/uncertainty of the network about its color estimate.

2.4 Reweight-CC 可视化结果



增强的像素点

- 中等亮度的消色像素（a和f）
- 镜面反射的表面（g和h）
- 高频出现的色彩元素
  - 肤色（a-c）
  - 绿色的植物（d，e和i）

抑制的像素点

- 墙上的红漆（b）
- 橙子(i)
- 海军蓝汽车(j)
- 猩红色的油漆桶 (k)

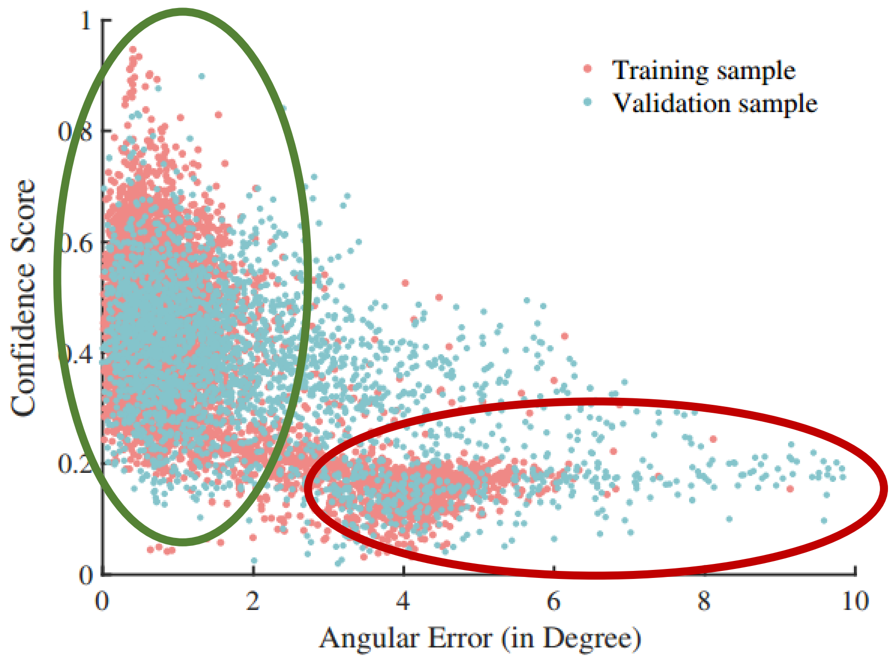
2.4 Reweight-CC 可视化结果

- 从图像上采样12个方形sub-images区域
- Resize to 224 × 224
- 确定Color Checker的位置并**删除包含Color Checker的sub-images**
- 计算剩下的sub-images的local estimated color vector
- 计算整张图的global estimated color vector

$$\hat{\mathbf{L}}_{global} = \sum_{i}^{12} \frac{\hat{c}_i \cdot \hat{\mathbf{L}}_{local,i}}{\sum_i \hat{c}_i}$$

2.4 Reweight-CC 置信度

准确预测的图像的confidence分值高



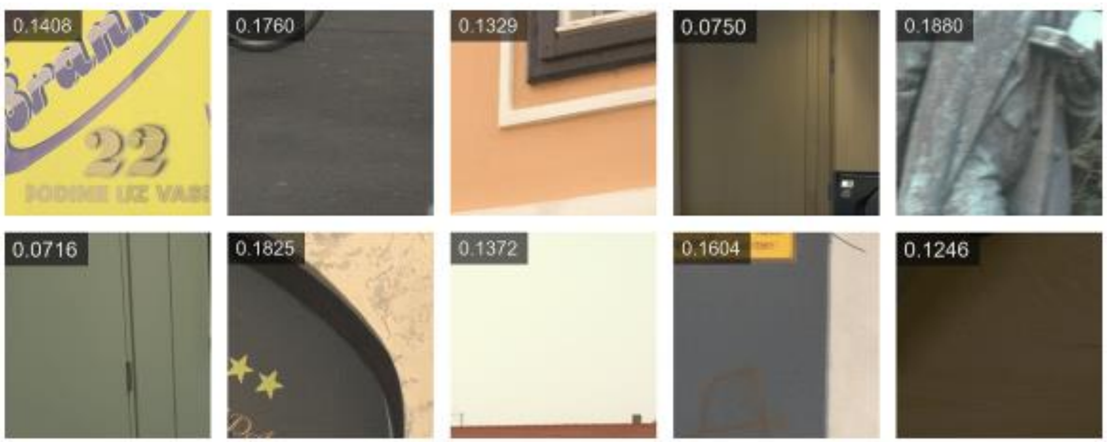
错误预测的图像的confidence分值低

Figure 6: The confidence scores produced by the confidence estimation branch versus angular errors of the illuminant color estimation. 3-Hierarchy architecture was used in this experiment, with $\lambda_0 = 8 \times 10^{-5}$ and $\beta = 0.6$. The model was trained on the MultiCam dataset, which contains 57,974 samples for training and 28,985 for validation after data augmentation and sub-image cropping. For clarity, only 10% samples were randomly picked up and plotted.

2.4 Reweight-CC 置信度

高置信度场景

包含经常出现的物体（草坪，中性面，砖屋顶等）的场景



低置信度场景

图案模糊的场景（少见的logo，单色表面等）

2.4 Reweight-CC 评测 参数量

Table 2: The number of model parameters and operations (multiply-adds) in one forward propagation, assuming the input image has fixed size of 224×224.

Method	#Param.	#Ops.
DS-Net[a] [51]	$\approx$17.3M	$\approx 6.0 \times 10^{10}$
Semantic CC [15]	$\approx$13.9M	$\approx 4.1 \times 10^{9}$
FC4 (AlexNet) [13]	$\approx$3.8M	$\approx 4.8 \times 10^{9}$
Deep Outdoor CC [52]	$\approx$3.7M	$\approx 2.3 \times 10^{9}$
Ours, 1-Hierarchy, w/o conf. est.	7.6K	5.9×10^{7}
Ours, 1-Hierarchy, with conf. est.	13.0K	5.9×10^{7}
Ours, 2-Hierarchy, w/o conf. est.	32.7K	1.9×10^{8}
Ours, 2-Hierarchy, with conf. est.	52.7K	1.9×10^{8}
Ours, 3-Hierarchy, w/o conf. est.	112.6K	4.7×10^{8}
Ours, 3-Hierarchy, with conf. est.	189.4K	4.7×10^{8}

[a] DS-Net accepts 44×44 patches in the original paper. Enlarging the size of input will exponentially increase the amounts of operations for DS-Net. Nonetheless, to keep the comparison fair we fixed the sizes of inputs for all methods.

best

second

Method	ColorChecker Dataset Mean	Med.	Tri.	Best 25%	Worst 25%	NUS-8 Camera Dataset Mean	Med.	Tri.	Best 25%	Worst 25%
White-Patch [1]	7.55	5.68	6.35	1.45	16.12	10.62	10.58	10.49	1.86	19.45
Gray-World [4]	6.36	6.28	6.28	2.33	10.58	7.70	6.71	6.90	2.51	14.05
1st-order Gray-Edge [6]	5.33	4.52	4.73	1.86	10.03	4.14	3.20	3.39	0.90	9.00
Shades-of-Gray [53]	4.93	4.01	4.23	1.14	10.20	3.67	2.73	2.91	0.82	8.21
Bayesian [54]	4.82	3.46	3.88	1.26	10.49	3.40	2.57	2.73	0.77	7.41
Spatio-spectral Statistics [55]	3.59	2.96	3.10	0.95	7.61	3.20	2.22	2.43	0.72	7.36
Pixels-based Gamut [9]	4.20	2.33	2.91	0.50	10.72	2.96	2.33	2.47	0.80	6.18
Bianco CNN [12]	2.63	1.98	2.10	0.72	3.90	2.92	2.04	2.24	0.62	6.61
Cheng et al. 2015 [11]	2.42	1.65	1.75	0.38	5.87	2.24	1.46	1.68	0.48	6.08
CCC [56]	1.95	1.22	1.38	0.35	4.76	2.38	1.48	1.69	0.45	5.85
DS-Net [51]	1.90	1.12	1.33	0.31	4.84	2.23	1.57	1.72	0.47	5.15
FC4 (SqueezeNet) [13]	1.65	1.18	1.27	0.38	3.78	2.18	1.48	1.64	0.46	5.03
FFCC (model Q) [24]	2.01	1.13	1.38	0.30	5.14	2.06	1.39	1.53	0.39	4.80
Ours, 1-Hierarchy, w/o conf. est.	2.41	2.02	2.00	0.59	5.10	2.84	1.92	2.04	0.80	5.82
Ours, 1-Hierarchy, with conf. est.	2.32	1.96	1.96	0.60	4.65	2.84	1.88	1.90	0.75	5.39
Ours, 2-Hierarchy, w/o conf. est.	2.18	1.73	1.82	0.53	4.70	2.32	1.64	1.67	0.46	5.44
Ours, 2-Hierarchy, with conf. est.	2.10	1.68	1.77	0.49	4.32	2.27	1.61	1.63	0.48	5.16
Ours, 3-Hierarchy, w/o conf. est.	1.98	1.38	1.52	0.51	4.52	2.18	1.59	1.74	0.48	5.35
Ours, 3-Hierarchy, with conf. est.	1.85	1.31	1.37	0.44	4.14	2.20	1.53	1.60	0.44	5.07

2.4 Reweight-CC 评测 参数量

best

second

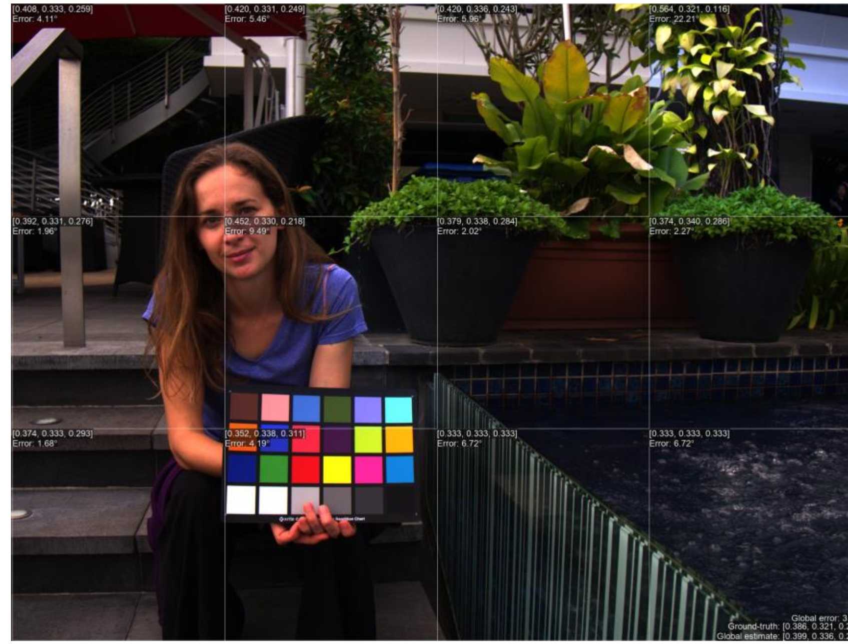
MultiCam Dataset

Method	Mean	Med.	Tri.	Best 25%	Worst 25%
Gray-World [4]	7.85	6.51	6.86	1.76	16.10
Pixel-based Gamut* [9]	6.12	4.57	4.99	1.32	13.45
Spatio-spectral Statistics* [55]	5.94	5.01	5.22	1.98	11.29
White-Patch [1]	5.78	4.37	4.73	1.31	12.64
1st-order Gray-Edge [6]	5.46	4.25	4.52	1.37	11.67
Shades-of-Gray [53]	5.39	4.19	4.48	1.13	11.68
Bayesian* [54]	5.30	4.10	4.42	1.25	12.14
2st-order Gray-Edge [6]	5.30	4.01	4.33	1.30	11.50
Bianco CNN* [12]	3.85	2.82	3.00	1.04	8.50
FFCC (model Q)* [24]	3.15	2.43	2.60	0.67	6.88
SqueezeNet-FC4* [13]	3.03	2.26	2.35	0.79	6.50
Ours, 1-Hierarchy, w/o conf. est.*	3.96	3.00	3.29	0.91	8.15
Ours, 1-Hierarchy, with conf. est.*	3.92	2.91	3.21	0.86	7.85
Ours, 2-Hierarchy, w/o conf. est.*	3.49	2.77	2.89	0.88	7.33
Ours, 2-Hierarchy, with conf. est.*	3.44	2.74	2.85	0.86	7.09
Ours, 3-Hierarchy, w/o conf. est.*	3.20	2.42	2.58	0.83	7.04
Ours, 3-Hierarchy, with conf. est.*	3.20	2.40	2.55	0.85	6.83

2.4 Reweight-CC 结果

Image from NUS-8, by OlympusEPL6

Original Image



Local Color Constancy Image

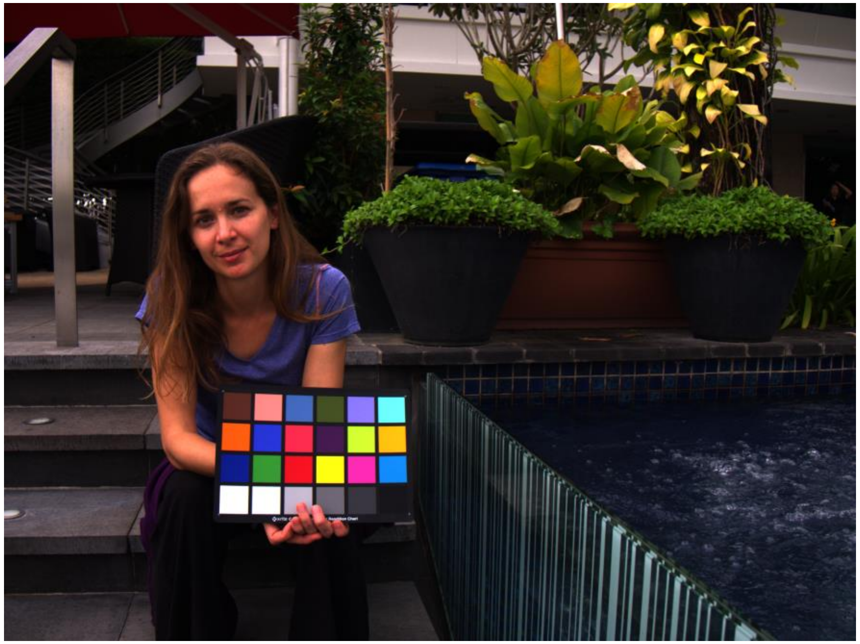
Trained by MultiCam dataset
Hierarchical level 3

Ground Truth: 0.386 0.321 0.293
Global Estimate: 0.399 0.336 0.265
Global Error: 3.30

2.4 Reweight-CC 结果

Image from NUS-8, by OlympusEPL6

Original Image



Global Color Constancy Image

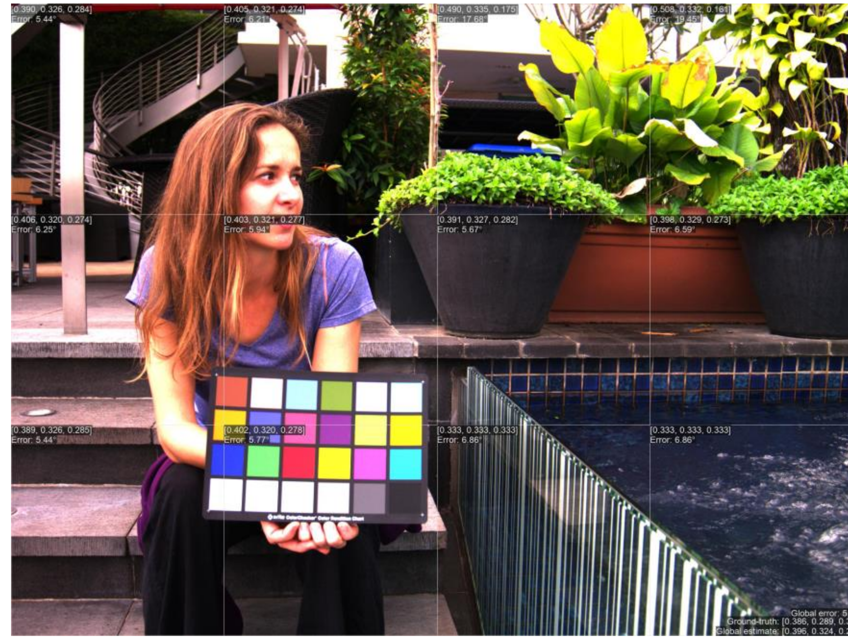
Trained by MultiCam dataset
Hierarchical level 3

Ground Truth: 0.386 0.321 0.293
Global Estimate: 0.399 0.336 0.265
Global Error: 3.30

2.4 Reweight-CC 结果

Image from NUS-8, by PanasonicGX1

Original Image



Local Color Constancy Image

Trained by MultiCam dataset
Hierarchical level 3

Ground Truth: 0.386 0.289 0.325
Global Estimate: 0.396 0.324 0.280
Global Error: 5.75

2.4 Reweight-CC 结果

Image from NUS-8, by PanasonicGX1

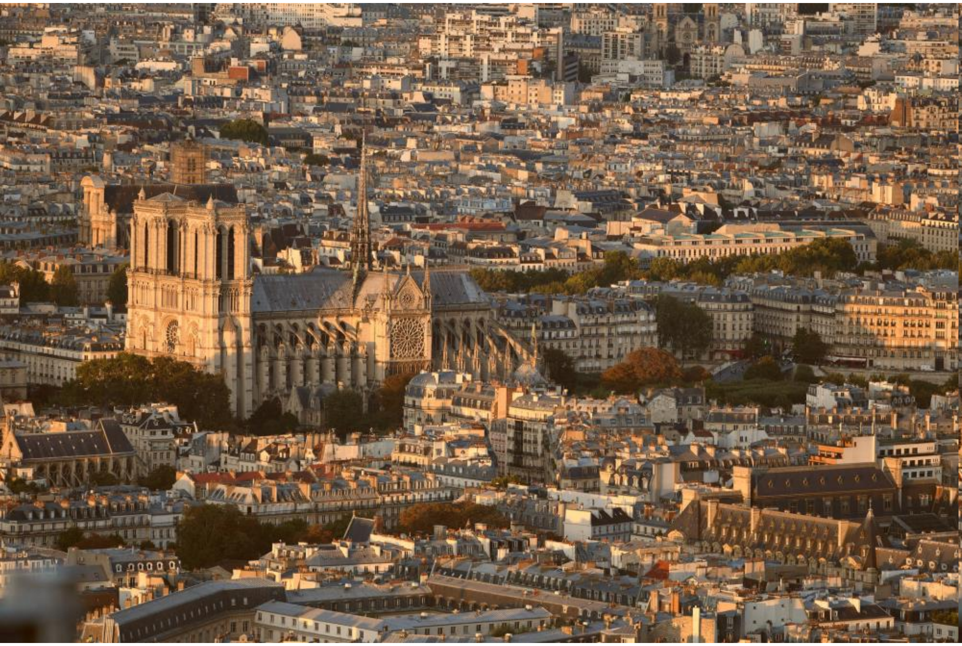
Original Image

Global Color Constancy Image

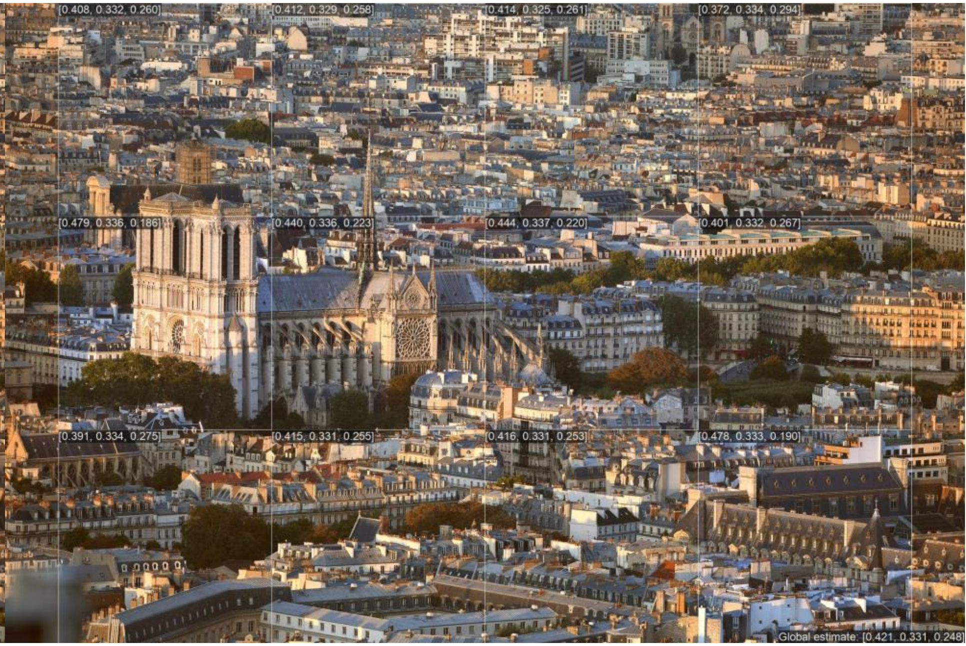
Trained by MultiCam dataset
Hierarchical level 3

Ground Truth: 0.386 0.321 0.293
Global Estimate: 0.399 0.336 0.265
Global Error: 3.30

2.4 Reweight-CC 结果



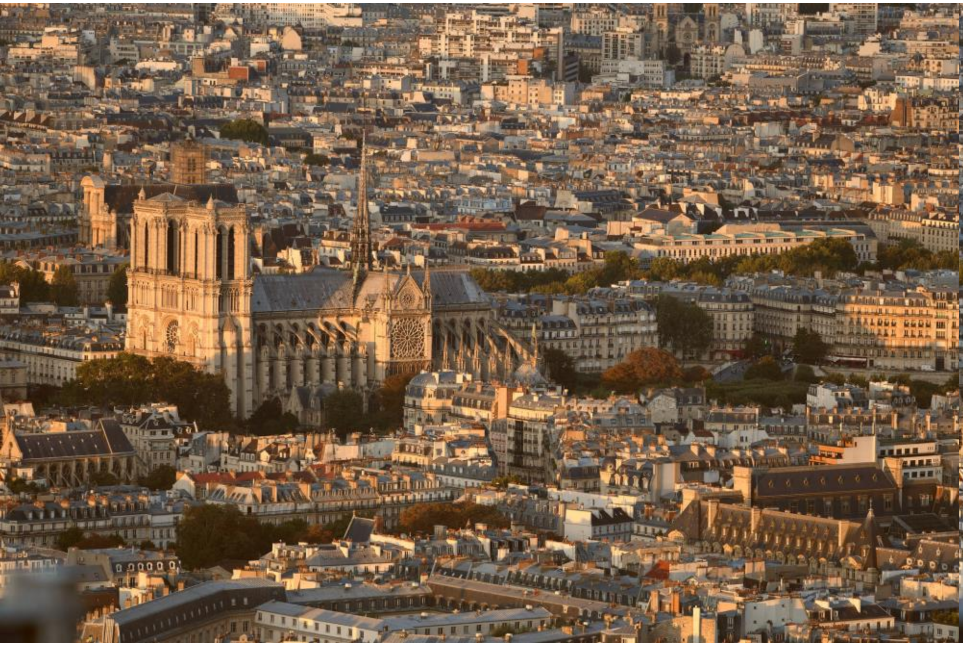
Original Image



Local Color Constancy Image

Trained by MultiCam dataset
Hierarchical level 3

2.4 Reweight-CC 结果

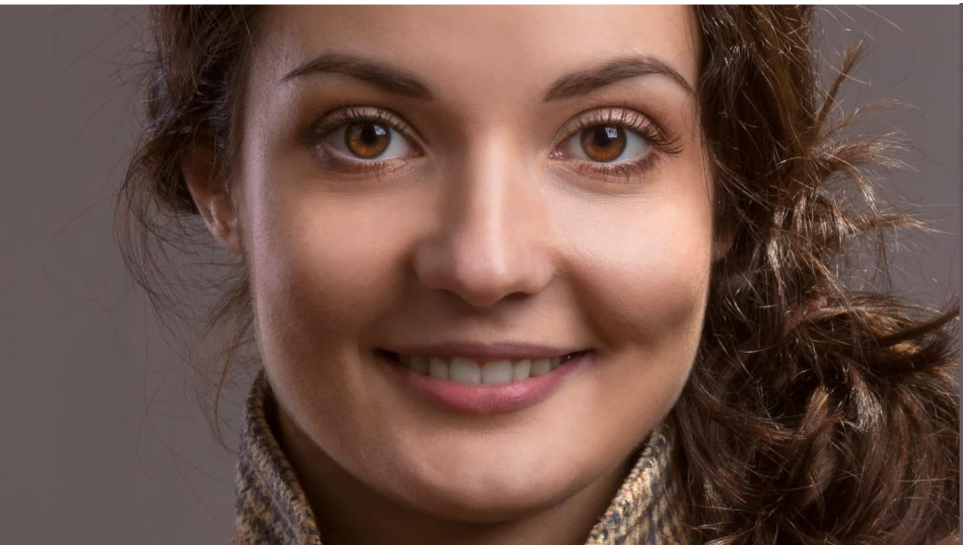


Original Image

Global Color Constancy Image

Trained by MultiCam dataset
Hierarchical level 3

2.4 Reweight-CC 结果

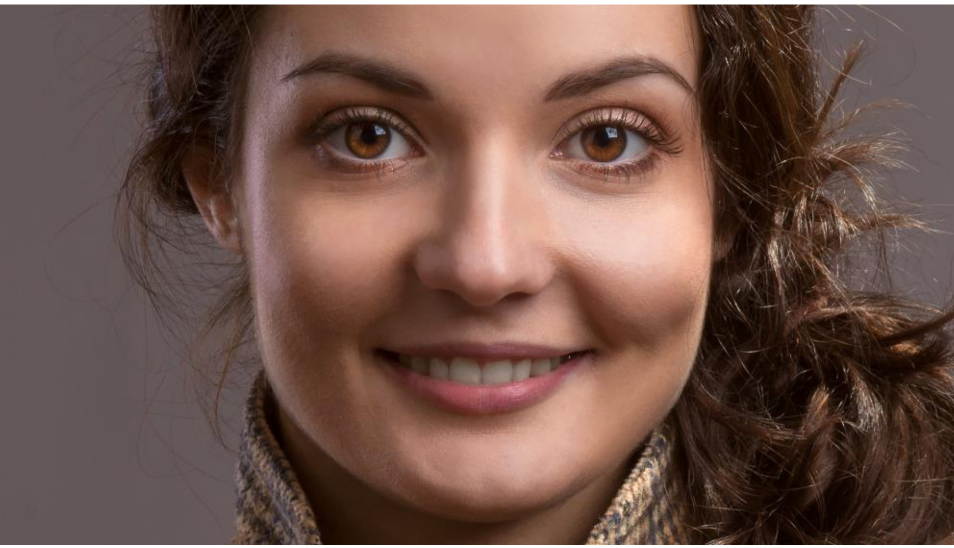


Original Image

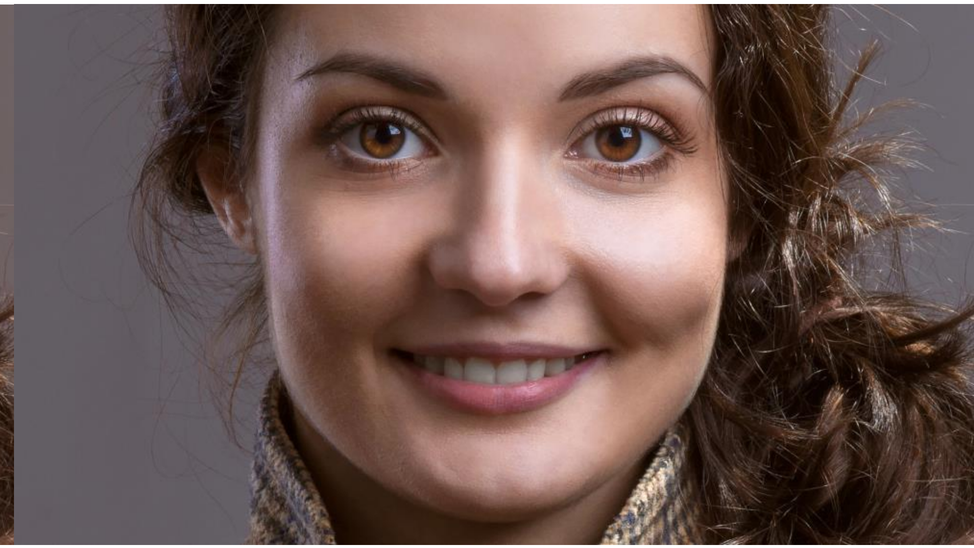
Local Color Constancy Image

Trained by MultiCam dataset
Hierarchical level 3

2.4 Reweight-CC 结果



Original Image



Global Color Constancy Image

Trained by MultiCam dataset
Hierarchical level 3