Automating the Design of Image Processing Pipelines for Novel Color Filter Arrays: Local, Linear, Learned (L^3) Method

Qiyuan Tian[a], Steven Lansel[b], Joyce Farrell[a], Brian Wandell[a,c]

[a]Department of Electrical Engineering, Stanford University

[b]Olympus Corporation of the Americas

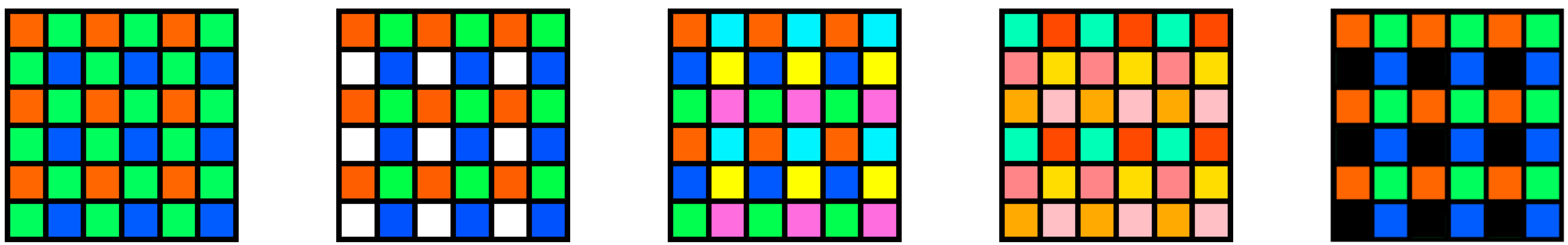
[c]Department of Psychology, Stanford University

2/4/2014

Motivation

- **Pixel density**: Now beyond most spatial resolution requirements

- **Opportunity**: Extra sensor pixels enable CFA designs that increase dynamic range, improve color and additional multispectral information



Bayer RGBW RGBCMY Medical RGBX

- **Challenge**: The design of image processing that effectively exploits the CFA properties is computationally difficult and labor-intensive

Automating Image Processing Design

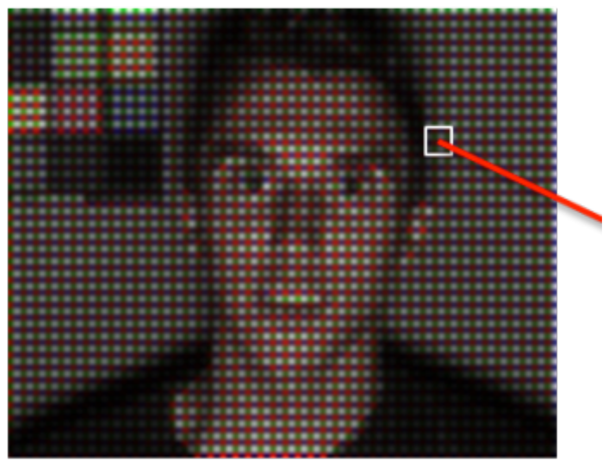
Radiometric spectral scenes

Scene target
statistic
(e.g., CIE XYZ)

Automating Image Processing Design

Radiometric spectral scenes

ISET optics/sensor simulation

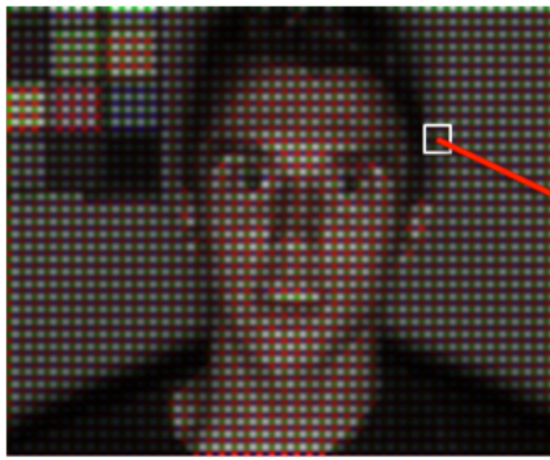


Scene target statistic (e.g., CIE XYZ)

Automating Image Processing Design

Radiometric spectral scenes

ISET optics/sensor simulation



Patch classification

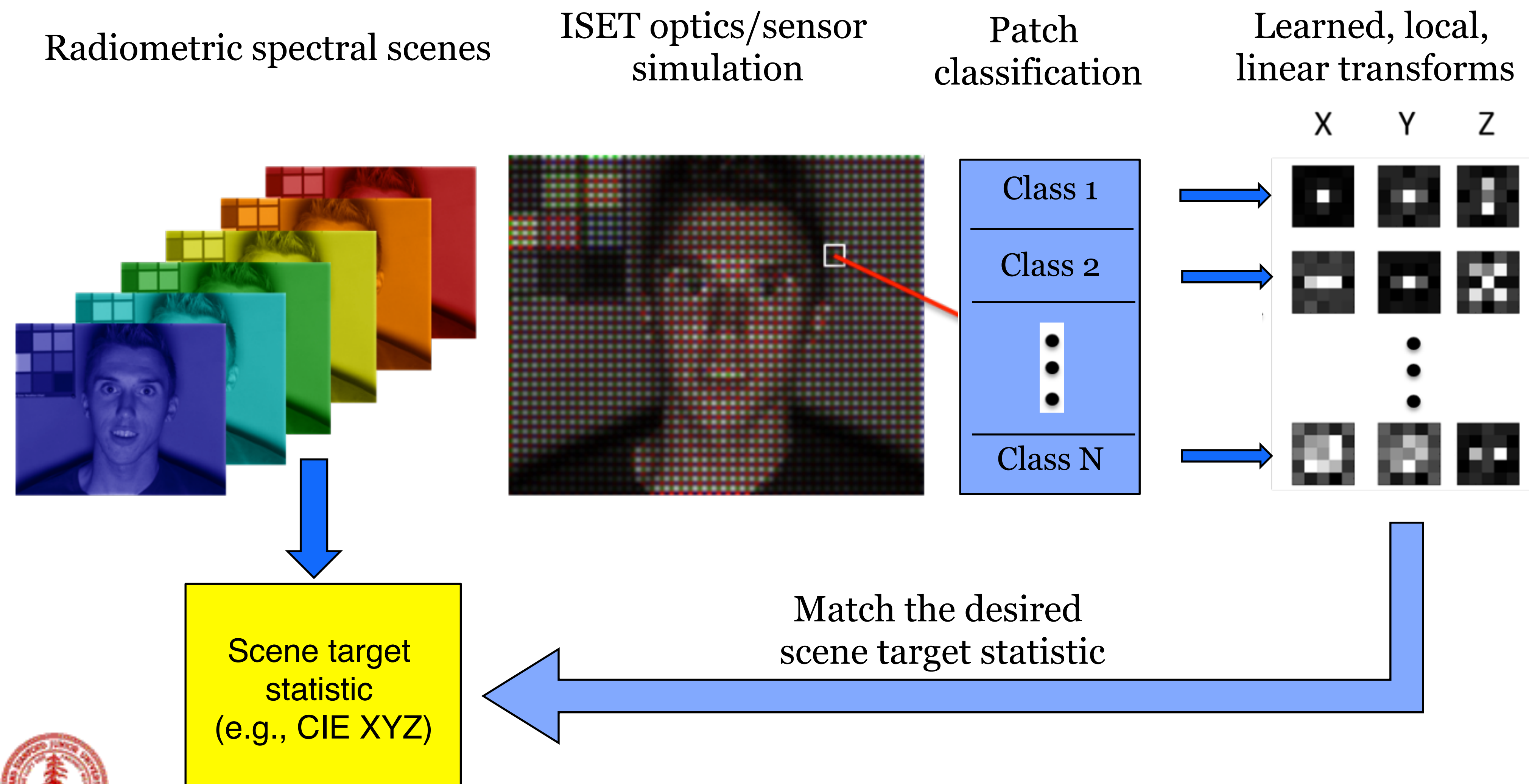
Class 1

Class 2

Class N

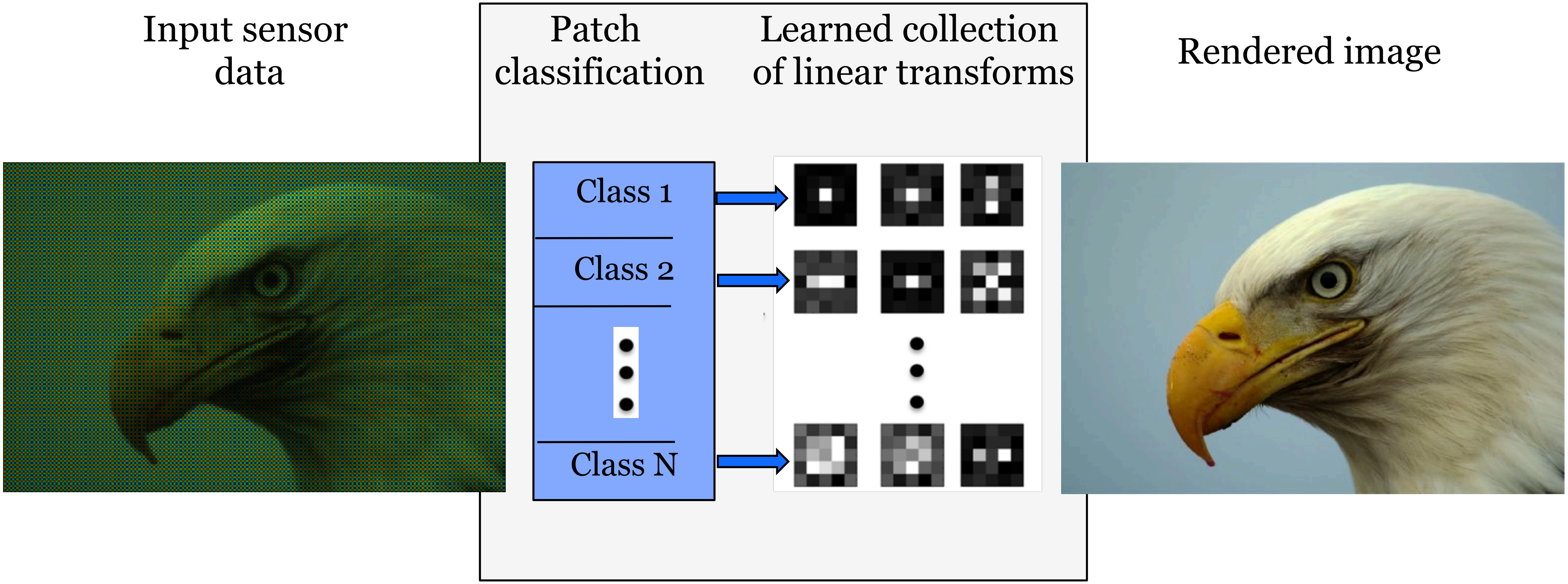
Scene target statistic (e.g., CIE XYZ)

Automating Image Processing Design



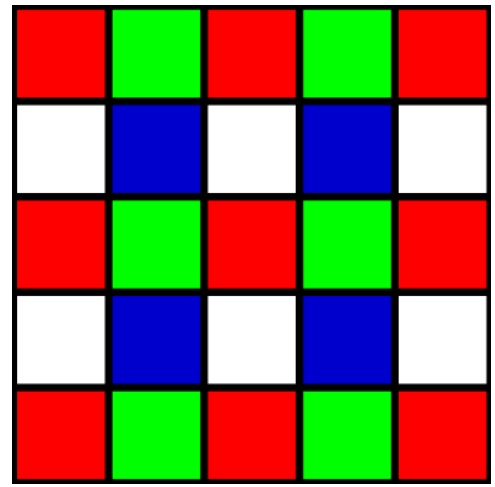
Automating Image Processing Design

- The image pipeline is the classification combined with the learned, local, linear transforms (hence, the name L^3)
- The pipeline performs demosaicking, sensor correction, denoising, and illuminant correction.



Example: Designing an RGBW pipeline

- Generating training data
- Camera simulation
- Patch classification system
- Learning the local filters



Radiometric scenes

ImagEval Consulting

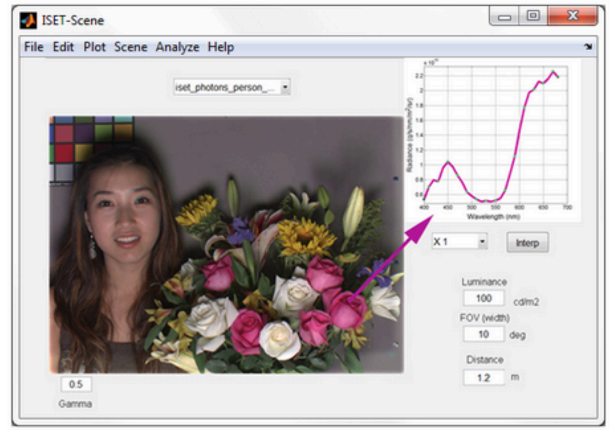
Home Products Downloads Documentation About Us Contact Us Blog

We help you evaluate image quali

Scene Database

High dynamic range spectral data are essential to accurately simulate the effects of optics and sensor. Using a very quality initial data set makes it possible to assess the properties of many other sensors whose dynamic range are limited compare to the scene data.

The ISET package includes images from a collection of high-dynamic range, multispectral (HDRS) scene. Each scene is represented as a multidimensional array describing the spectral radiance (photons/sec/nm/sr/m2) at each pixel in the sampled scene. The spectral radiance image data are assumed to arise from a single image plane at a specified distance from the optics. The spectral power of the scene illuminant is included with each data file. ISET provides software tools to calculate spectral reflectances and change scene illumination.



Multispectral and hyperspectral image data can be downloaded from this and other websites. ISET provides scripts to illustrate how to read and create scene radiance data from the spectral image databases listed below.

2004 Multispectral Images (400 – 700 nm): Faces, Fruit, Vegetables, Toys, Flowers, Macbeth Color Checker

2004 HDR Multispectral Images (400-700 nm): Backlight person, Person in Shadow, Stanford Memorial Church Interior

2008 Multispectral Images (400 – 700 nm): Faces

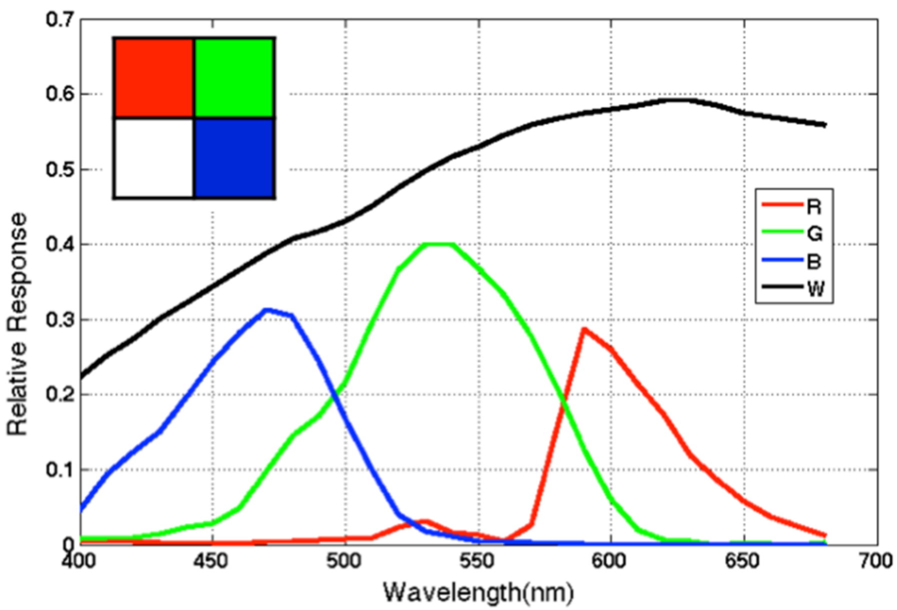
2009 Multispectral Images: (380 – 1068 nm): Fruit, Vegetables, Books, Calibration Targets

2012 Hyperspectral Scene Database : (415 – 915 nm): Faces, Landscape, Stanford Memorial Church Facade

Generating the Training Data

- **Simulate camera optics and sensor**
  Using Image Systems Engineering Toolbox (ISET) [1]
- **Training scenes**: Radiometric natural scenes [2]
- **Desired output values**: CIE XYZ values

Summary of ISET camera simulation



Pixel width/ height (μm)	2.2
Fill factor	0.45
Dark voltage (V/sec)	1×10^{-5}
Read noise (mV)	1.34
Dark signal nonuniformity (mV)	1.4
Photoreceptor nonuniformity (%)	2.2×10^{-3}
Conversion Gain (μV/e)	200
Voltage swing (V)	1.8
Well capacity (electrons)	9000
Analog gain	7.98

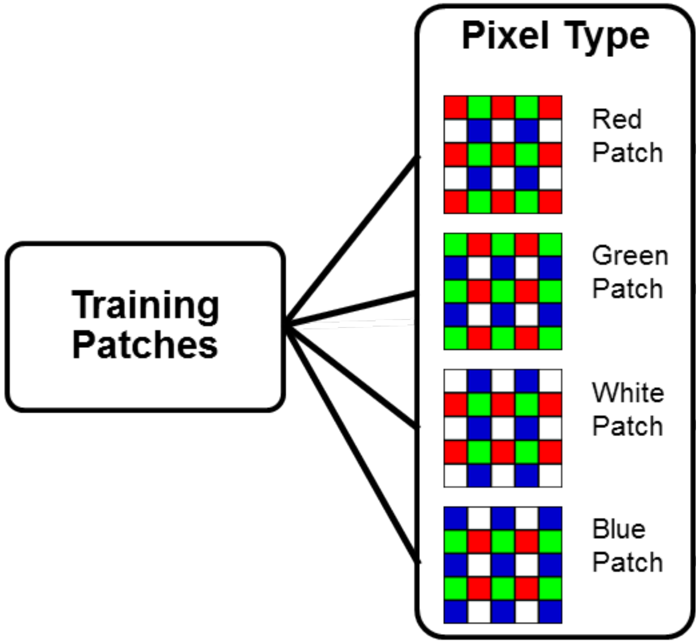
[3]

Defining the Patch Classification System

Training Patches

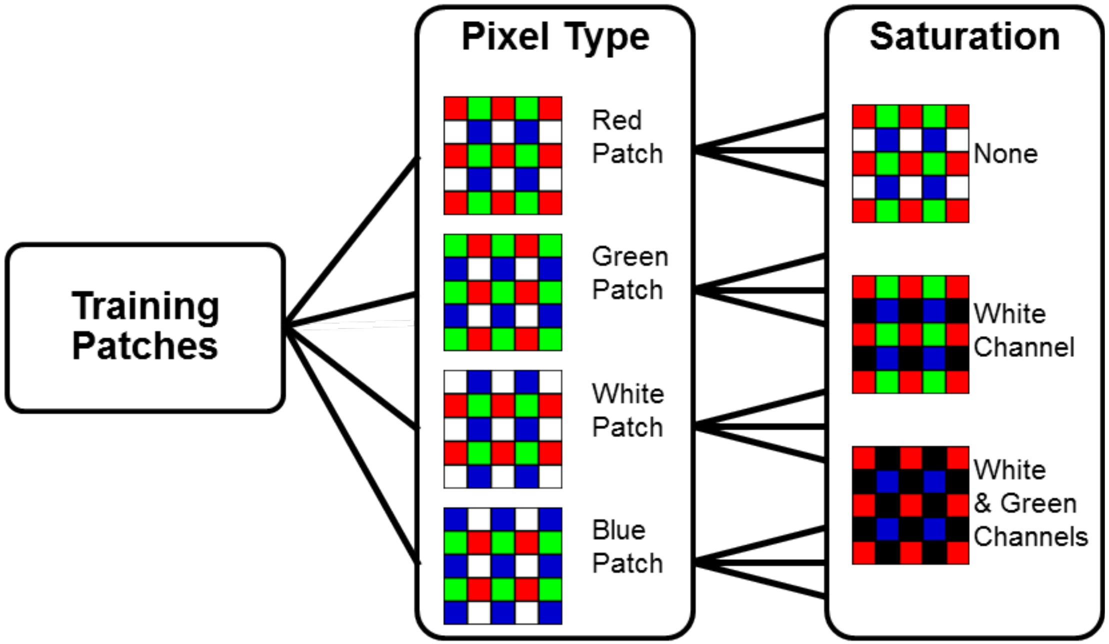
- Half local patch classes are not used due to impossible combination
- Each class has three operators to estimate CIE X, Y, Z values respectively
- ~240 local patch classes, ~720 operators

Defining the Patch Classification System



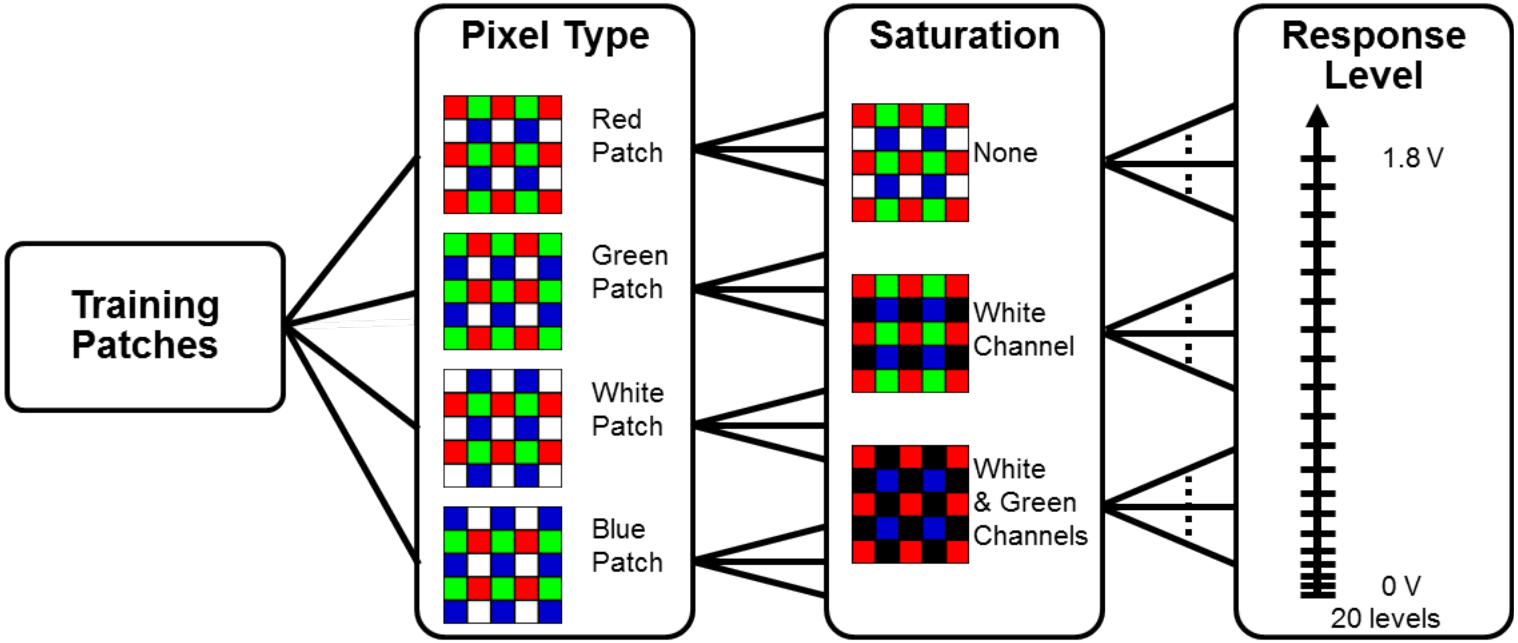
- Half local patch classes are not used due to impossible combination
- Each class has three operators to estimate CIE X, Y, Z values respectively
- ~240 local patch classes, ~720 operators

Defining the Patch Classification System



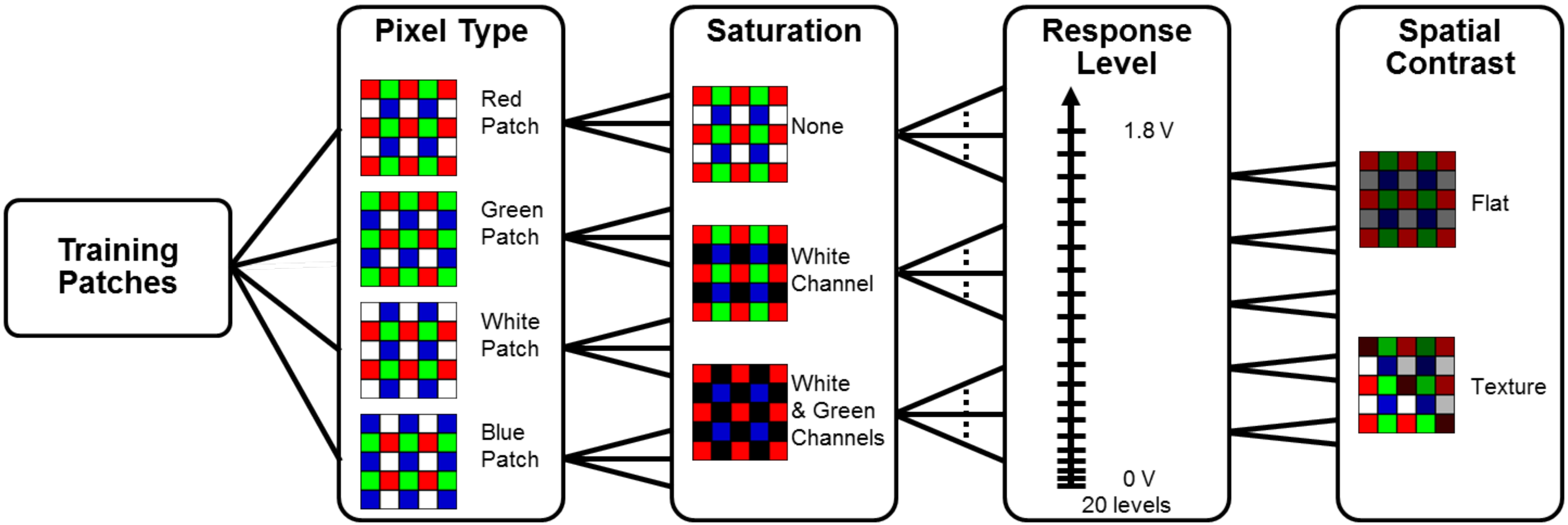
- Half local patch classes are not used due to impossible combination
- Each class has three operators to estimate CIE X, Y, Z values respectively
- ~240 local patch classes, ~720 operators

Defining the Patch Classification System



- Half local patch classes are not used due to impossible combination
- Each class has three operators to estimate CIE X, Y, Z values respectively
- ~240 local patch classes, ~720 operators

Defining the Patch Classification System



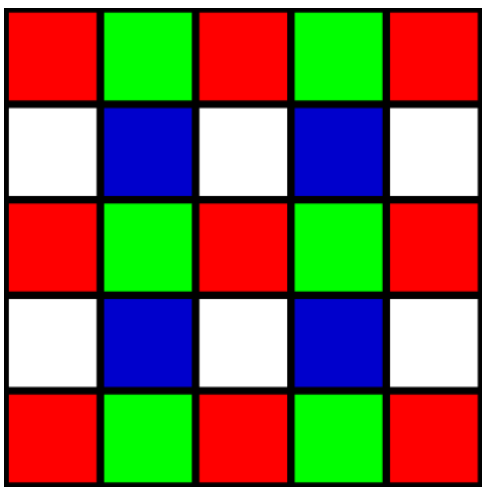
- Half local patch classes are not used due to impossible combination
- Each class has three operators to estimate CIE X, Y, Z values respectively
- ~240 local patch classes, ~720 operators

Learning the Local Linear Filter

Wiener filter example

In each channel minimize the sum of squared errors in between the desired output and the estimate; this filter accounting for noise covariance

Red-pixel centered patch



Filters that solve for X-channel

Dark

0.14	0.23	0.16	0.23	0.14
0.90	0.11	1.00	0.10	0.88
0.15	0.24	0.17	0.24	0.15
0.89	0.10	1.00	0.10	0.89
0.14	0.23	0.16	0.23	0.14

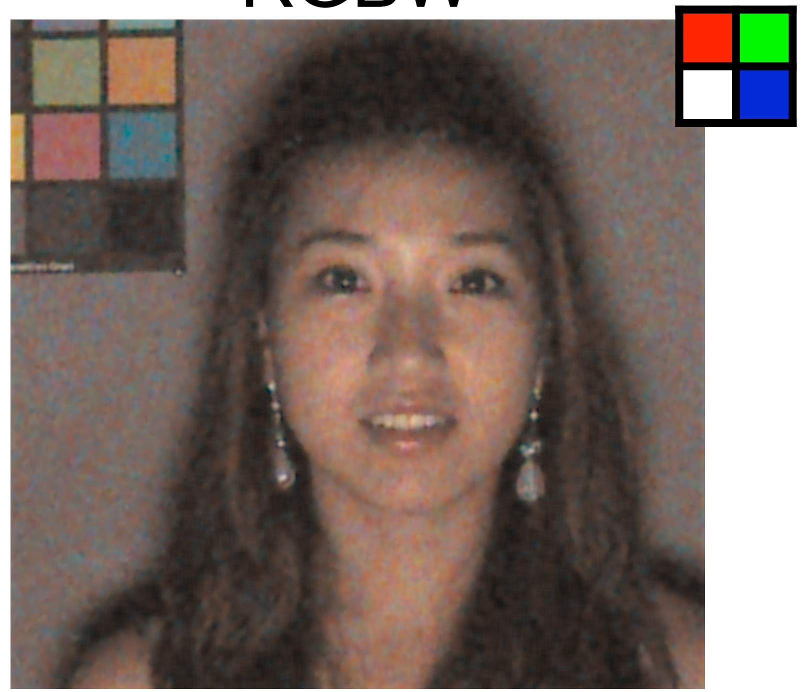
Bright

-0.10	-0.07	0.32	-0.05	-0.10
0.00	0.16	0.00	0.09	0.00
0.18	0.14	1.00	0.19	0.18
0.00	0.07	0.00	0.02	0.00
-0.10	-0.05	0.29	-0.01	-0.12

Preliminary L^3 Results in Dark Lighting Condition

Bayer

RGBW



<u>Simulation conditions</u>

Scene luminance: **1 cd/m²**

Exposure: **100 ms**

F-number: **f/4**

Improving the Wiener Filter
Accounting for Perception

- **Ridge regression**:

  weighted squared error = bias error^2 + α_i • variance error

- **Effects of increasing α_i>1**:

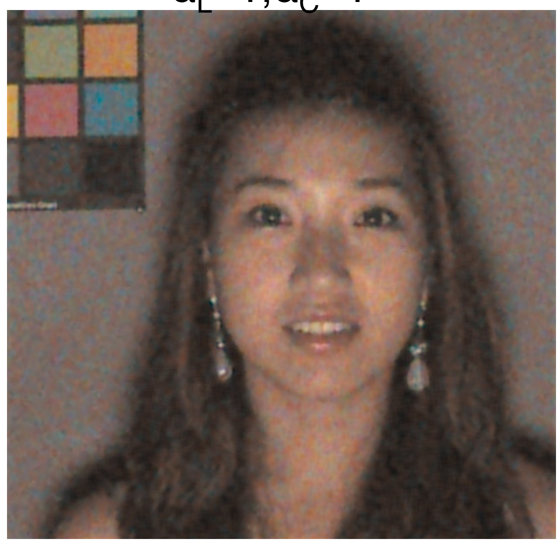
  decrease variance-> reduce noise

  increase bias-> spatial blur and desaturated color

- **Choose α_i in opponent color space**

α_L=16,α_C=1

α_L=1,α_C=1



α_L=1,α_C=16

Spatial blur ←

→ **Color desaturation**

Rendered Results

Bayer

RGBW

1 cd/m²   80 cd/m²

Flat regions:
$\alpha_L=16, \alpha_C=4$

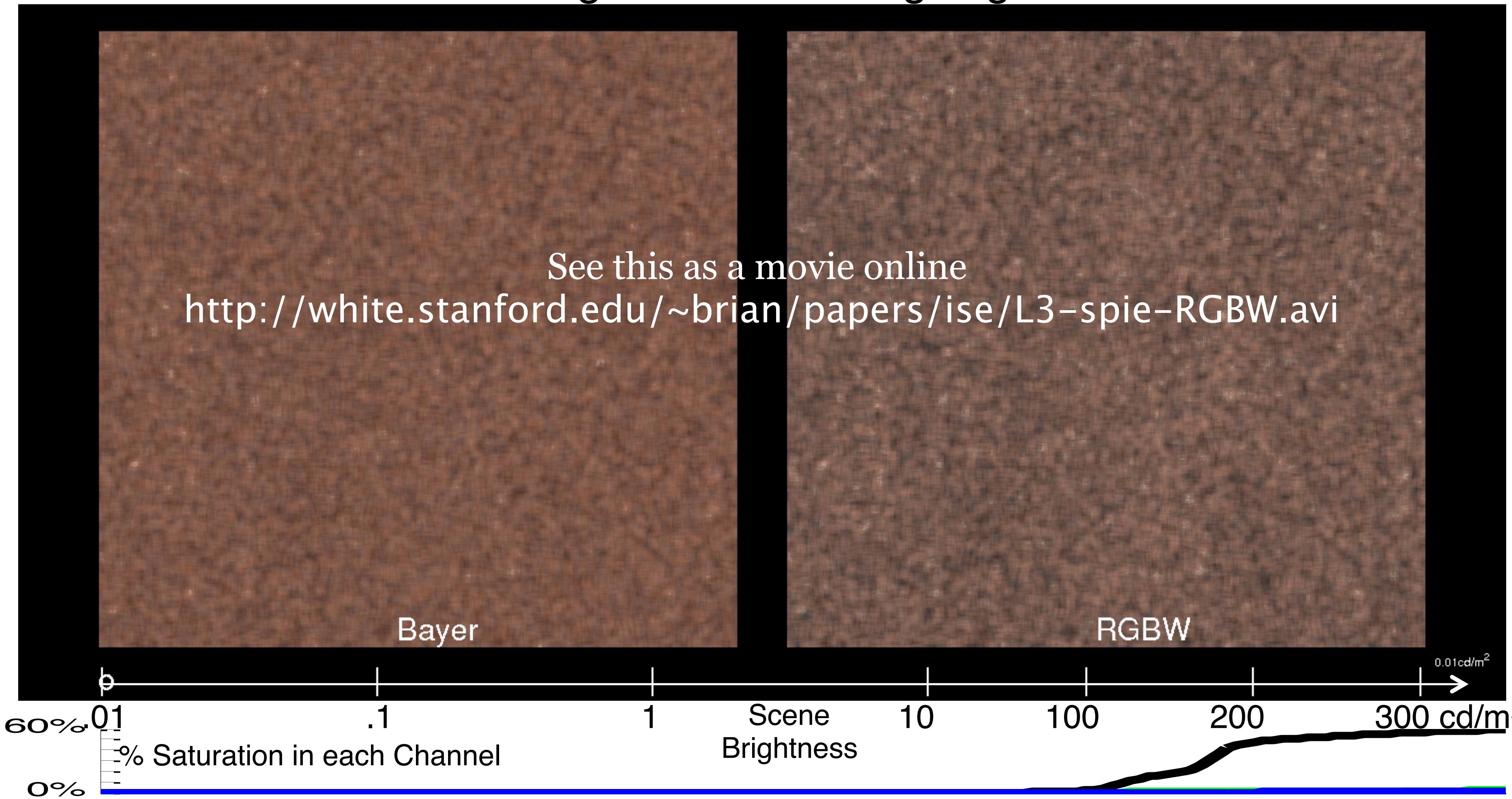
Texture regions:
$\alpha_L= 1, \alpha_C=4$

Exposure: **100ms**
F-number: **f/4**

Q. Tian et al. Automating the Design of Image Processing Pipelines for Novel Color Filter Arrays: Local, Linear, Learned (L^3) Method

RGBW for Low Light Photography

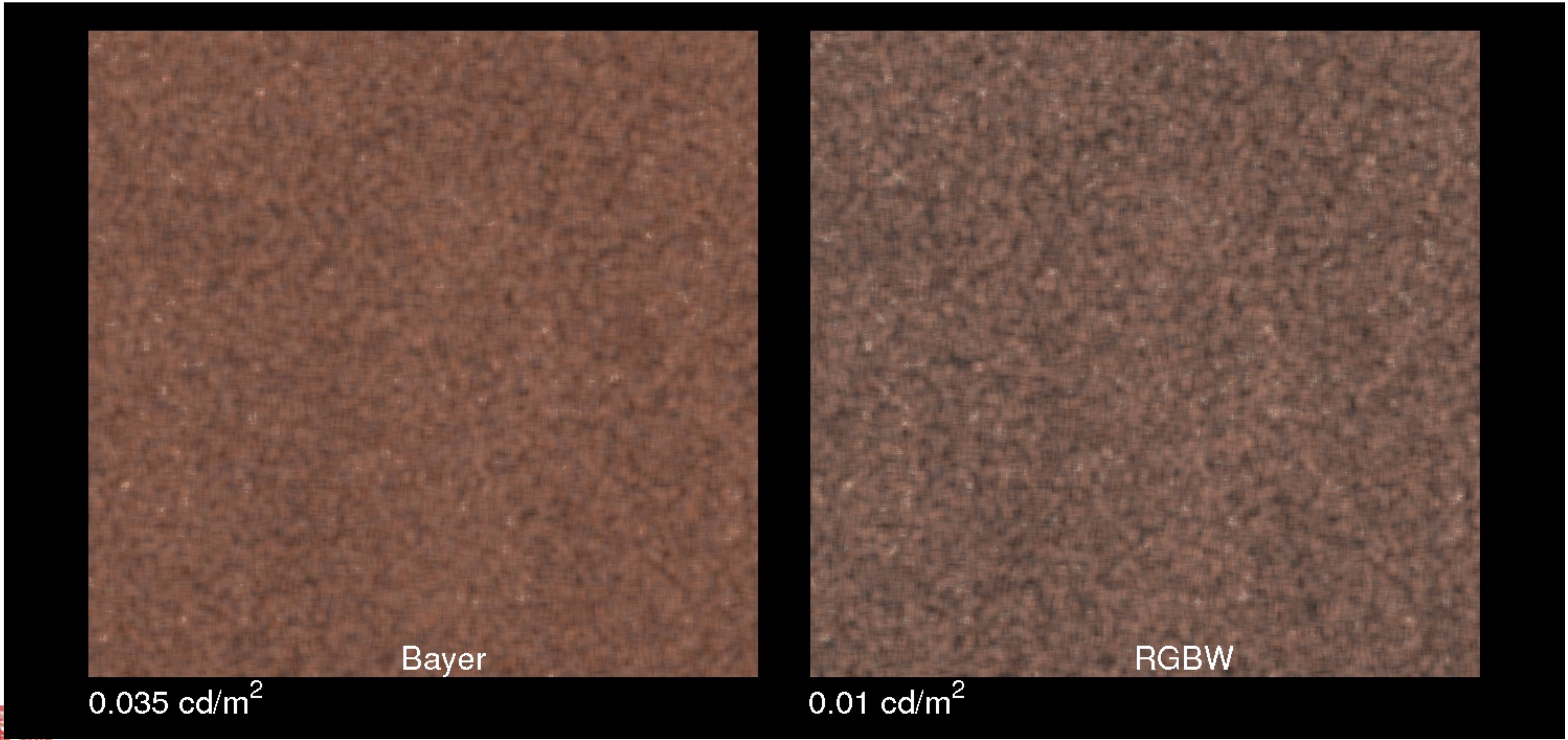
- RGBW: better in low light & same in high light



Q. Tian et al. Automating the Design of Image Processing Pipelines for Novel Color Filter Arrays: Local, Linear, Learned (L^3) Method

RGBW Giving ~2 f-stops of Light

- Bayer scene is 3.5× brighter than RGBW scene
- Result is approximately same image quality

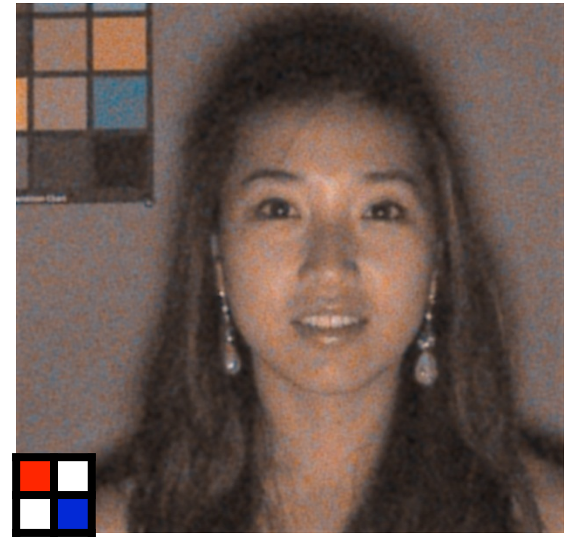


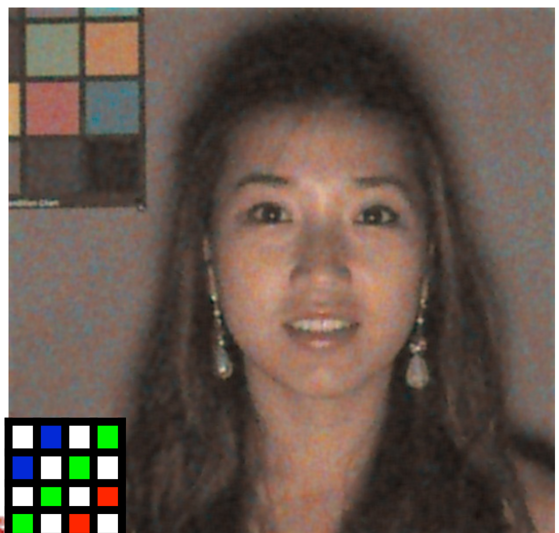
Automated Designs for Additional CFAs

Bayer

RGBW [2]

Aptina CLARITY+ [4]





Kodak [5]

Wang et al. [6]

Simulation conditions

Scene luminance: **1cd/m^2**
Exposure: **100 ms**
F-number: **f/4**

Conclusion

- Developed an automated method for designing an image processing pipeline for any CFA
  - Training data
  - Camera simulation
  - Learning local linear transformations (L^3)

- Illustrated the pipeline development for an RGBW CFA
  - Showed the efficacy of this design for low light photography
  - Measured two stop improvement at low light, similar performance in high light

- Created processing pipelines for several other CFAs

References

[1] Farrell, J., Catrysse, P., and Wandell, B., "The digital camera is an imaging system," in *Imaging Systems*. Optical Society of America, 2010, p. JTuA26.

[2] Parmar, M., Lansel, S., &Farrell, J. E. (2012), " An LED based lighting system for acquiring multispectral scenes", In Proc. SPIE, vol 8299.

[3] M. Parmar and B. A.Wandell, "Interleaved imaging: an imaging system design inspired by rod-cone vision," B. G. Rodricks and S. E. Susstrunk, Eds., vol. 7250. SPIE, January 18, 2009, p. 725008.

[4] Aptina's Clarity+TM Solution: An Aptina™ Technology White Paper.

[5] Kumar, M., Morales, E. O., Adams, J. E., & Hao, W., New digital camera sensor architecture for low light imaging. In Image Processing (ICIP), 2009 16th IEEE International Conference on (pp. 2681-2684). IEEE.

[6] Wang, J., Zhang, C., & Hao, P., New color filter arrays of high light sensitivity and high demosaicking performance. In Image Processing (ICIP), 2011 18th IEEE International Conference on (pp. 3153-3156). IEEE.

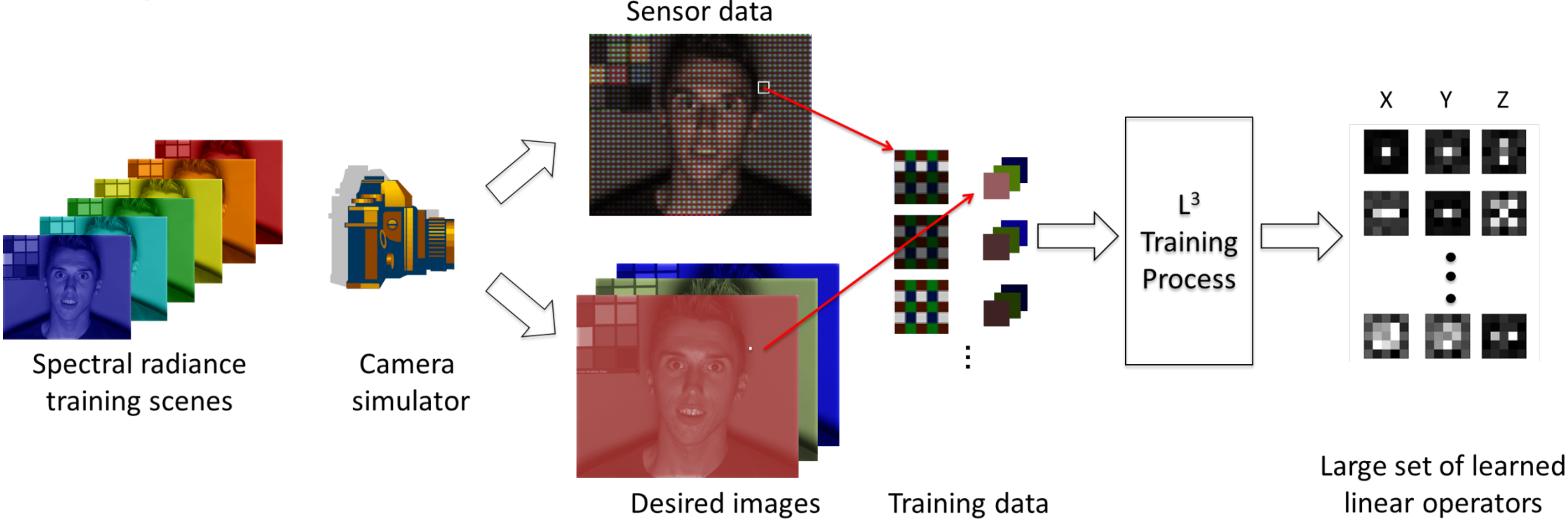
Thanks for your attention !
Questions?

Local, linear, learned (L^3)
Automating Image Processing Design

- **Local**: The full response space is separated into many local classes. The locality includes similarity in color, response level, pattern, and position.
- **Linear:** We seek a linear transformation appropriate for each local response class. The linear transformations differ from one another, so that the whole image processing pipeline is quite nonlinear
- **Learned:** We use training data and machine learning methods to optimize the collection of linear transformations that comprise the image processing.
- **The image pipeline:** The large collection of local, learned, linear transforms performs demosaicking, sensor correction, denoising, and illuminant correction.

L^3 Automating the Image Processing Design

- **Local, Linear and Learned**: fully automated and universal image processing pipeline (combined demosaicking, denoising and color transform) for any CFA based on machine learning
- **"Local"** : Refers to locality in space, color, response level, pattern, etc.



L^3 overview