

Outline

- 1 Introduction
- 2 Introduction to Digital Camera Pipeline
- 3 Quality and Complexity
- 4 Color Processing Inside Camera
- 5 Discussion

Welcome to the Tutorial!

- **Keigo Hirakawa**
 - Research Associate @ Harvard University
 - statistical signal processing; color image acquisition/display; computer vision; embedded systems
- **Patrick Wolfe**
 - Associate Professor @ Harvard University
 - statistical signal processing; high-dimensional data sets; audio/speech/image signals and networks
- **Truong Nguyen**
 - Professor @ University of California, San Diego
 - video processing; image compression; communications; electrocardiogram; filter bank; low-complexity processing

Goals for the Tutorial

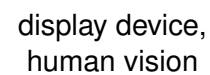
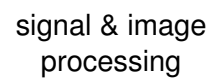
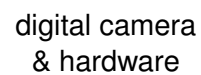
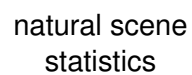
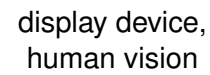
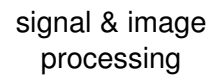
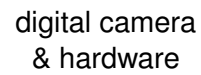
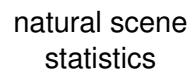
- Provide comprehensive overview of digital camera pipeline
- Discuss current and future challenges
- Convince to you how color has been long ignored
⇒ Formalize color science in camera pipeline context
- Develop tools for analyzing sensor data
- Bridge gaps between industry and academia
- Mathematically formalize problems motivated by needs of industry

Roadmap for the Tutorial

- **Overview & Color Processing** (Hirakawa, 1st hour)
 - overview of camera pipeline
 - market trends, architecture, challenges for pipeline designs
 - color processing inside camera
- **Sampling, Artifacts & Noise Issues** (Wolfe, 2nd hour)
 - sampling, color filter array, interpolation
 - crosstalk artifacts
 - noise (Poisson vs Gaussian), interpolation+denoising
- **Post-Pipeline: Enhancement, Display & Video** (Nguyen, 3rd hour)
 - image and video enhancement
 - image super resolution
 - frame rate conversion
 - enhancement for fluidic lens

- copies of our presentation slides
- list of relevant references.

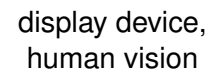
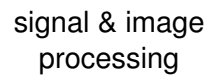
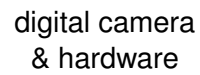
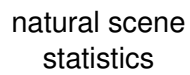
- 1 Introduction
- 2 Introduction to Digital Camera Pipeline
- 3 Quality and Complexity
- 4 Color Processing Inside Camera
- 5 Discussion



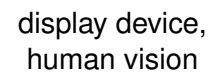
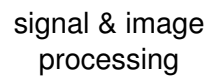
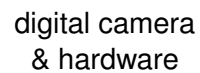
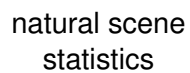
data generating
model, light source

noise model, sensor resolution

interpolation,
denoising,
restoration, etc



display resolution,
subjective analysis



DATA LOST HERE!!
↓
impose limits on
vision

Avenues for Improved Color Imaging

● Signal Characterization

- accurate model of sensor noise
- joint spatio-spectral statistical model for color image signal
- light source and reflectance models

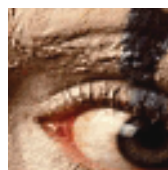
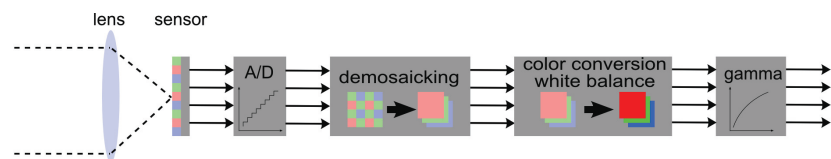
● Image Acquisition

- quantitative analysis of information loss
- sampling scheme preserving integrity of signal
- accurate model of noise/distortions, hardware limitations

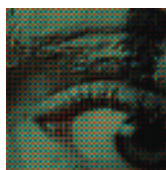
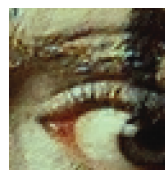
● Image Processing

- fast and efficient implementation
- enabling new functionality, extend usability
- human visual system, color science, enhancements
- compression, display, analysis of information loss

Most *Basic* Camera Pipeline



subject

sensor
output

demosaicking

color
correct/
white-
balance

Outline

- 1 Introduction
- 2 Introduction to Digital Camera Pipeline
- 3 **Quality and Complexity**
- 4 Color Processing Inside Camera
- 5 Discussion

Current Trends and Camera Pipeline

Consumer Trends

- lower cost
- lower cost
- lower cost
- better image quality
- higher resolution
- faster frame rate/processing
- longer battery life

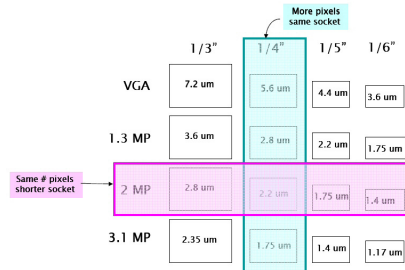
Industry: reduce pixel sensor size

- → higher resolution
- → reduced footprint
- → reduced cost
- → reduced power
- → increased noise
- → increased crosstalk
- → increased computation
- → reduced dynamic range
- → reduced frame rate

More DSP problems arise from reduced pixel sensor size!

Current Trends and Camera Pipeline

Performance and complexity are not opposing “design axes”!



- Greater distortion during sampling results in worse quality.
- Greater distortion during sampling results in higher computational complexity.
- Higher complexity (resolution) results in greater distortion.

Instead, we examine resolution-distortion trade-offs...

(figure adopted from presentation by Dr. Elaine W. Jin, Aptina Imaging)

Poisson/Measurement Process

Poisson/Measurement Process

$$y_{\tau}(n) = \underbrace{\tau^2 \{h_{\tau} * x\}}_{\text{subsampled image signal}}(\tau n)$$

$$\underbrace{z_{\tau}(n)}_{\text{measurement}} | \underbrace{x(n)}_{\text{light intensity}} = \mathcal{P}(\lambda y_{\tau}(n))$$

- Randomness is inherent in the signal.
- Photon count/arrival process modeled as Poisson process.
- This is *NOT* additive Gaussian noise!
- Fewer photons, low SNR

Distortion Measurement

- **Whittaker-Shannon** reconstruction
 $\mathcal{W}(x) = \{h_\tau * x\}$; sampling interval τ .
- Optimal estimate if nothing else known about the signal.
- Distortion as measured by
 - Whittaker-Shannon *reconstructibility* in L^2 sense.
 - measured against *continuous* ideal signal $g(\mathbf{t})$.

Distortion Analysis

$$J(x; \tau) = \left\| \underbrace{\frac{\mathcal{W}\{z_\tau\}(\mathbf{t})}{\tau^2 \lambda}}_{\text{reconstruction}} - \underbrace{x(\mathbf{t})}_{\text{ideal signal}} \right\|^2$$

Distortion Measurement

$$\begin{aligned}
 & E[J(x; \tau) | x] \\
 &= E \left[\left\| \frac{\mathcal{W}\{z_\tau/\lambda - y_\tau\}(\mathbf{t})}{\tau^2} - \left(x(\mathbf{t}) - \frac{\mathcal{W}\{y_\tau\}(\mathbf{t})}{\tau^2} \right) \right\|^2 \middle| x \right] \\
 &= \frac{1}{\tau^4} \sum_{\mathbf{n} \in \mathbb{Z}^2} \text{Var}(z_\tau(\mathbf{n})/\lambda | x) + \left\| \{(\delta - h_\tau) * x\}(\mathbf{t}) \right\|^2 \\
 &= \frac{1}{\tau^4} \sum_{\mathbf{n} \in \mathbb{Z}^2} y_\tau(\mathbf{n})/\lambda + \left\| \{(\delta - h_\tau) * x\}(\mathbf{t}) \right\|^2 \\
 &= \frac{\nu_x}{\tau^2 \lambda} + \left\| \{(\delta - h_\tau) * x\}(\mathbf{t}) \right\|^2
 \end{aligned}$$

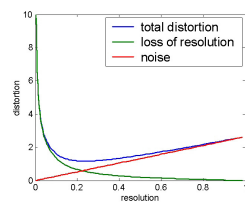
Distortion Measurement

$$E[J(x; \tau)|x] = \underbrace{\frac{\nu_x}{\tau^2 \lambda}}_{\text{Poisson noise}} + \underbrace{\left\| \{(\delta - h_\tau) * x\}(\mathbf{t}) \right\|^2}_{\text{loss of resolution}}$$

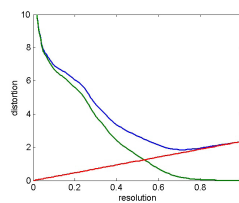
Distortion is separable:

- noise reduced by lower resolution (larger τ)
- distortion increased by lower resolution (larger τ)
- **distortion not monotonic!**
- **increasing resolution can do more harm than good!**
- **things are good for bright images; situation worse for dark images!**

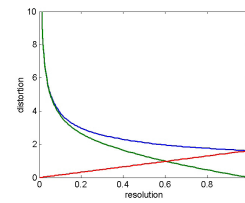
Expected Resolution-Distortion Trade-Offs



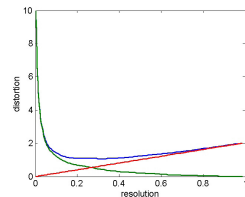
lena



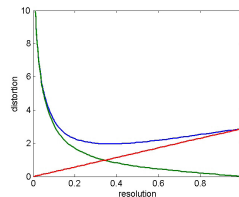
barbara



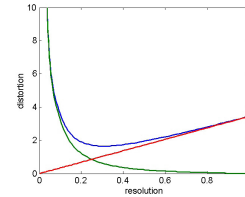
peppers



house



boat



fingerprint

Noise and Other Trends

- Noise scales linearly with resolution
- Resolution gain is image-content dependent
- Sensor defects scale linearly with resolution
- Crosstalk problem—we'll address that in the second lecture.

The Realities of Megapixel Count...

Consumer perception that pixel count is a proxy for image quality hurts everyone!

- increased cost to manufacture
- decreased image quality
- competition among manufacturers over pixel count
- increased cost for the consumers

Providing Alternatives to Pixel Count

I3A (Int'l Imaging Industry Association) **CPIQ** (Camera Phone Image Quality)

- initiatives

- develop a consumer-oriented camera phone rating system to replace the mega-pixel number on the box
- rating system based on perceptually calibrated objective and subjective metrics

- goals

- consumers will have meaningful system to aid in camera phone purchases
- camera phone manufacturers can focus on enhancing the user experience instead of pixel count

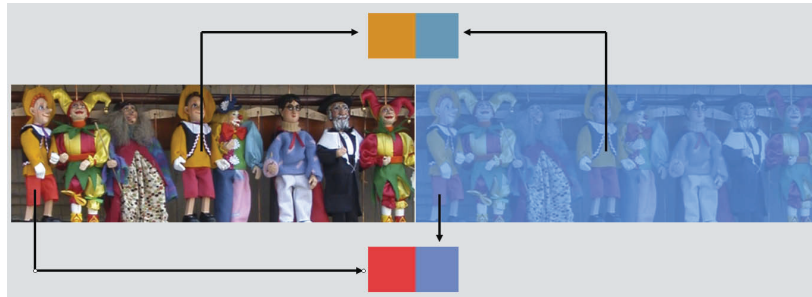
(slide adopted from presentation by Dr. Elaine W. Jin, Aptina Imaging)

Outline

- 1 Introduction
- 2 Introduction to Digital Camera Pipeline
- 3 Quality and Complexity
- 4 Color Processing Inside Camera**
- 5 Discussion

Color Science

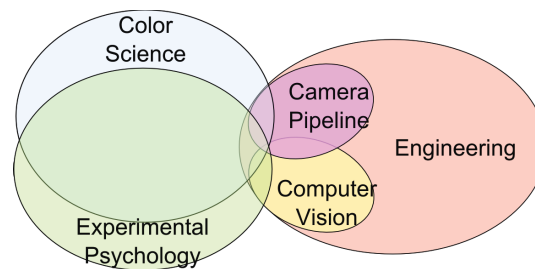
Why is color image processing so difficult?



- color science
- human vision/experimental psychology
- engineering

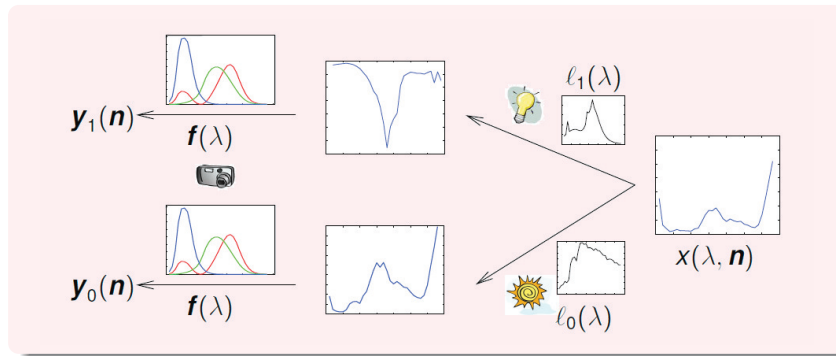
figure credit: Hamilton Chong, Harvard University

Fragmentation in Color Science



- Severe disagreement in terminologies.
- Major fragmentation between engineering and color science/psychology.
- Major overlap between color science and experimental psychology.
- Most overlap between “engineering” and color science/psychology are motivated by non-engineering applications.
- Lack of incorporating color science/psychology in camera pipeline designs.

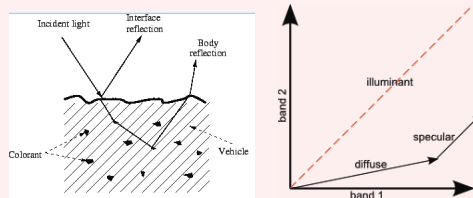
Image Formation



- $\mathbf{y}(\mathbf{n}) = \begin{bmatrix} y^{\{1\}}(\mathbf{n}) \\ y^{\{2\}}(\mathbf{n}) \\ y^{\{3\}}(\mathbf{n}) \end{bmatrix} = \int \mathbf{f}(\lambda) \ell(\lambda) x(\lambda, \mathbf{n}) d\lambda$
- (Same scene) \times (Different illuminants) = (Different Colors)

Image Formation

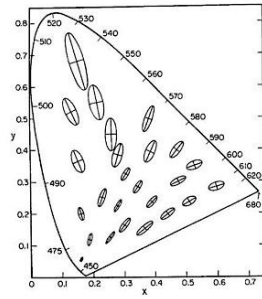
Dichromatic reflection model



$$\mathbf{y}(\mathbf{n}) = \int [\mathbf{f}(\lambda) \ell(\lambda) [(1 - \alpha(\mathbf{n})) + \alpha(\mathbf{n}) x(\lambda, \mathbf{n})] d\lambda$$

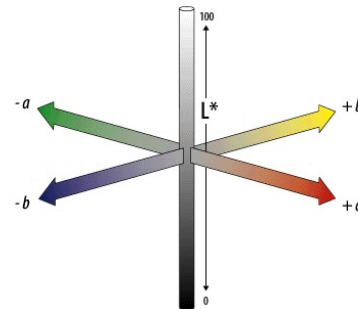
credit: Color Group, University of East Anglia

Human Visual System: Color Sensitivity



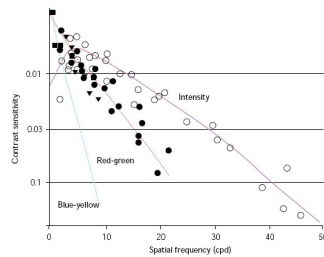
MacAdam Ellipses
contours of "just noticeable" sensitivity

figure credit: Adobe Inc



CIE-LAB Space
perceptually uniform response

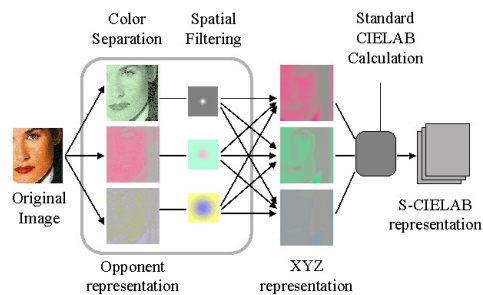
Human Visual System: Spatial Sensitivity



Contrast Sensitivity Function
spectral-dependency of spatial sensitivity

figure credit: X. Zhang (HP Labs) & B. Wandell (Stanford University)

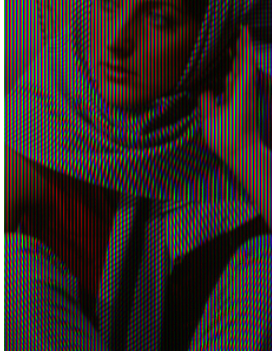
S-CIELAB Model



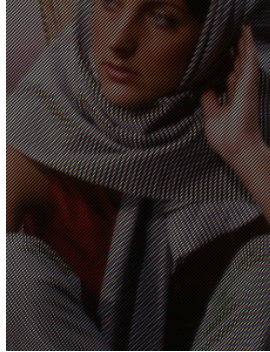
S-CIE-LAB Space
spatio-spectral perceptual uniformity

Why should engineers care about human vision?

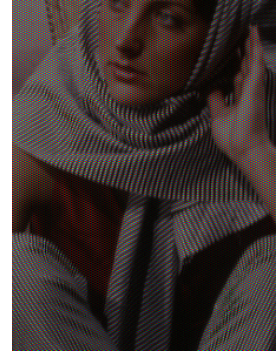
Here is a **display color filter array** example:



your TV/laptop LCD



one-laptop-per-child LCD



my laptop LCD

K. Hirakawa, P.J. Wolfe, "Fourier Domain Display Color Filter Array Design for Enhanced Image Fidelity," IEEE ICIP, 2007

Color Processing Inside Camera

Various types of color processing inside camera pipelines:

- **Estimation**
 - color corrected for sensor sensitivities
 - color balance compensated for illuminant
- **Enhancement**
 - color and contrast adjusted for human preference
- **Technical Specs**
 - anticipate interactions with display
 - assume certain viewing conditions
 - adhere to compression standards
 - color management

Flavors of “Color Mapping” Problem


 $y_0(n)$

 $y_1(n)$

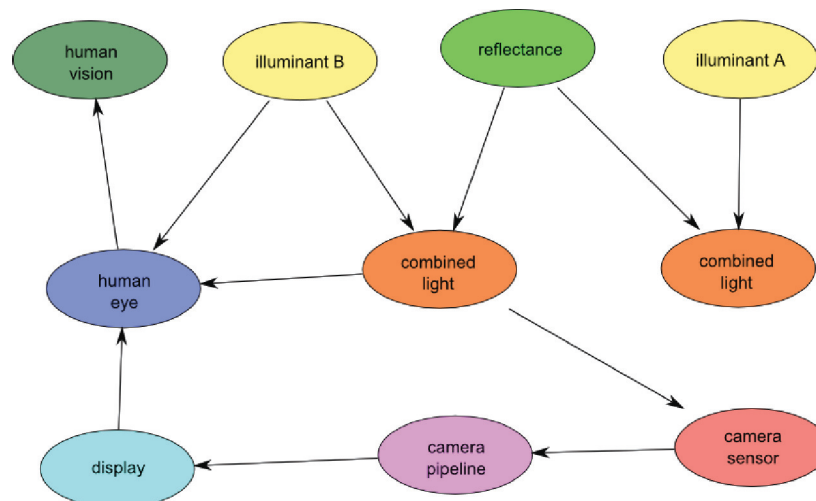
 $y_2(n)$

Implications of the following are *different*!

- Human eye (chromatic adaptation) maps colors to **discount illumination**
- White-balance maps colors to **match what the photographer saw**
- Color constancy maps colors to **canonical illuminant**
- Hyperspectral imaging maps colors to **reflectance values**
- Illuminant estimation finds **color of the light source**

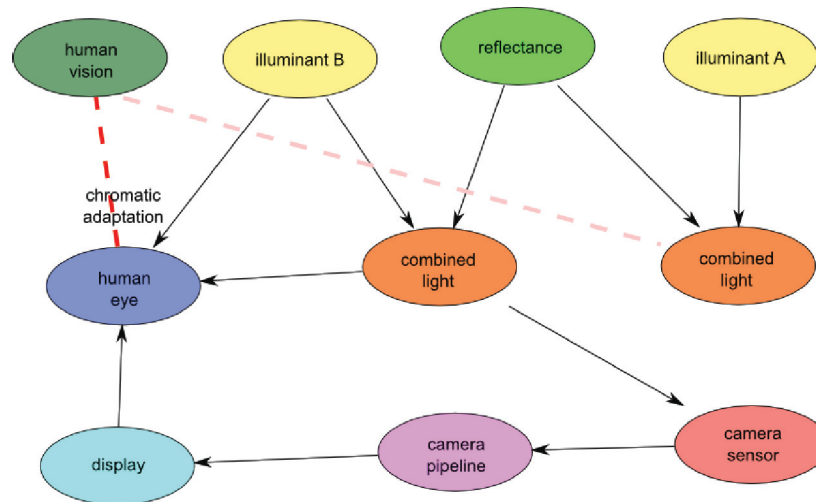
Not even an agreement on what the “true color” is!

Flavors of “Color Mapping” Problem



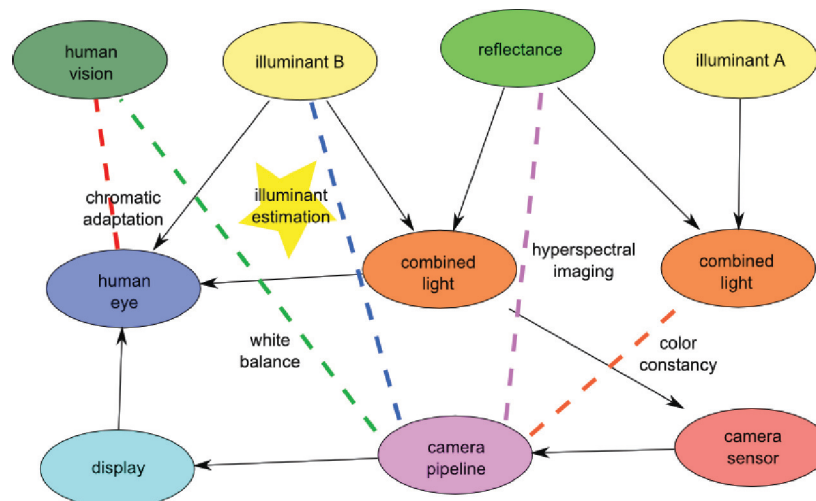
Be careful how you state your “color mapping” problem!

Flavors of “Color Mapping” Problem



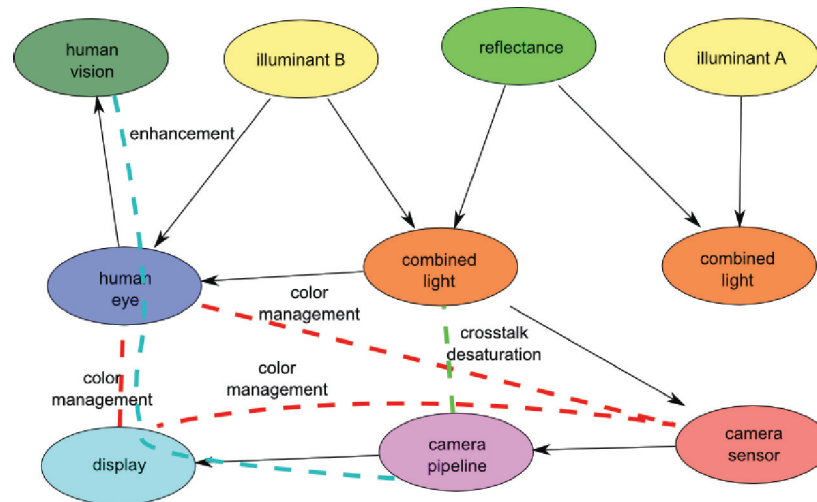
Be careful how you state your “color mapping” problem!

Flavors of “Color Mapping” Problem



Be careful how you state your “color mapping” problem!

Calibration, Color Management, Enhancement



Other color mapping problems to consider.

Practical Approaches to White Balance

Nature of Mapping

Chromatic adaptation map

- J. Von Kries, 1905
- D. Jameson, L.M. Hurvich, 1964
- A. Poirson, B.A. Wandell, 1996
- X. Zhang, B.A. Wandell, 1997
- E.J. Chichilnisky, B.A. Wandell, 1999
- V.C. Smith, J. Pokorny, 1996
- K. Hirakawa, T.W. Parks, 2005

Color constancy approximation map

- E.H. Land, 1974
- G. Buchsbaum, 1980
- G. West, M. H. Brill, 1982
- G. Finlayson, M. Drew, B. Funt, 1993
- G. Finlayson, S. Hordley, A. Alsam, 2006
- H. Chong, S. Gortler, T. Zickler, 2007

Illuminant Estimation

- E. Land, 1974. [Grey World]
- D. Brainard, W. Freeman, 1993. [Bayesian/Priors]
- G. Finlayson, S. Hordley, P. Hubel, 2001. [Color by Correlation]
- V. Cardei, B. Funt, K. Barnard, 2002. [Neural Networks]
- J. van de Weijer, T. Gevers, A. Gijsenij, 2007. [Grey Edge]
- A. Chakrabarti, K. Hirakawa, T. Zickler, 2008. [Spatial Model]

- chromatic adaptation/color constancy "maps" are parameterized by illuminant
- "two-step" approach: estimate illuminant, then map colors

Structure of Mapping

Color mapping schemes are parameterized by **illuminant**

- Chromatic adaptation maps

$$\mathbf{x} = \underbrace{\begin{bmatrix} w^{\{1\}} \\ w^{\{2\}} \\ w^{\{3\}} \end{bmatrix}^{-1}}_{\text{diagonal}} \mathbf{y}, \quad \mathbf{x} = \underbrace{\mathbf{y} - k \text{lum}(\mathbf{y}) \mathbf{M} \begin{bmatrix} w^{\{1\}} \\ w^{\{2\}} \\ w^{\{3\}} \end{bmatrix}}_{\text{subtractive}}$$

- Color constancy approximation maps

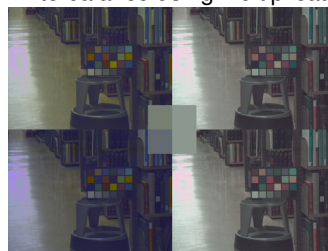
$$\mathbf{x} = \underbrace{\begin{bmatrix} w^{\{1\}} \\ w^{\{2\}} \\ w^{\{3\}} \end{bmatrix}^{-1}}_{\text{diagonal}} \mathbf{y}, \quad \mathbf{x} = \underbrace{\mathbf{M}^{-1} \begin{bmatrix} w^{\{1\}} \\ w^{\{2\}} \\ w^{\{3\}} \end{bmatrix}^{-1}}_{\text{linear}} \mathbf{M} \mathbf{y}$$

Structure of Mapping

Examples of white balance using multiplicative and subtractive mapping.

greyworld

Finlayson 2001



multiplicative

subtractive



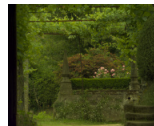
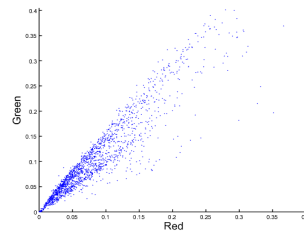
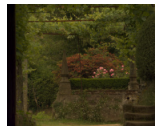
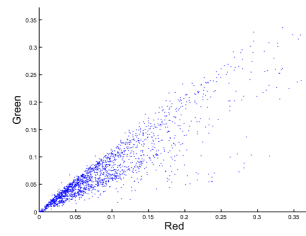
multiplicative

subtractive

- nature of mapping is just as important as illuminant estimation!!
- subtractive white balance is robust to illuminant estimation.

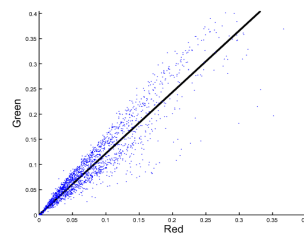
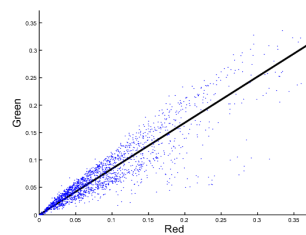
K. Hirakawa, T.W. Parks, "Chromatic Adaptation and White Balance Problem," ICIP 2005

Estimating Illuminant from an Image



Uncertainties in reflectance values make illuminant estimation problem difficult.

Estimating Illuminant from an Image



Uncertainties in reflectance values make illuminant estimation problem difficult.

Estimating Illuminant from an Image

- **White Patch Retinex** (V.C. Cardei and B. Funt, 1999)
 - brightest patch = white
 - $x^{\{i\}}(\mathbf{n}) = \frac{y^{\{i\}}(\mathbf{n})}{\max_{\mathbf{n}'} y^{\{i\}}(\mathbf{n}')}$
- **Grey World** (G. Buchsbaum, 1980)
 - mean intensity in all channels = grey/neutral
 - $x^{\{i\}}(\mathbf{n}) = \frac{y^{\{i\}}(\mathbf{n})}{\text{mean}(\{y^{\{i\}}(\mathbf{n}')\}_{\mathbf{n}'})}$
- **Grey Segmented Patch** (R. Gershon *et al.*, 1988)
 - mean segmented patch = grey/neutral
- **Grey Edge** (J. van de Weijer *et al.*, 2007)
 - mean edge map = grey/neutral

Estimating Illuminant from an Image

- **Dichromatic model** (G.D. Finlayson, G. Schaefer, 2002)
 - within object, color variation along illuminant
 - $\mathbf{y}(\mathbf{n}) = \int [\mathbf{f}(\lambda)\ell(\lambda)[(1 - \alpha(\mathbf{n})) + \alpha(\mathbf{n})\mathbf{x}(\lambda, \mathbf{n})] d\lambda$
- **Gamut Constrained Color Constancy** (D. Forsyth, 1990; G.D. Finlayson, 2003)
 - set of intensities under canonical illuminant is convex
- **Bayesian Methods** (D. Brainard and W.T. Freeman, 1993; P.V. Gehler *et al.*, 2008)
- **Learning-Based Methods**
 - linear filters: (A.C. Hurlbert and T. Poggio, 1988)
 - neural networks: (V. Cardei *et al.*, 2002)



Weaknesses in the model:

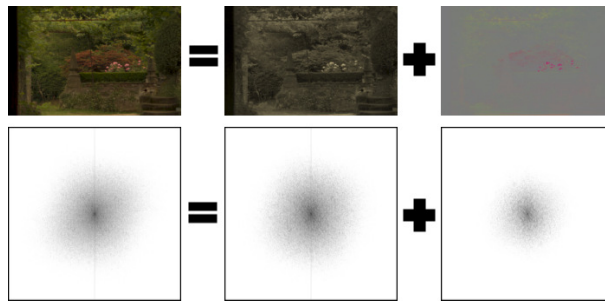
- ignore **spatial correlations** in reflectances !
- algorithms sensitive to **dominant color**!



Weaknesses in the model:

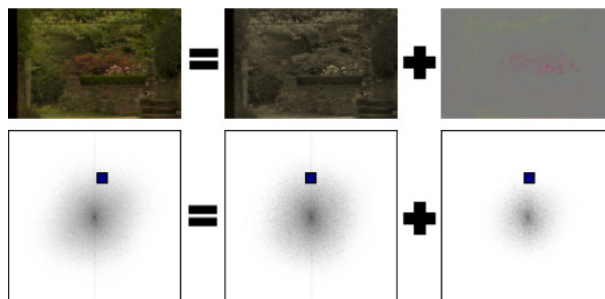
- ignore **spatial correlations** in reflectances !
- algorithms sensitive to **dominant color**!

Fourier Supports



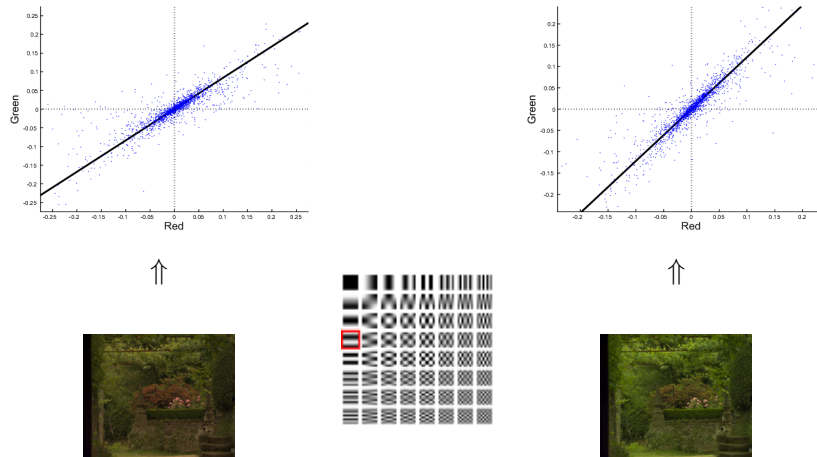
The strengths of correlation between reflectance values measured at various wavelengths (λ) are **spatial frequency-dependent**.

Fourier Supports



The strengths of correlation between reflectance values measured at various wavelengths (λ) are **spatial frequency-dependent**.

Estimating Illuminant from an Image



Spatio-Spectral Perspective

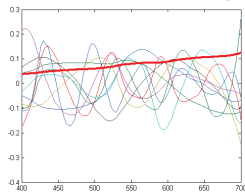
The strengths of correlation between reflectance values measured at various wavelengths (λ) are **spatial frequency-dependent**.

Why should this be so? Some intuitions...

Spatio-Spectral Perspective

The strengths of correlation between reflectance values measured at various wavelengths (λ) are **spatial frequency-dependent**.

Why should this be so? Some intuitions...

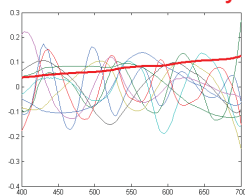


- predominance of “flat” reflectance spectrum
- object boundaries dominated by change of intensity, not color

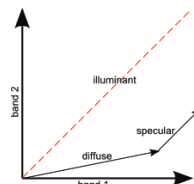
Spatio-Spectral Perspective

The strengths of correlation between reflectance values measured at various wavelengths (λ) are **spatial frequency-dependent**.

Why should this be so? Some intuitions...



- predominance of “flat” reflectance spectrum
- object boundaries dominated by change of intensity, not color

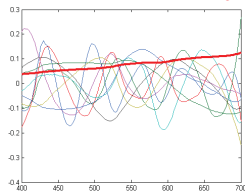


- dichromatic model: light is reflected without absorption
- diffuse is low-frequency, specular is high-frequency

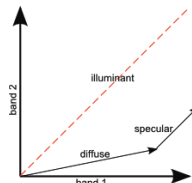
Spatio-Spectral Perspective

The strengths of correlation between reflectance values measured at various wavelengths (λ) are **spatial frequency-dependent**.

Why should this be so? Some intuitions...



- predominance of "flat" reflectance spectrum
- object boundaries dominated by change of intensity, not color



- dichromatic model: light is reflected without absorption
- diffuse is low-frequency, specular is high-frequency



edges/textures dominated by changes in brightness

Navigation icons: back, forward, search, etc.

Statistical Model

Idea: spatial decorrelation suffices as spectral decorrelation!

- DCT/wavelets/filterbank \Rightarrow basis $\{\mathbf{D}_k\}_{k=0}^{K-1}$

$$\mathcal{D}_k^T \mathbf{X} = \begin{bmatrix} \mathbf{D}_k^T & & \\ & \mathbf{D}_k^T & \\ & & \mathbf{D}_k^T \end{bmatrix} \begin{bmatrix} \mathbf{X}^{\{1\}} \\ \mathbf{X}^{\{2\}} \\ \mathbf{X}^{\{3\}} \end{bmatrix} = \begin{bmatrix} \mathbf{D}_k^T \mathbf{X}^{\{1\}} \\ \mathbf{D}_k^T \mathbf{X}^{\{2\}} \\ \mathbf{D}_k^T \mathbf{X}^{\{3\}} \end{bmatrix} \Bigg\} 3 \times 1$$

- Transform coefficients modeled as multivariate Gaussian

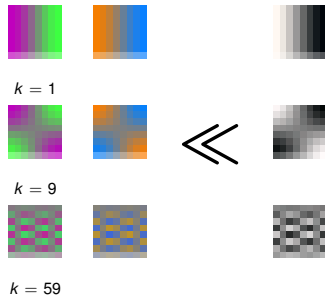
$$\mathcal{D}_k^T \mathbf{X} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \Lambda_k), \quad k > 0$$

- Model for complete reflectance image (Λ_k trained from training set):

$$P(\mathbf{X}) \propto \prod_{k>0} \frac{1}{\det(\Lambda_k)^{1/2}} \exp\left(-\frac{1}{2}(\mathcal{D}_k^T \mathbf{X})^T \Lambda_k^{-1} \mathcal{D}_k^T \mathbf{X}\right)$$

Insights from Λ_k

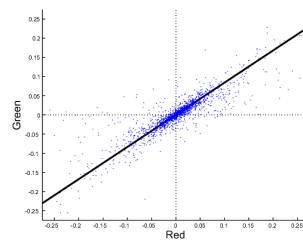
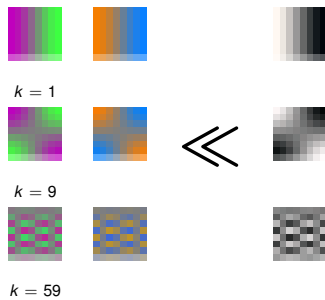
Eigenvectors of Λ_k



Expected from scatter plots

Insights from Λ_k

Eigenvectors of Λ_k



Expected from scatter plots

Estimation Algorithm

- \mathbf{w} diagonal transform to reflectance image

- $\mathbf{Y}_j \xrightarrow{\mathbf{w}} \hat{\mathbf{X}}_j(\mathbf{w}) = \begin{bmatrix} w^{\{1\}} \mathbf{Y}_j^{\{1\}} \\ w^{\{2\}} \mathbf{Y}_j^{\{2\}} \\ w^{\{3\}} \mathbf{Y}_j^{\{3\}} \end{bmatrix}$

Model-Fitting

- $\mathbf{w} = \arg \max_{\mathbf{w}'} \sum_j \log P(\hat{\mathbf{X}}_j(\mathbf{w}'))$
- optimize under constraint $\|\mathbf{w}\|^2 = 3$

Estimation Algorithm

- \mathbf{w} diagonal transform to reflectance image

- $\mathbf{Y}_j \xrightarrow{\mathbf{w}} \hat{\mathbf{X}}_j(\mathbf{w}) = \begin{bmatrix} w^{\{1\}} \mathbf{Y}_j^{\{1\}} \\ w^{\{2\}} \mathbf{Y}_j^{\{2\}} \\ w^{\{3\}} \mathbf{Y}_j^{\{3\}} \end{bmatrix}$

Model-Fitting

- $\mathbf{w} = \arg \max_{\mathbf{w}'} \sum_j \log P(\hat{\mathbf{X}}_j(\mathbf{w}'))$
- optimize under constraint $\|\mathbf{w}\|^2 = 3$

Experimental Validation



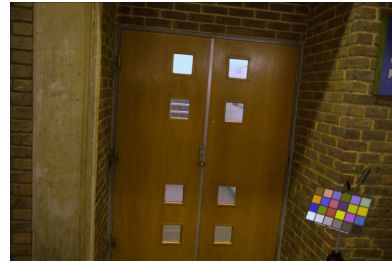
input



greyworld

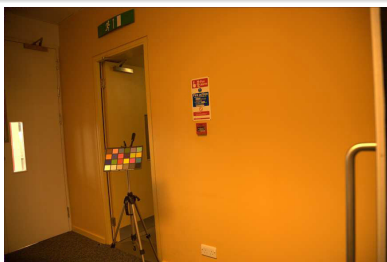


greyedge

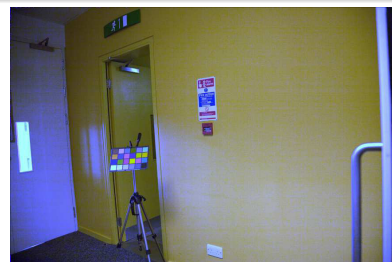


spatial

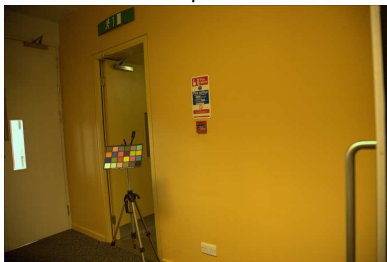
Experimental Validation



input



greyworld

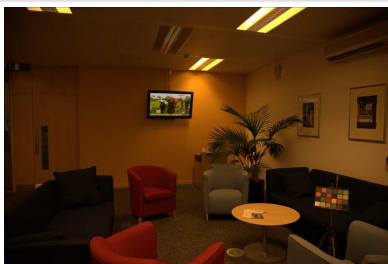


greyedge



spatial

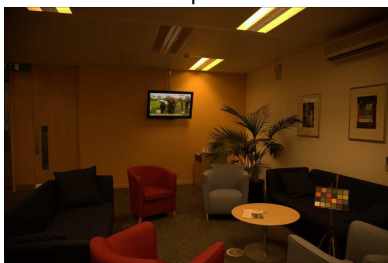
Experimental Validation



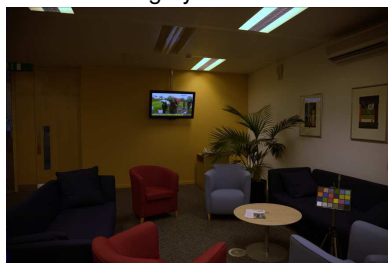
input



greyworld



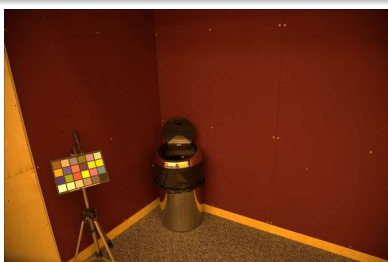
greyedge



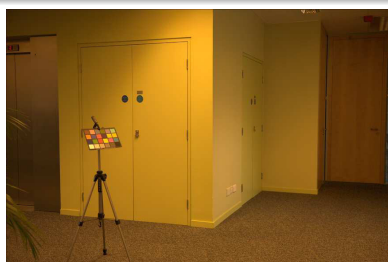
spatial

Navigation icons: back, forward, search, etc.

Experimental Validation



input



input



input



input

Navigation icons: back, forward, search, etc.

Experimental Validation



spatial



spatial



spatial



spatial

Outline

- 1 Introduction
- 2 Introduction to Digital Camera Pipeline
- 3 Quality and Complexity
- 4 Color Processing Inside Camera
- 5 Discussion

Hot Topics and Outstanding Problems

- Public perception of “megapixel count”
- Sensor noise problems—low light regimes
- Optical diffraction and minority carrier diffusion (crosstalk)
- Compressed dynamic range of CMOS sensors
- Sensor binning and artifacts
- Bad pixel amelioration
- Image and color enhancement
- Storage/transmission of high data-rate multimedia
- Standards and their adaptation

Thank you!

- hirakawa@stat.harvard.edu
- www.accidentalmark.com/research
- sisl.seas.harvard.edu