A Surface Reflectance Model and Applications from Image Sequences

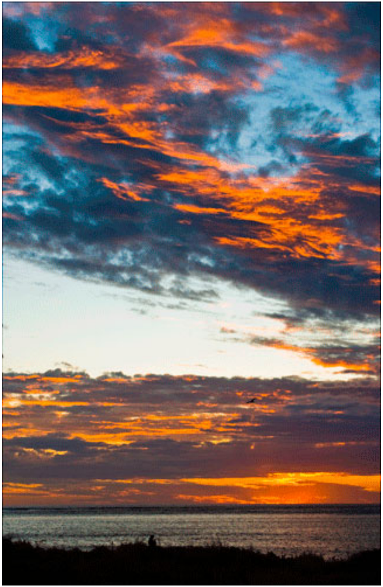
Véronique Prinet
with Dani Lischinski and Michael Werman

The Hebrew University of Jerusalem, Israel

April 16^{th} 2015, CASIA

Table of contents

Motivation



Motivation

- Most scenes are illuminated by several (global or local) illuminants (*eg.* outdoor scene at sunset).
- Color cast or strong specularities can cause vision algorithms (*eg.* segmentation, recognition) to produce erroneous results.

Related work

[Shafer 85], [Klinker & al. 88], [Lee & Bajcsy 92], [Tan & al. 2003], [Robles-Kelly & al. 2010], [Yang & al. 2011]

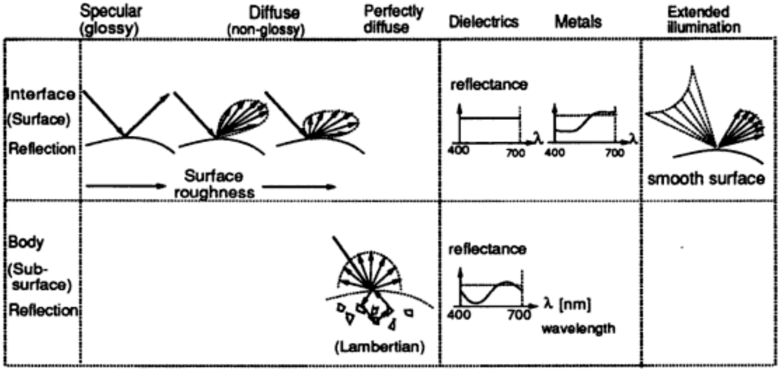


Fig. 1. Reflection models

Related work

- Gray-Edge [Weijer & al. 2007], Generalised Gamut Mapping [Gijsenij & al. 2010],
- Multi-illuminant [Gijsenij & al. 2012, Ebner 2004],
- White balance correction [Hsu & al. 2008].
- Illuminant in video sequences [Wang & al. 2011],

Image Formation Model

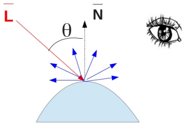
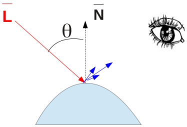
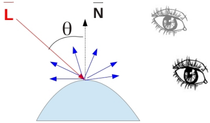
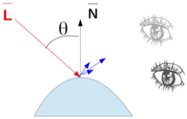
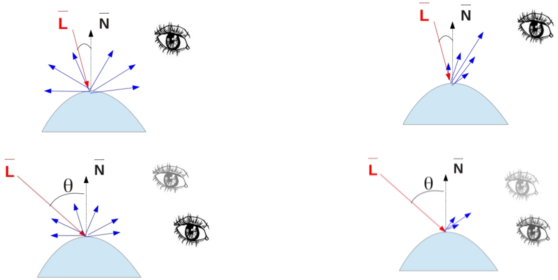
	Diffuse component $\mathbf{D}(\mathbf{p})$	Specular component $m(\mathbf{p})\mathbf{L}$
Light motion		
Camera motion		

Image Formation Model

Diffuse component $\mathbf{D}(\mathbf{p})$ Specular component $m(\mathbf{p})\mathbf{L}$

Light motion

Camera motion



Introduction & Background

Illuminant Chromaticity from Image Sequence [ICCV 2013]

- One global illuminant
- Two dominant illuminants
- Results

Specularity Enhancement from Image Sequence [ICIP 2013]

- Approach
- Results

Conclusions and Future Work

Image Formation Model & Hypothesis

Dichromatic reflection model for dieletric objects [Shafer 85] in time space.

$$\mathbf{J}(\mathbf{p}) = \mathbf{D}(\mathbf{p}) + \mathbf{S}(\mathbf{p})$$

Image Formation Model & Hypothesis

Dichromatic reflection model for dieletric objects [Shafer 85] in time space.

$$\mathbf{J}(\mathbf{p}) = \mathbf{D}(\mathbf{p}) + m_s(\mathbf{p})\mathbf{L}(\mathbf{p})$$

Image Formation Model & Hypothesis

Dichromatic reflection model for dieletric objects [Shafer 85] in time space.

$$\mathbf{J}(\mathbf{p}, t) = \mathbf{D}(\mathbf{p}, t) + m_s(\mathbf{p}, t)\mathbf{L}(\mathbf{p})$$

Image Formation Model & Hypothesis

Dichromatic reflection model for dieletric objects [Shafer 85] in time space.

$$\mathbf{J}(\mathbf{p}, t) = \mathbf{D}(\mathbf{p}, t) + m_s(\mathbf{p}, t)\mathbf{L}(\mathbf{p})$$

Hypothesis

- Incident light **L** is uniformly distributed over time.
- Object body reflectance **D** is time invariant.

Single Global Illuminant Γ

$$\mathbf{J}(\mathbf{p},t) = \mathbf{D}(\mathbf{p},t) + m_s(\mathbf{p},t)\mathbf{L}$$

$$\Rightarrow \quad \mathbf{J}(\mathbf{p}+\Delta\mathbf{p}, t+\Delta t) - \mathbf{J}(\mathbf{p},t) \;=\; \Delta m_s(\mathbf{p},t)\;\mathbf{L}$$

$$\Rightarrow \quad \frac{L_c}{\sum_c L_c} = \frac{\Delta J_c(\mathbf{p},t)}{\sum_c \Delta J_c(\mathbf{p},t)} \qquad c \in \{r,g,b\}$$

with $\mathbf{L} = (L_r, L_g, L_b)$, $\mathbf{J} = (J_r, J_g, J_b)$ and
$\Delta\mathbf{J}(\mathbf{p},t) = \mathbf{J}(\mathbf{p}+\Delta\mathbf{p}, t+\Delta t) - \mathbf{J}(\mathbf{p},t)$.

Single Global Illuminant Γ

$$\mathbf{J}(\mathbf{p},t) = \mathbf{D}(\mathbf{p},t) + m_s(\mathbf{p},t)\mathbf{L}$$

$$\Rightarrow \quad \mathbf{J}(\mathbf{p}+\Delta\mathbf{p}, t+\Delta t) - \mathbf{J}(\mathbf{p},t) \; = \; \Delta m_s(\mathbf{p},t)\,\mathbf{L}$$

$$\Rightarrow \qquad \frac{L_c}{\sum_c L_c} = \frac{\Delta J_c(\mathbf{p},t)}{\sum_c \Delta J_c(\mathbf{p},t)} \qquad c \in \{r,g,b\}$$

with $\mathbf{L} = (L_r, L_g, L_b), \mathbf{J} = (J_r, J_g, J_b)$ and $\Delta\mathbf{J}(\mathbf{p},t) = \mathbf{J}(\mathbf{p}+\Delta\mathbf{p}, t+\Delta t) - \mathbf{J}(\mathbf{p},t)$.

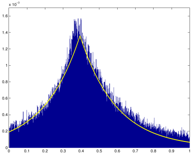
The dispacement field, $\Delta\mathbf{p}$, is assumed to be known.

Single Global Illuminant $\boldsymbol{\Gamma}$

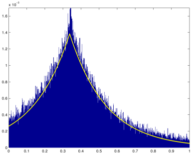
Define $x_c(\mathbf{p}) = \frac{\Delta J_c(\mathbf{p},t)}{\sum_c \Delta J_c(\mathbf{p},t)}$, $\mathbf{x} = \{x_c(\mathbf{p}), c \in \{r, b, g\}, \forall \mathbf{p} \in \mathcal{E}\}$.

We observe empirically that $h(x_c)$ has the following (Laplace) distribution:

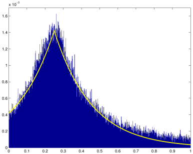
$h(x_r)$



$h(x_g)$



$h(x_b)$



This suggests to cast the problem of $\boldsymbol{\Gamma} = \{\Gamma_c\}$ estimation in a MAP framework, where:

$$\hat{\boldsymbol{\Gamma}} = \arg\max_{\boldsymbol{\Gamma}} P(\boldsymbol{\Gamma}|\mathbf{x})$$

Two Dominant Illuminants $\{\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2\}$

- *Assume now*:

$$\mathbf{L}(\mathbf{p}) = k_1(\mathbf{p})\,\boldsymbol{\Gamma}_1 + k_2(\mathbf{p})\,\boldsymbol{\Gamma}_2.$$

- For locally uniform incident light $\mathbf{L}(\mathbf{p}) = cst = \mathbf{L}^s, \forall \mathbf{p} \in s$, where s is a small space-time patch, then:

$$\Gamma_c^s = \frac{L_c^s}{\sum_c L_c^s} = \alpha^s\,\Gamma_{1,c} + (1-\alpha^s)\,\Gamma_{2,c}.$$

with $\alpha^s = k_1^s/(k_1^s + k_2^s)$.

- *Re-parametrisation*:

$$\alpha^s = b_c + a_c \Gamma_c^s \qquad \forall c \in \{r, g, b\},$$

with $a_c = \frac{1}{\Gamma_{1,c} - \Gamma_{2,c}}$ and $b_c = \frac{-\Gamma_{2,c}}{\Gamma_{1,c} - \Gamma_{2,c}}$.

Two Dominant Illuminants $\{\mathbf{\Gamma}^1, \mathbf{\Gamma}^2\}$

- *Define the quadratic cost function*:

$$\begin{aligned} E(\alpha, \mathbf{a}, b) &= \sum_s (\alpha^s - \sum_c a_c \Gamma_c^s - b)^2 + \epsilon ||\mathbf{a}||^2 \\ E(\alpha, \mathbf{w}) &= ||\bar{\alpha} - K\mathbf{w}||^2 \end{aligned}$$

with $\bar{\alpha} = (\alpha^1, ..., \alpha^S, 0)$, $\mathbf{w} = (a_r, a_g, a_b, b)^t$,

$$K = \begin{bmatrix} \Gamma_r^1 & \Gamma_g^1 & \Gamma_b^1 & 1 \\ & .... & & \\ \epsilon^{1/2} a_r & \epsilon^{1/2} a_g & \epsilon^{1/2} a_b & 0 \end{bmatrix}.$$

- *Solve for* $\bar{\alpha}$. We can show that, for the optimal value of $\mathbf{w}$ ([Levin & al. 2008]):

$$E(\alpha) = \bar{\alpha}^t M \bar{\alpha}.$$

- *Solve for* $\hat{\mathbf{\Gamma}}_1, \hat{\mathbf{\Gamma}}_2$. Finally:

$$(\hat{\mathbf{\Gamma}}_1, \hat{\mathbf{\Gamma}}_2) = \arg\min \sum_s ||\alpha_s \, \mathbf{\Gamma}_1 \; + \; (1 - \alpha_s) \, \mathbf{\Gamma}_2||^2$$

Experimental setting

- Dataset :
  - One illuminant : benchmark dataset (GrayBall) and in-house datasets (13 sequences, acquired under normal and extrem lighting conditions);
  - Two illuminants : in-house dataset (3 sequences).
- Ground truth : from gray-card placed in the scene during acquisition.
- Temporal window : 3-5 frames ;
  Tiles size (2 illuminants): 100×100 pixels.
- Evaluation : $err = \arccos(\mathbf{\Gamma}^{Est}.\mathbf{\Gamma}^{GT})$

Results : Single Illuminant

Figure : Video dataset recorded under normal lighting conditions.

	Average	*Best 1/3*	*Worst 1/3*
GE-1 [Weijer & al. 2007]	6.572	**2.1787**	11.271
GE-2 [Weijer & al. 2007]	7.150	2.958	11.723
GGM [Gijsenij & al. 2010]	7.013	6.208	9.166
IIC [Tan & al. 2003]	8.303	3.984	12.540
Our approach	**5.389**	2.402	**8.784**

Table : Average angular errors (in degrees) .

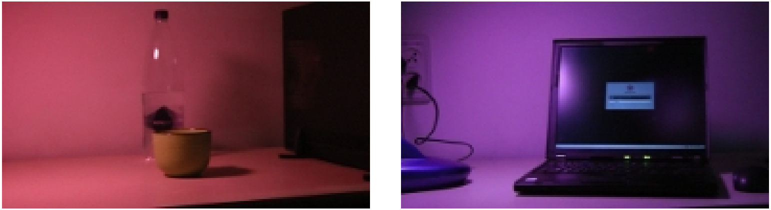


Figure : Video dataset simulating extreme lighting conditions: reddish and bluish.

	Reddish	*Bluish*
GE-1 [Weijer & al. 2007]	8.907	13.052
GE-2 [Weijer & al. 2007]	10.246	13.657
GGM [Gijsenij & al. 2010]	15.544	25.505
IIC [Tan & al. 2003]	—	19.675
Our approach	**7.708**	**6.236**

Table : Average angular errors (in degrees).

Figure : Sample frames from the Grayball database (out of a total of 11,136 images).

	Mean	*Median*
GrayWorld	7.9	7.0
GGM [Gijsenij & al. 2010]	6.9	5.8
GE-2 [Wang & al. 2011]	5.4	4.1
Ours	5.4	4.6

Table : Angular errors.

Application: White balance correction

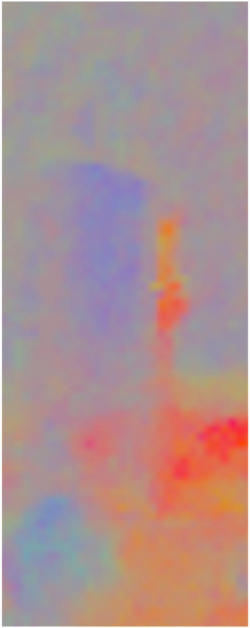
(a) Input sequence *(b) Our approach* *(c) Ground truth*

Two illuminants

Figure : Three sequences captured with two lights sources.

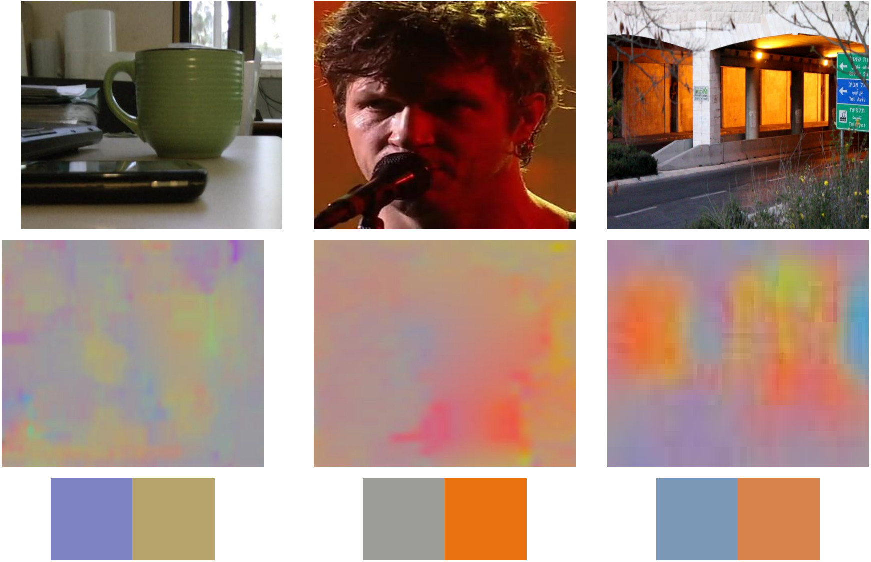
	Ours		[Gijsenij & al. 2012]		*Local GW*	
	Γ_1	Γ_2	Γ_1	Γ_2	Γ_1	Γ_2
Seq. (a)	**9.65**	5.14	31.69	**4.8**	12.94	10.49
Seq. (b)	**5.74**	**4.76**	9.69	9.82	5.89	8.81
Seq. (c)	**7.35**	6.49	17.9	5.65	7.63	**5.74**

Input $\mathbf{J}(\mathbf{p}, t)$

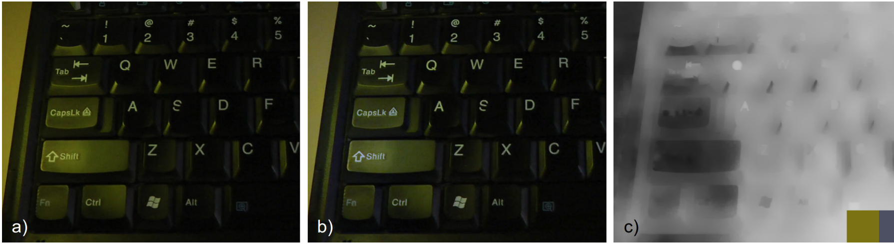


$\boldsymbol{\Gamma}^s$

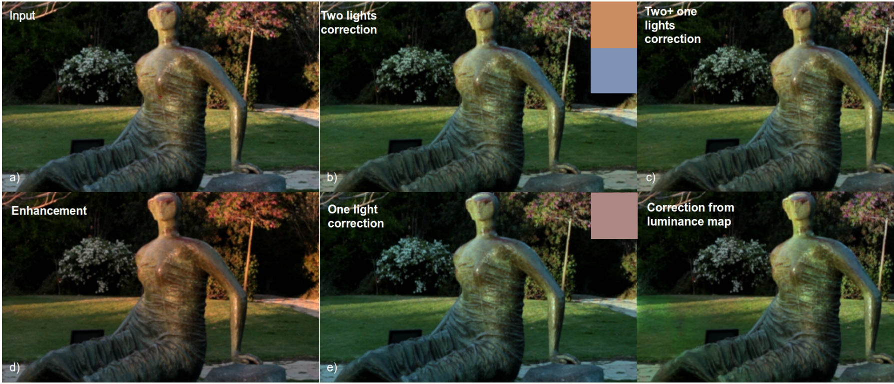
α^s


Tel Aviv

Application: Color cast correction



Application: Relighting




a)
b)
c)
d)

Specularity map $m(\mathbf{p}, t)$

Recall:

$$J_c(\mathbf{p} + \boldsymbol{\Delta}\mathbf{p}, t + \Delta t) - J_c(\mathbf{p}, t) \\ = m(\mathbf{p} + \boldsymbol{\Delta}\mathbf{p}, t + \Delta t) - m(\mathbf{p}, t)$$

- Displacement $\Delta\mathbf{p}$ is known or estimated accurately,
- White illuminant: $\mathbf{L}(\mathbf{p}, t) = \mathbf{L} = (1, 1, 1)$,
- Body reflectance $\mathbf{D}$ is invariant over time.

Approach 1: a 1^{rst} order integration

$\underline{\Delta \mathbf{p} \neq 0}$

$$\frac{d}{dt} m(\mathbf{p}(t), t) \; = \frac{d}{dt} J_c(\mathbf{p}(t), t)$$

$$\Rightarrow m(\mathbf{p}(t), t) = J_c(\mathbf{p}(t), t) - D_{t_0,c}$$

with initial conditions: $D_{t=0,c}(\mathbf{p}) = J_c(\mathbf{p}(t_0), t_0) - m(\mathbf{p}(t_0), t_0)$.

Approach 2: a 2^{nd} order integration

Minimise:

$$\int_t \int_{\mathbf{p}} d\mathbf{p}\, dt\, \left\{\Delta J_c(\mathbf{p}, t) - \big(m(\mathbf{p} + \Delta\mathbf{p}, t + \Delta t) - m(\mathbf{p}, t)\big)\right\}^2$$

with respect to $m(.)$.

$$\Rightarrow \Delta^2 J_c(\mathbf{p}, t) = m_{pp}(\mathbf{p}, t)\Delta\mathbf{p}^2 + m_{tt}(\mathbf{p}, t)\Delta t^2 + 2\, m_{pt}(\mathbf{p}, t)\Delta\mathbf{p}\Delta t$$

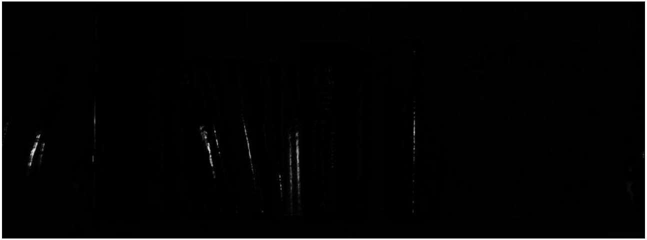
- discretize using (central) finite differences
- appropriate boundary/initial conditions

  $\Rightarrow$ solve a linear system of the form: $A\bar{\mathbf{m}} = B$.

Results: Static Scenes

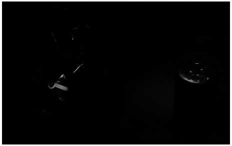
Results: Static Scenes

Results: Static Scenes

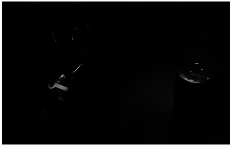


Results: Static Scenes

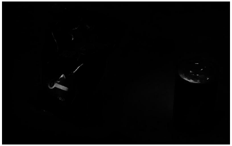
Camera motion



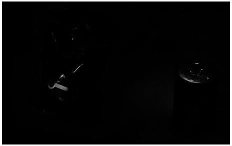
Camera motion



Camera motion



Camera motion



Conclusion

Summary

- *Physically-based model* for illuminant estimation or specularity detection,
- *Leverage temporal information*.

Limitations

- Physically-based model is an *approximation* of the real world,
- *Displacement flow accuracy* is a limiting factor for local estimation of incident light,
- Light sources should not be too close (in space and spectral domain).

Future work

- Improvement of the white balance correction task (two light sources),
- Extension to multiple light sources,
- Estimation of scene structure and intrinsic images.

Thanks for your attention.