



Improving the Visible with the Invisible

Incorporating Near-Infrared Cues in Computational Photography and Computer Vision



Sabine Süsstrunk

Images and Visual Representation Group (IVRG)

The “Internet of Things”

- Humans generated **more data** in 2009 than in the **previous 5'000** years combined.¹

¹Dave Evans , Chief Futurist for Cisco Systems, 2010.

The “Internet of Things”

- Humans generated *more data* in 2009 than in the *previous 5'000* years combined.¹
- By 2012, **90% of all data** will be video.¹

¹Dave Evans , Chief Futurist for Cisco Systems, 2010.

The “Internet of Things”

- Humans generated *more data* in 2009 than in the *previous 5'000* years combined.¹
- By 2012, *90% of all data* will be video.¹
- **24 hours** of video are uploaded **every minute** on youtube.²

¹Dave Evans , Chief Futurist for Cisco Systems, 2010. ²Google, 2010.

The “Internet of Things”

- Humans generated *more data* in 2009 than in the *previous 5'000* years combined.¹
- By 2012, *90% of all data* will be video.¹
- *24 hours* of video are uploaded *every minute* on youtube.²
- Over *5 billion* photos on flickr.³

¹Dave Evans , Chief Futurist for Cisco Systems, 2010. ²Google, 2010. ³CNN, 2010

The “Internet of Things”

- Humans generated *more data* in 2009 than in the *previous 5'000* years combined.¹
- By 2012, *90% of all data* will be video.¹
- *24 hours* of video are uploaded *every minute* on youtube.²
- Over *5 billion* photos on flickr.³
- ***Each month***, 2.5 billion photos are uploaded to Facebook.³

¹Dave Evans , Chief Futurist for Cisco Systems, 2010. ²Google, 2010. ³CNN, 2010

The Challenge of Imaging

***How do we render, interpret,
organize, and communicate
all this visual data?***

The Challenge of Imaging

*How do we render, interpret,
organize, and communicate
all this visual data?*

we = our machines

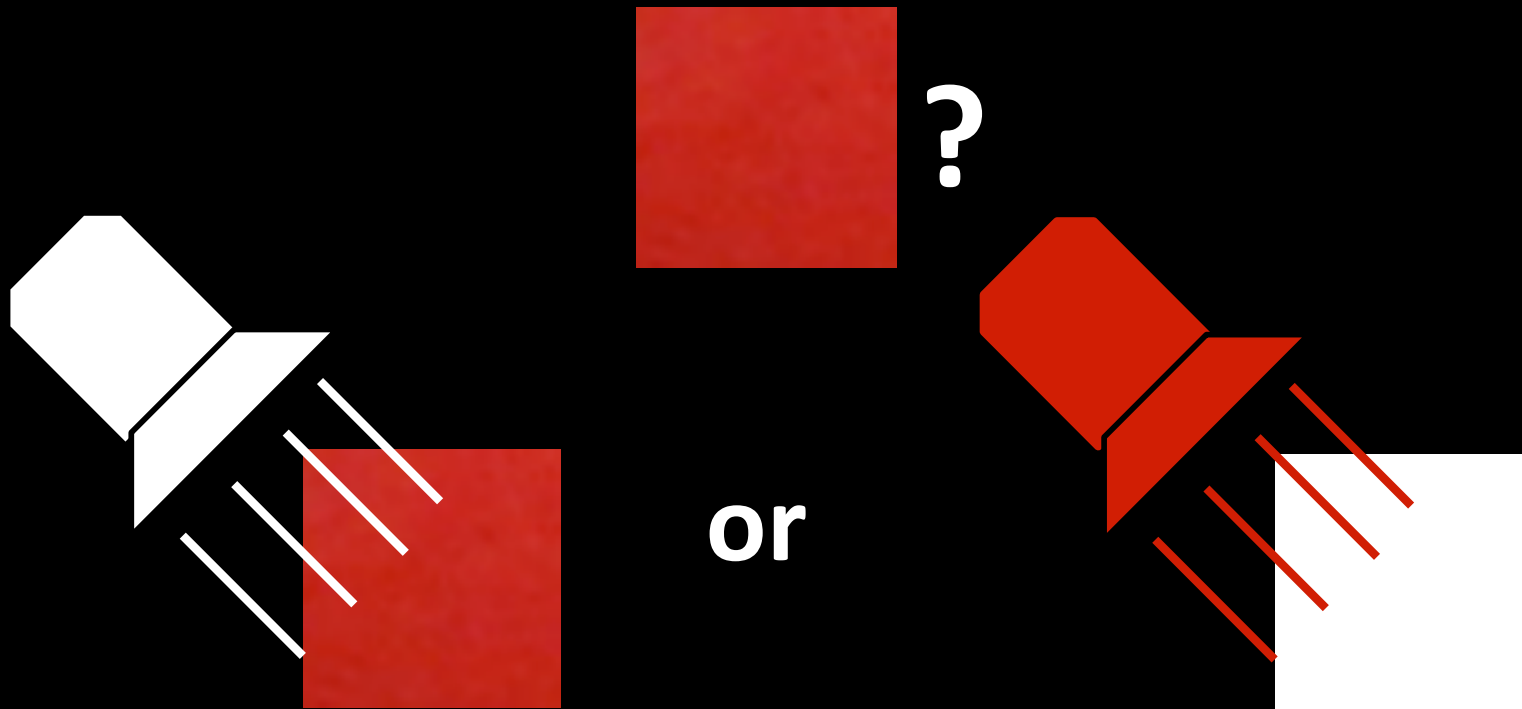
Fundamental Imaging Problem

- What is the contribution to the pixel value of the object *reflectance* versus the *light source*?



Fundamental Imaging Problem

- What is the contribution to the pixel value of the object *reflectance* versus the *light source*?

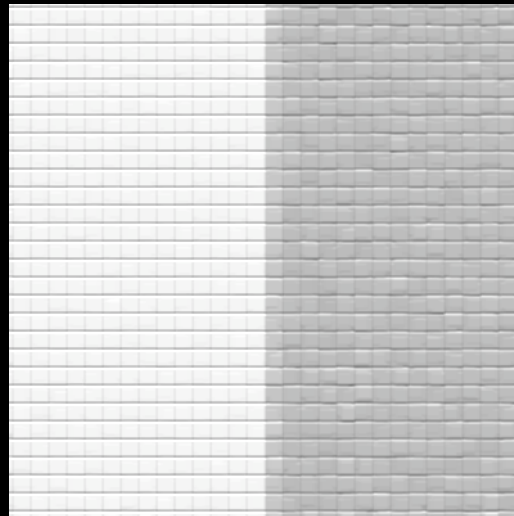


A *red* surface under white light?

A white surface under *red* light?

Fundamental Imaging Problem

- What is the contribution to the pixel value of the object *reflectance* versus the *light source*?



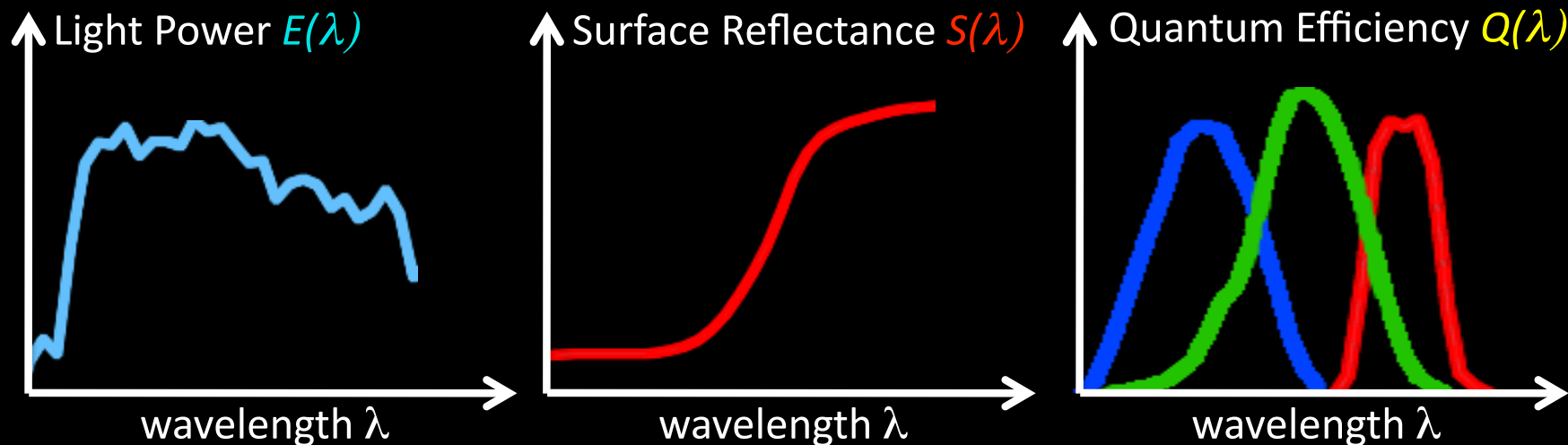
Two shades of tile under one light?
One shade of tile under *two* lights?

Image Formation

p_k : Camera Response

$$p_k = \int_{\lambda=400nm}^{700nm} E(\lambda) S(\lambda) Q_k(\lambda) d\lambda$$

$$k = \{B, G, R\}$$



Scene Understanding

- Separating the effect of illuminant and surface reflectance of the camera response (pixel value) is an *ill-posed* problem.

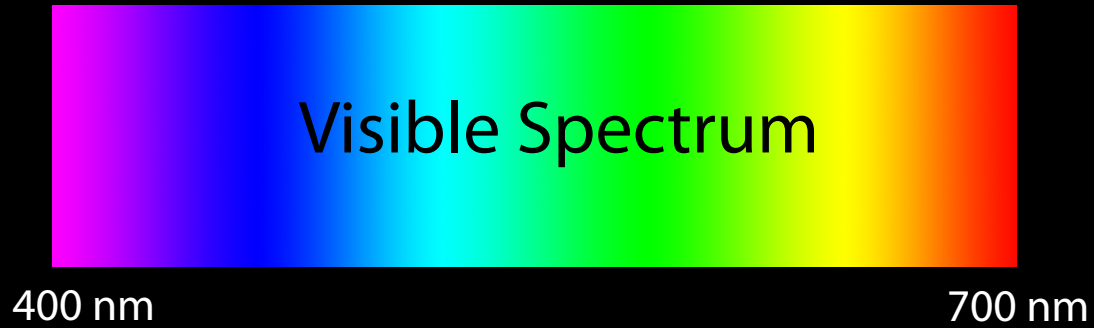
Scene Understanding

- Separating the effect of illuminant and surface reflectance of the camera response (pixel value) is an *ill-posed* problem.
- To estimate the two, we either make assumptions about the scene...

Scene Understanding

- Separating the effect of illuminant and surface reflectance of the camera response (pixel value) is an *ill-posed* problem.
- To estimate the two, we either make assumptions about the scene...
- *...or we capture more physical information about the scene.*

Current Imaging: RGB



Visible (Red, Green, and Blue: RGB)

Scientific Imaging: Near-Infrared

Near-Infrared

700 nm

1100 nm



Near-infrared (NIR)

Future Imaging: RGB + NIR

Visible Spectrum

Near-Infrared

400 nm

700 nm

1100 nm



Visible (Red, Green, and Blue: RGB)



Near-infrared (NIR)

Benefits of RGB + NIR

- We can **exploit** the correlation and de-correlation of visible and NIR image frequencies and intensities to:
 - **enhance** the visual quality of images/videos.
 - **extract** more accurate information about the scene.

Outline

- Acquisition of RGB + NIR
- Enhance imaging:
 - Haze removal
 - Skin smoothing
- Extract more information about the scene:
 - Illuminant and shadow detection
 - Material-based object segmentation
 - Scene recognition
- Conclusion

Acknowledgment

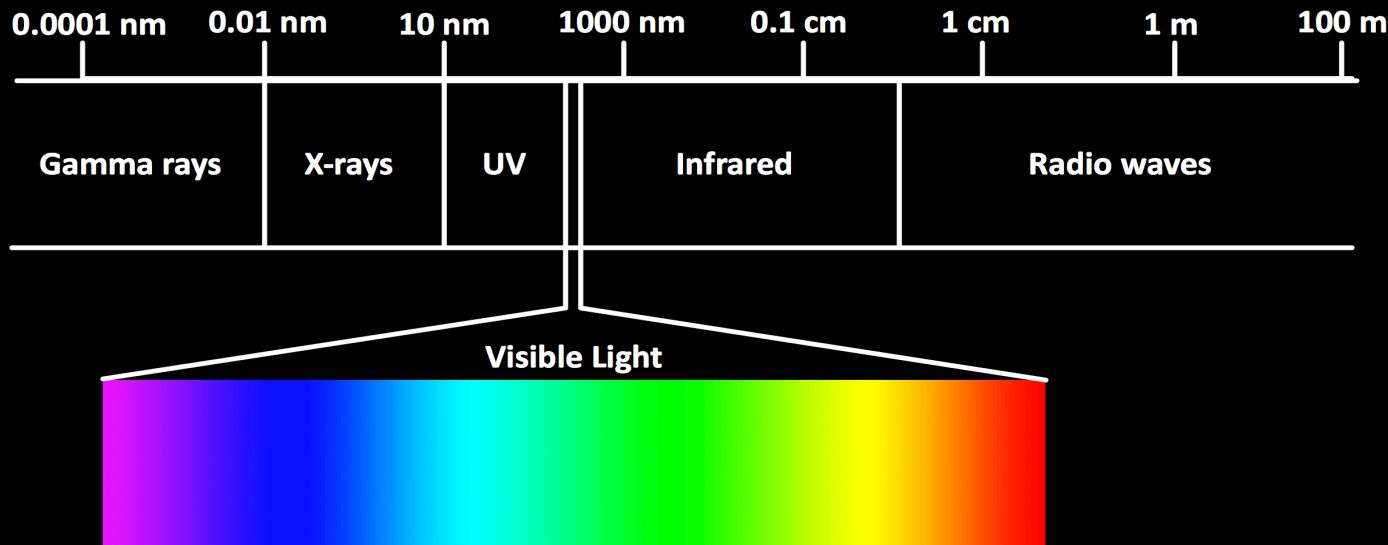
- Joint work with
 - ***Nathalie Barbuscia***
 - ***Matthew Brown***
 - ***Damien Firmenich***
 - ***Clément Fredembach***
 - ***Arthur Germain***
 - ***Yue M. Lu***
 - ***Neda Salamati***
 - ***Zahra Sadeghipoor***
 - ***Lex Schaul***
 - ***Daniel Tamburrino***
- Funding
 - ***Swiss National Science Foundation***
 - ***Xerox Foundation***



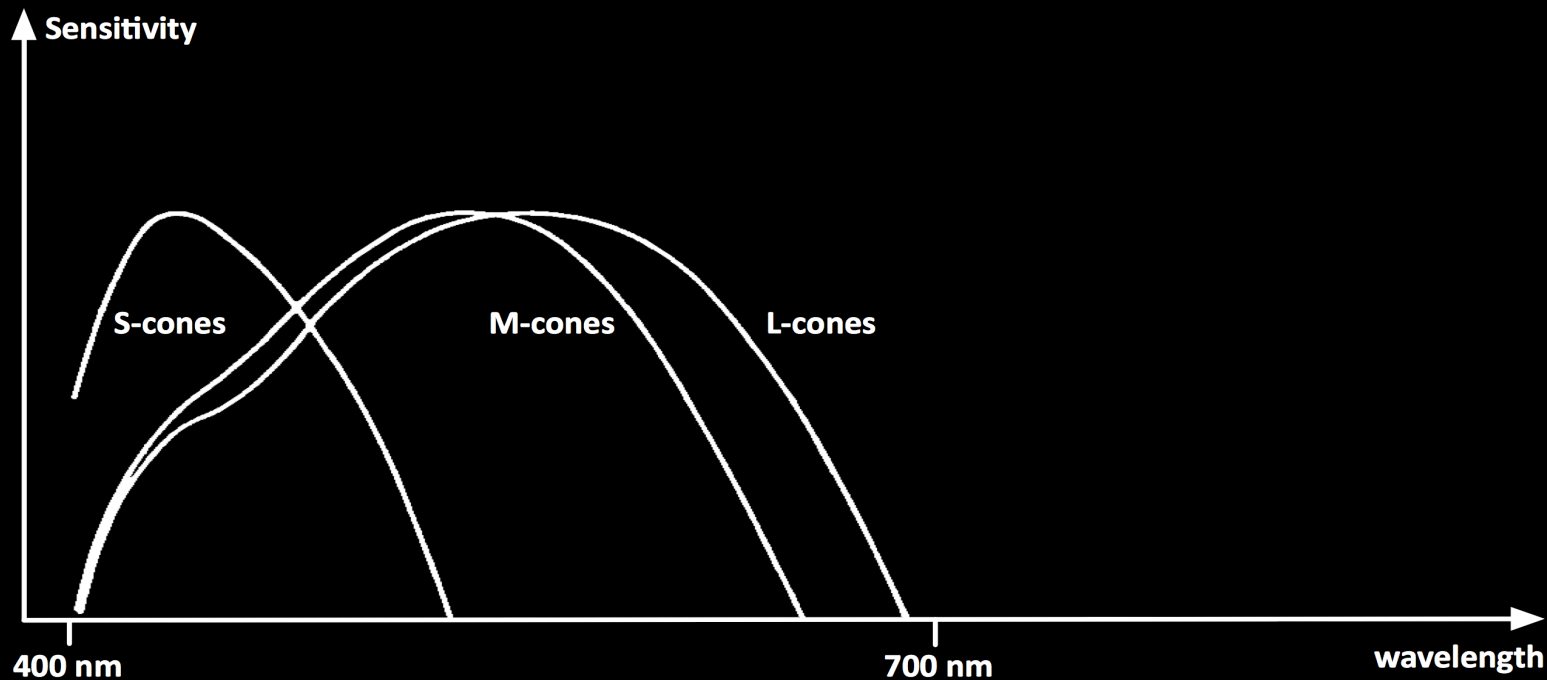
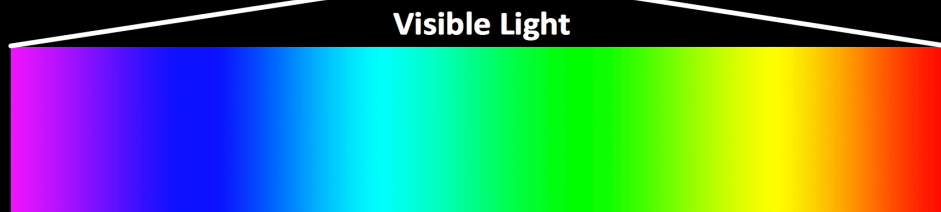
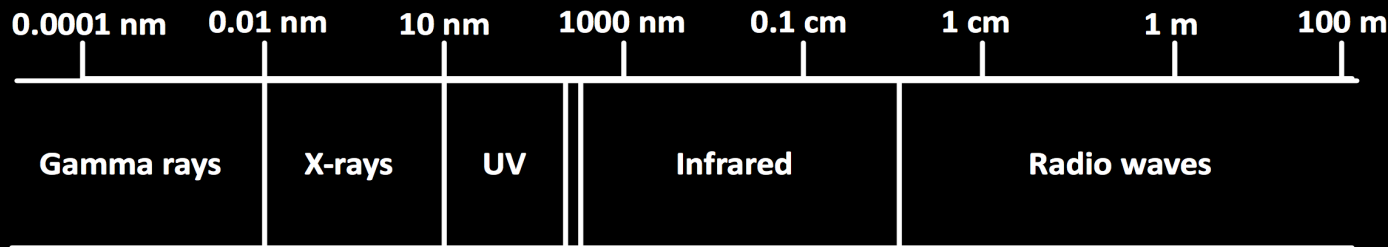
Outline

- ***Acquisition of RGB + NIR***
- Enhance imaging:
 - Haze removal
 - Skin smoothing
- Extract more information about the scene:
 - Illuminant and shadow detection
 - Material-based object segmentation
 - Scene recognition
- Conclusion

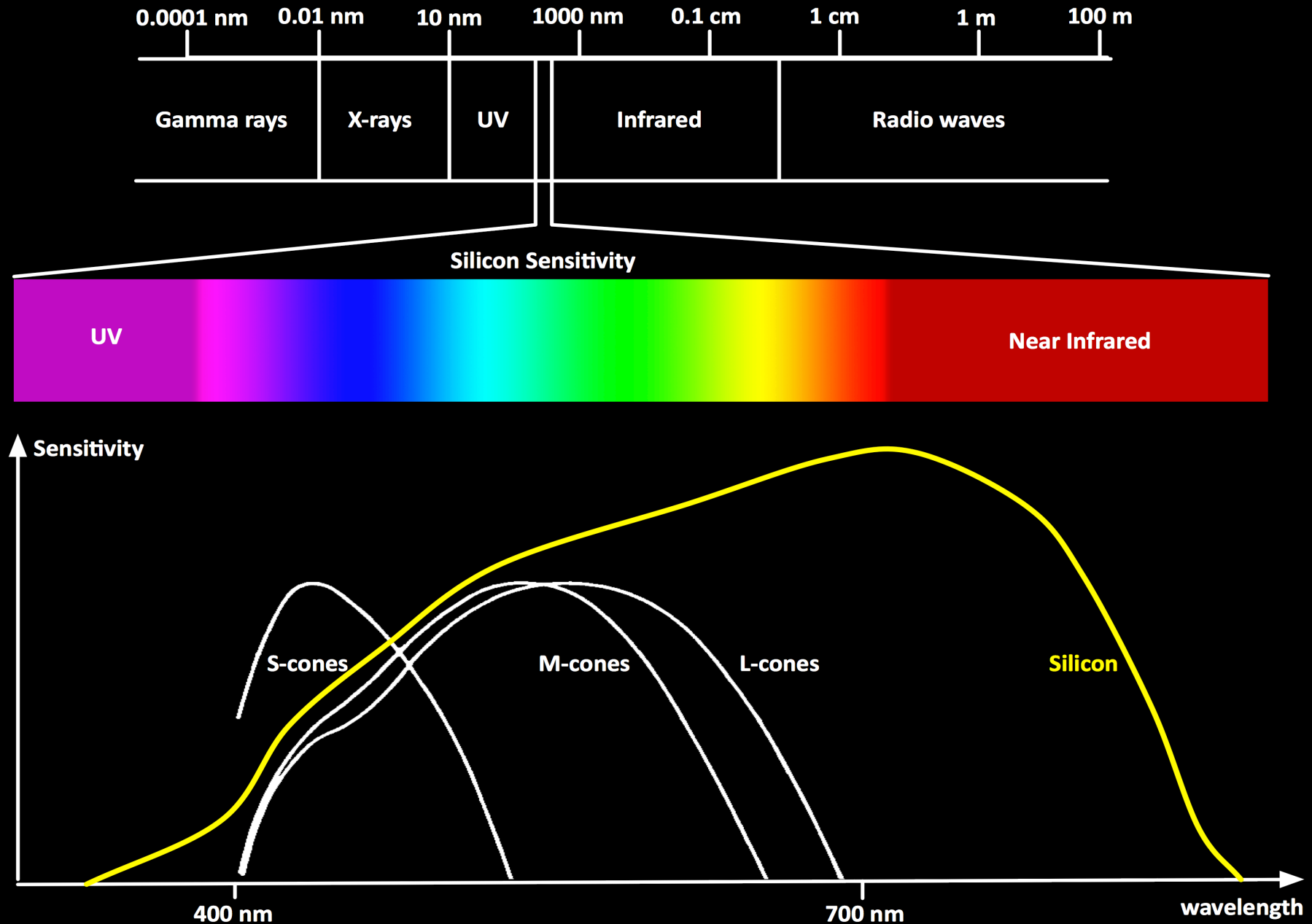
Human vision and the spectrum



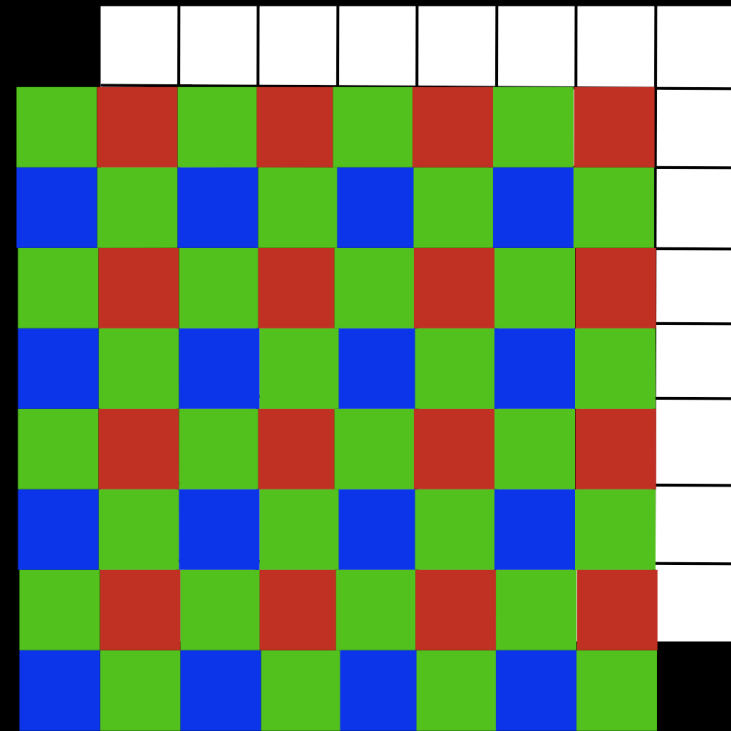
Human vision and the spectrum



The sensitivity of silicon

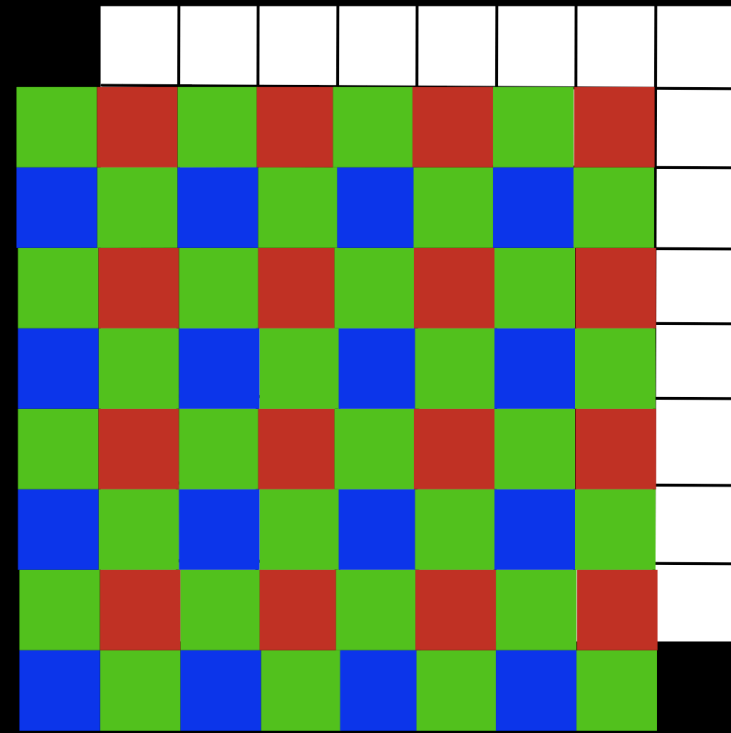
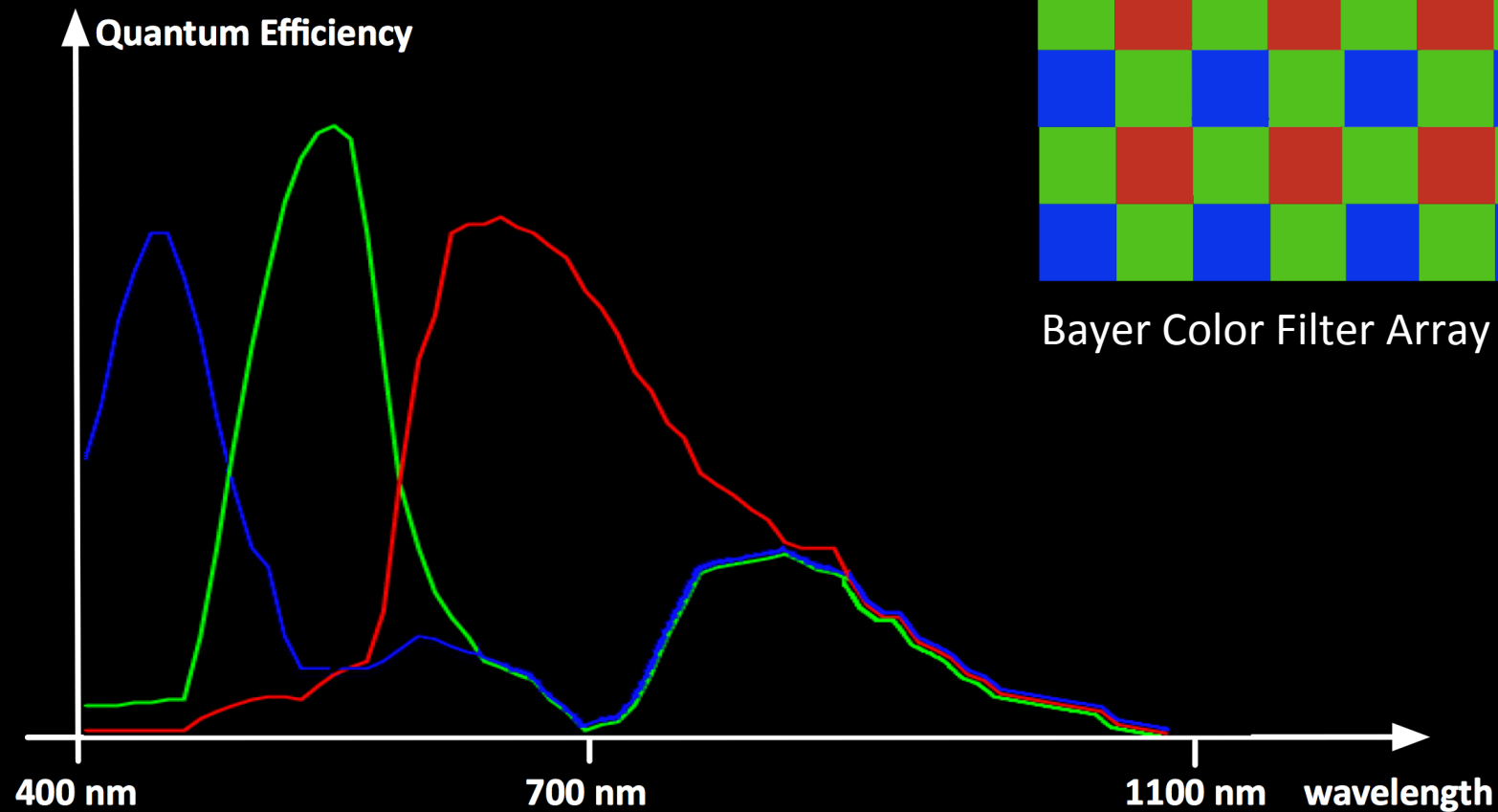


Camera Design



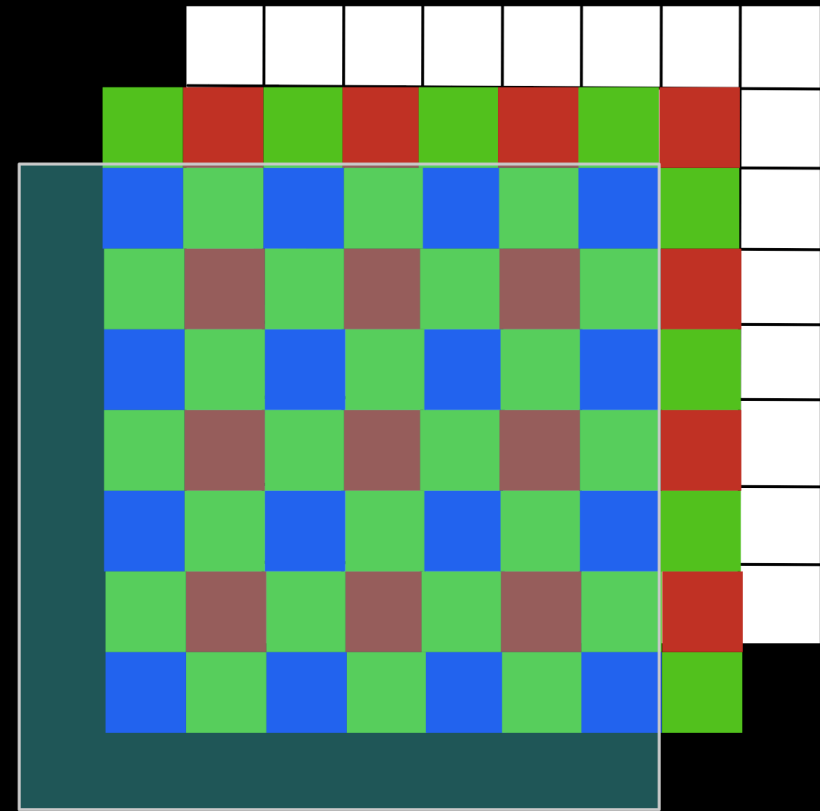
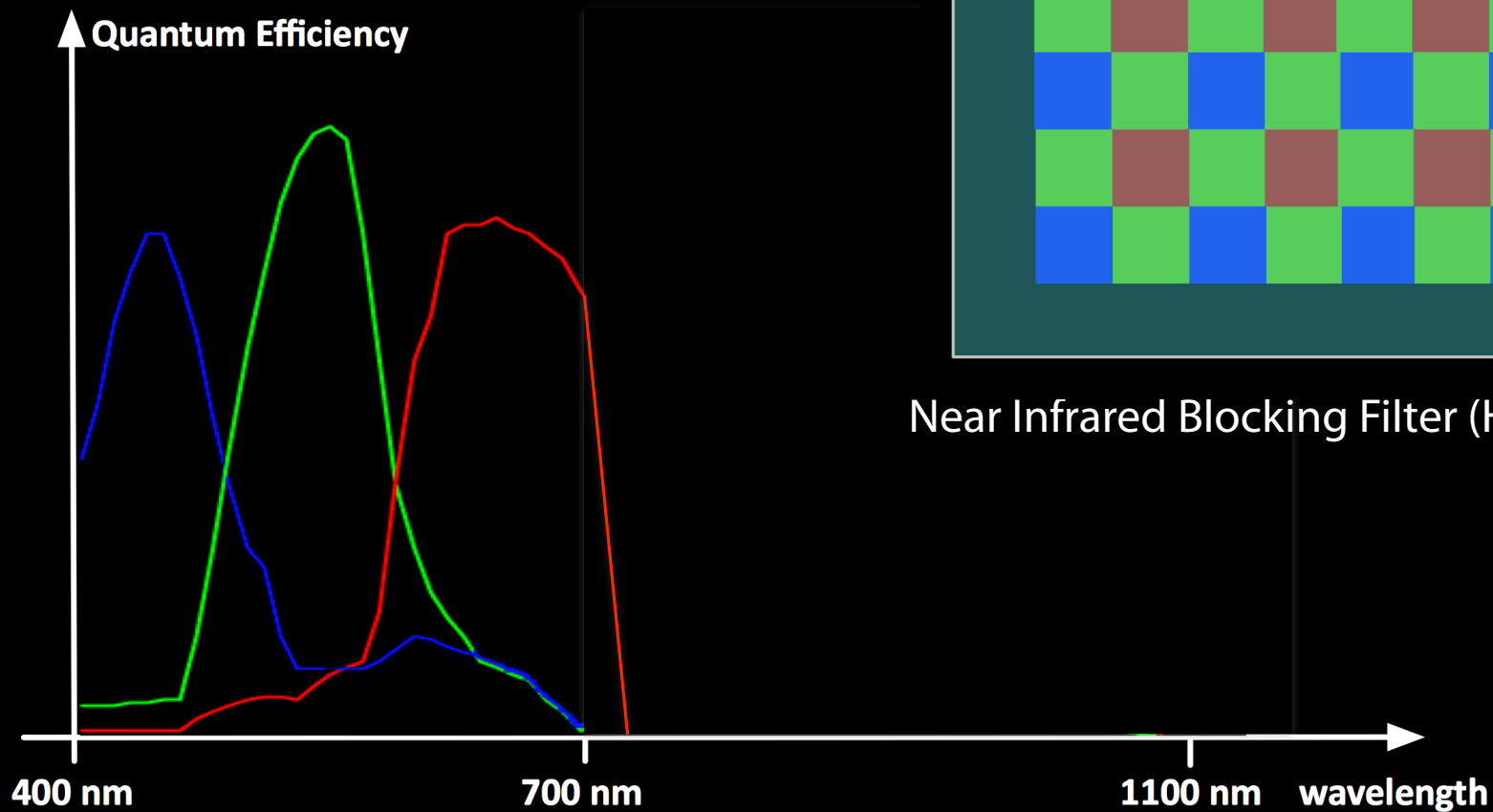
Bayer Color Filter Array (CFA)

Camera Sensitivity



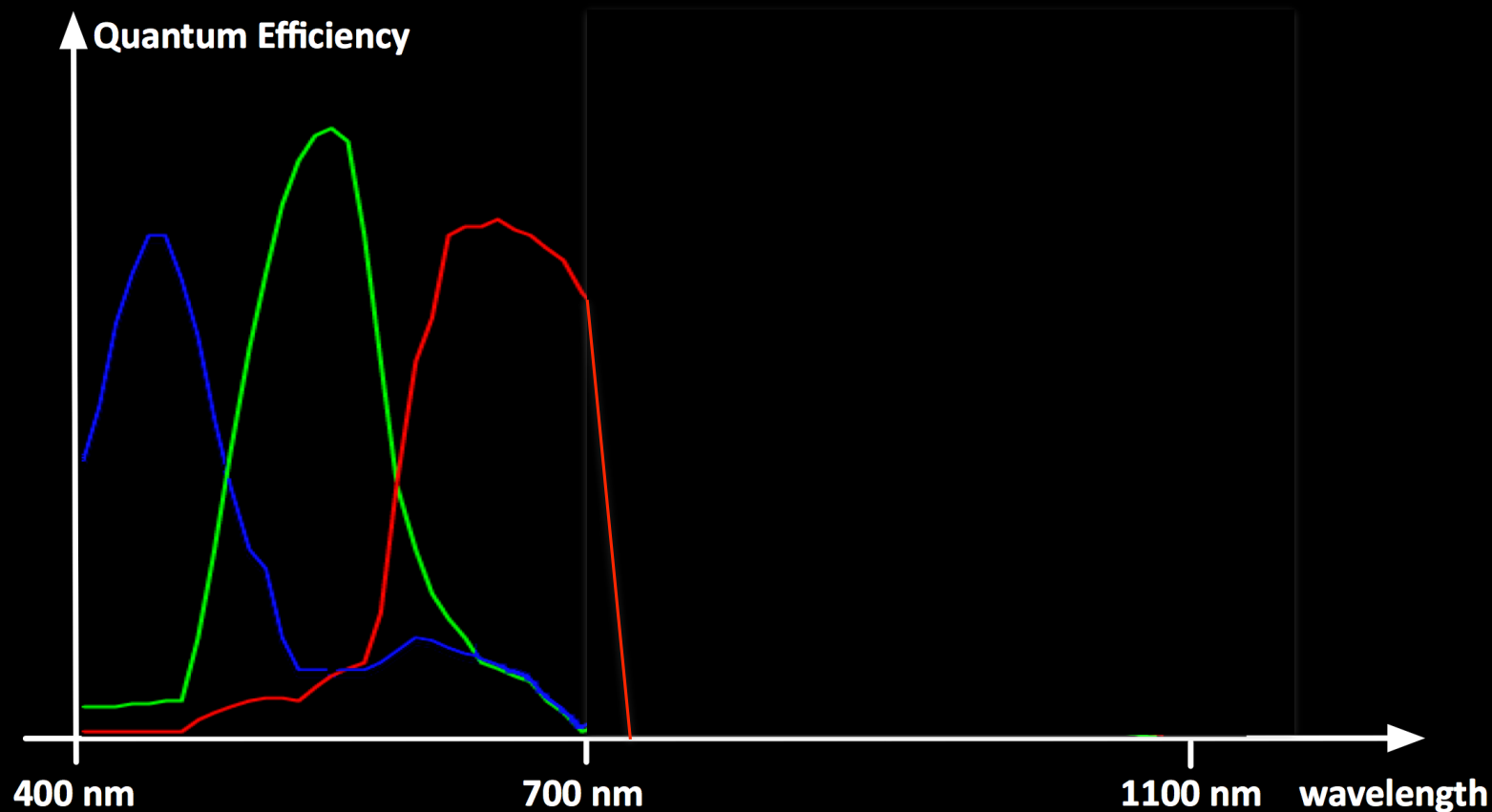
Bayer Color Filter Array (CFA)

Camera Sensitivity



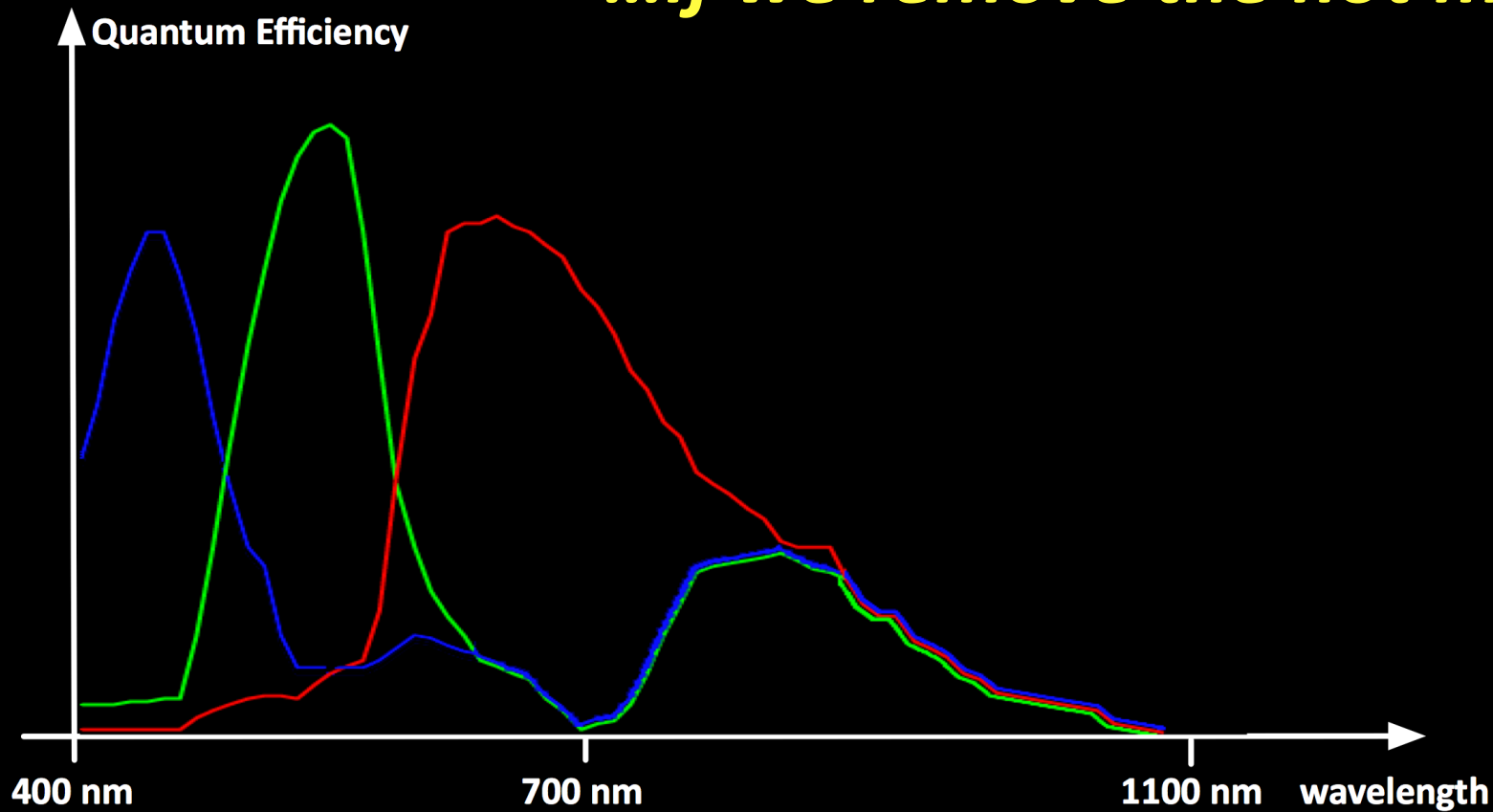
Near Infrared Blocking Filter (Hot Mirror)

All digital cameras are inherently sensitive to NIR...

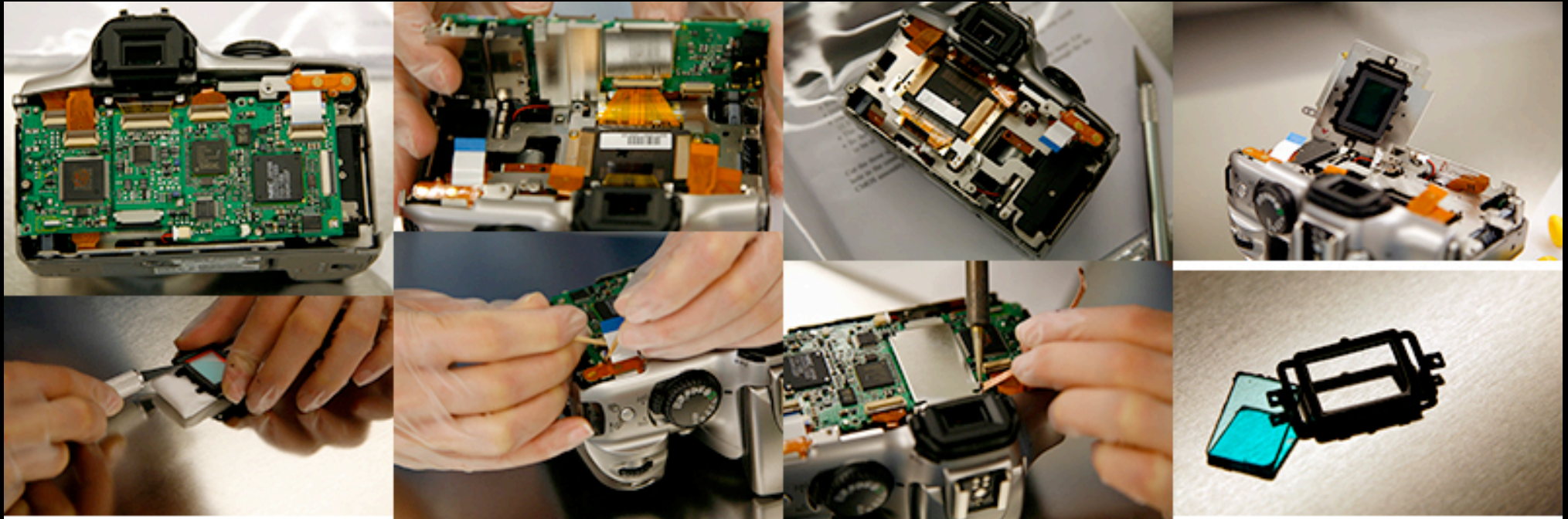


All digital cameras are inherently sensitive to NIR...

...if we remove the hot mirror!

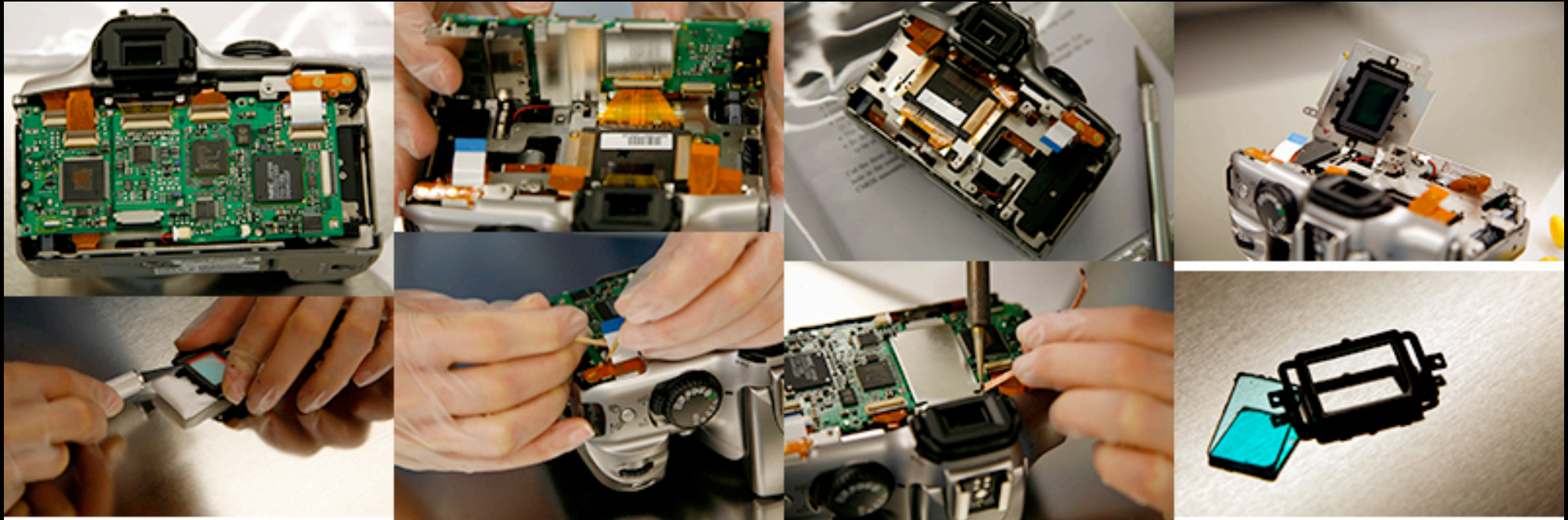


Camera construction



C. Fredembach and S. Süsstrunk, Colouring the near infrared, *IS&T/SID 16th Color Imaging Conference*, 2008.
MERL Best Student Paper Award

Camera construction



...or you buy one on-line!

Capturing visible images



+



=

Visible
Image

Visible Image



Capturing NIR images



+



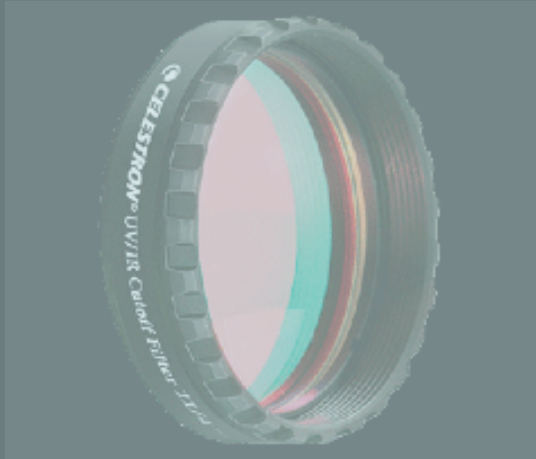
=

NIR
Image

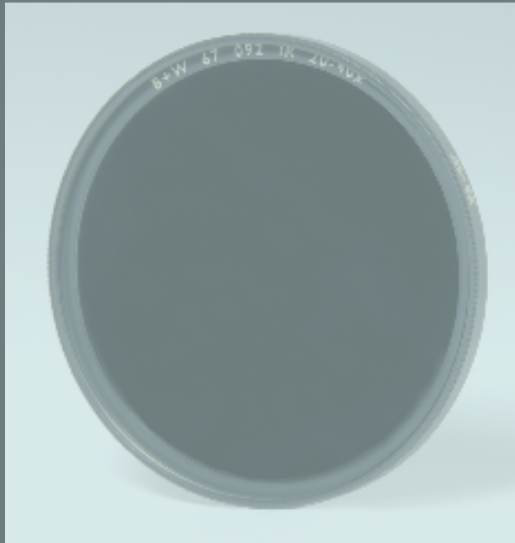
NIR Image



Capturing full spectrum images



+



+



Full Spectrum Image



Visible Image



Current Work

- *Simultaneous* capture of RGB and NIR on a *single* sensor.



“Multispectral” Color Filter Array

- Exploit spatial and spectral correlation to acquire **4-channel RGB+NIR:**

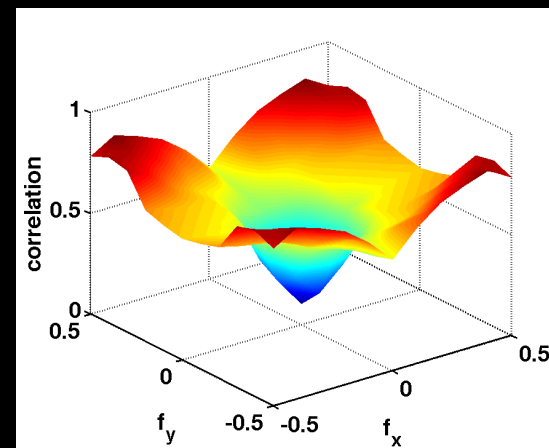
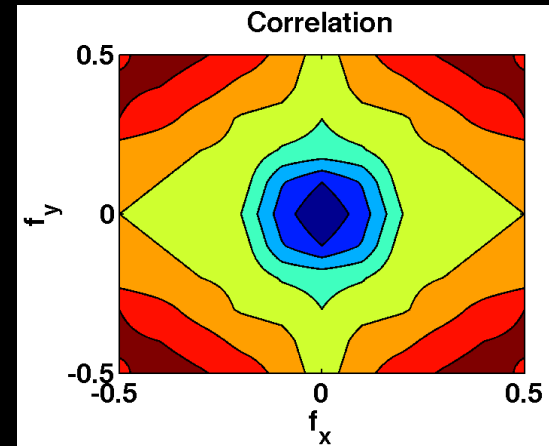


“Multispectral” Color Filter Array

- Exploit spatial and spectral correlation to acquire **4-channel RGB+NIR**:



50 RGB/NIR image pairs



“Multispectral” Color Filter Array

- The new color filter array (CFA) for **4-channel RGB + NIR** with **no hot mirror**:

$P_r = 0.62$ $P_g = 0.99$ $P_b = 0.01$	$P_r = 0.85$ $P_g = 0.37$ $P_b = 0.99$	$P_r = 0.99$ $P_g = 0.26$ $P_b = 0$	$P_r = 0.46$ $P_g = 0.88$ $P_b = 0.99$
$P_r = 0.99$ $P_g = 0.66$ $P_b = 0.11$	$P_r = 0$ $P_g = 0.64$ $P_b = 0.07$	$P_r = 0.37$ $P_g = 0.01$ $P_b = 0.97$	NIR only
$P_r = 0$ $P_g = 0.01$ $P_b = 0.99$	$P_r = 0.95$ $P_g = 0.67$ $P_b = 0$	$P_r = 0.99$ $P_g = 0.18$ $P_b = 0.99$	$P_r = 0.05$ $P_g = 0.99$ $P_b = 0$
$P_r = 0.47$ $P_g = 0.99$ $P_b = 0.98$	$P_r = 0.51$ $P_g = 0$ $P_b = 0.21$	$P_r = 0$ $P_g = 0.99$ $P_b = 0.98$	$P_r = 0.99$ $P_g = 0$ $P_b = 0.10$

Original RGB



Reconstructed RGB



► Av. CPSNR = 34.30 (38.47 for RGB only)

Original NIR



Reconstructed NIR



► Av. PSNR = 34.30

Outline

- Acquisition of RGB + NIR
- ***Enhance imaging:***
 - ***Haze removal***
 - Skin smoothing
- Extract more information about the scene:
 - Illuminant and shadow detection
 - Material-based object segmentation
 - Scene recognition
- Conclusion

EPFL

= Ecole Polytechnique Fédéral de Lausanne

= Swiss Federal Institute of Technology, Lausanne



Normal view...



Enhanced view



Haze Removal



Light scattering in the atmosphere produces haze
(Raleigh's law):

$$s_{\text{particle}} < \lambda/10 : \quad E_s \propto \frac{E_0}{\lambda^4}$$

Haze Removal



Light scattering in the atmosphere produces haze
(Raleigh's law):

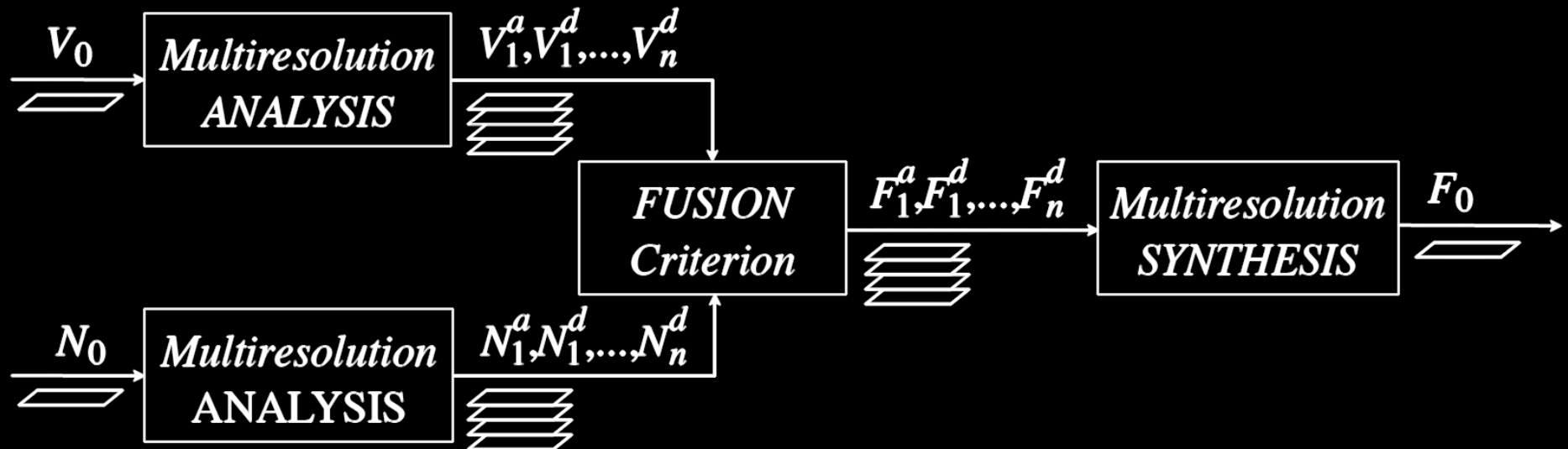
$$s_{\text{particle}} < \lambda/10 : \quad E_s \propto \frac{E_0}{\lambda^4}$$

Intuition

- The *high frequencies* of the visible image can be *exchanged* by the high frequencies of the NIR image.

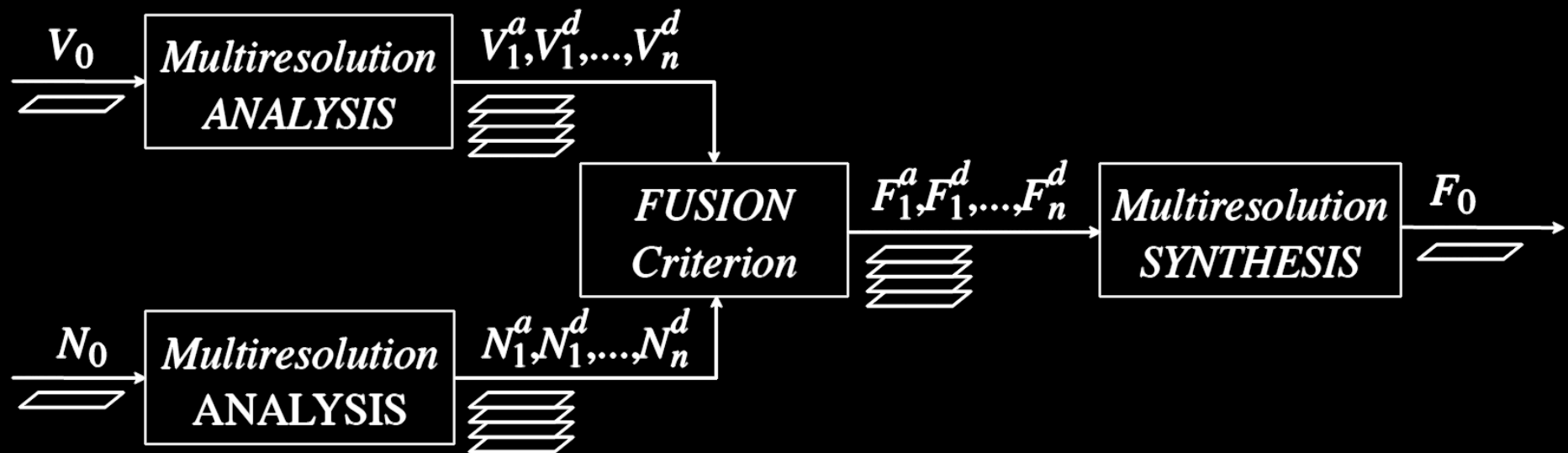
Algorithm

- Multi-resolution analysis and fusion of both visible luminance **V** and NIR intensities **N** (criterion: contrast):



Algorithm

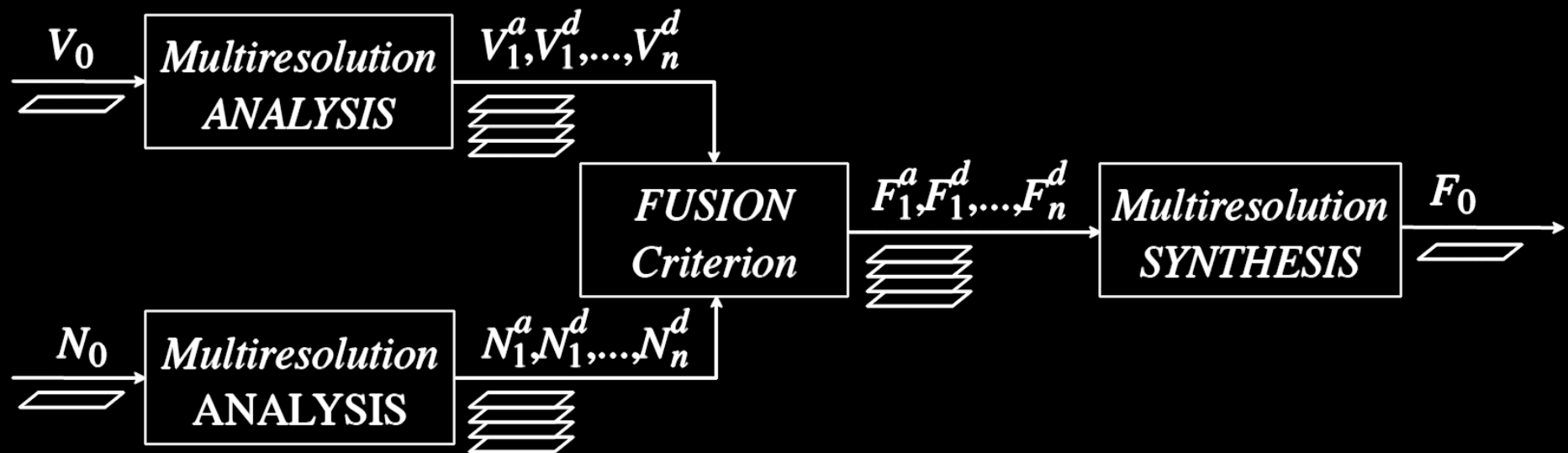
- Multi-resolution analysis and fusion of both visible V and NIR N images.



- Analysis Criterion: **Weber Contrast**
$$C = \frac{Y - Y_s}{Y_s}$$

Algorithm

- Multi-resolution analysis and fusion of both visible V and NIR N images.



- Analysis Criterion:
$$I_k^d = \frac{I_k^a - I_{k+1}^a}{I_{k+1}^a}$$

Synthesis

- Keep the visible low frequency information, and then retain the **highest contrast** from either the visible or the NIR image.

$$F_0 = V_n^a \prod_{k=1} (\max(V_k^d, N_k^d) + 1)$$

Results



Results



Original



Unsharp Masking

L. Schaul, C. Fredembach, and S. Süsstrunk, Color image de-hazing using the near-infrared, *IEEE International Conference on Image Processing (ICIP)*, 2009.

Results



Original



De-hazing using NIR

L. Schaul, C. Fredembach, and S. Süsstrunk, Color image de-hazing using the near-infrared, *IEEE International Conference on Image Processing (ICIP)*, 2009.

Results



R. Fattal, Single image dehazing, *Proc. International Conference on Computer Graphics and Interactive Techniques*, 2008.

K. He, J. Sun, and X. Tang, Single image haze removal using dark channel prior, *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2009.

High Dynamic Range



X. Zhang, T. Sim, and X. Miao, Enhancing Photographs with Near Infrared Images, *Proc. IEEE Computer Vision and Pattern Recognition (CVPR)*, 2008.

High Dynamic Range



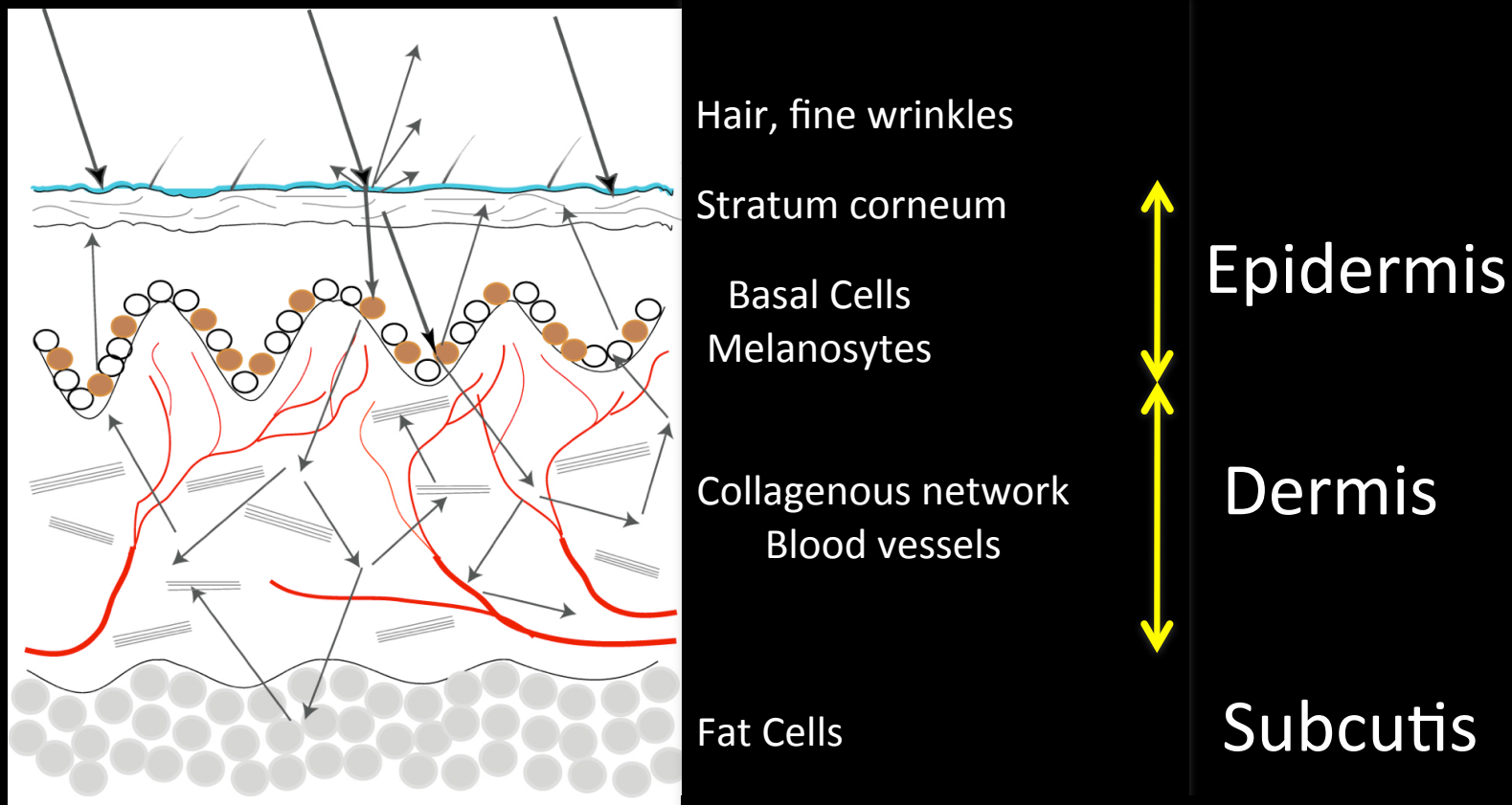
X. Zhang, T. Sim, and X. Miao, Enhancing Photographs with Near Infrared Images, *Proc. IEEE Computer Vision and Pattern Recognition (CVPR)*, 2008.

Outline

- Acquisition of RGB + NIR
- Enhance imaging:
 - Haze removal
 - *Skin smoothing*
- Extract more information about the scene:
 - Illuminant and shadow detection
 - Material-based object segmentation
 - Scene recognition
- Conclusion

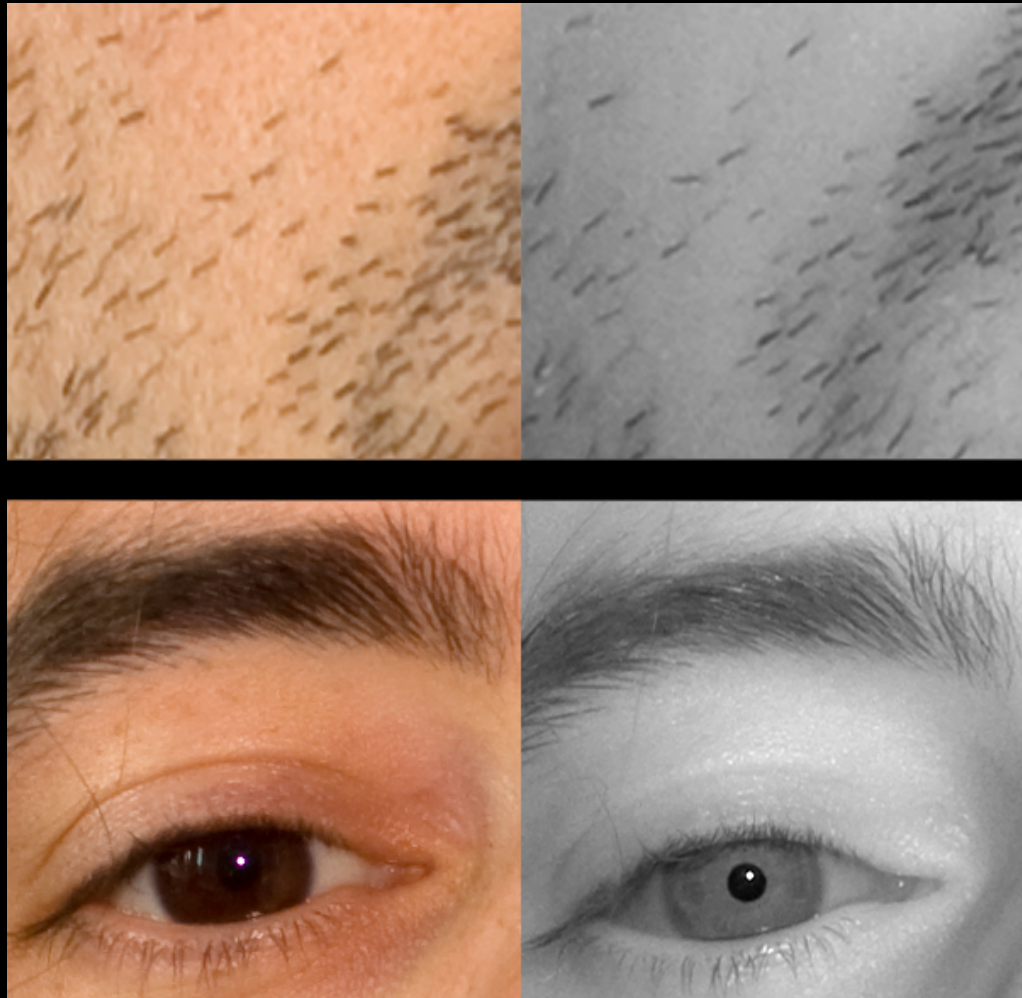
Penetration of radiation in skin

- **Absorption** and **scattering** are inversely proportional to wavelength.



Intuition

- If NIR penetrates deeper into skin, then **surface** imperfections are **less** visible.



Intuition

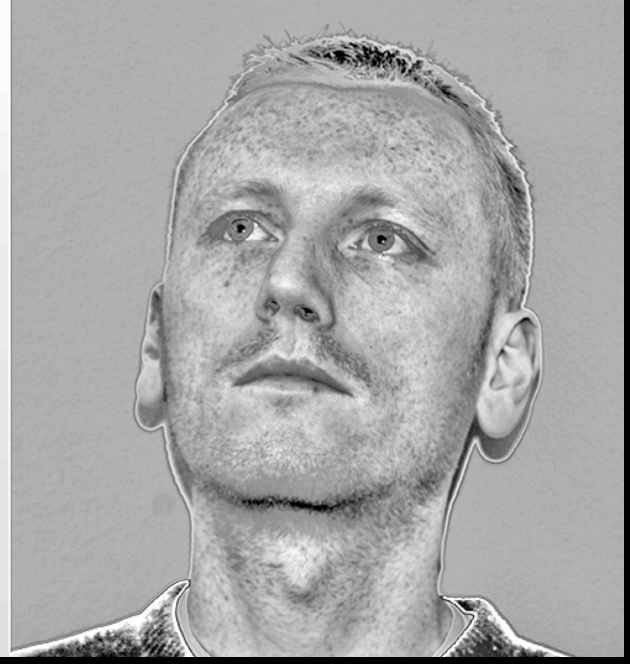
- Unwanted skin imperfections (freckles, pores, warts, wrinkles) are usually *small*.



Original

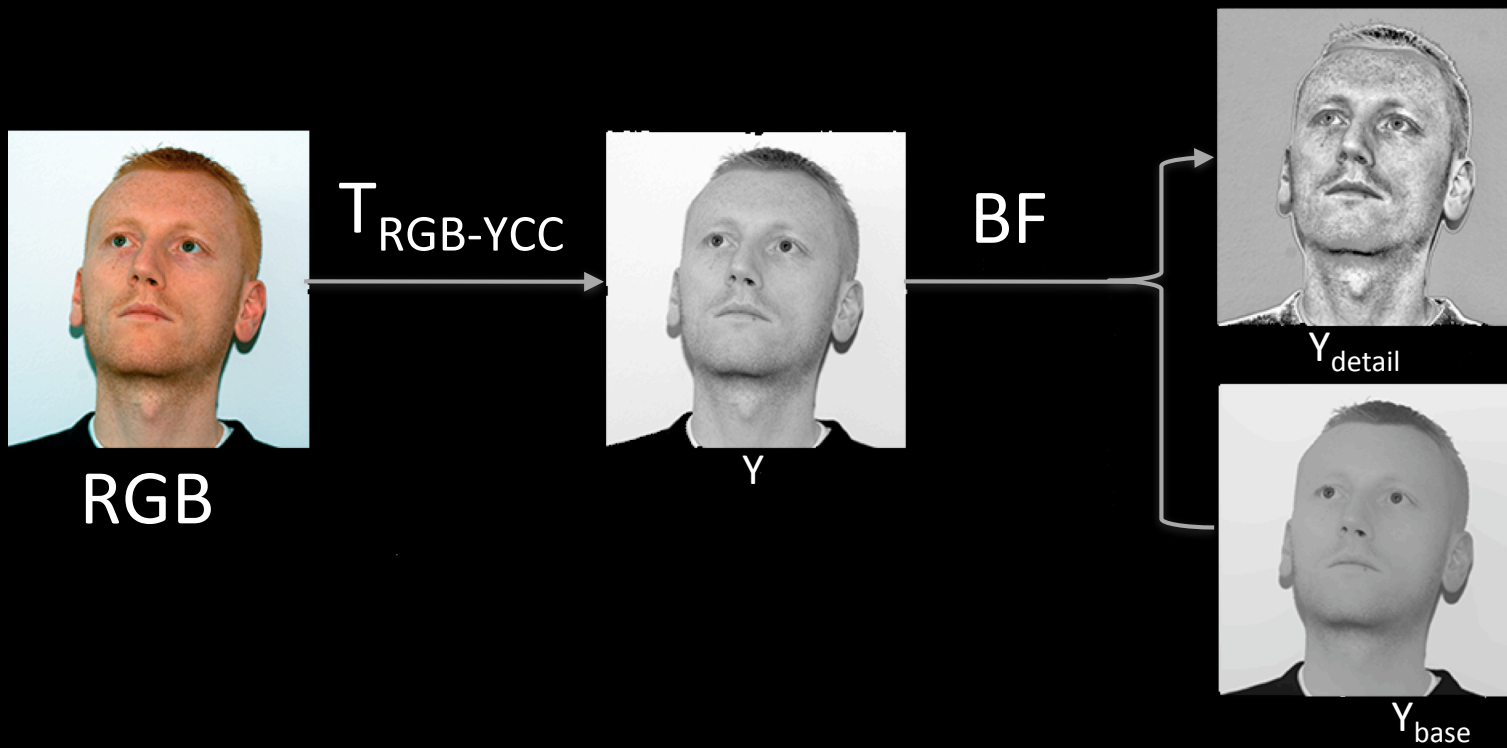


Base



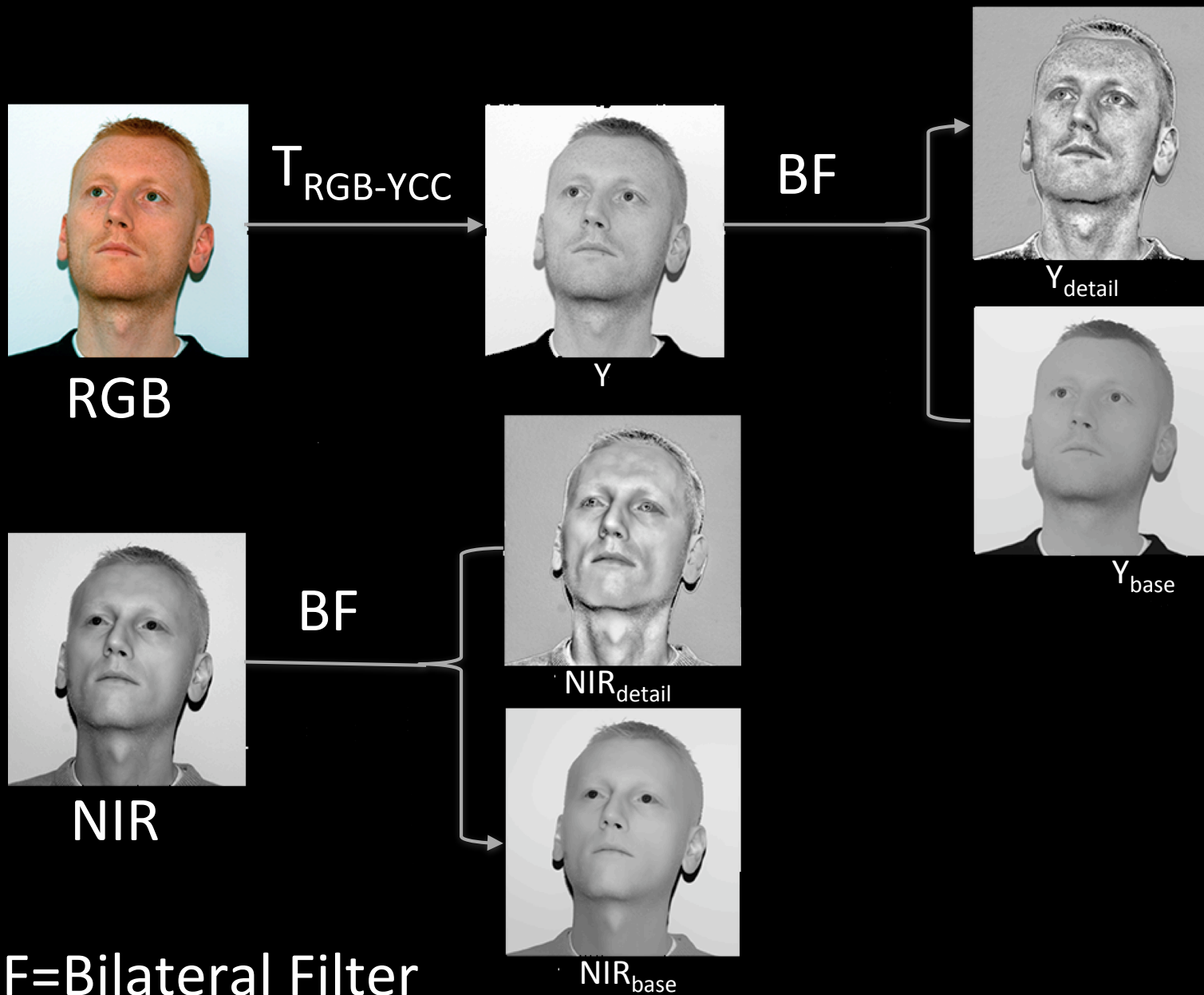
Detail

Algorithm

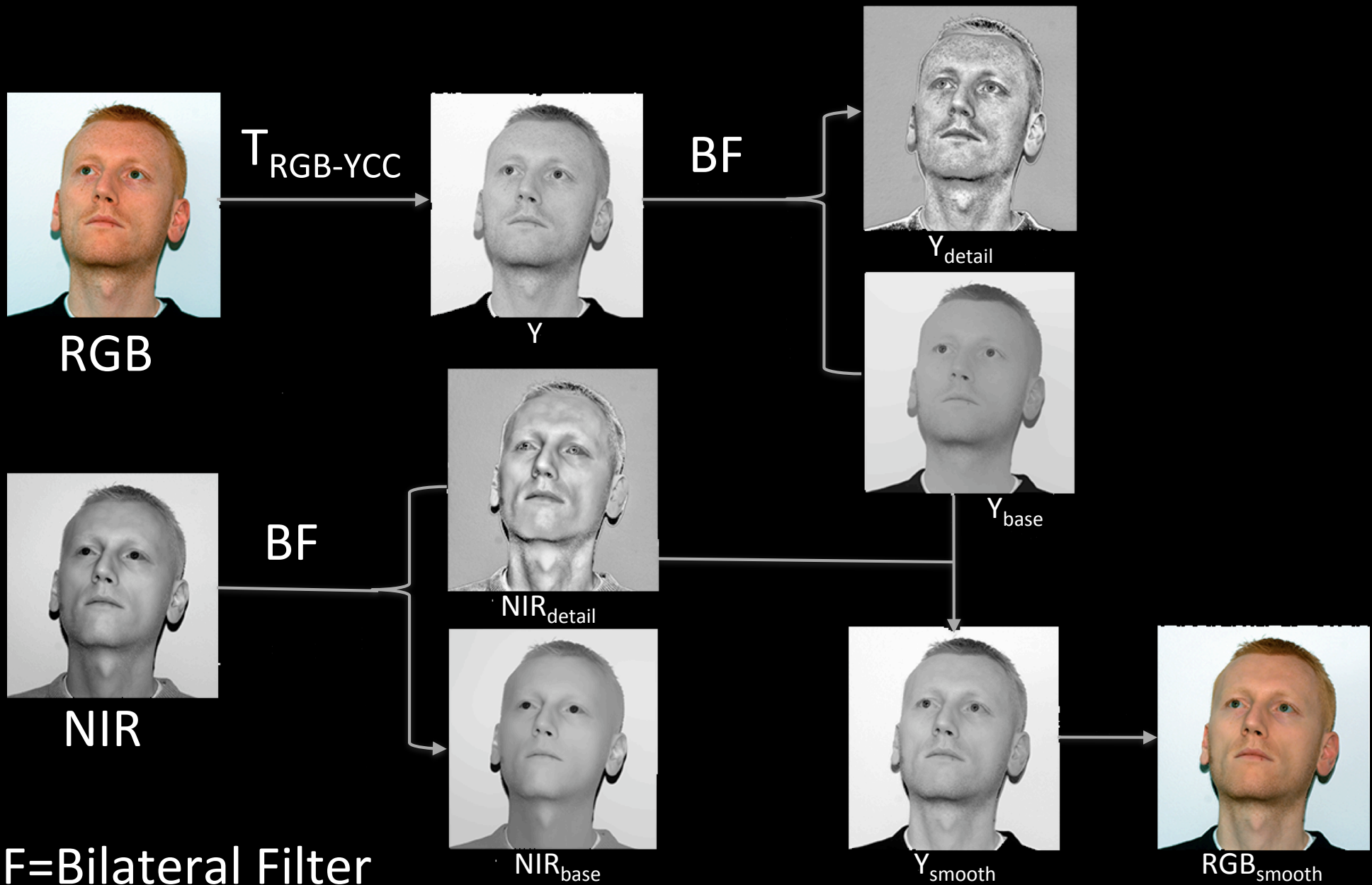


BF=Bilateral Filter

Algorithm



Algorithm



Skin Smoothing



Visible



NIR



Visible + NIR

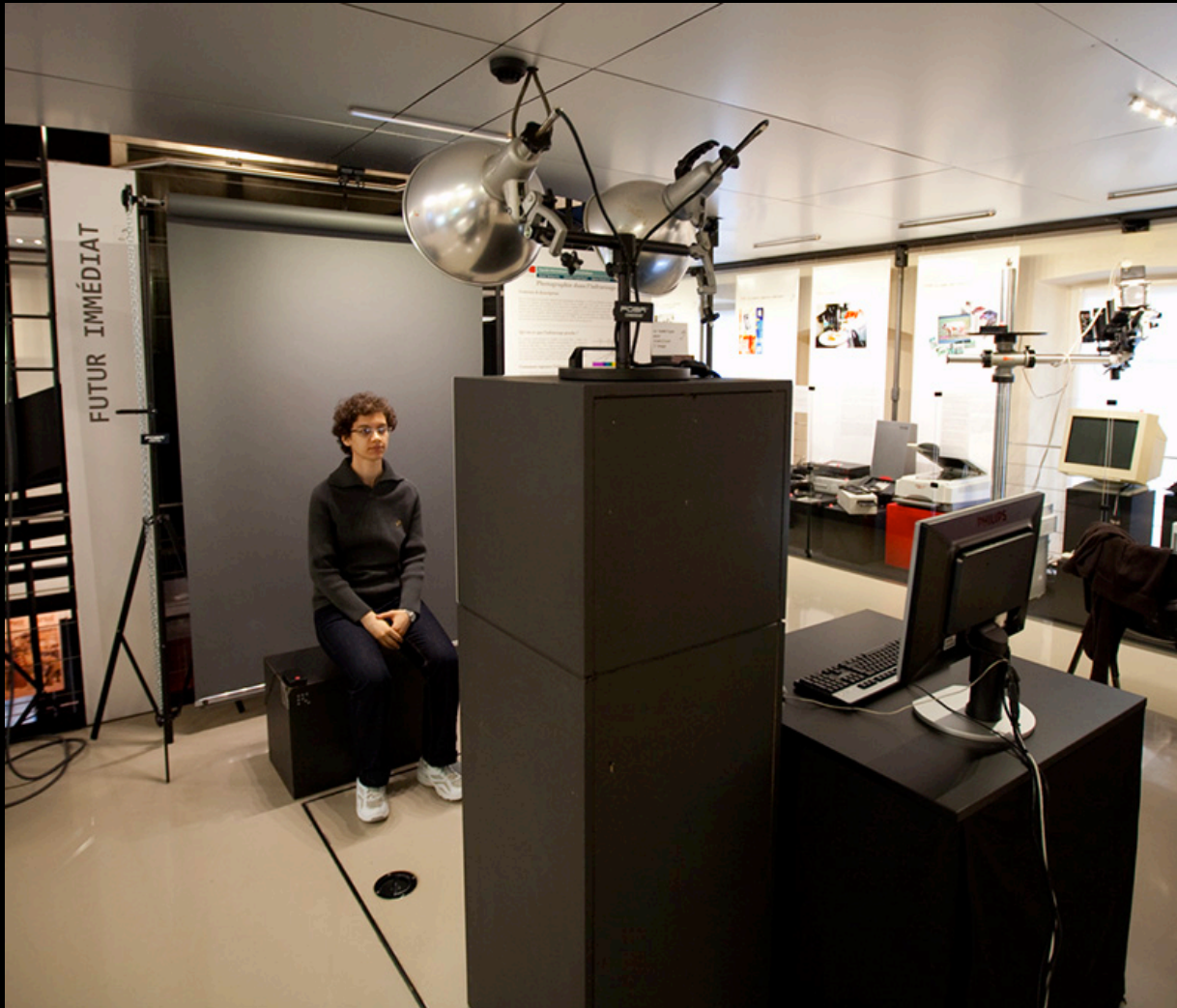
C. Fredembach, N. Barbuscia, and S. Süsstrunk, Combining visible and near-infrared images for realistic skin smoothing, *IS&T/SID 17th Color Imaging Conference*, 2009.

Skin Smoothing



C. Fredembach, N. Barbuscia, and S. Süsstrunk, Combining visible and near-infrared images for realistic skin smoothing, *IS&T/SID 17th Color Imaging Conference*, 2009.

Camera Museum in Vevey



S. Süsstrunk, C. Fredembach, and D. Tamburrino, Automatic skin enhancement with visible and near-infrared image fusion, *ACM Multimedia*, 2010. **Best Demo Paper Award**

Camera Museum in Vevey

Visible



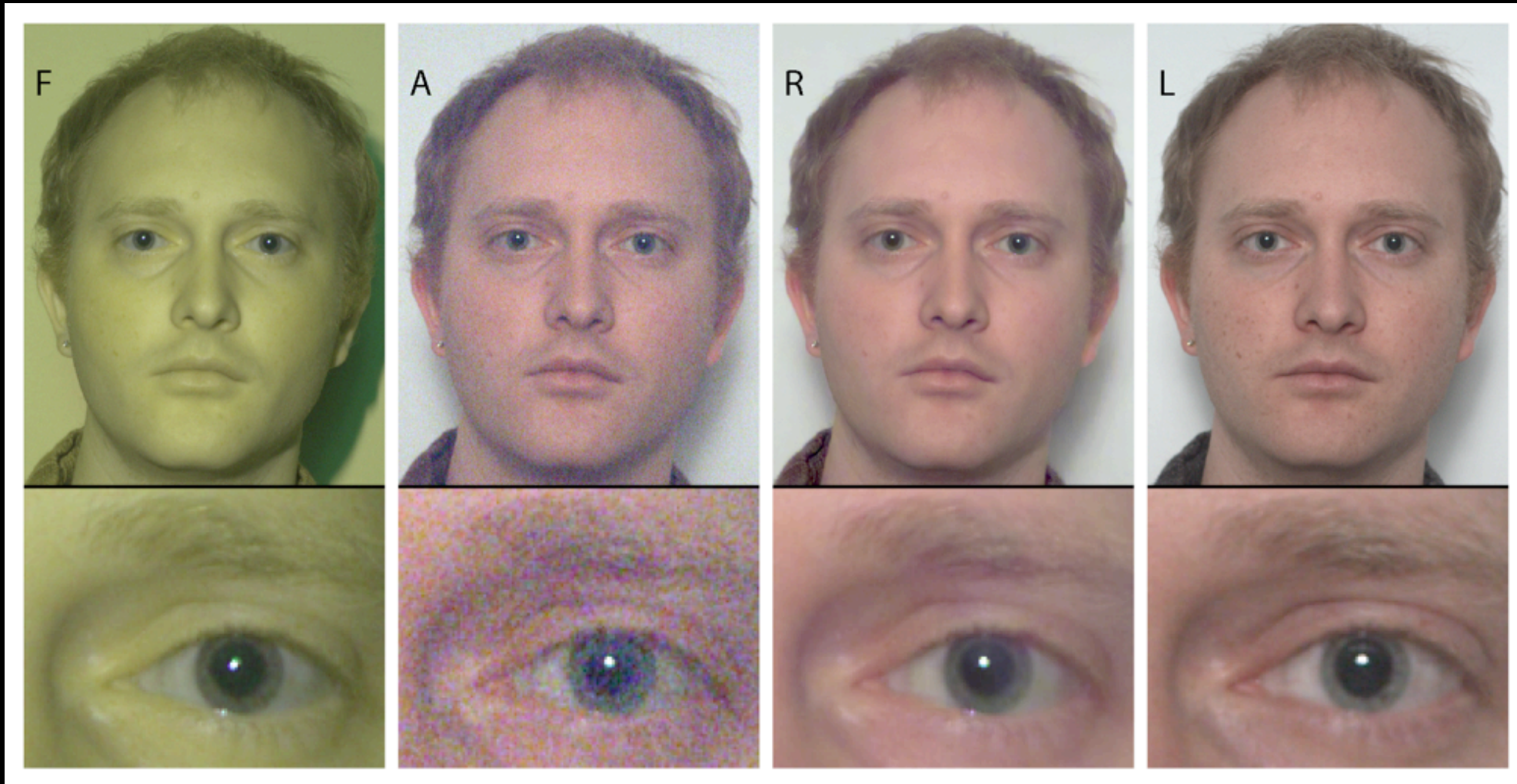
Preference: 25.4% (1,235)

Visible + NIR



Preference: 74.6% (3,623)

Dark Flash Photography



NIR Flash

Short Exp.

Reconstructed

Ground Truth

Video Re-Lighting



C. Wu, R. Samadani, P. Gunawardane, Same frame rate IR to enhance visible video conference lighting, *Proc. IEEE International Conference on Image Processing (ICIP)*, 2011.

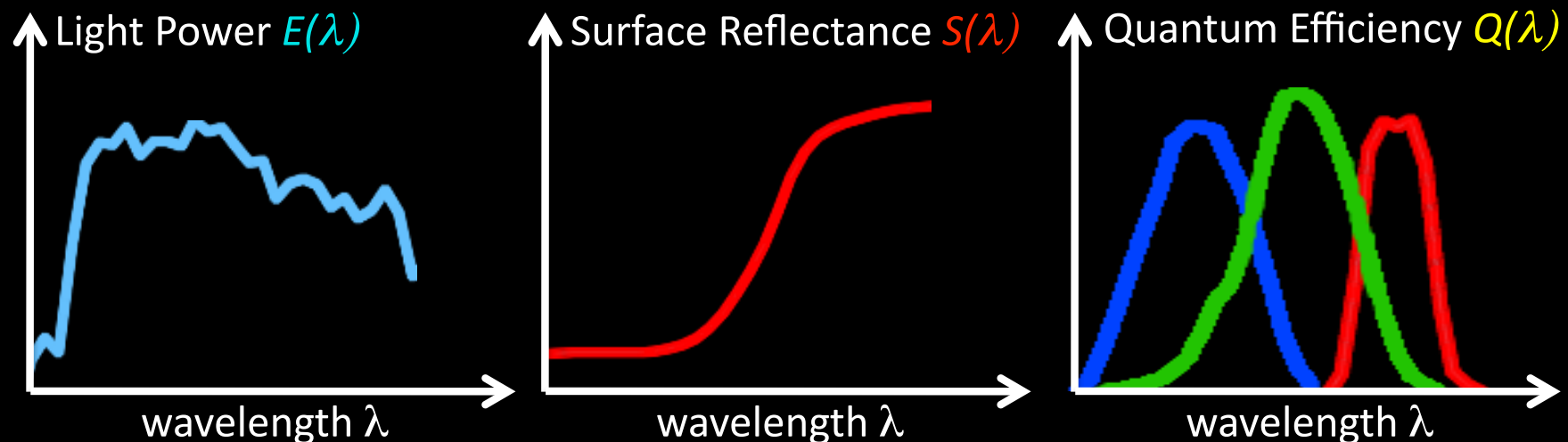
Outline

- Acquisition of RGB + NIR
- Enhance imaging:
 - Haze removal
 - Skin smoothing
- ***Extract more information about the scene:***
 - ***Illuminant and shadow detection***
 - Material-based object segmentation
 - Scene recognition
- Conclusion

Recall

- Separating the effect of illuminant and surface reflectance of the camera response (pixel value) is an *ill-posed* problem.

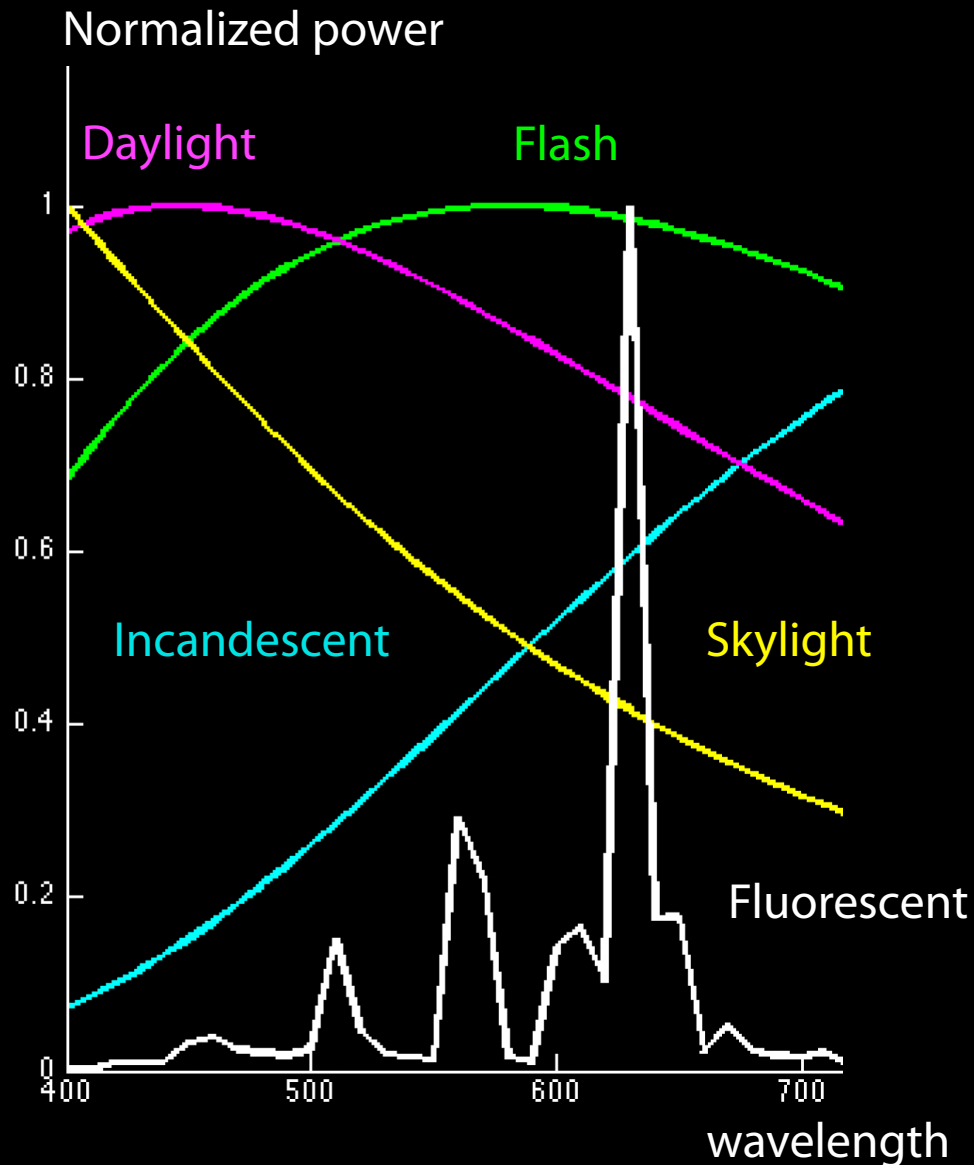
$$p_k = \int_{\lambda=400nm}^{700nm} E(\lambda) S(\lambda) Q_k(\lambda) d\lambda$$



Intuition

- ***Doubling*** the spectral bandwidth considered (400-1100nm) increases estimation accuracy for color constancy algorithms.

Light Sources in the visible spectrum



Daylight



Skylight



Flashlight



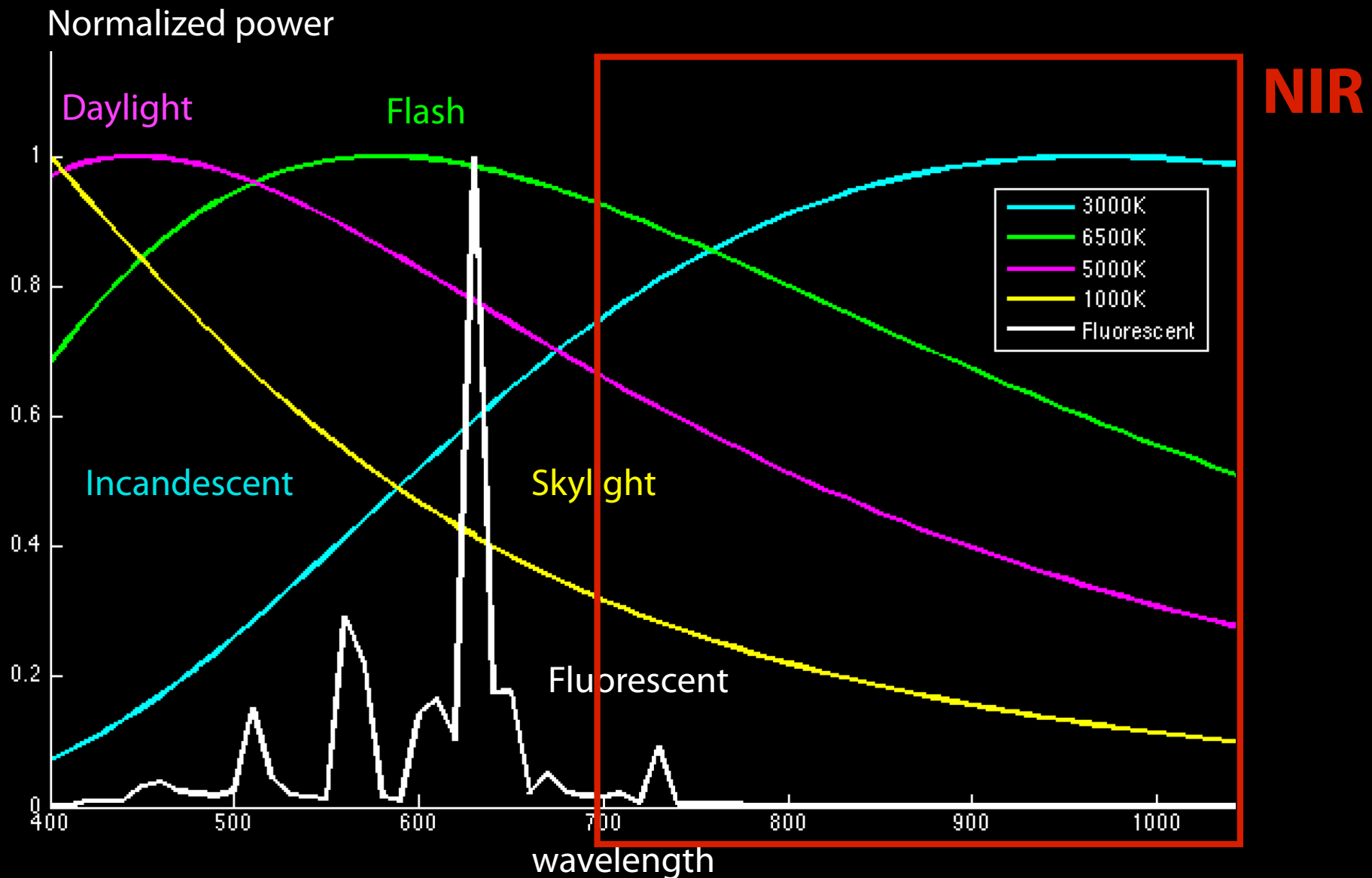
Incandescent



Fluorescent



Light Sources (visible + NIR)



Light Source Estimation



Overcast Sky
Estimated Color Temperature:
7500K



Sunny Day
Estimated Color Temperature:
5000K



Incandescent Lamp
Estimated Color Temperature:
3000K



Shadow Detection



Visible

NIR

Binary Shadow Mask

C. Fredembach and S. Süsstrunk, Automatic and accurate shadow detection from (potentially) a single image using near-infrared information, EPFL Tech Report 165527, 2010.

Outline

- Acquisition of RGB + NIR
- Enhance imaging:
 - Haze removal
 - Skin smoothing
- Extract more information about the scene:
 - Illuminant and shadow detection
 - ***Material-based object segmentation***
 - Scene recognition
- Conclusion

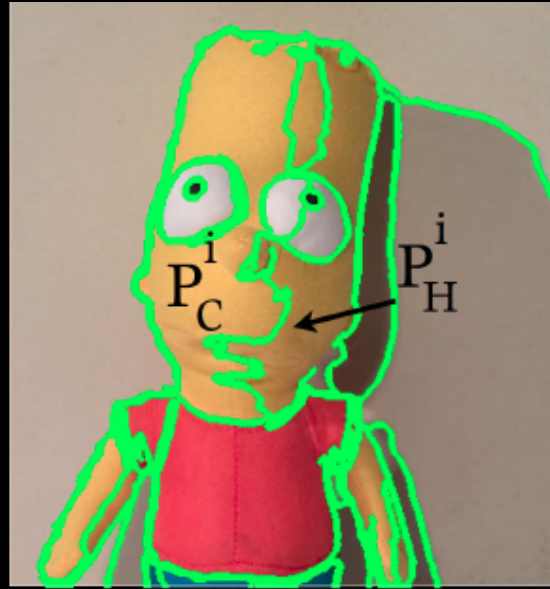
Material Reflectances



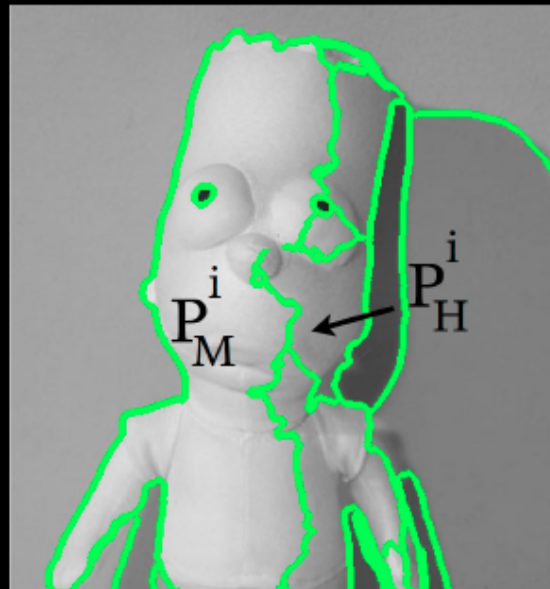
Yellow Arrows: same color in RGB, different intensity in NIR

Green Arrow: different colors in RGB, same intensity in NIR

Image Segmentation



P_C : **Color** segments
 P_H : **Shadow** segments



P_M : **Material** segments
 P_H : **Shadow** segments

Intrinsic Image (RGB + NIR)

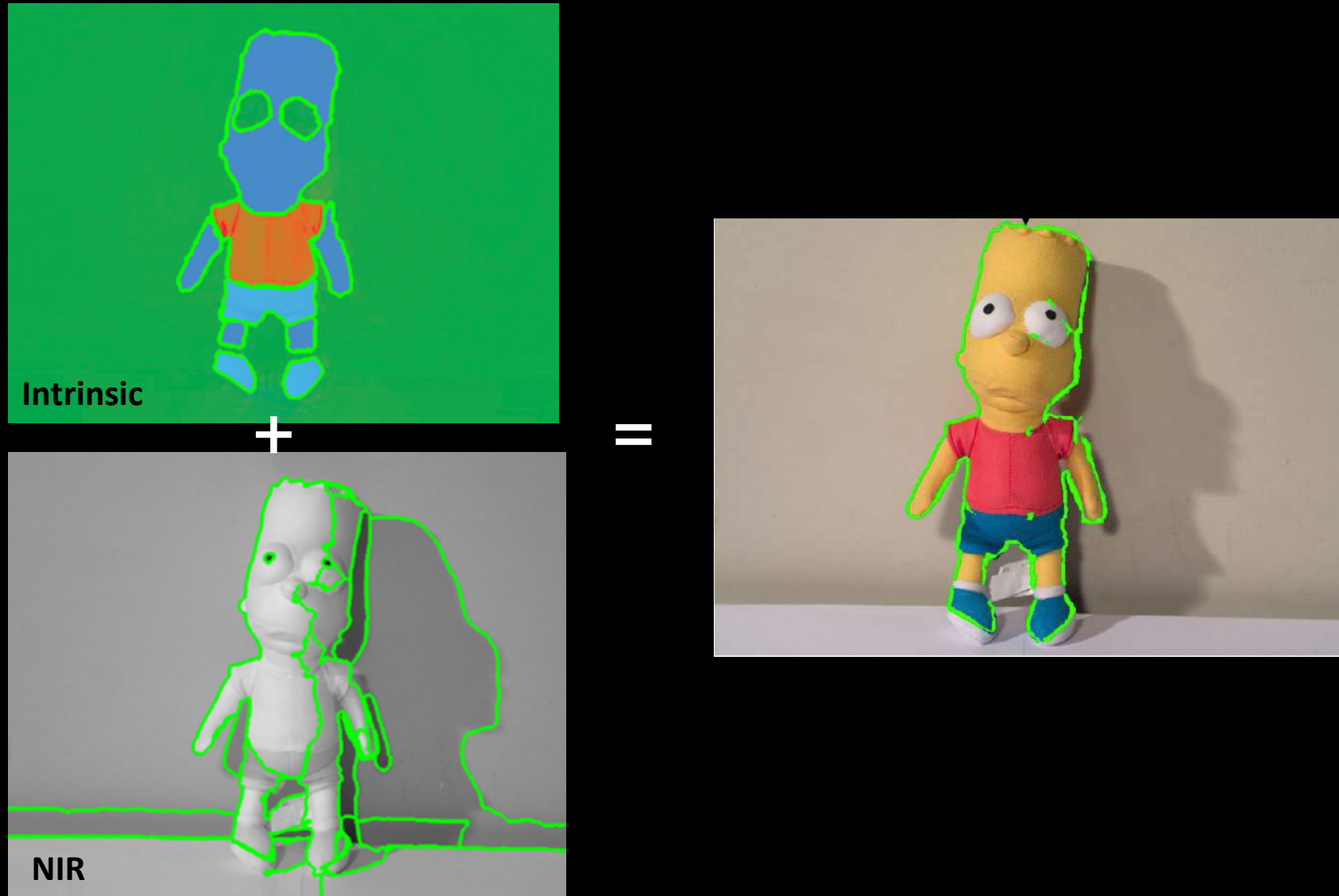


+

=

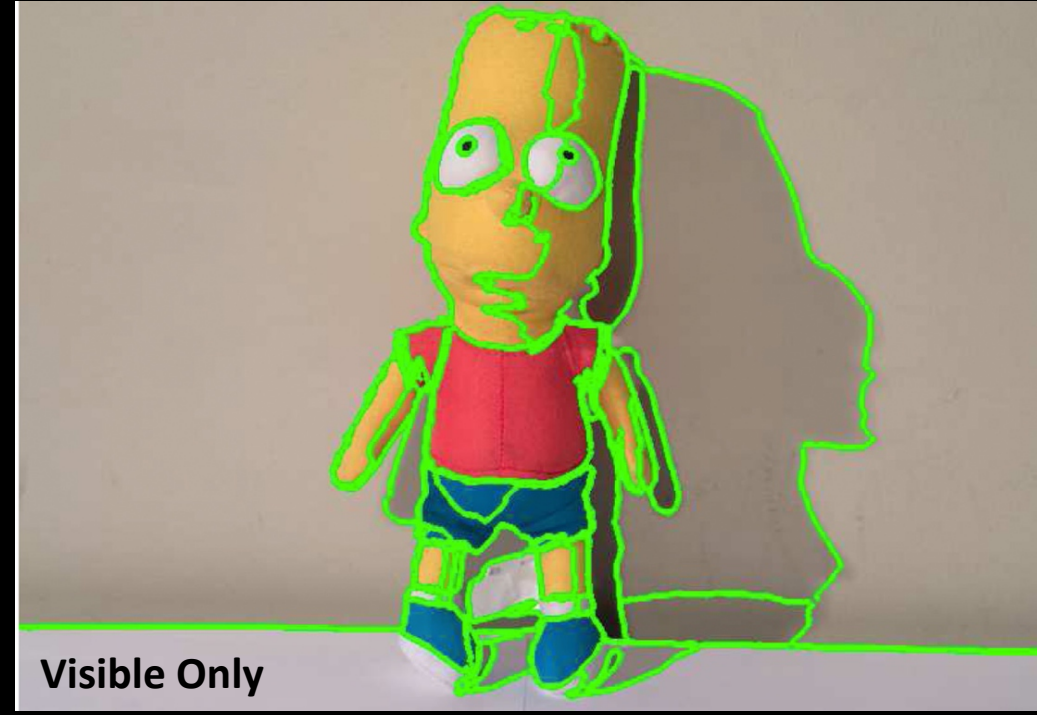


Image Segmentation



N. Salamati and S. Süsstrunk, Material-Based Object Segmentation using Near-Infrared Information, *IS&T/SID 18th Color Imaging Conference*, 2010. **Best Interactive Paper Award**

Comparison with visible only



Outline

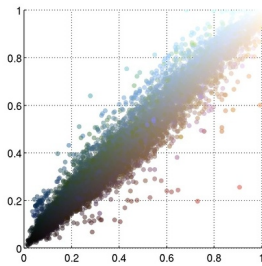
- Acquisition of RGB + NIR
- Enhance imaging:
 - Haze removal
 - Skin smoothing
- Extract more information about the scene:
 - Illuminant and shadow detection
 - Material-based object segmentation
 - ***Scene recognition***
- Conclusion

Image Database (477 RGB+NIR)

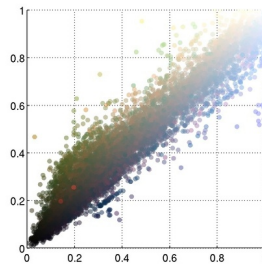


Correlation between RGB and NIR

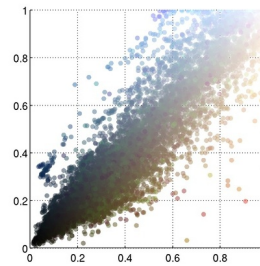
- 10'000 RGB-NIR pixel pairs, joint entropy:



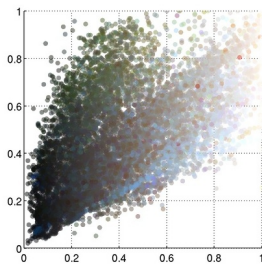
(a) R-G (4.13)



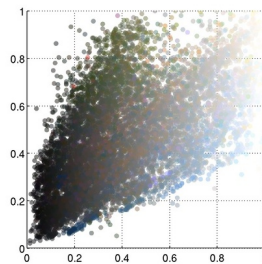
(b) B-G (4.26)



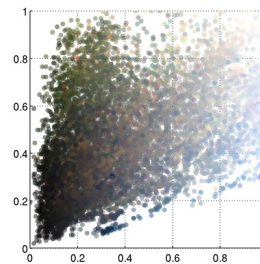
(c) R-B (4.60)



(d) R-NIR (5.29)



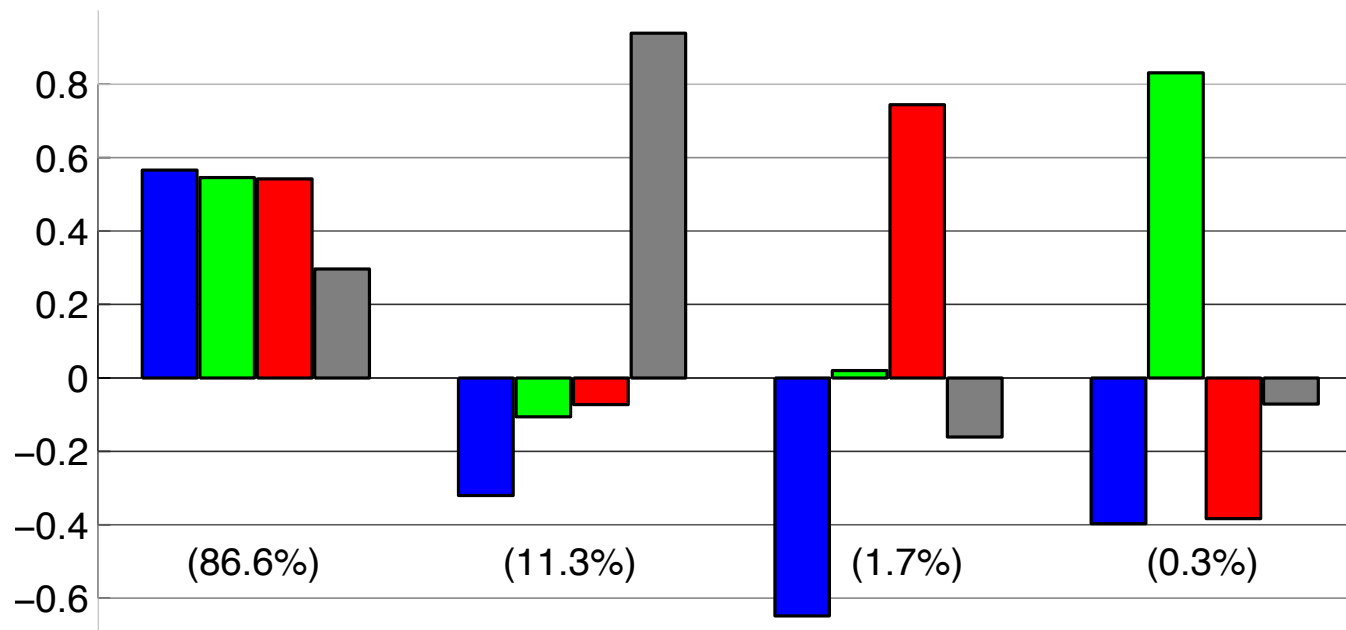
(e) G-NIR (5.30)



(f) B-NIR (5.44)

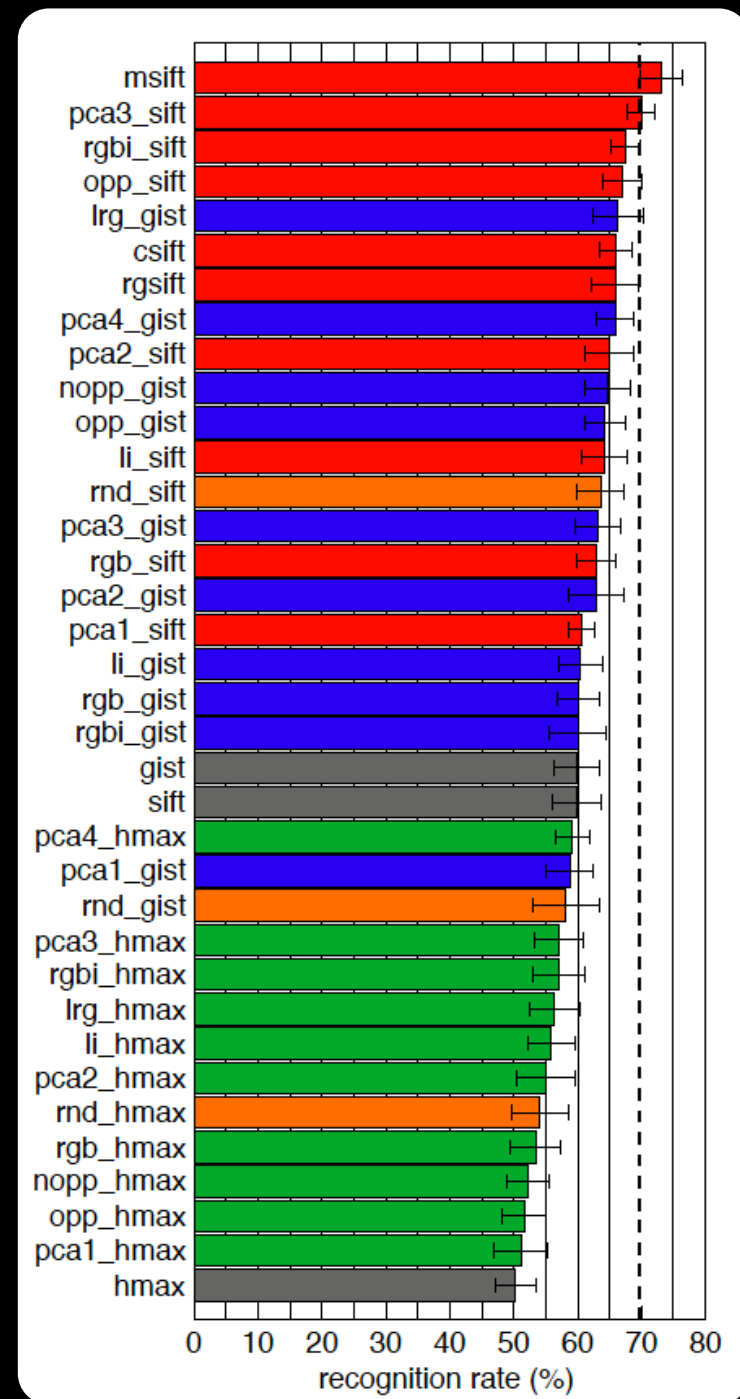
Decorrelation...

- PCA on 477 4-channel images (RGB + NIR):



Multi-spectral SIFT

- 477 RGB+NIR images
 - 9 scene categories
- **SIFT**, **GIST**, **HMAX** descriptors.
- 11 random images per class for training, rest for testing.
 - Repeated 10 times
- **MSIFT** (=pca 4) statistically significantly **outperforms** the others.



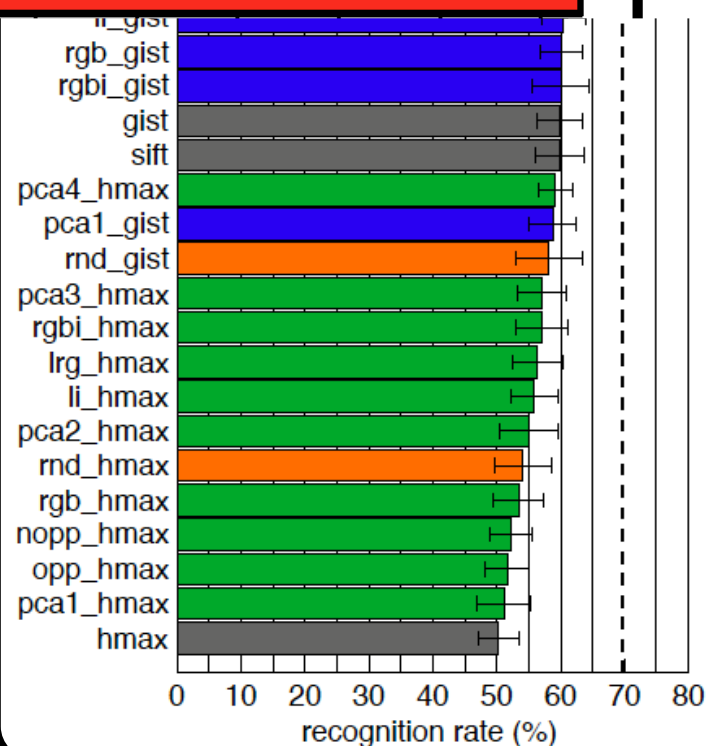
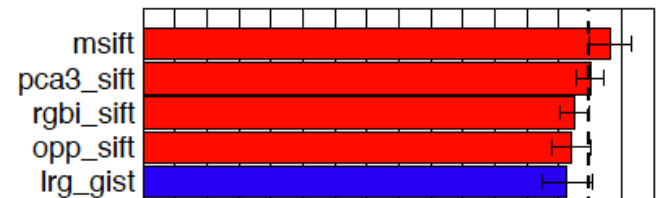
Multi-spectral SIFT

msift
pca3_sift
rgbi_sift
opp_sift

training, rest for testing.

– Repeated 10 times

- **MSIFT** (=pca 4) statistically significantly **outperforms** the others.



Better Features and Classifiers

Classifier	Descriptor1	Descriptor2	Fusion type	Accuracy
FV + SVM, with NIR	$COL_{r,g,b}$	$SIFT_{l,n}$	EF	$84.3 \pm (1.7)$
			LF	$87.4 \pm (2.7)$
		$SIFT_n$	EF	$84.1 \pm (2.6)$
			LF	$86.2 \pm (2.0)$
	$COL_{r,g,b,n}$	$SIFT_{l,n}$	EF	$84.4 \pm (2.8)$
			LF	$85.5 \pm (2.1)$
		$SIFT_n$	EF	$84.1 \pm (2.9)$
			LF	$86.5 \pm (2.4)$
	$COL_{p1,p2,p3}$	$SIFT_{l,n}$	EF	$83.3 \pm (3.4)$
			LF	$86.7 \pm (2.7)$
		$SIFT_n$	EF	$82.9 \pm (2.7)$
			LF	$87.5 \pm (2.4)$
	$COL_{p1,p2,p3,p4}$	$SIFT_{l,n}$	EF	$85.8 \pm (2.6)$
			LF	$86.5 \pm (2.8)$
		$SIFT_n$	EF	$83.2 \pm (3.0)$
			LF	$87.9 \pm (2.2)$
FV + SVM, no NIR	$COL_{r,g,b}$	$SIFT_l$	LF	$84.5 \pm (2.3)$
Brown and Süssstrunk [2]	—	—	—	$72.0 \pm (2.9)$

Better Features and Classifiers

Classifier	Descriptor1	Descriptor2	Fusion type	Accuracy
FV + SVM, with NIR	$COL_{r,g,b}$	$SIFT_{l,n}$	EF	$84.3 \pm (1.7)$
			LF	$87.4 \pm (2.7)$
		$SIFT_n$	EF	$84.1 \pm (2.6)$
			LF	$86.2 \pm (2.0)$
	$COL_{r,g,b,n}$	$SIFT_{l,n}$	EF	$84.4 \pm (2.8)$
			LF	$85.5 \pm (2.1)$
		$SIFT_n$	EF	$84.1 \pm (2.9)$
			LF	$86.5 \pm (2.4)$
	$COL_{p1,p2,p3}$	$SIFT_{l,n}$	EF	$83.3 \pm (3.4)$
			LF	$86.7 \pm (2.7)$
		$SIFT_n$	EF	$82.9 \pm (2.7)$
			LF	$87.5 \pm (2.4)$
	$COL_{p1,p2,p3,p4}$	$SIFT_{l,n}$	EF	$85.8 \pm (2.6)$
			LF	$86.5 \pm (2.8)$
		$SIFT_n$	EF	$83.2 \pm (3.0)$
			LF	$87.9 \pm (2.2)$
FV + SVM, no NIR	$COL_{r,g,b}$	$SIFT_l$	LF	$84.5 \pm (2.3)$
Brown and Ssstrunk [2]	–	–	–	$72.0 \pm (2.9)$

Conclusions

- Using an **extra** (non-visible) NIR channel has many benefits in multimedia applications.
 - ***Enhance the content***
 - ***Extract more accurate information about the scene.***

Other RGB-NIR Applications

- Microsoft Kinect [www.primesense.com]



- Projector uses 830nm (NIR) laser
- NIR camera is tuned to this frequency
- Projector-camera system to capture depth (stereo)

More Physics...

- Cameras (3D)
- Microphones
- GPS
- Accelerometer
- Digital Compass
- Ambient Light
- Temperature?
- Air Quality?
- ...



Photo Courtesy Chris Jordan

Conclusions

- Using an *extra* (non-visible) NIR channel has many benefits in multimedia applications.
 - *Enhance the content*
 - *Extract more accurate information about the scene.*
- There are many sensors providing **physical measurements** that can improve imaging applications.
 - ***And we have only begun to exploit their possibilities.***

Thank you!

sabine.susstrunk@epfl.ch

<http://ivrg.epfl.ch/publications/>

