RGB++

How "Side Information" Improves Computational Photography and Computer Vision

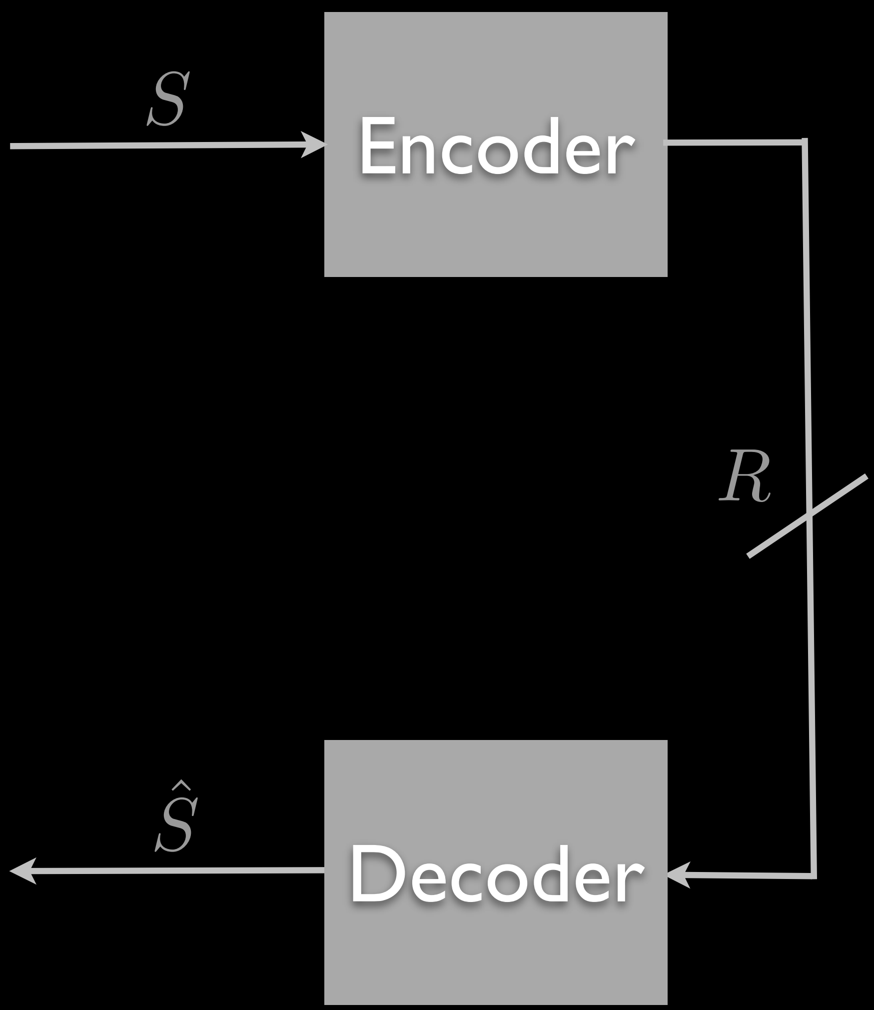
Sabine Süsstrunk
Images and Visual Representation Group (IVRG)

This talk is about linking Information Theory (or Signal Processing) with…

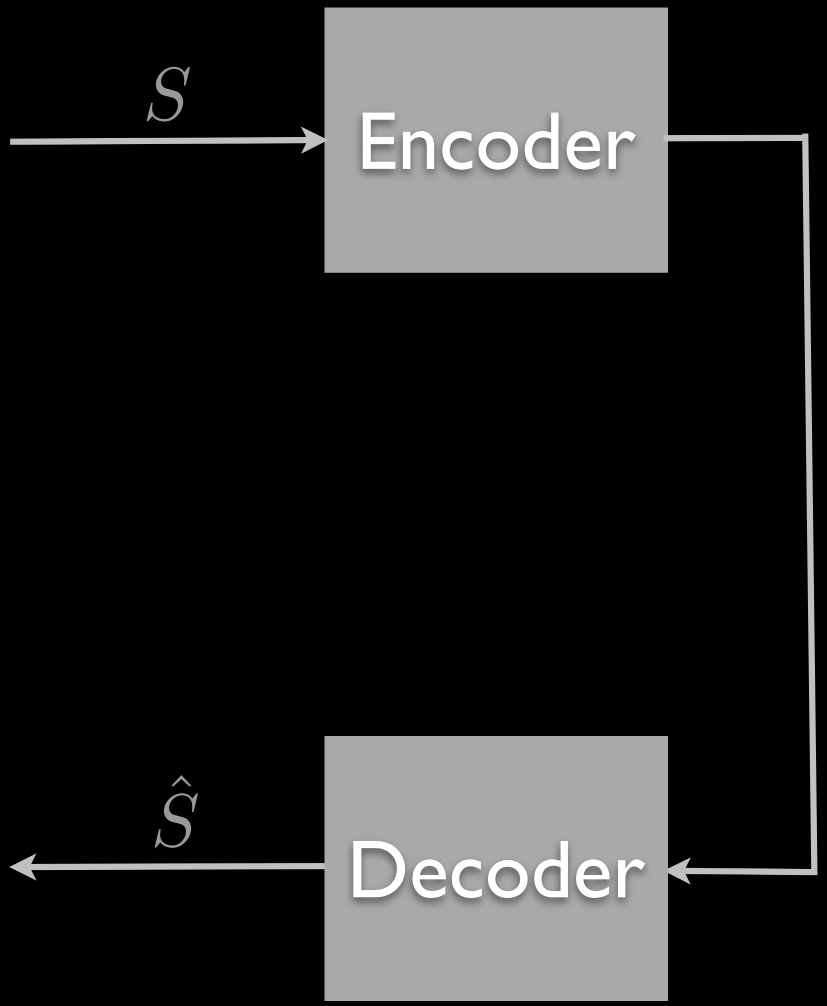
Photography

How the world has changed:
St. Peter's Square in 2005 and 2013

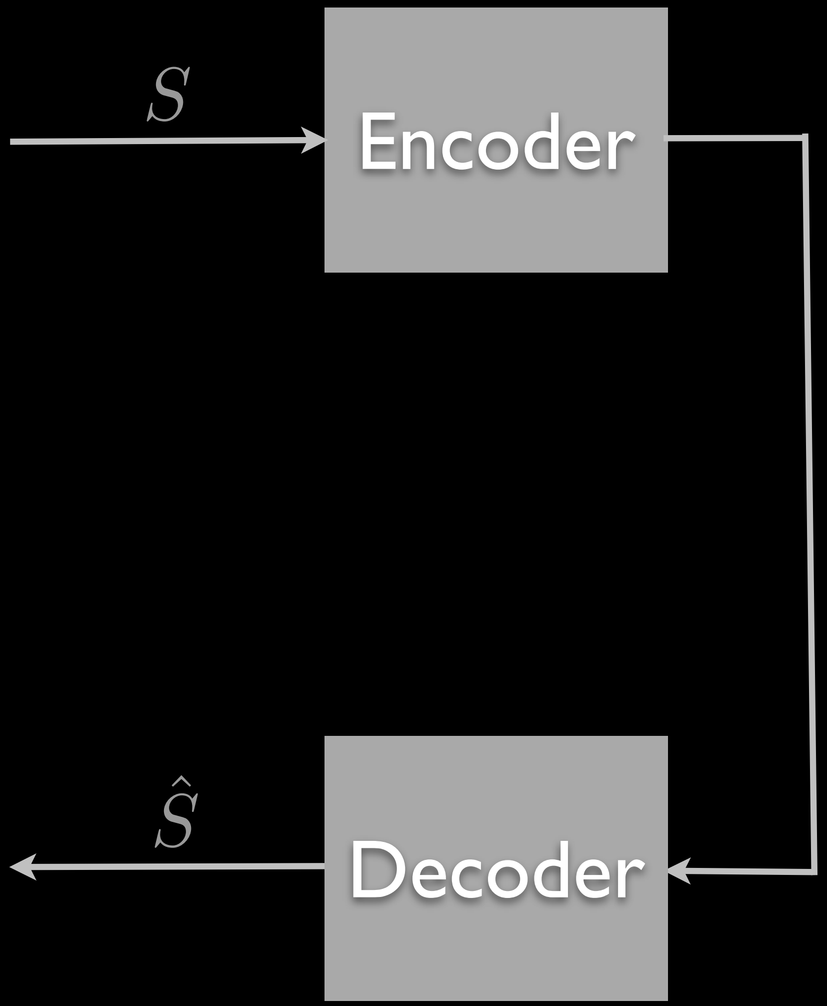
Source Coding



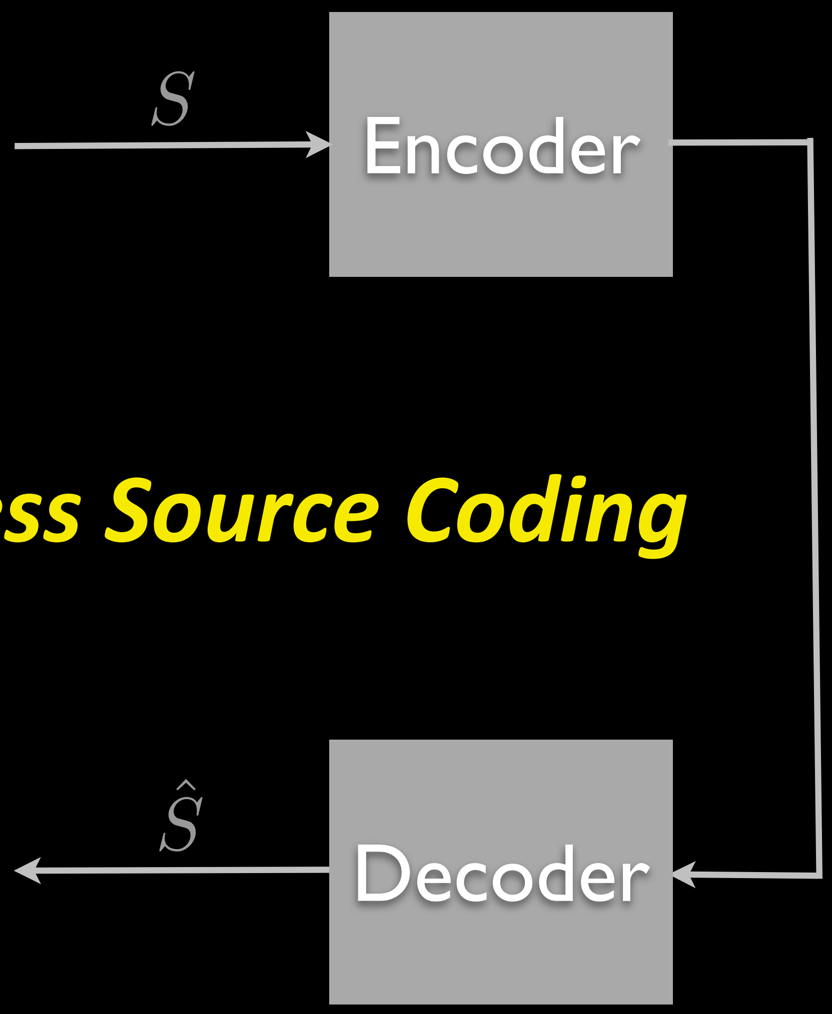
Source Coding



Source Coding



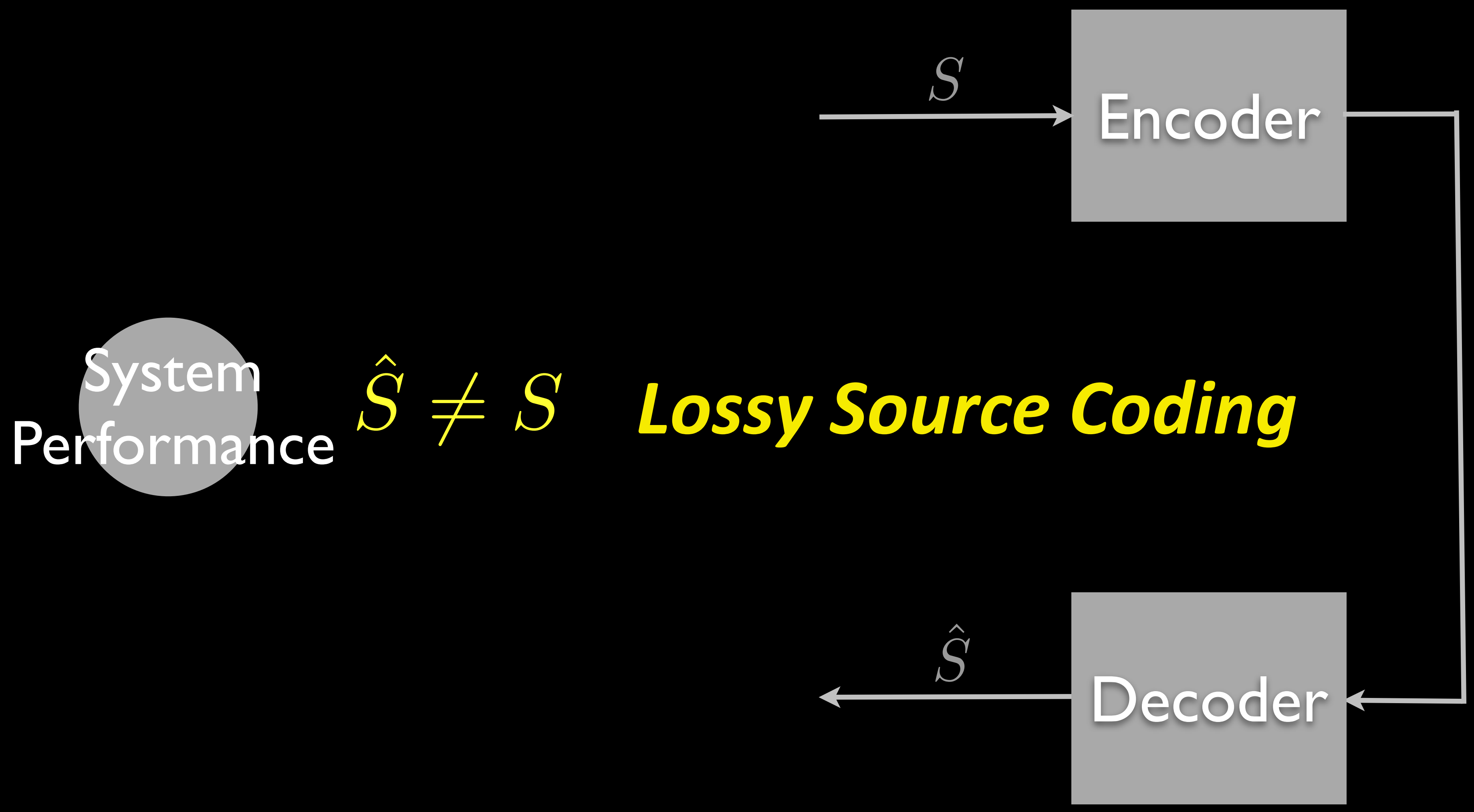
Source Coding



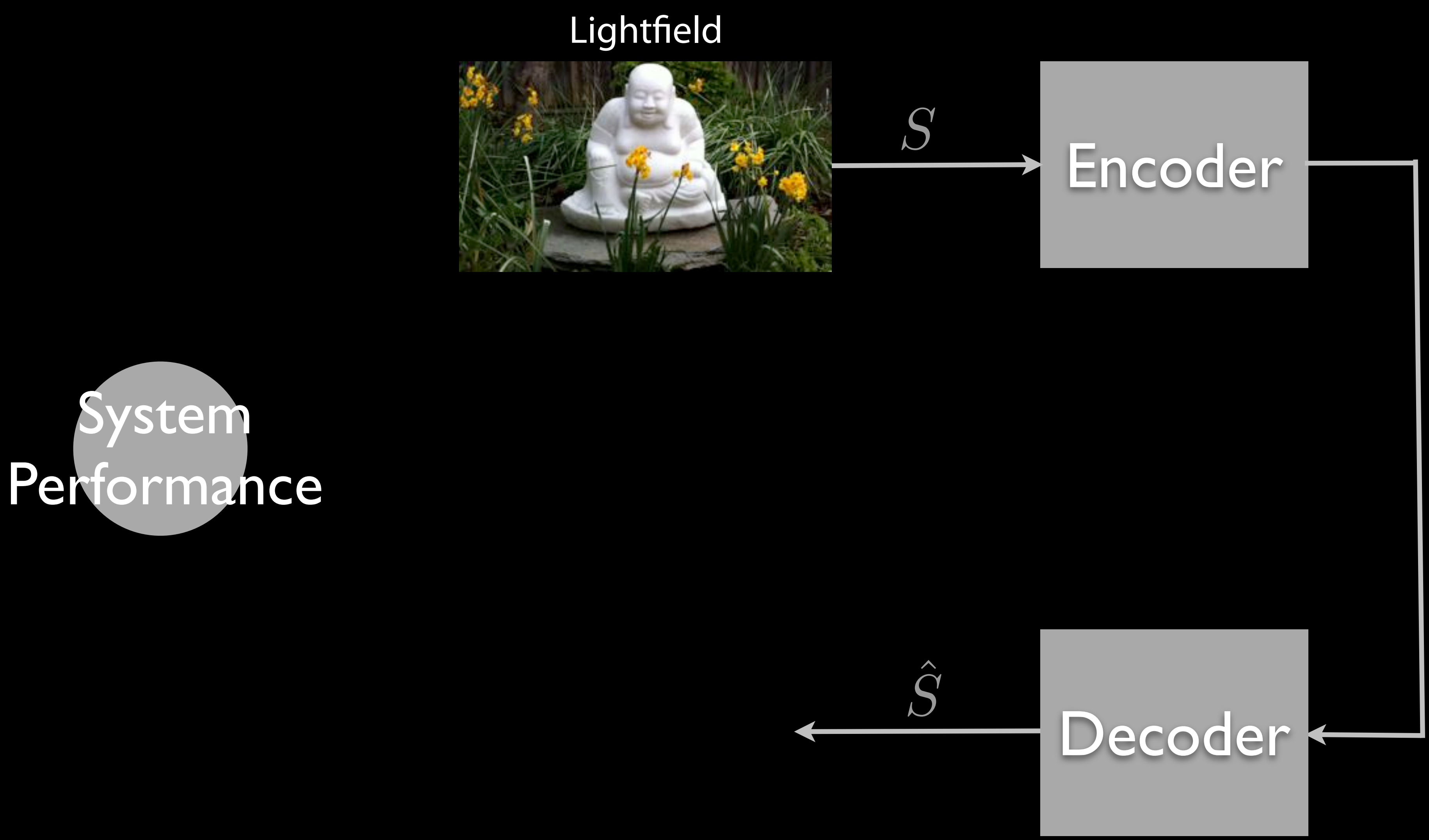
System Performance

$\hat{S} = S$ ***Lossless Source Coding***

Source Coding



Source: Lightfield



Encoder: Camera

Lightfield

S

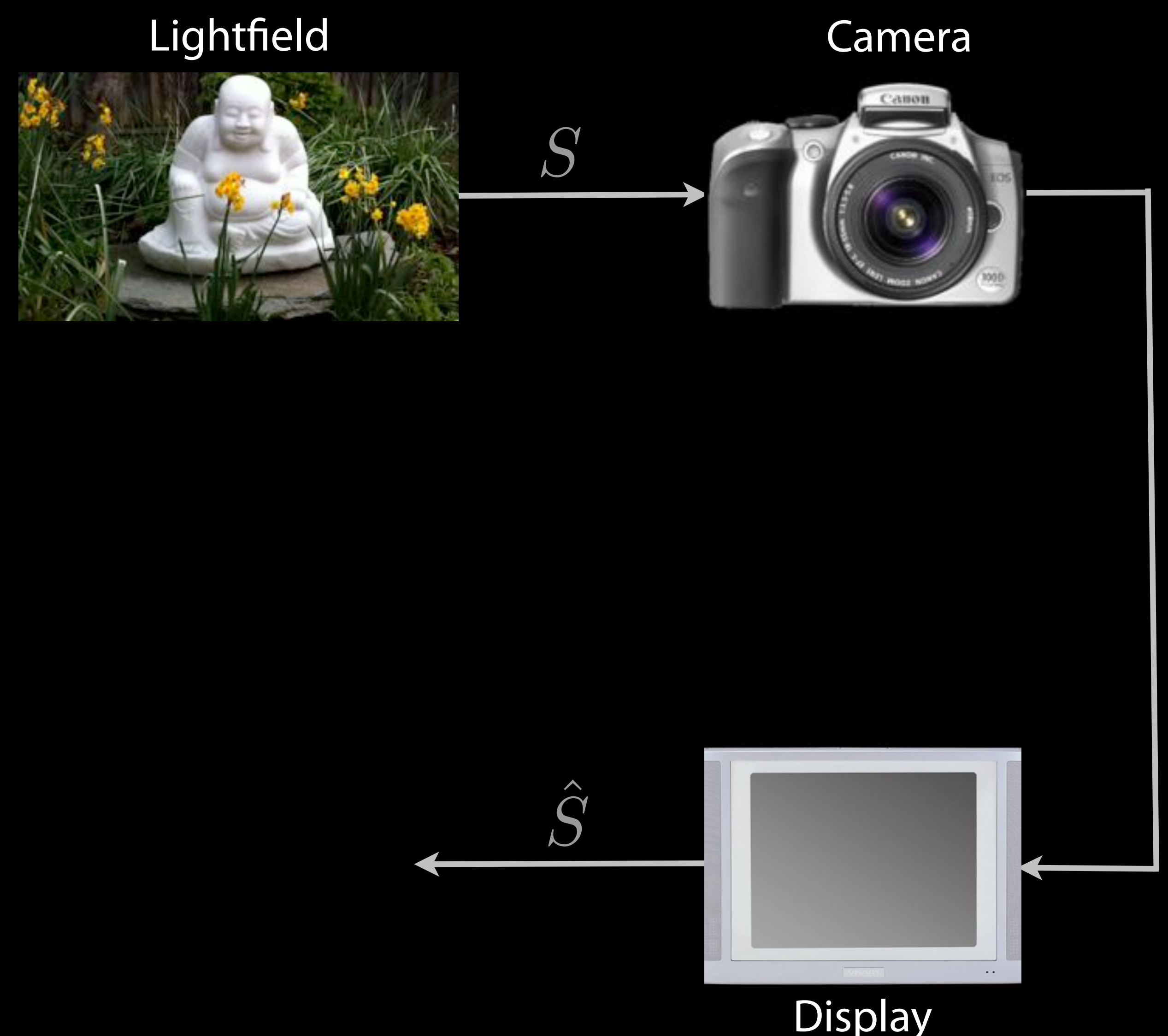
Camera

System Performance

$\hat{S}$

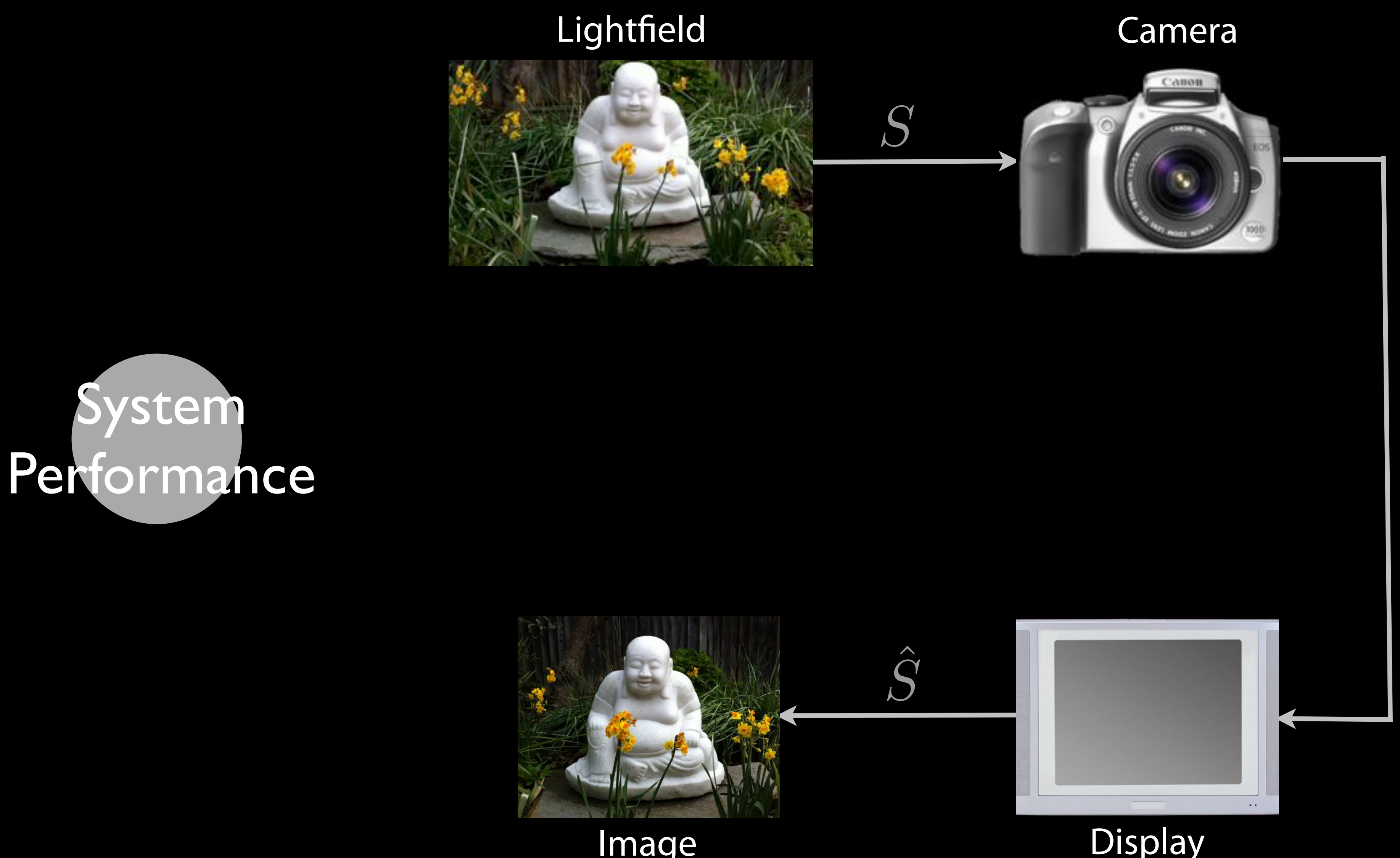
Decoder

Decoder: Display

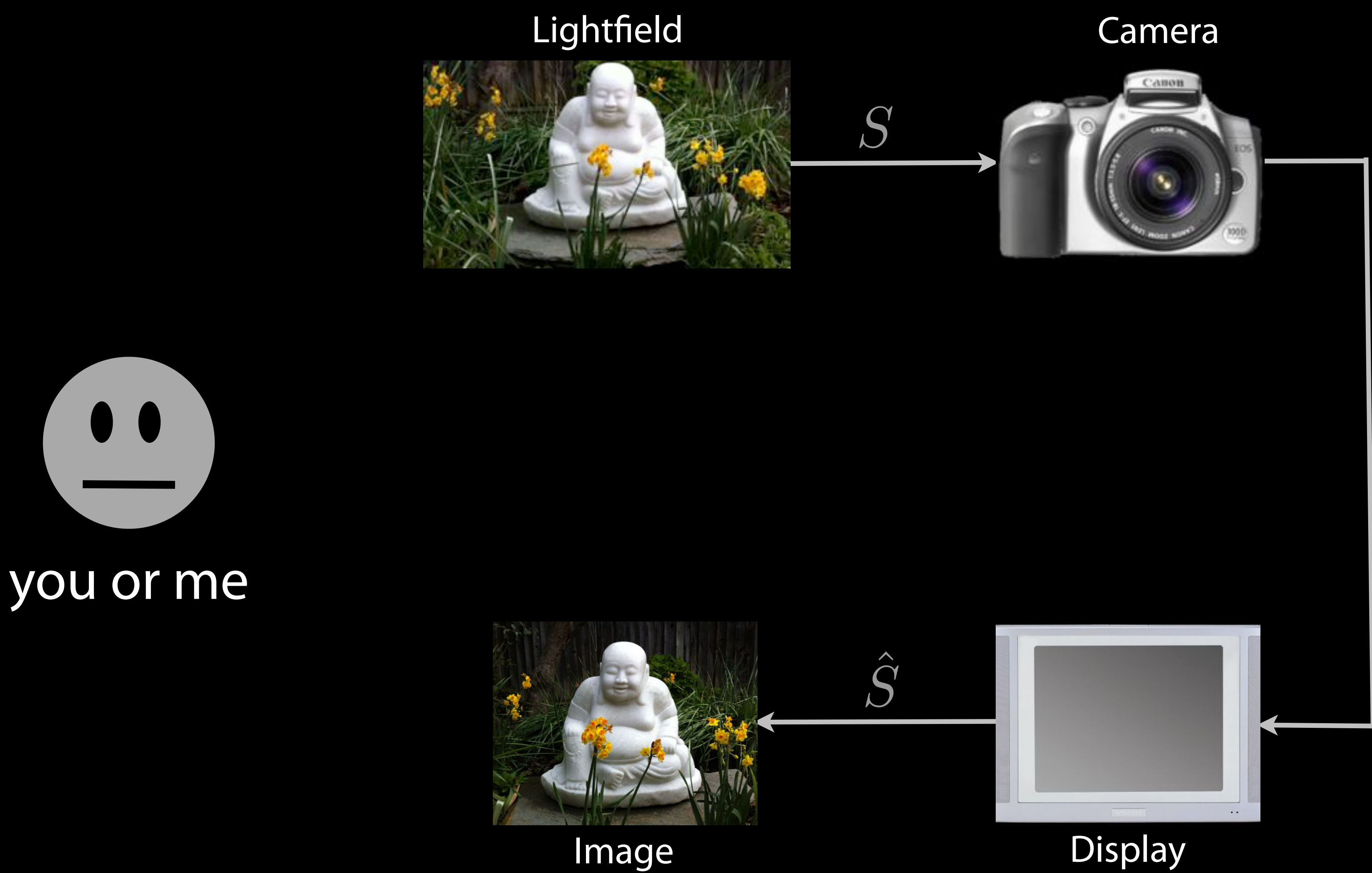


System Performance

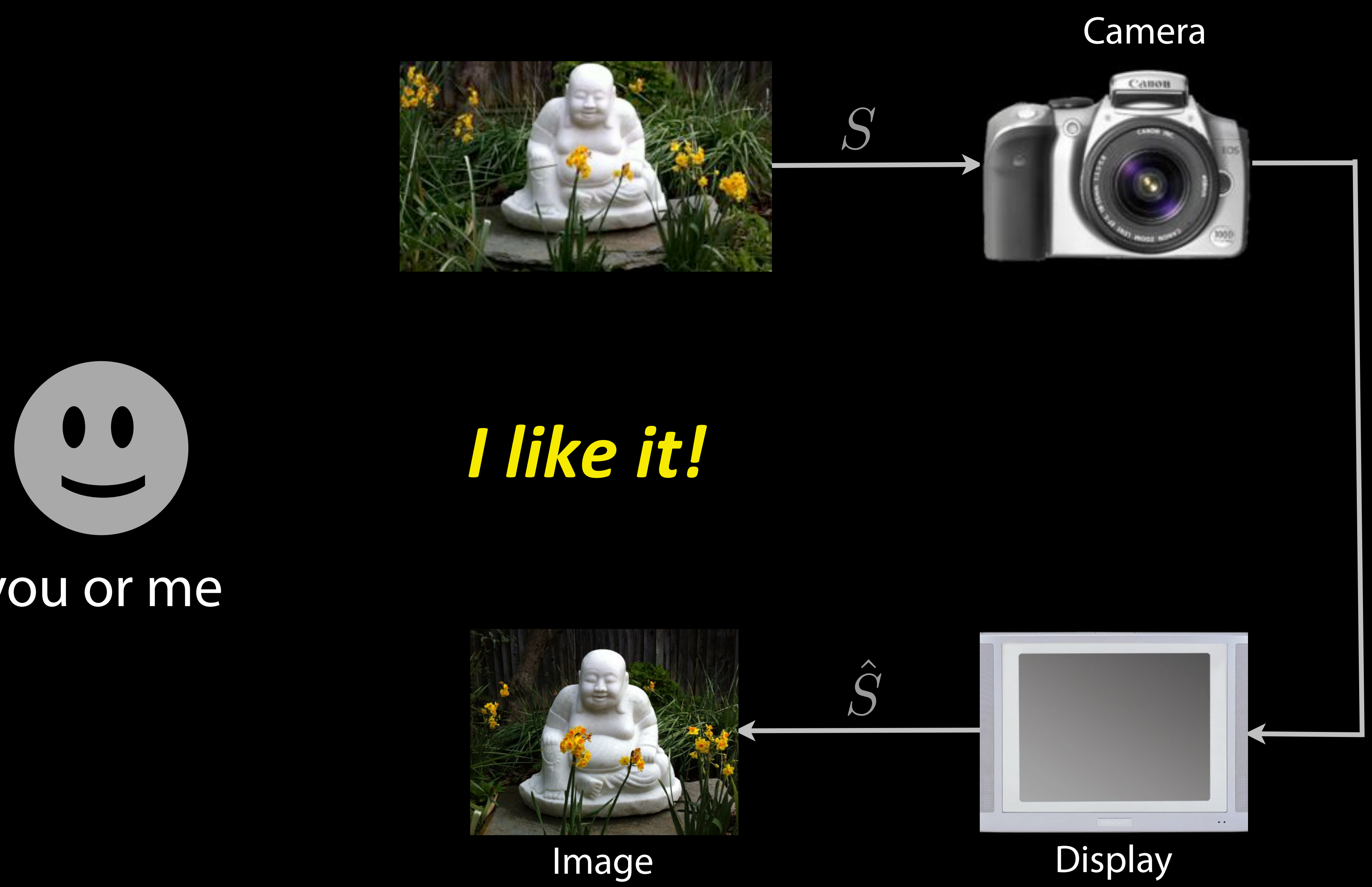
Estimated Source: Image



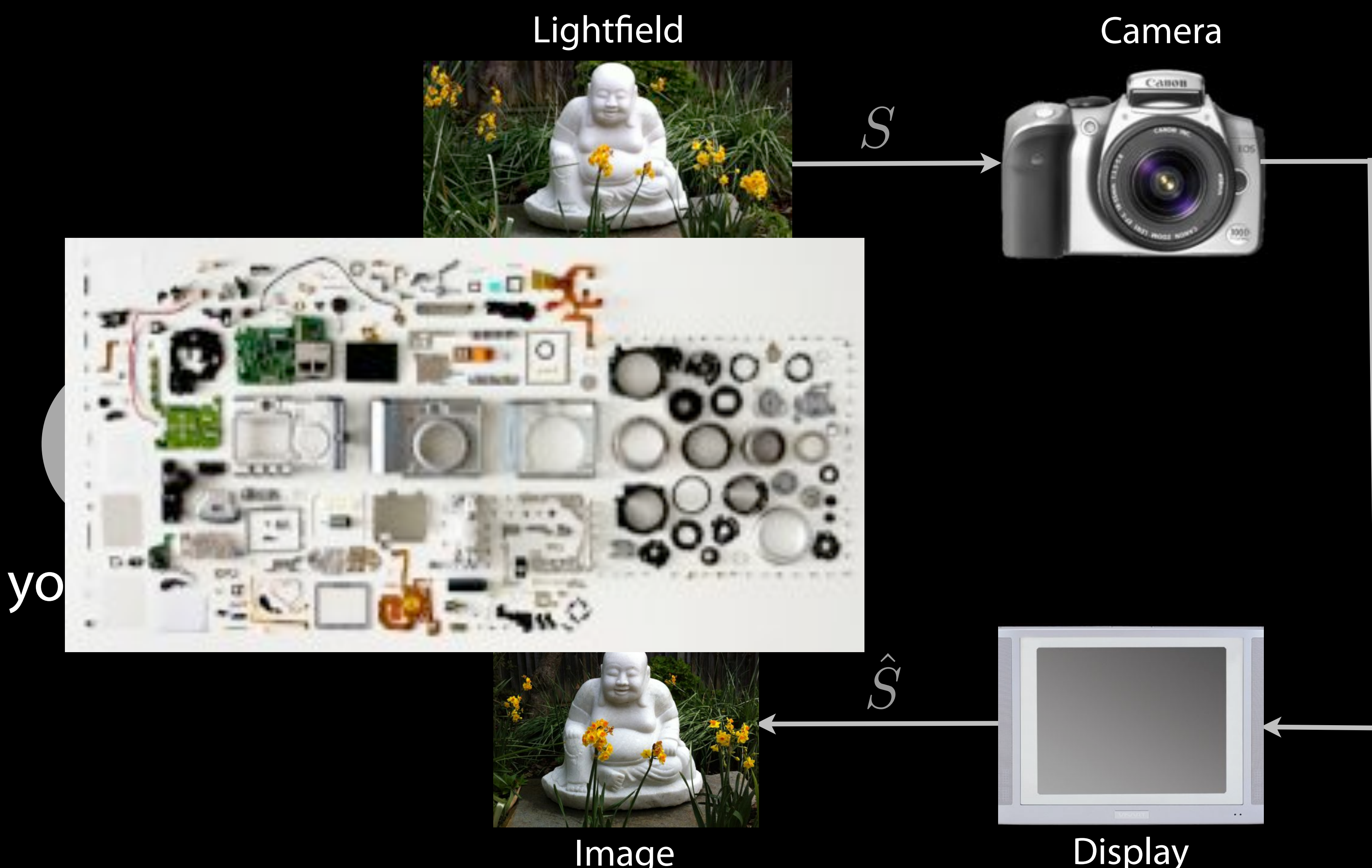
System Performance: you or me



Lossless Source Coding



Lossy Source Coding



http://www.jigsawexplorer.com/puzzles/disassembled-camera-jigsaw-puzzle/

Image Capture

Lightfield

Optics

Color Filter Array
+
Silicon Sensor

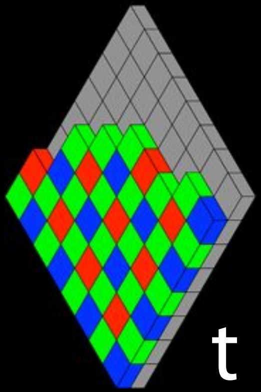
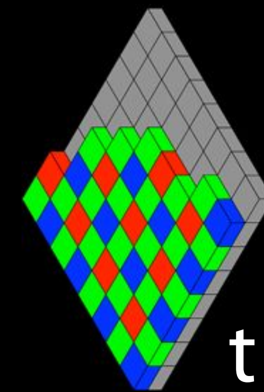


Image Capture

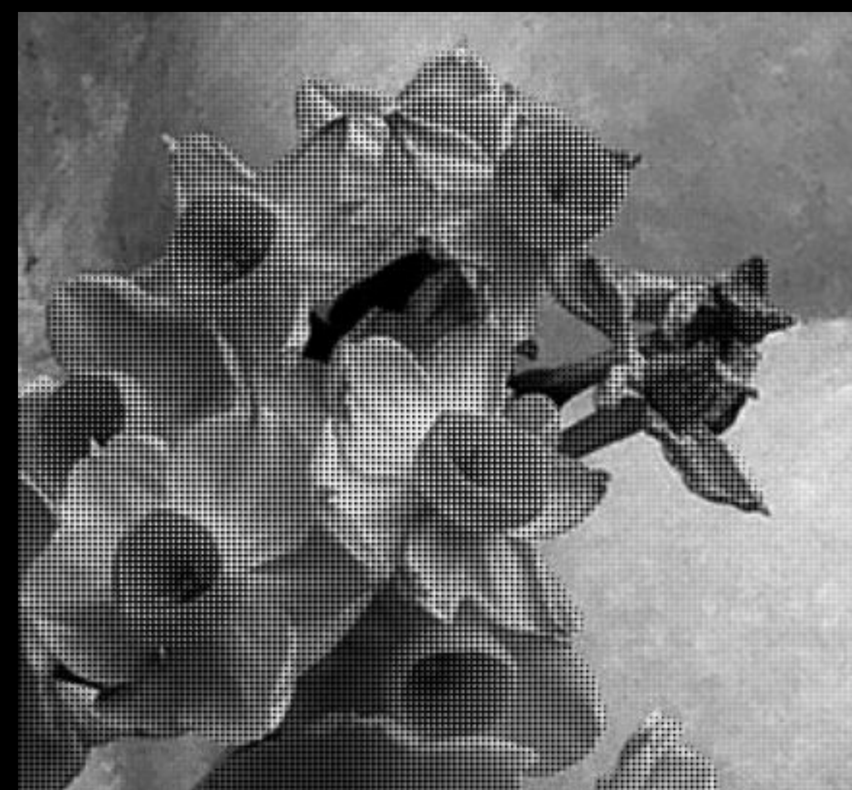
Lightfield

Optics

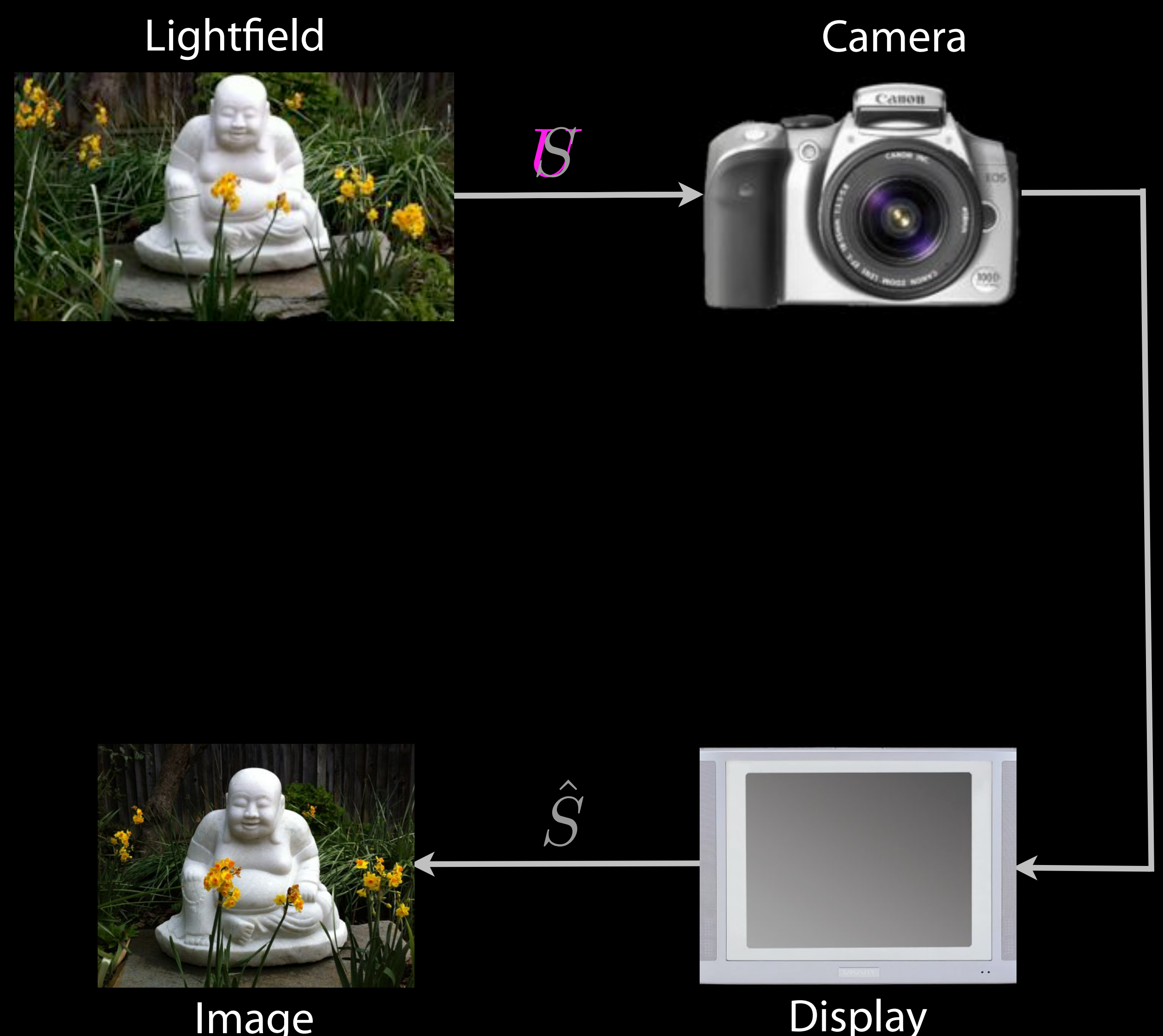
Color Filter Array
+
Silicon Sensor



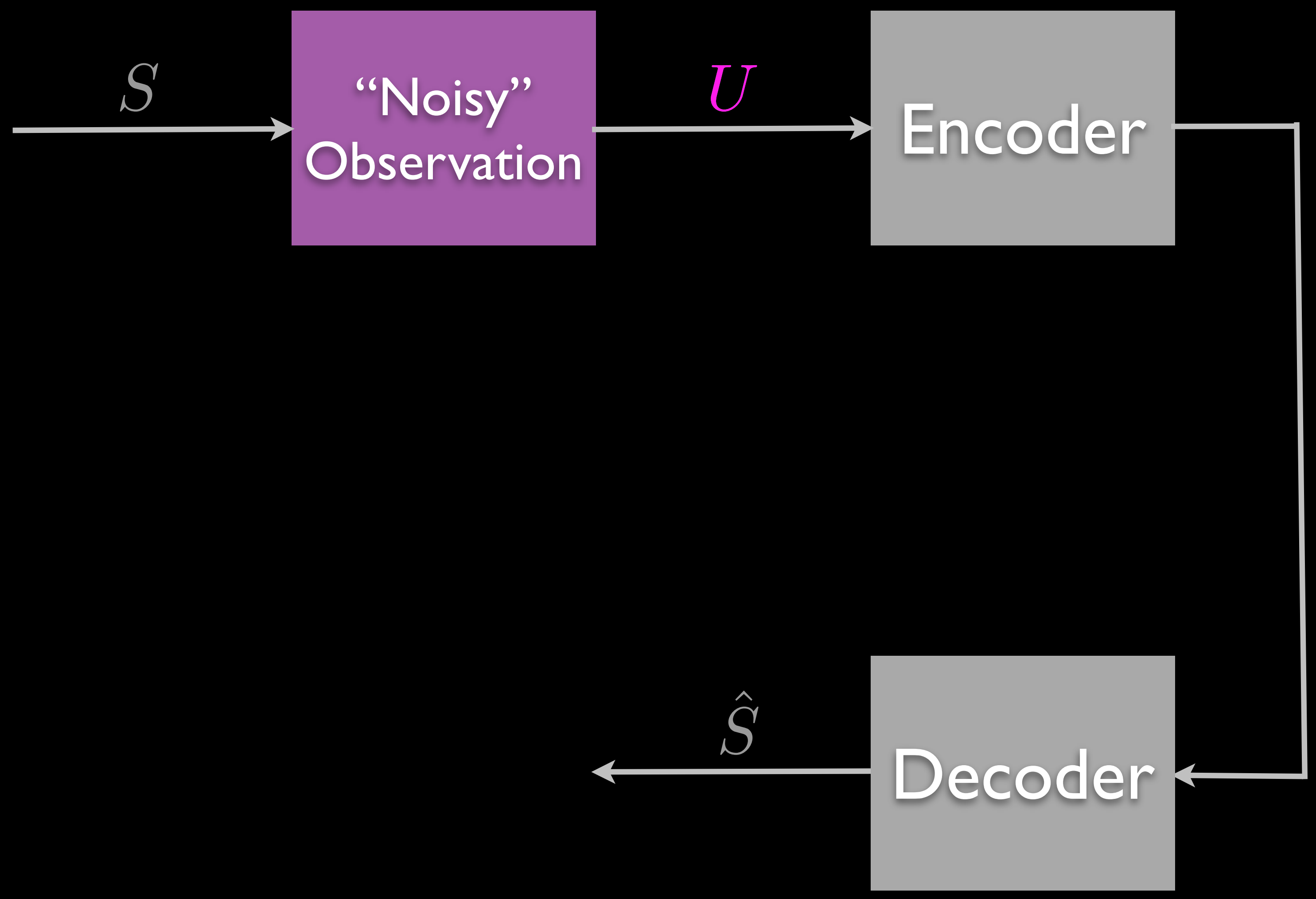
=



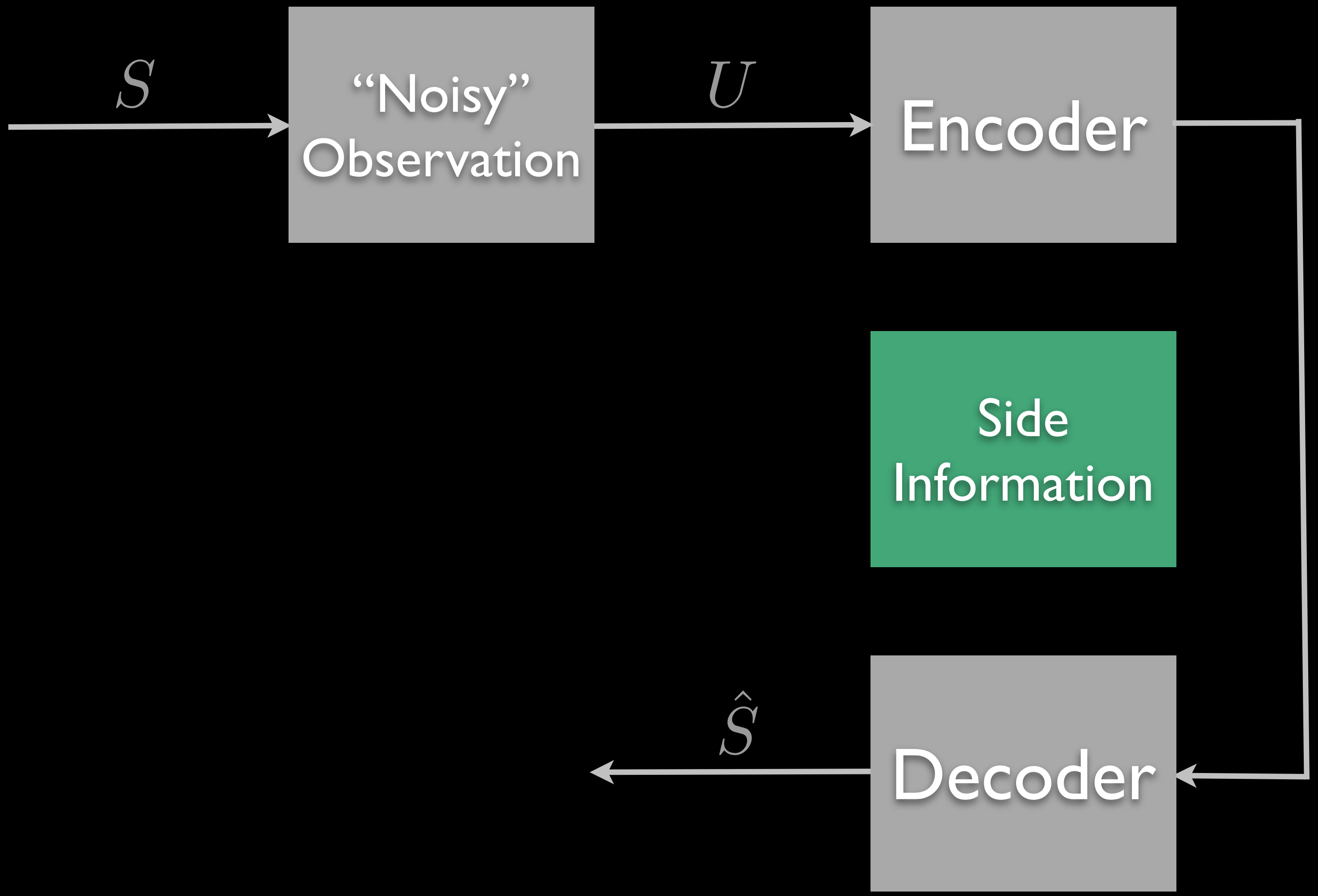
Source Coding



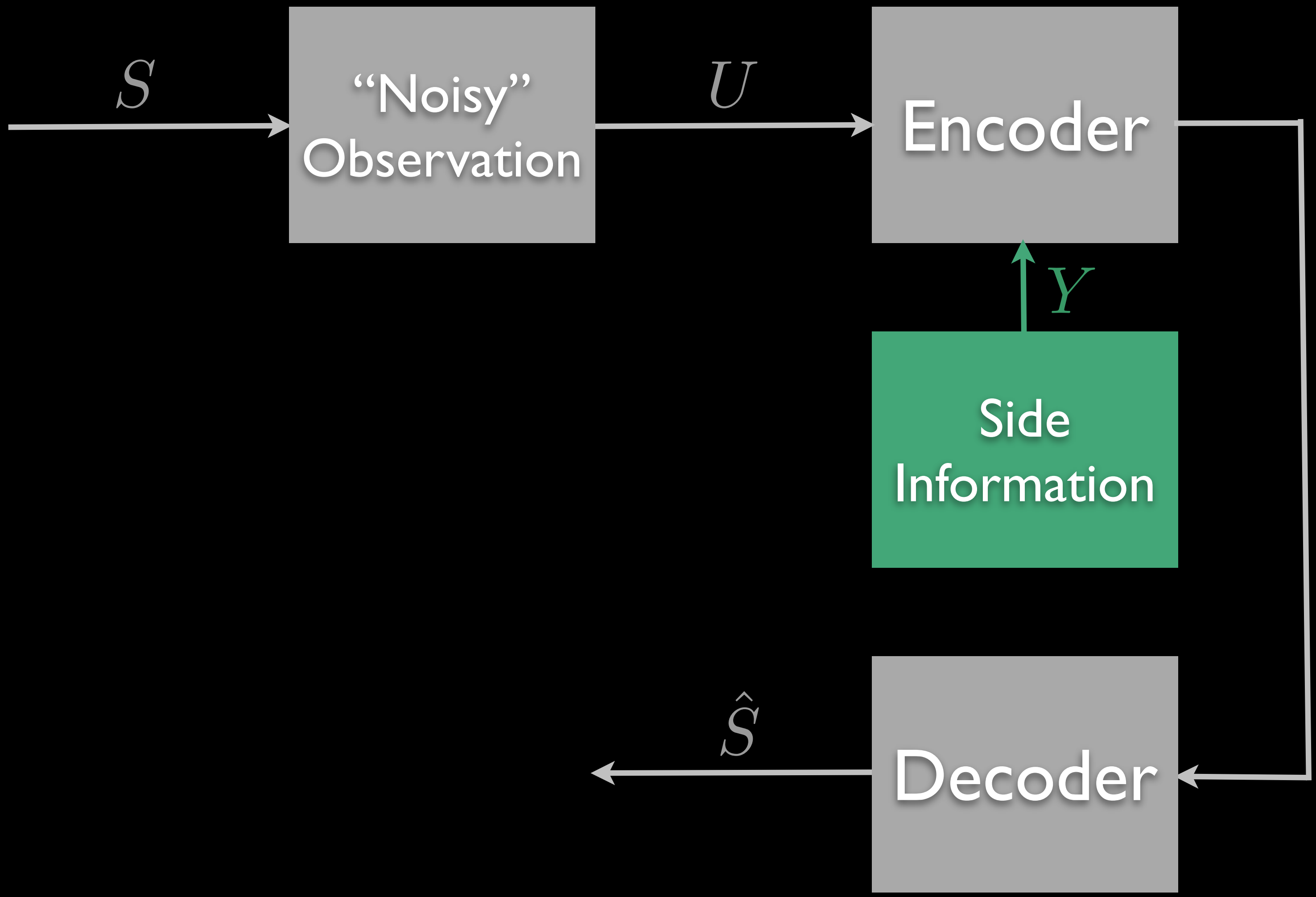
Remote Source Coding



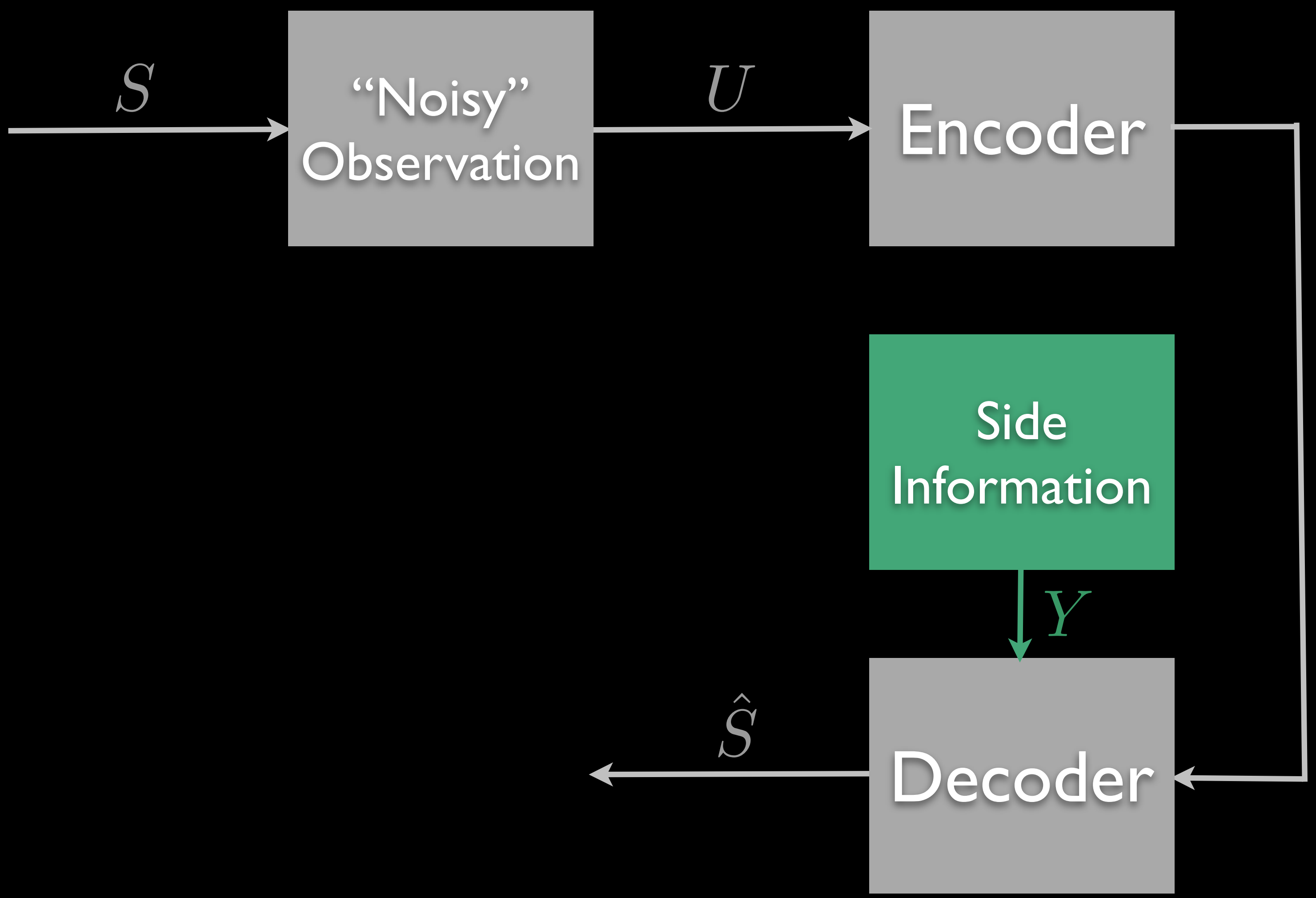
Side Information



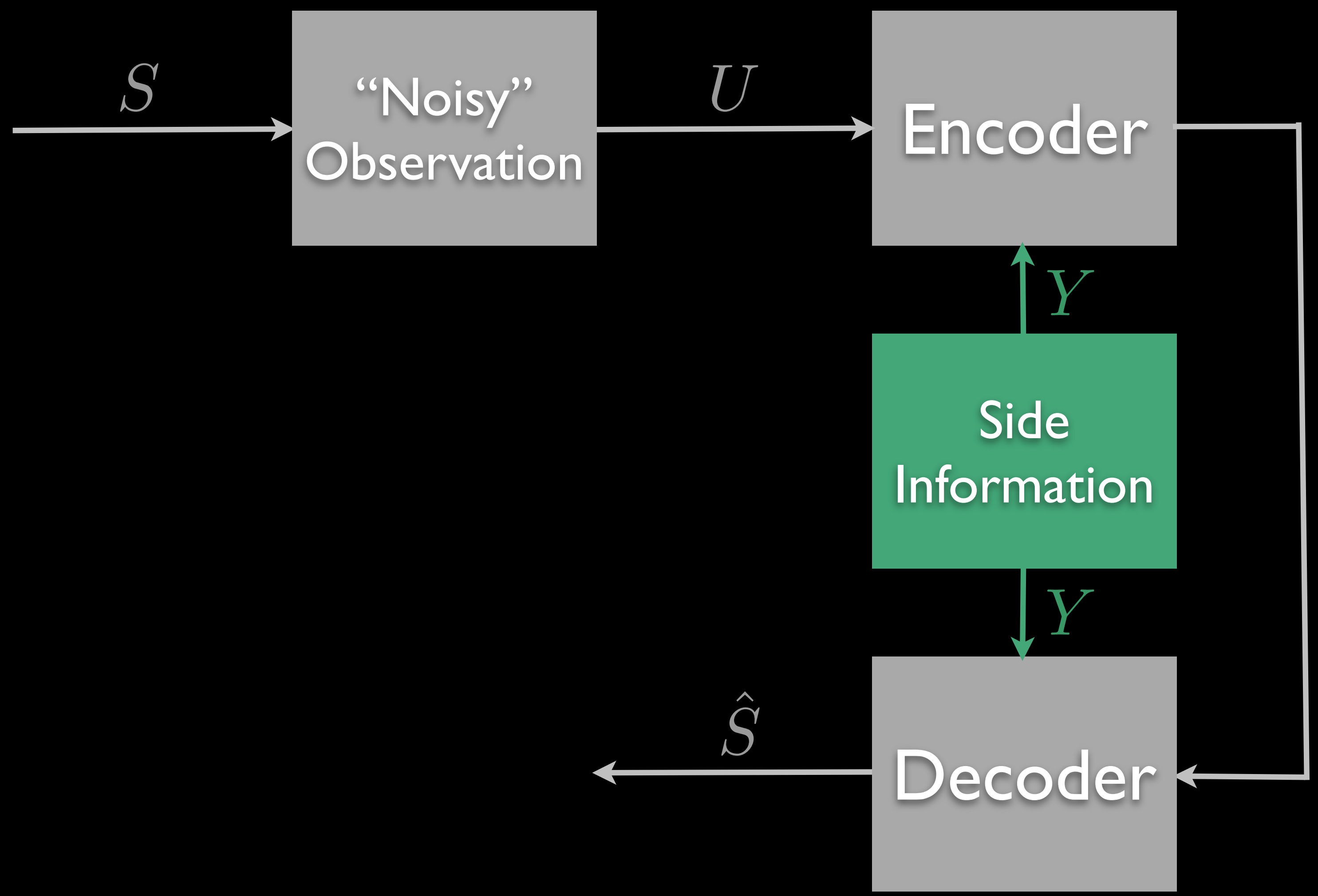
Side Information



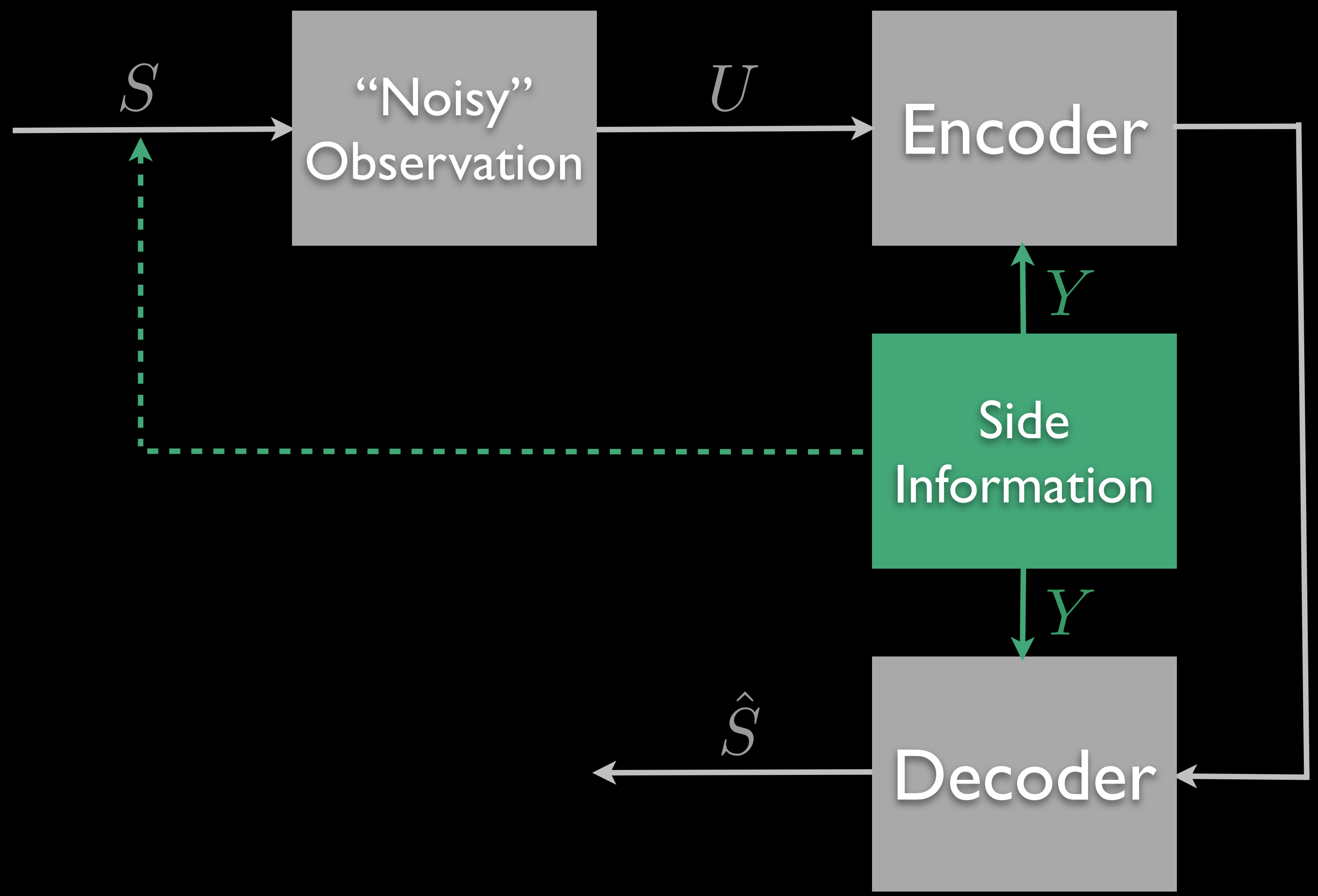
Side Information



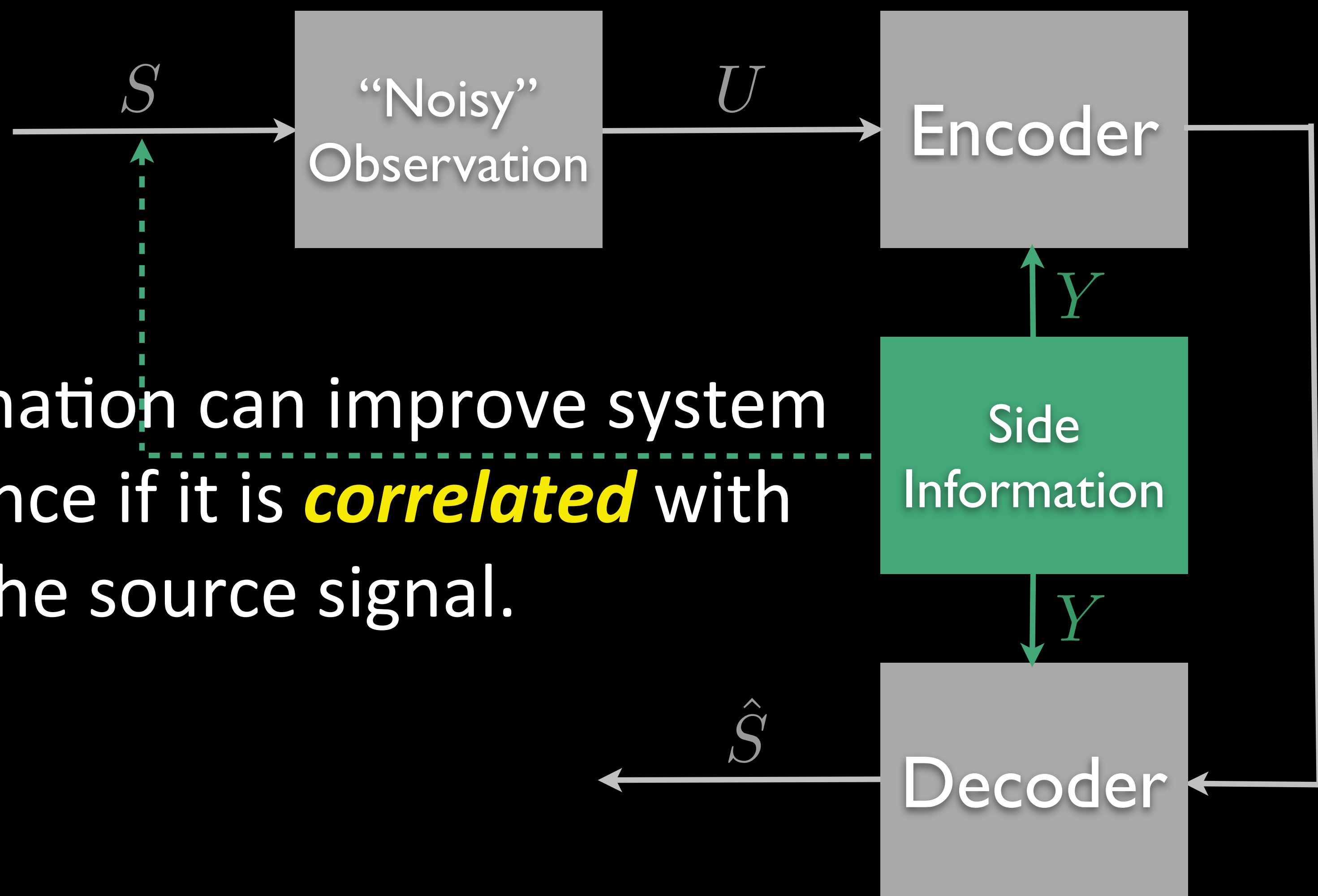
Side Information



Side Information

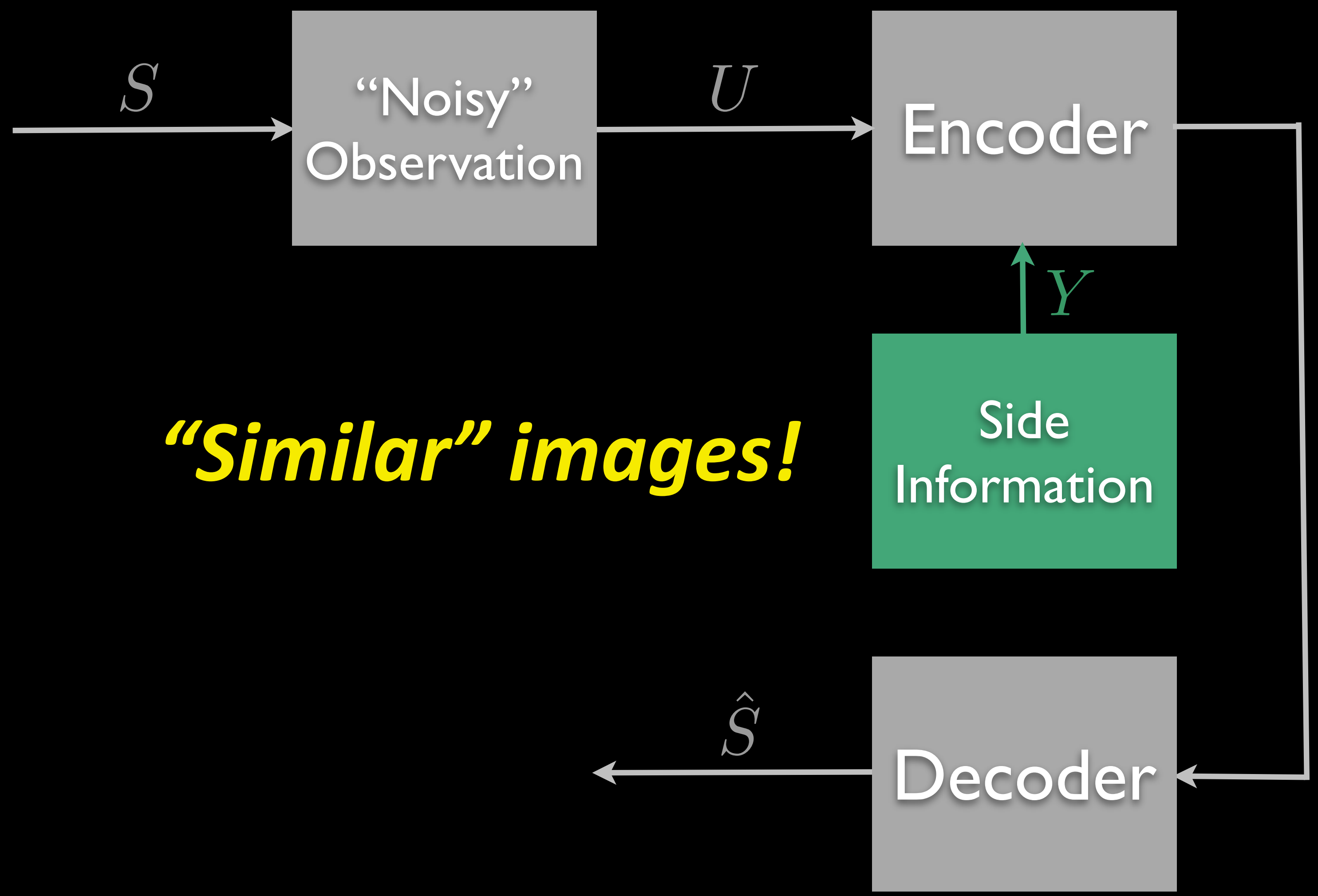


Side Information



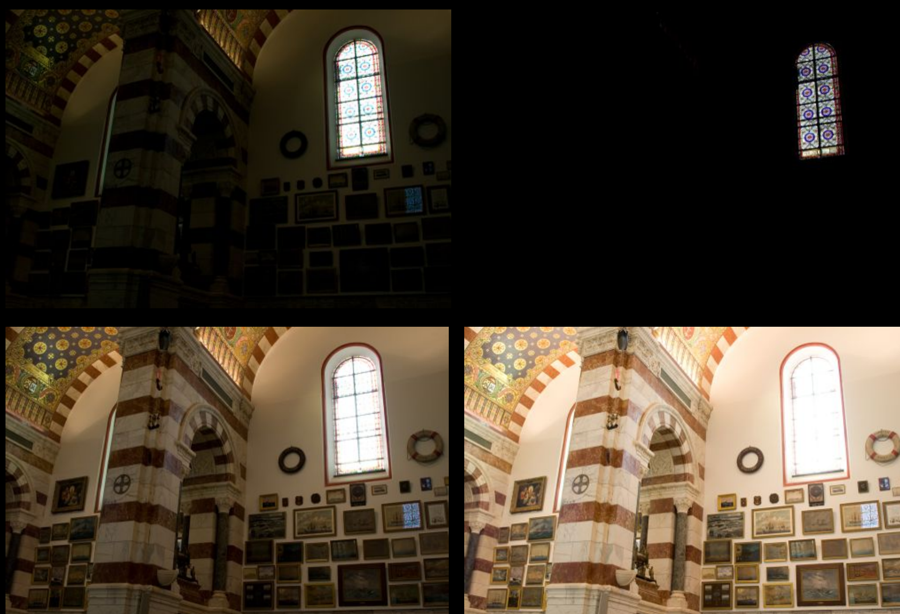
Side Information can improve system performance if it is ***correlated*** with the source signal.

Side Information



"Similar" images!

Multiple Exposures



Multiple exposures

High dynamic range rendering

L. Meylan, D. Alleysson, and S. Süsstrunk, A Model of Retinal Local Adaptation for the Tone Mapping of Color Filter Array Images. *JOSA A*, 2007.

Multiple Illuminations


Ambient Light


Flash


Flash/No Flash Composite

E. Eisemann and F. Durand. Flash Photography Enhancement via Intrinsic Relighting, *ACM SIGGRAPH*, 2004.

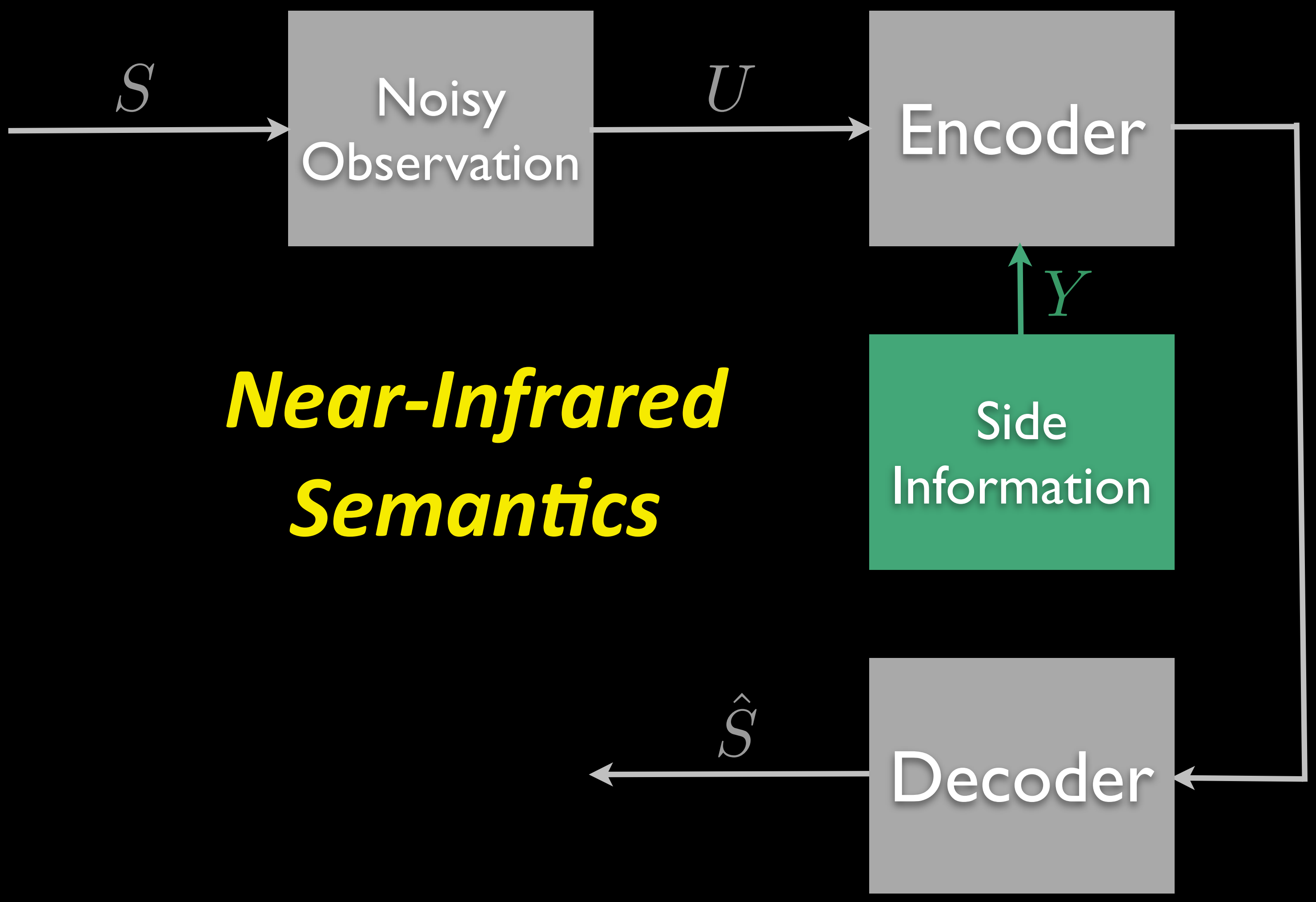
Multiple Time Instances

Multiple Exposures

“Good” Face Composite

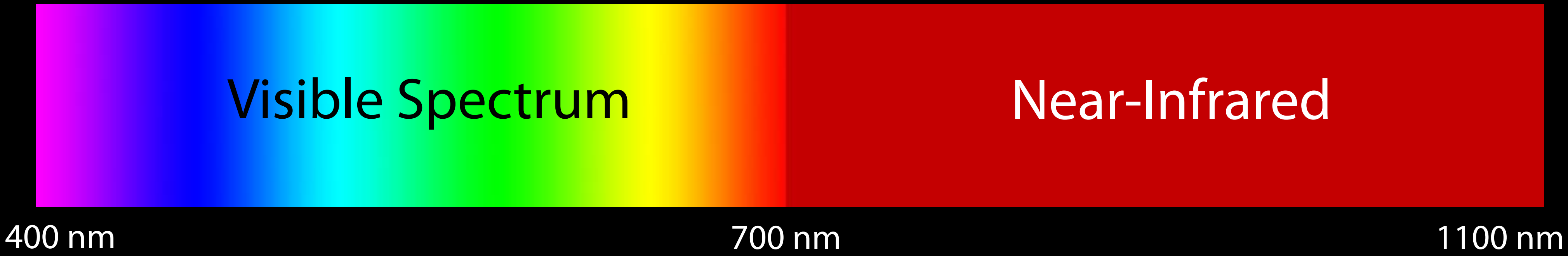
M. F. Cohen and R. Szeliski. The Moment Camera. *IEEE Computer*, 2006.

Our Research



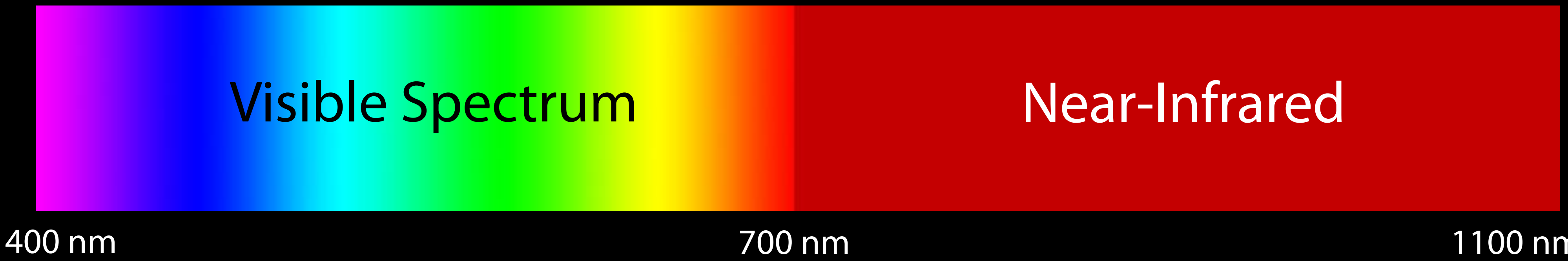
Near-Infrared Semantics

Multiple Spectra



Visible (Red, Green, and Blue: RGB)

Multiple Spectra



Visible (Red, Green, and Blue: RGB)

Near-infrared (NIR)

EPFL

= Ecole Polytechnique Fédéral de Lausanne
= Swiss Federal Institute of Technology, Lausanne

Normal view (RGB)...

Enhanced view (RGB+NIR)

Haze

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

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

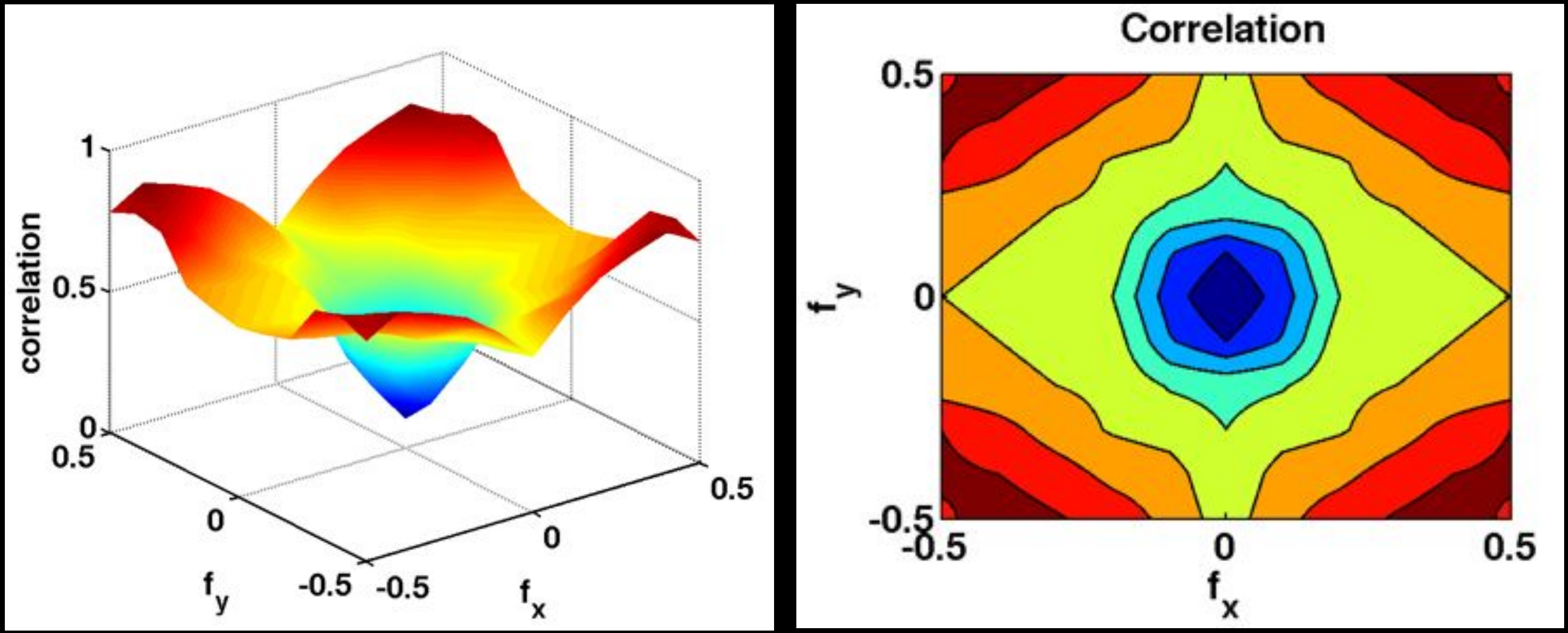
Haze

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.

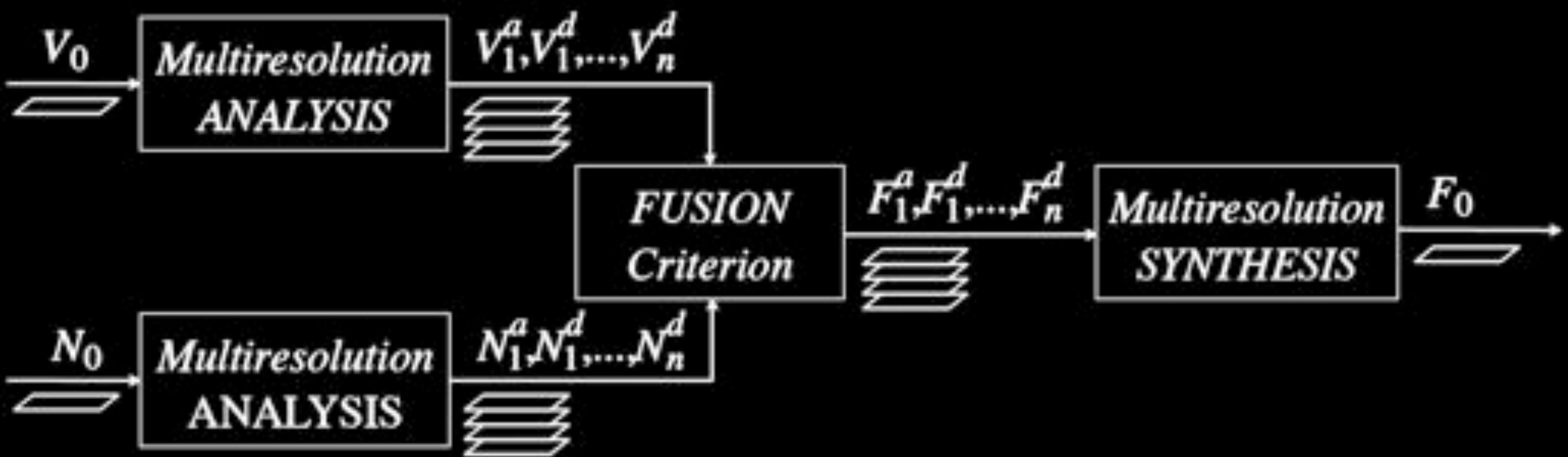


Spectral-spatial correlation of 50 RGB/NIR image pairs

Z. Sadeghipoor, Y. M. Lu, and S. Süsstrunk, Correlation-based Joint Acquisition and Demosaicing of Visible and Near-infrared Images, *ICIP*, 2011.

Algorithm

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



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

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

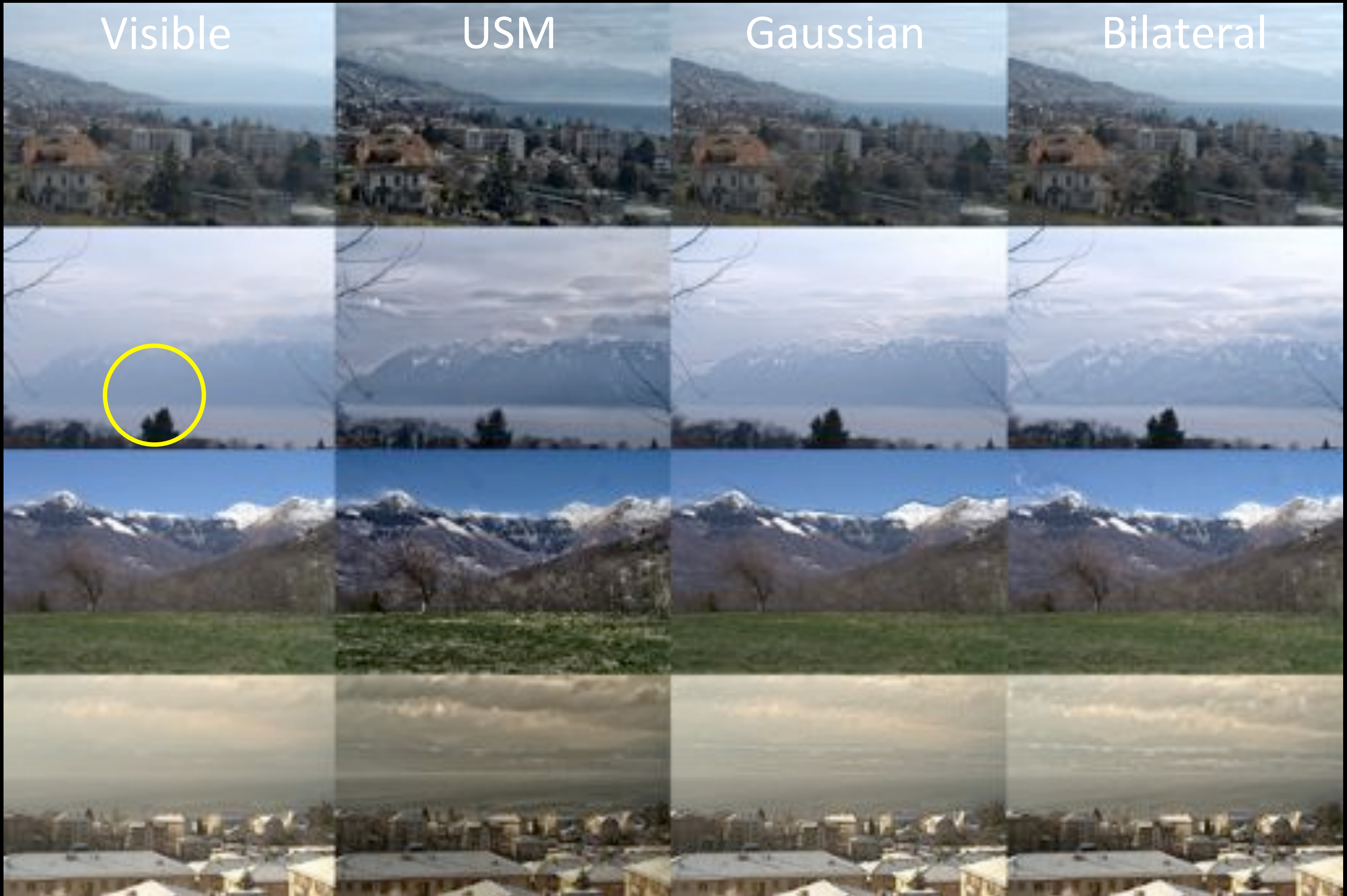
Synthesis

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

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

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

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.

System Performance?

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.

System Performance

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.

Material Absorption/Reflectances

- NIR radiation has different light scattering and ***absorption/ reflection*** behavior than visible light.

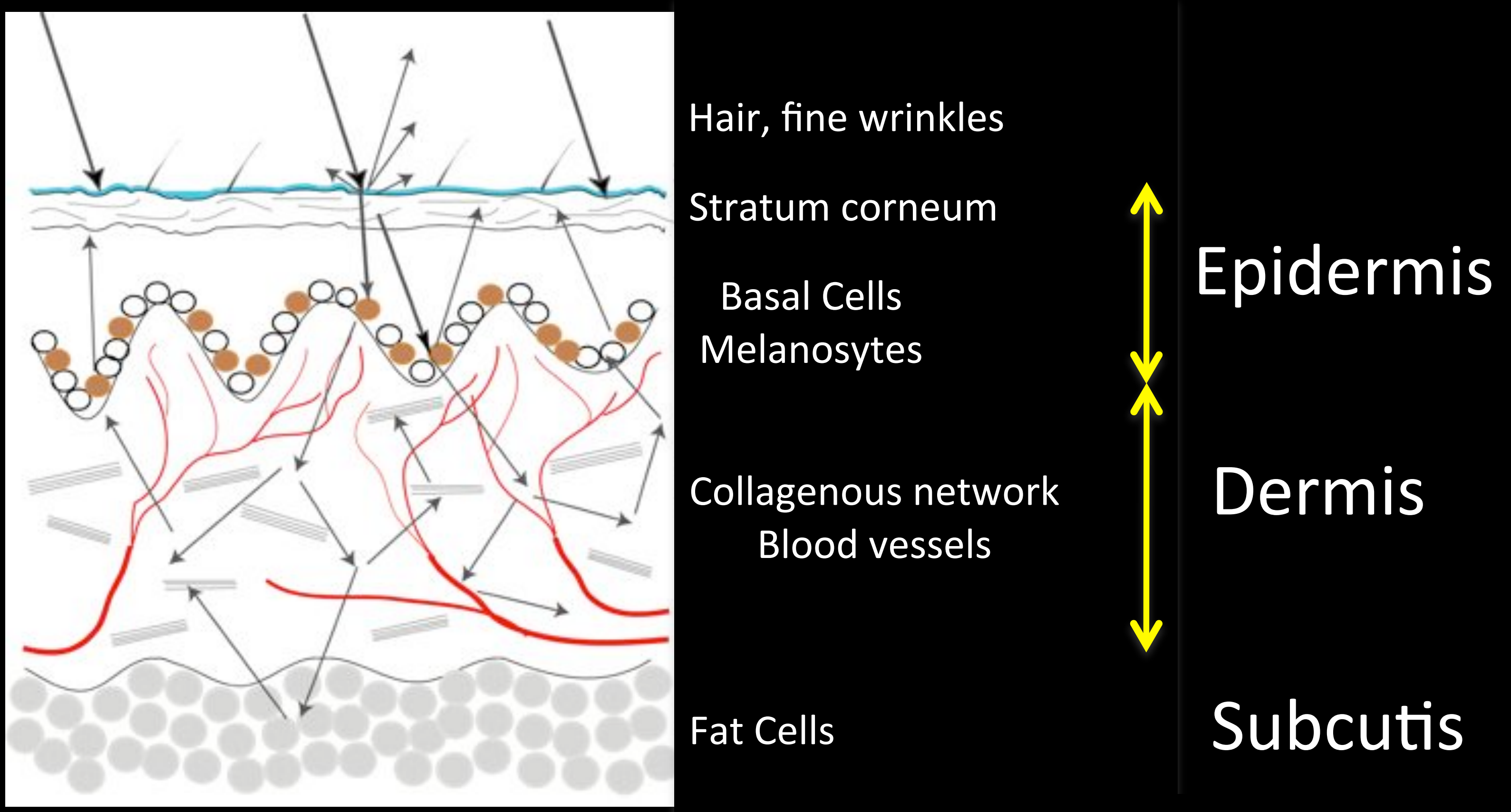
Yellow Arrow: same color in RGB, different intensity in NIR

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

N. Salamati, C. Fredembach, and S. Süsstrunk, Material Classification Using Color and NIR Images, *IS&T/SID 17th Color Imaging Conference, 2009.*

Penetration of radiation in skin

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



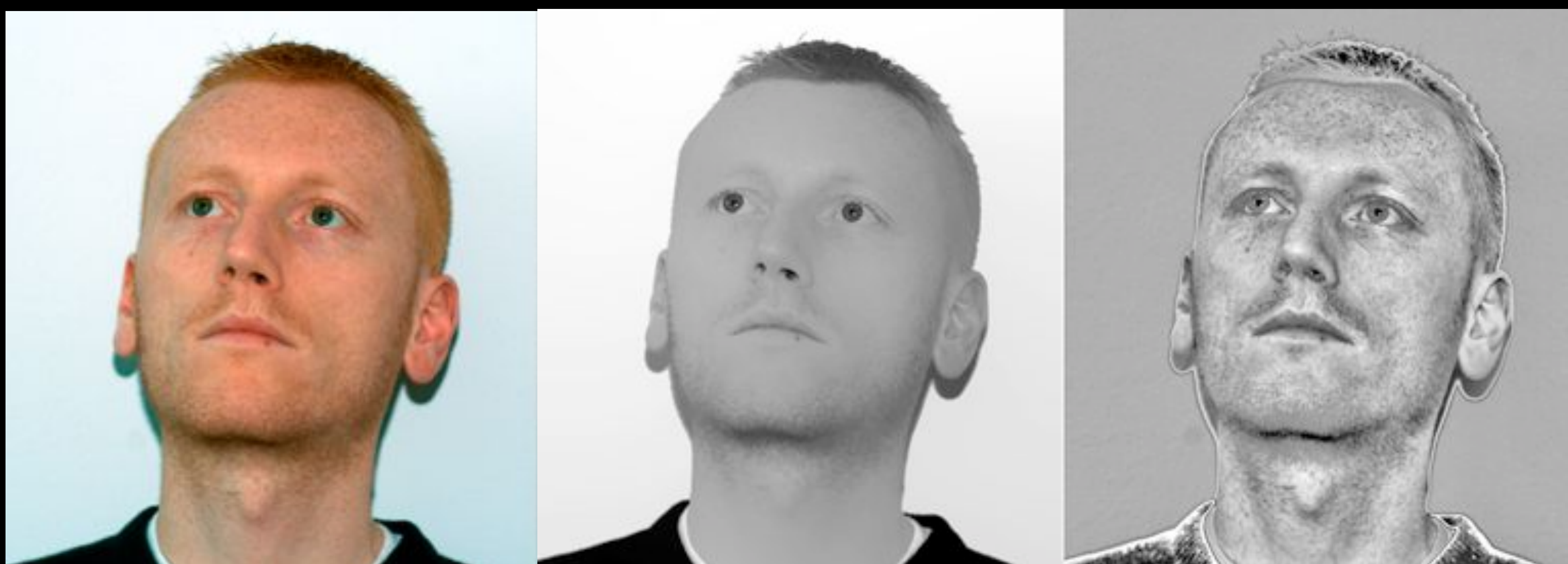
T. Igarashi, K. Nishino, and S.K. Nayar, The appearance of human skin. Technical Report: CUCS-024-05, 2005.

Intuition

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

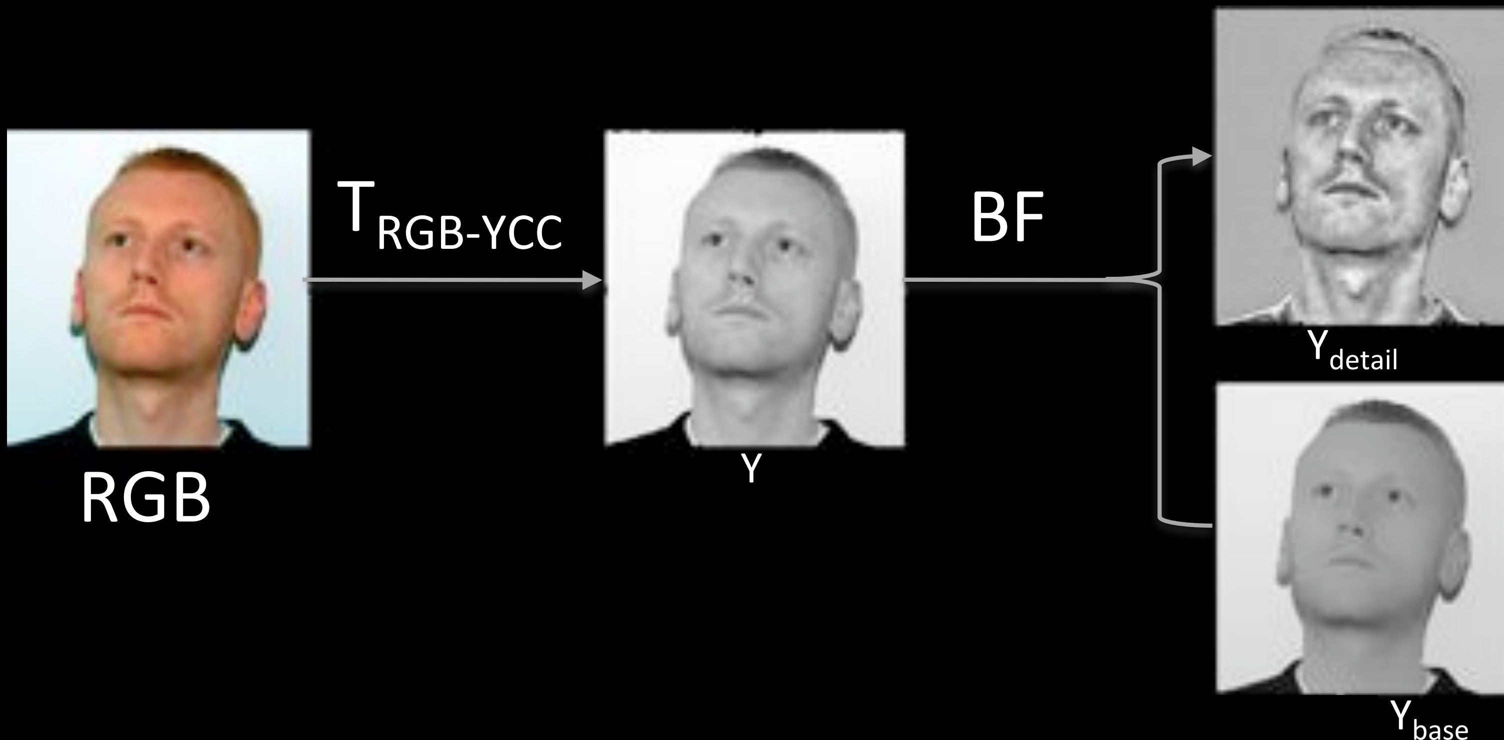
Intuition

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



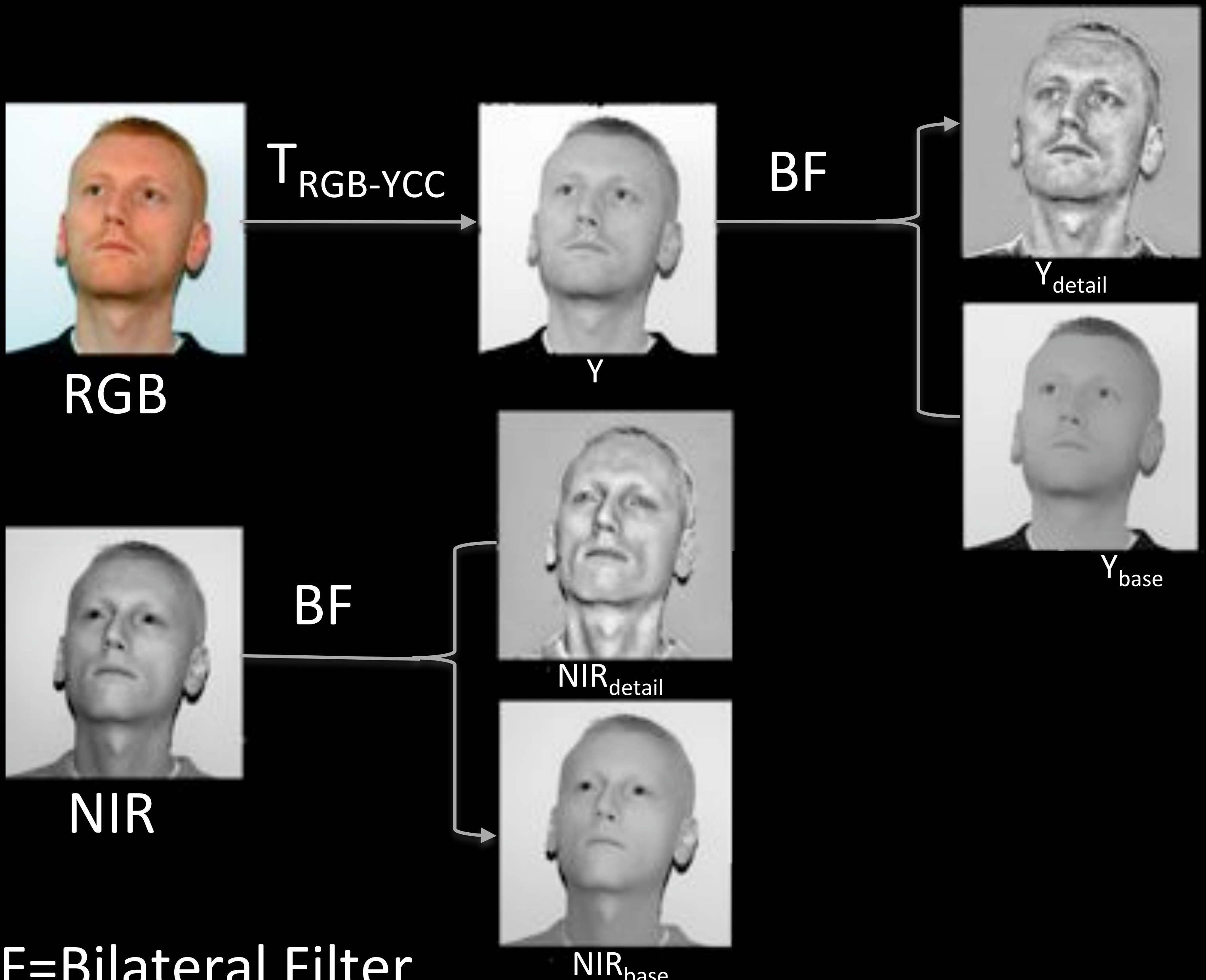
Original Base Detail

Algorithm



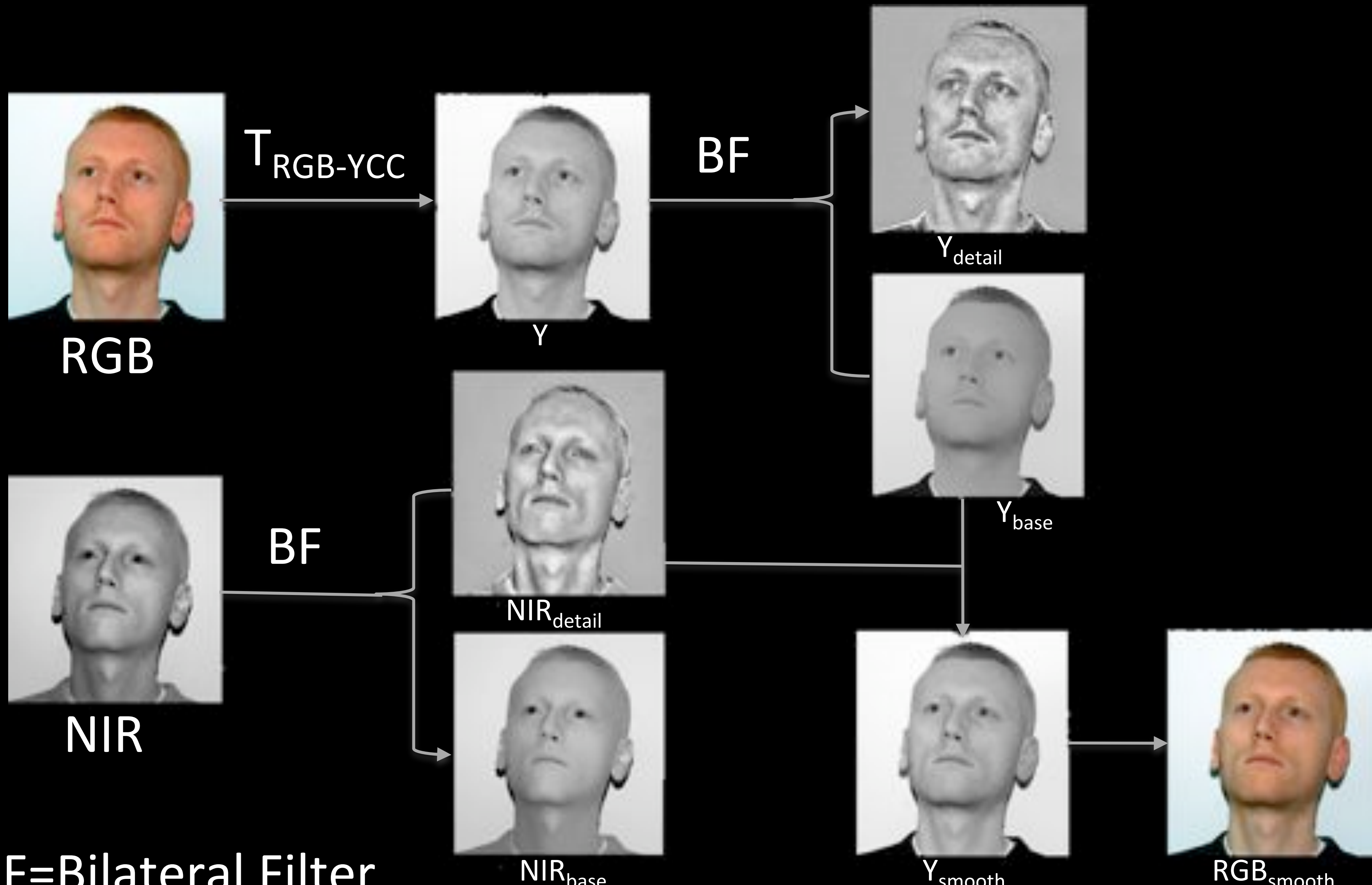
BF=Bilateral Filter

Algorithm



BF=Bilateral Filter

Algorithm



BF=Bilateral Filter

Skin Smoothing

Visible

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

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

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.

System Performance?

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

Camera Museum in Vevey

Preference: 26.6% (2,331)

Preference: 73.4% (6,424)

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

Color Constancy

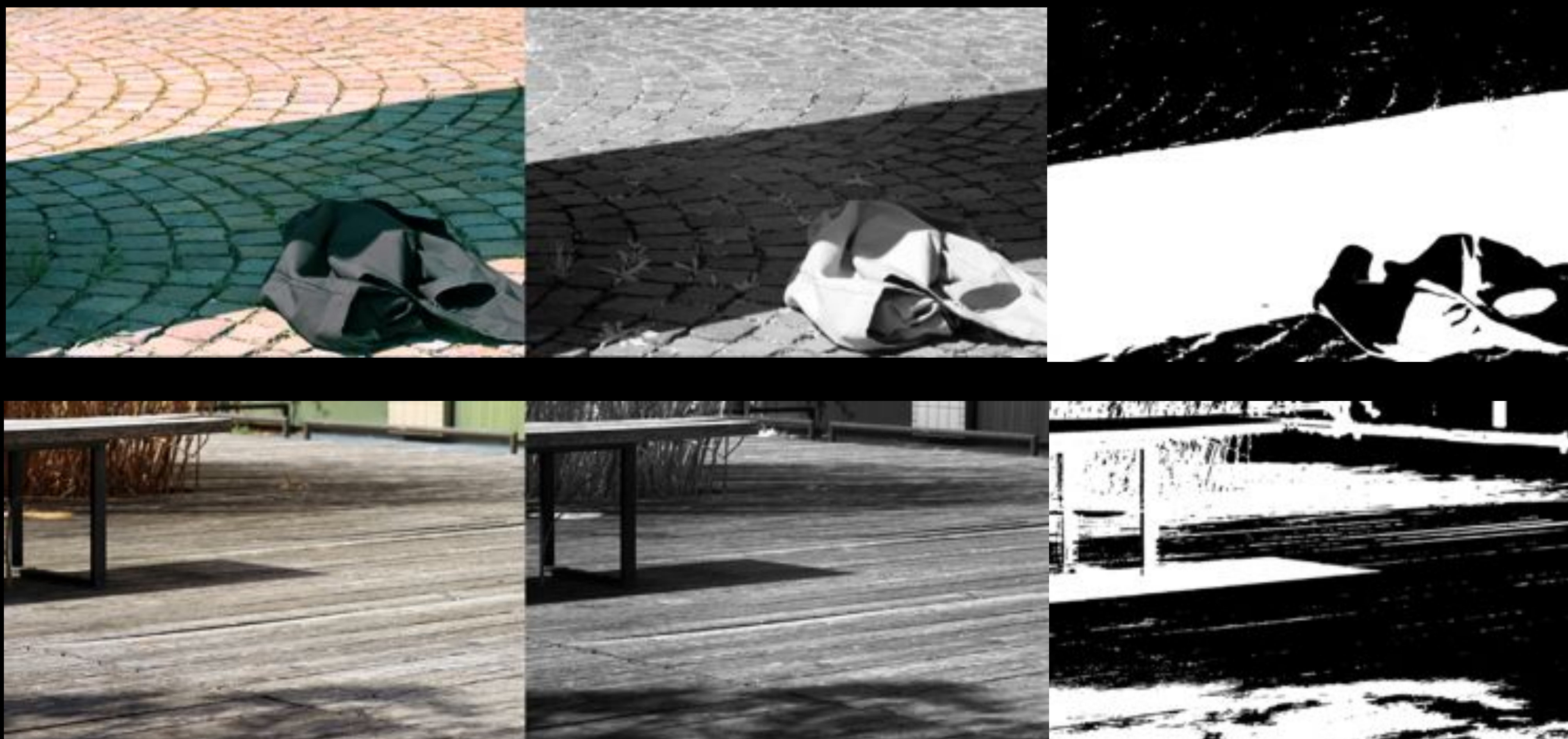
Overcast Sky
Estimated Color Temperature:
7500K

Sunny Day
Estimated Color Temperature:
5000K

Incandescent Lamp
Estimated Color Temperature:
3000K

C. Fredembach and S. Süsstrunk, Illuminant estimation and detection using near infrared, *IS&T/SPIE Electronic Imaging, Digital Photography V*, Vol. 7250, 2009.

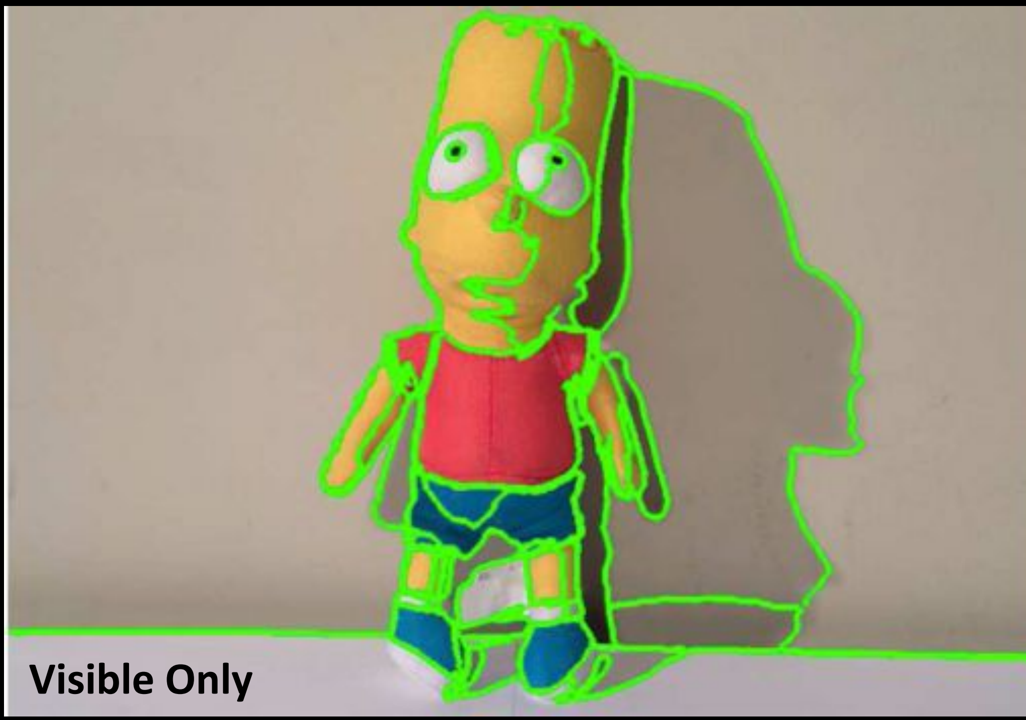
Shadow Detection



Visible NIR ***Binary Shadow Mask***

D. Rüfenacht, C. Fredembach, and S. Süsstrunk, "Automatic and Accurate Shadow Detection Using Near-infrared Information," submitted to *IEEE Transactions on Pattern Recognition and Machine Intelligence,* 2013.

Comparison with visible only

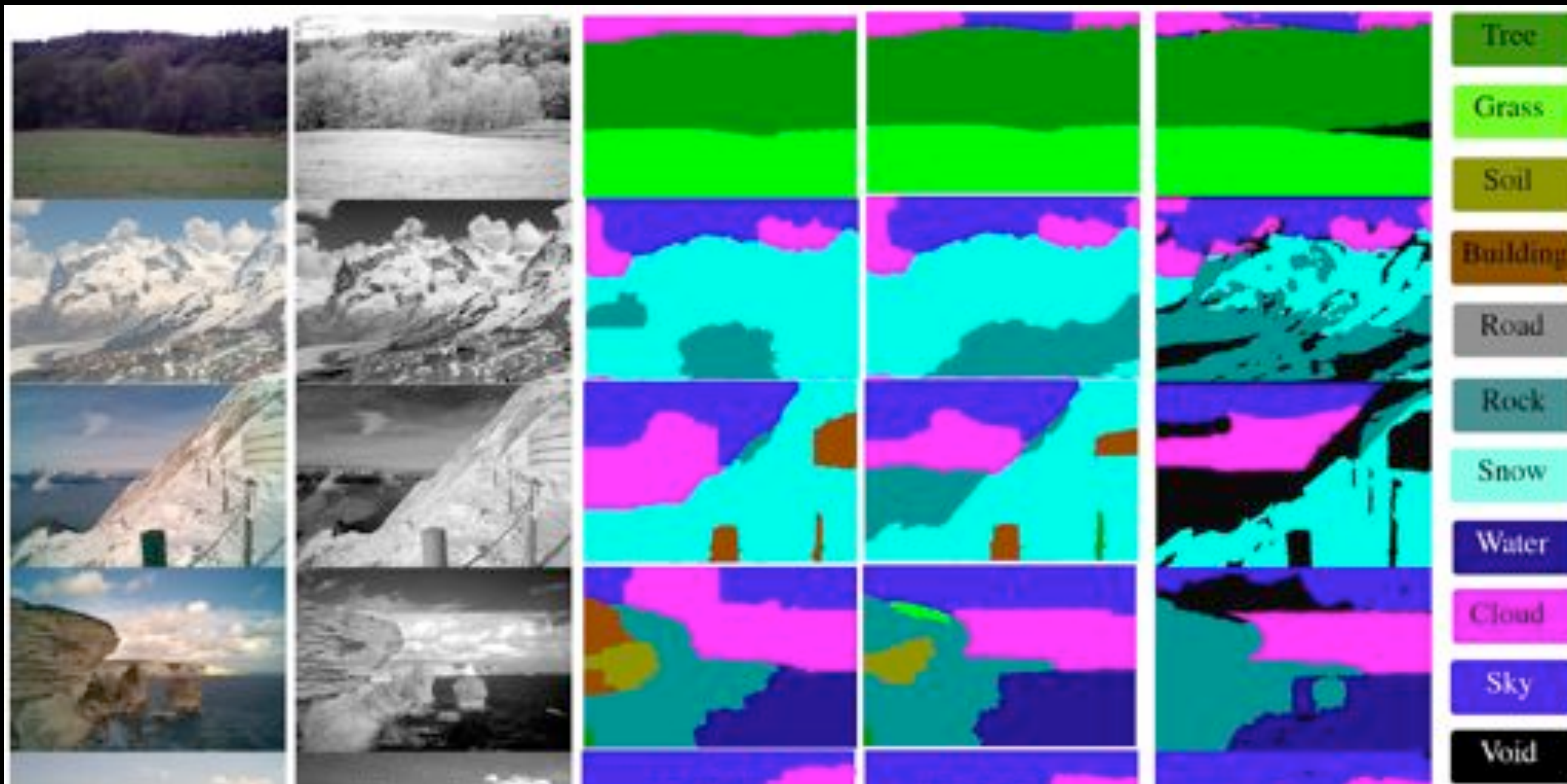


N. Salamati and S. Süsstrunk, Material-Based Object Segmentation using Near-Infrared Information, *IS&T/SID 18th Color Imaging Conference*, 2010. ***Best Interactive Paper Award***

Image Database (477 RGB+NIR)

M. Brown and S. Süsstrunk, Multi-spectral SIFT for Scene Category Recognition, *CVPR* 2011.

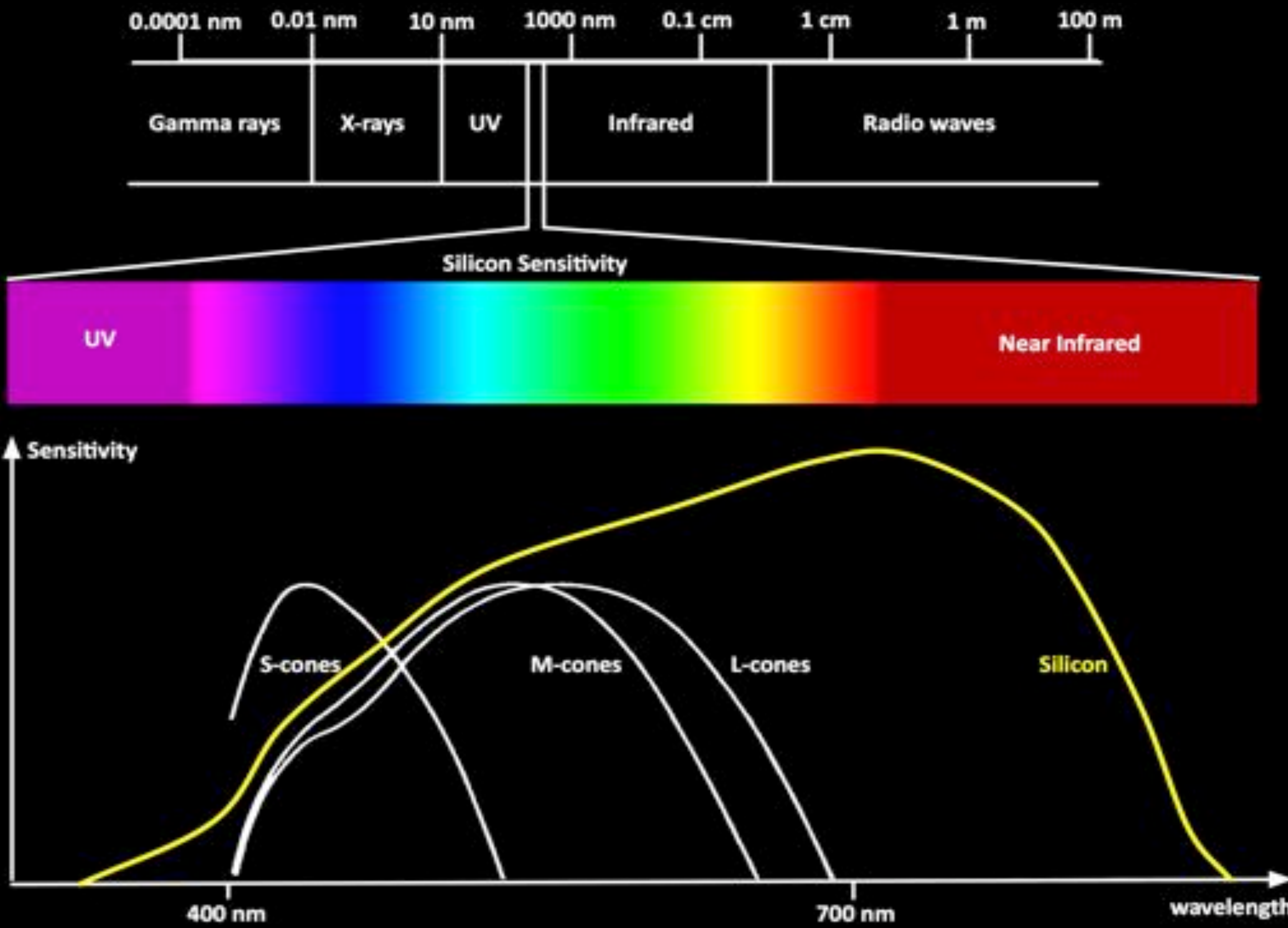
Semantic Image Segmentation



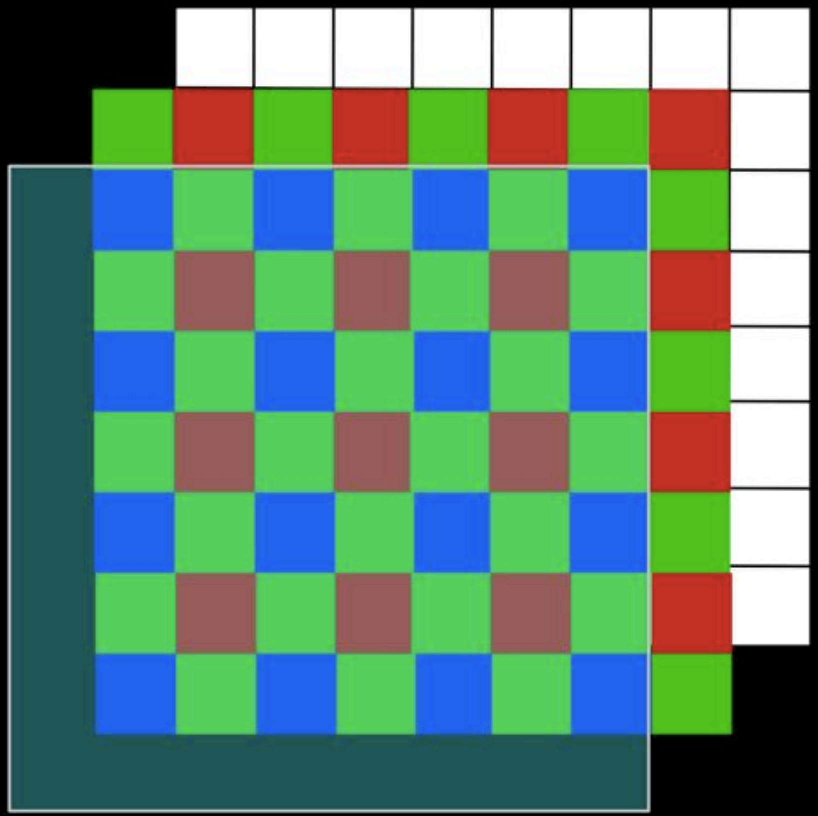
Visible NIR V Seg V+NIR Seg GT

N. Salamati, D. Larlus, G. Csurka, and S. Süsstrunk, Semantic Image Segmentation Using Visible and Near-Infrared Channels, in *ECCV's 4th Color and Photometry in Computer Vision Workshop* (2012).

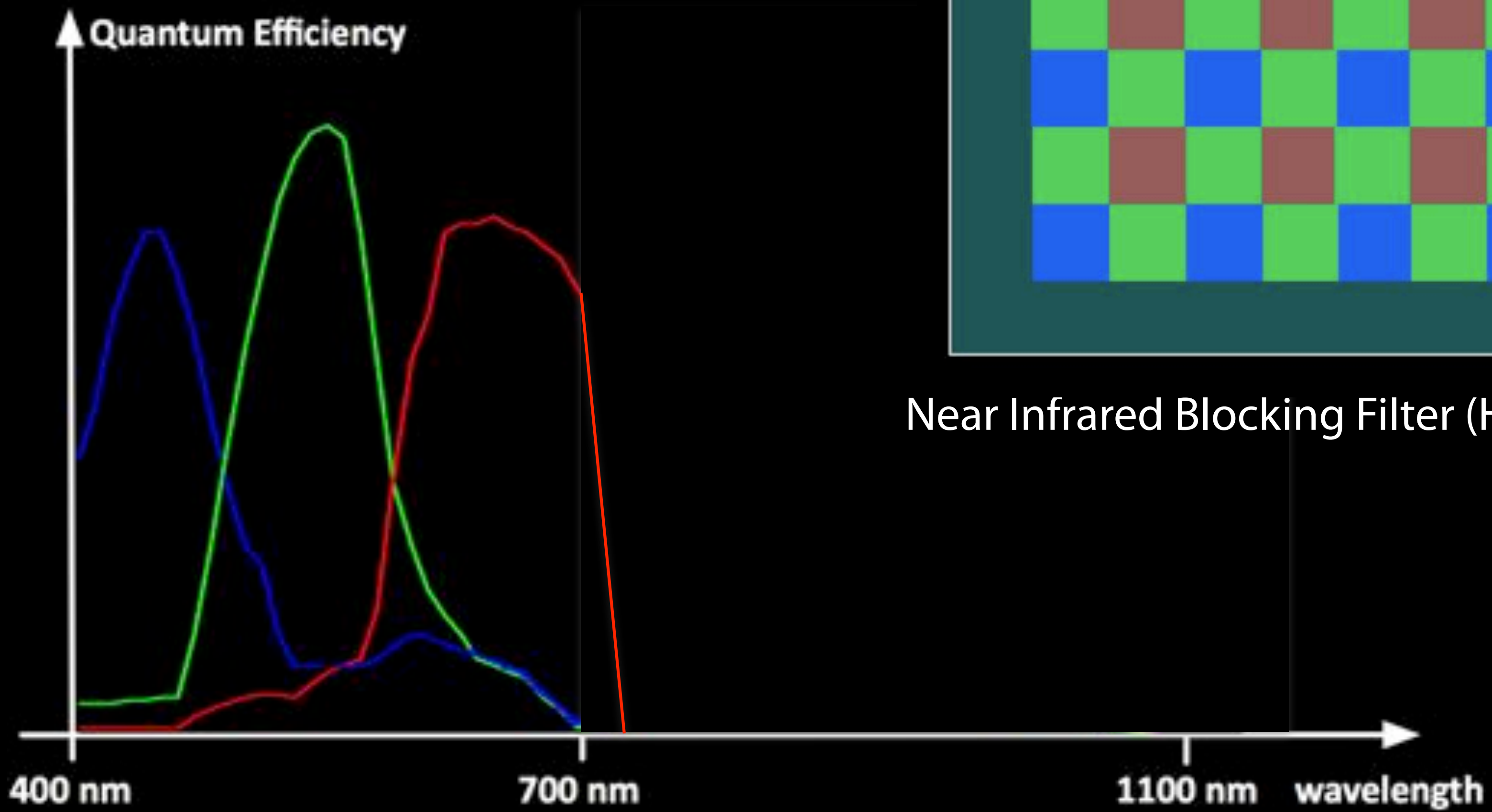
Why Near-Infrared?



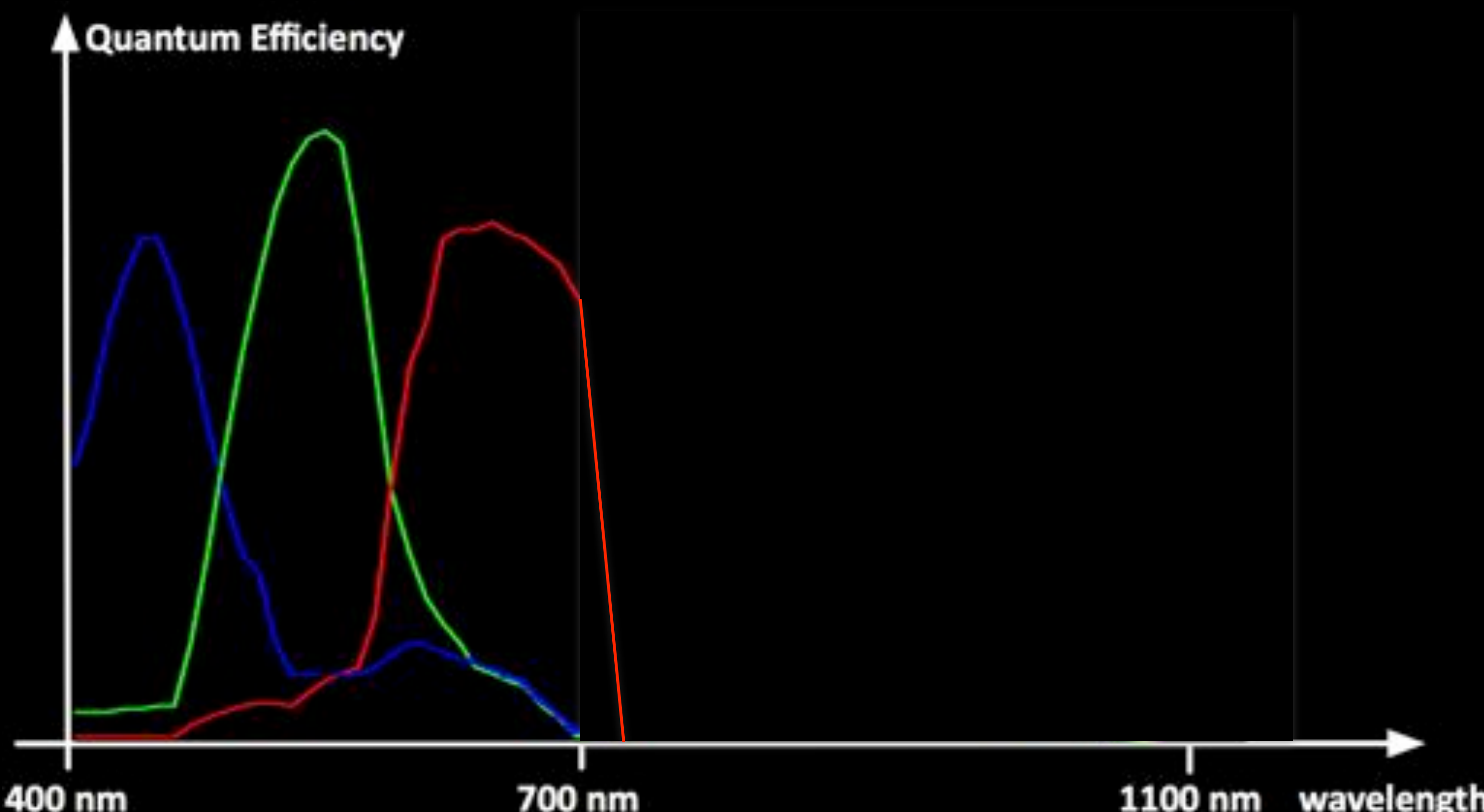
Camera Design



Near Infrared Blocking Filter (Hot Mirror)

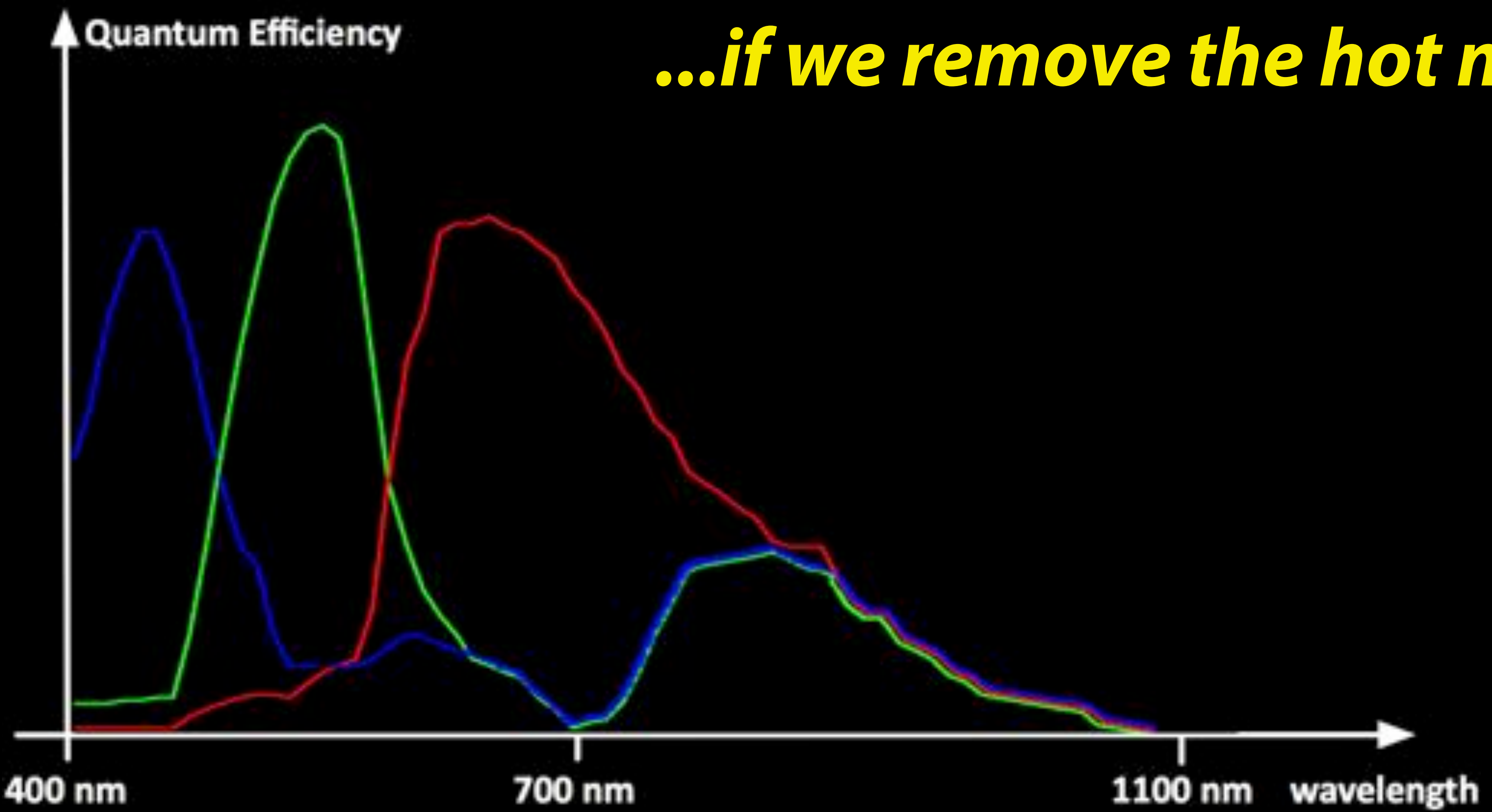


All digital cameras are inherently sensitive to NIR…

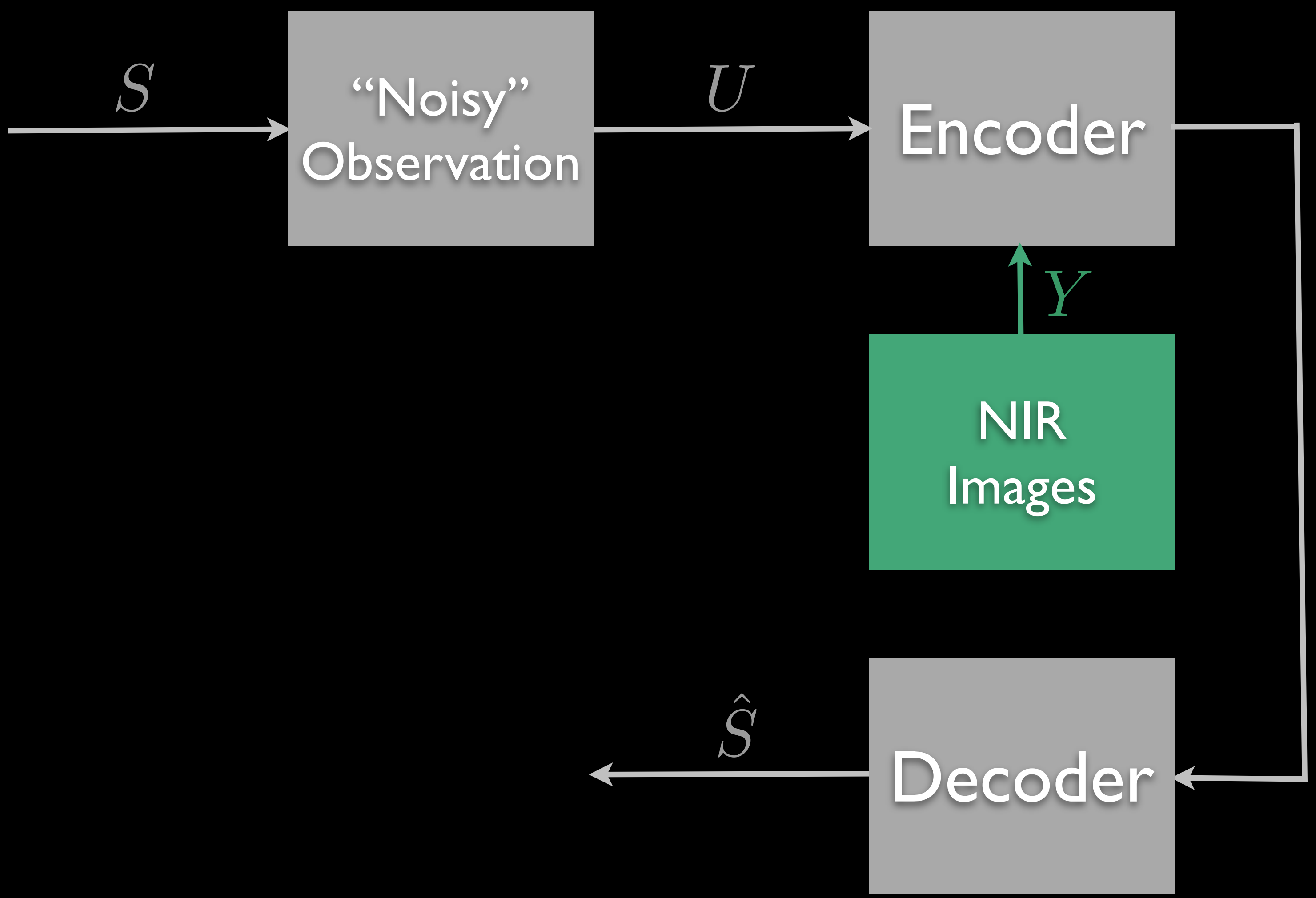


All digital cameras are inherently sensitive to NIR…

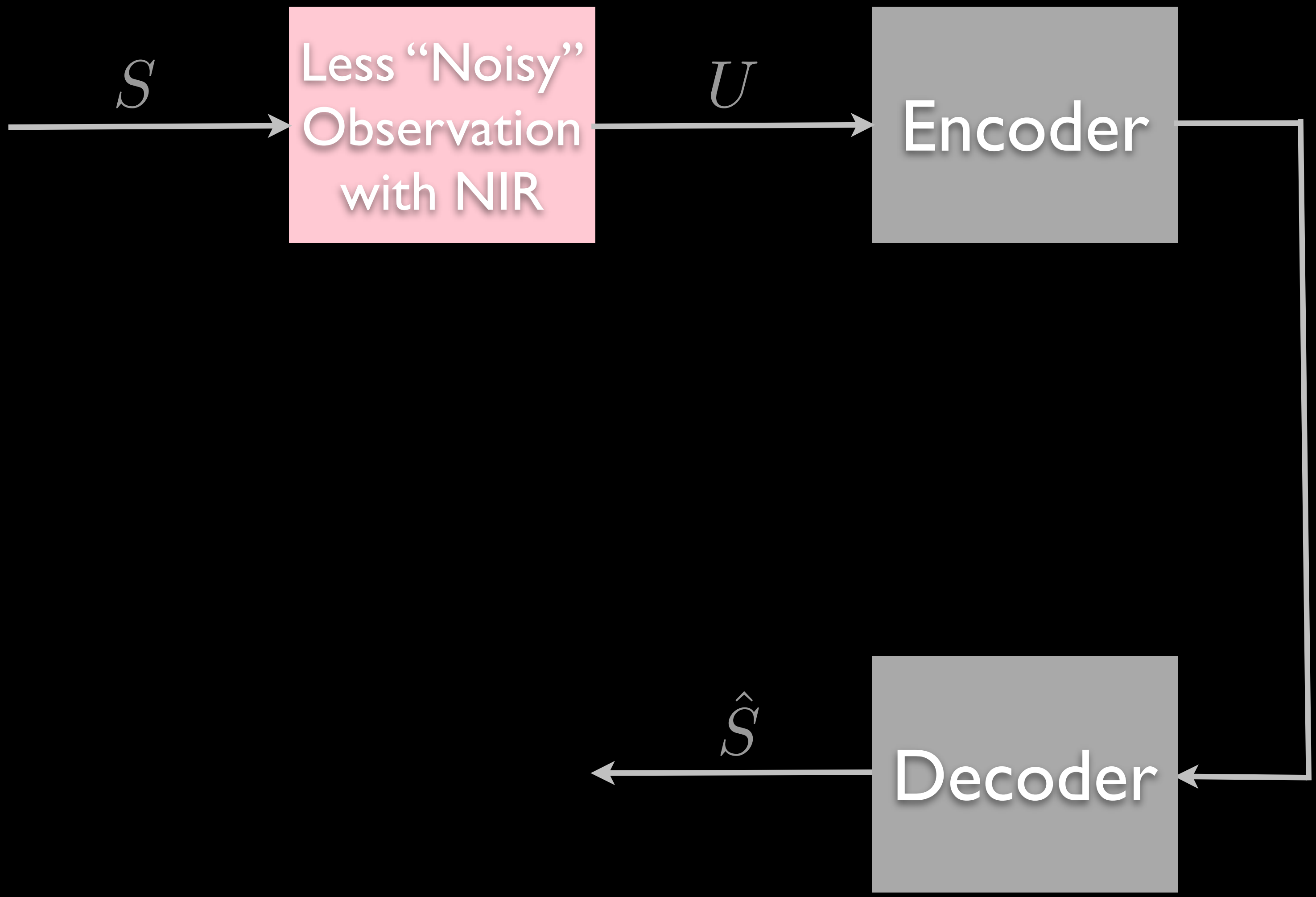
…if we remove the hot mirror!



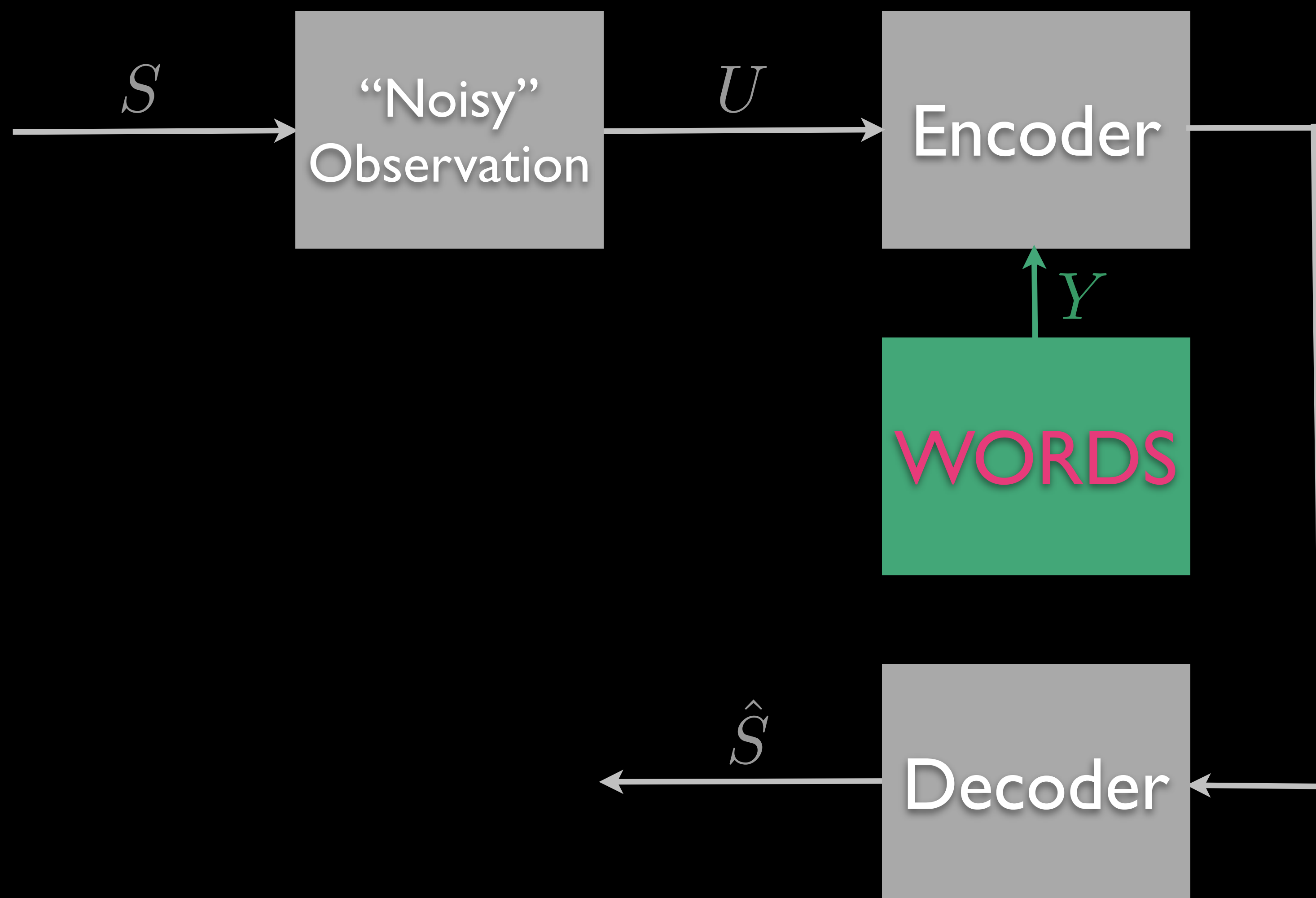
Side Information



RGBN Camera



Semantics



Which image do you like better?

sand

sunset

Which image do you like better?

dark

snow

Image rendering depends on ***context***

Camera Modes

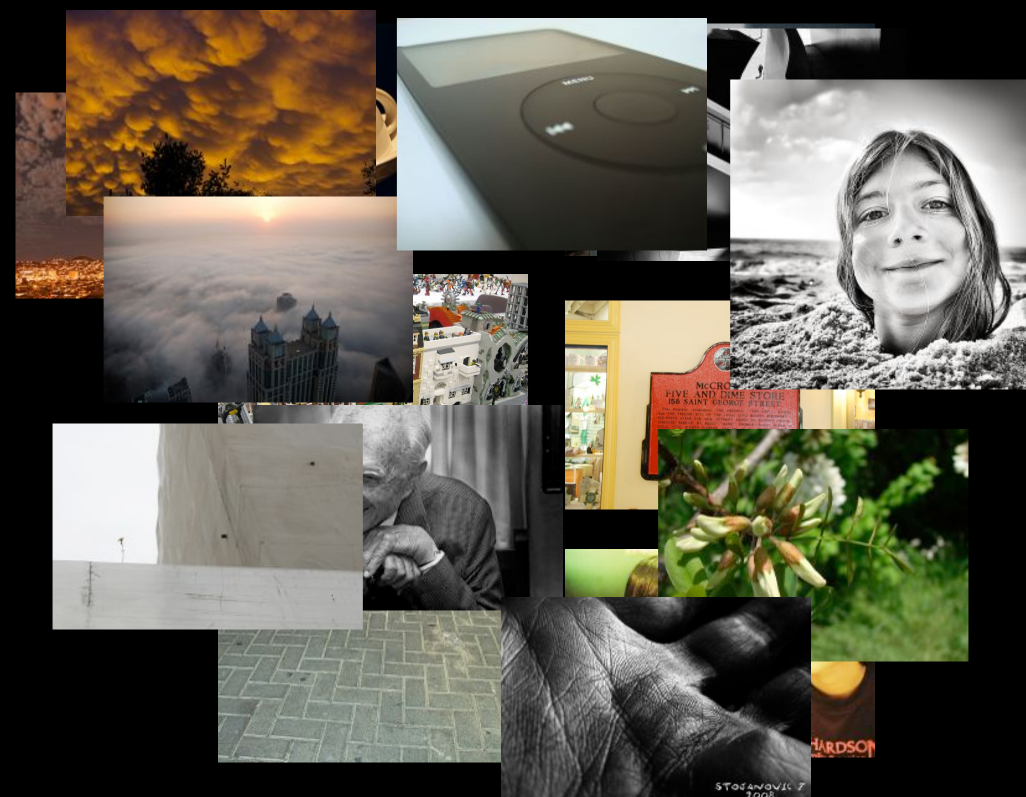
“automatic” mode

“foilage” mode

Canon PowerShot S100

Correlation?

?



How do we link image characteristics with words?

Wordl of the wikipedia page “word”

Image Database

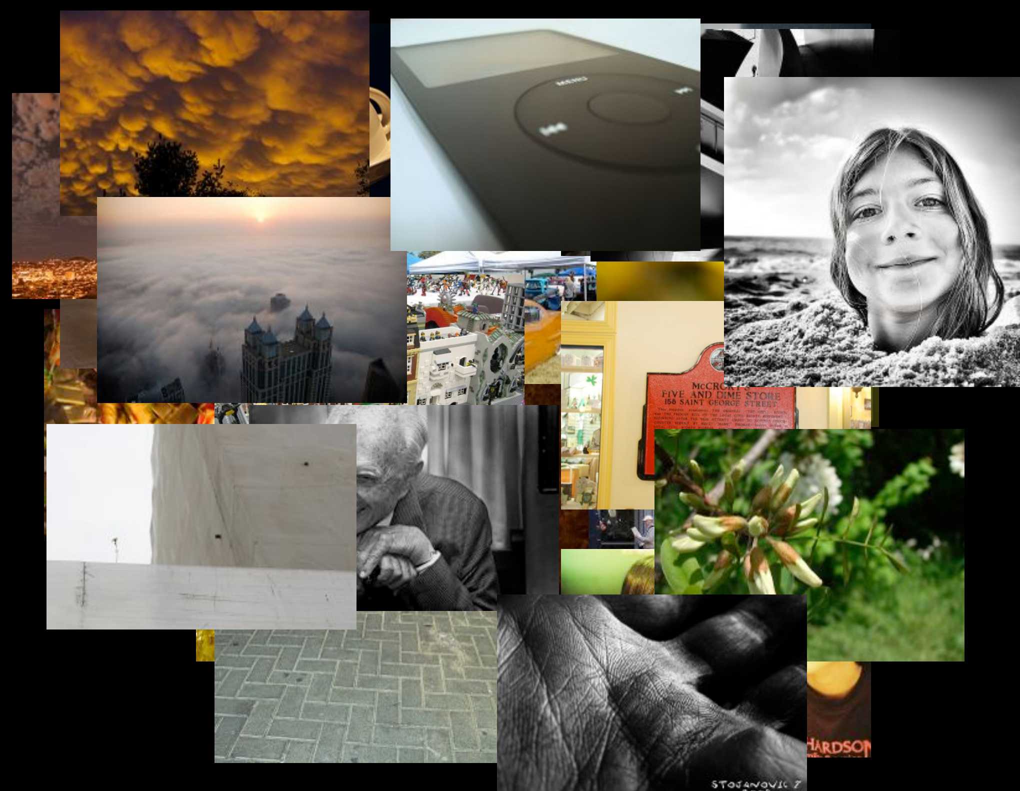
- MIR Flickr database, 1 Million annotated images.
- Selection based on Flickr's "interestingness" score.
- 1 MegaPixel, assume sRGB.

gold, oregoncoast, fortstevens, astoria, outside, lightroom, sigma, 1020mm, nikon, d40, diamondclassphotographer, grass, yellow, blue, sky, clouds, singlecloud, color, saturated, happy, field

MJ Huiskes, B Thomee, MS Lew. New trends and ideas in visual concept detection: the MIR Flickr retrieval evaluation initiative, *ACM Multimedia,* 2010.

Statistical Framework

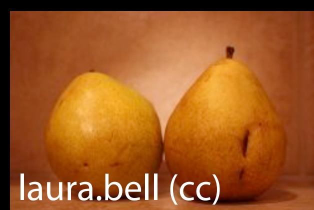
$\overline{gold}$



1M images + keywords

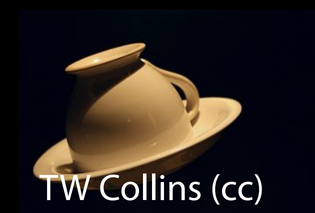
Statistical Framework

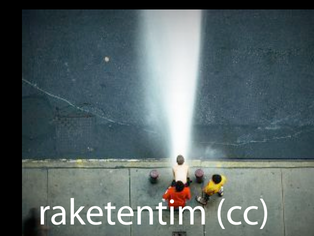
gold

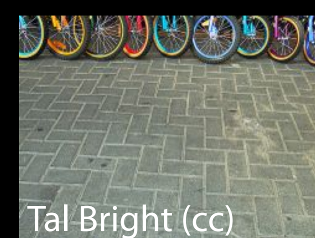


4

$\overline{gold}$



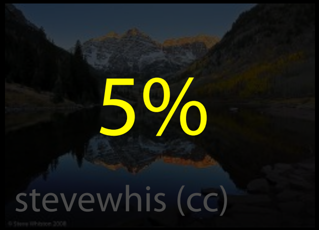


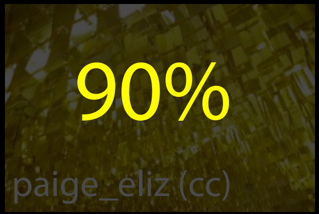


6

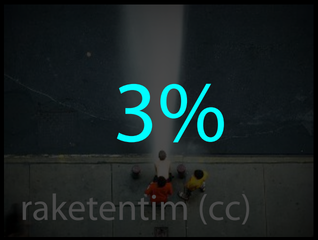
Statistical Framework

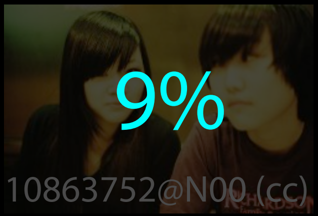
gold





$\overline{gold}$





percentage of yellow pixels

Statistical Framework

sorted list: 0% 2% 3% 5% 8% 9% 10% 30% 70% 90%

rank index: 1 2 3 4 5 6 7 8 9 10

ranksum: $T = 4 + 7 + 9 + 10 = 30$

Mann-Whitney-Wilcoxon ranksum test

$$\mu_T = \frac{n_w(n_w + n_{\overline{w}} + 1)}{2} \qquad \sigma_T^2 = \frac{n_w n_{\overline{w}}(n_w + n_{\overline{w}} + 1)}{12}$$

$n_w, n_{\overline{w}}$ cardinalities of both sets

$$z = \frac{T - \mu_T}{\sigma_T} = \frac{30 - 22}{4.69} \approx 1.71$$

F. Wilcoxon, Individual comparisons by ranking methods, Biometrics Bulletin, 1(6):80–83, 1945

Statistical Framework

sorted list: 0% 2% 3% 5% 8% 9% 10% 30% 70% 90%

rank index: 1 2 3 4 5 6 7 8 9 10

ranksum: $T = 4 + 7 + 9 + 10 = 30$

Mann-Whitney-Wilcoxon ranksum test

$$\mu_T = \frac{n_w(n_w + n_{\overline{w}} + 1)}{2} \qquad \sigma_T^2 = \frac{n_w n_{\overline{w}}(n_w + n_{\overline{w}} + 1)}{12}$$

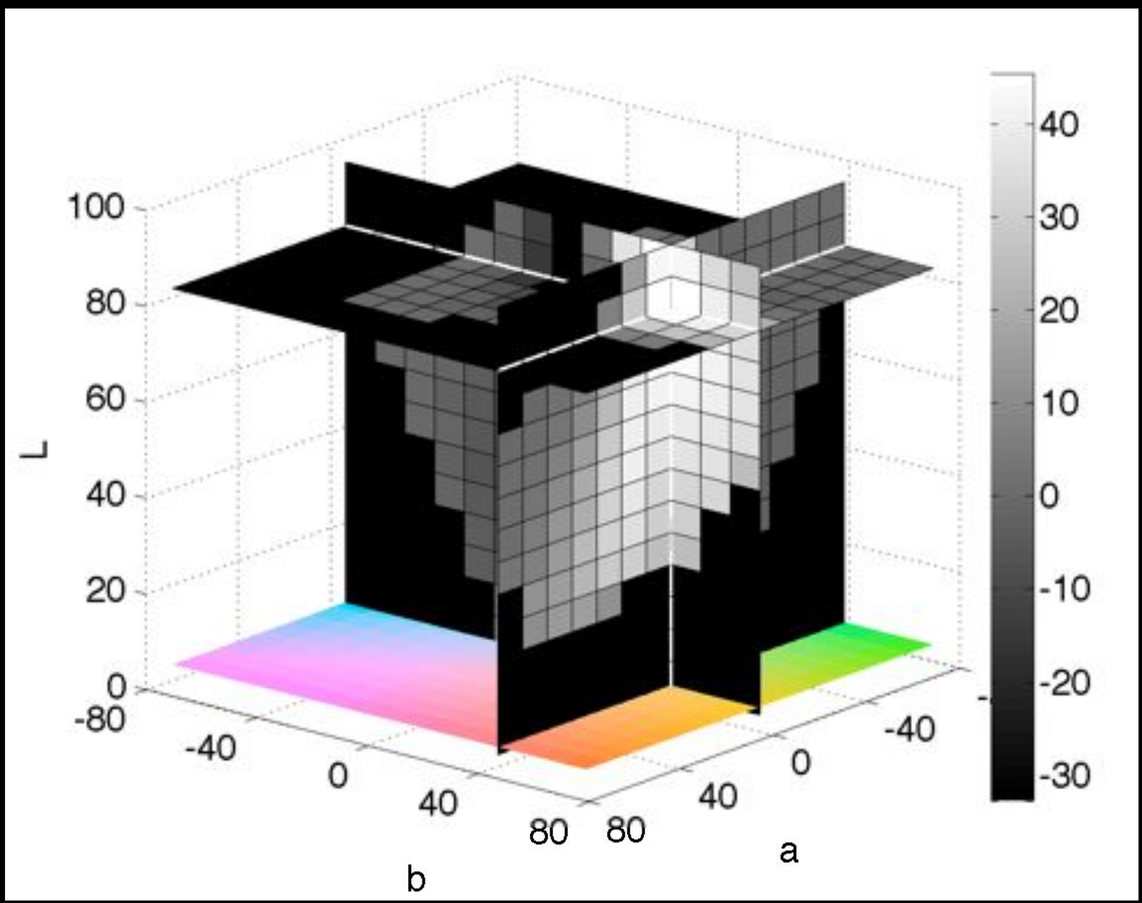
$n_w, n_{\overline{w}}$ cardinalities of both sets

$$z = \frac{T - \mu_T}{\sigma_T} = \frac{30 - 22}{4.69} \approx 1.71$$

$z > 0$ ➔ significantly more yellow pixels in *gold* images.

z Distribution

gold



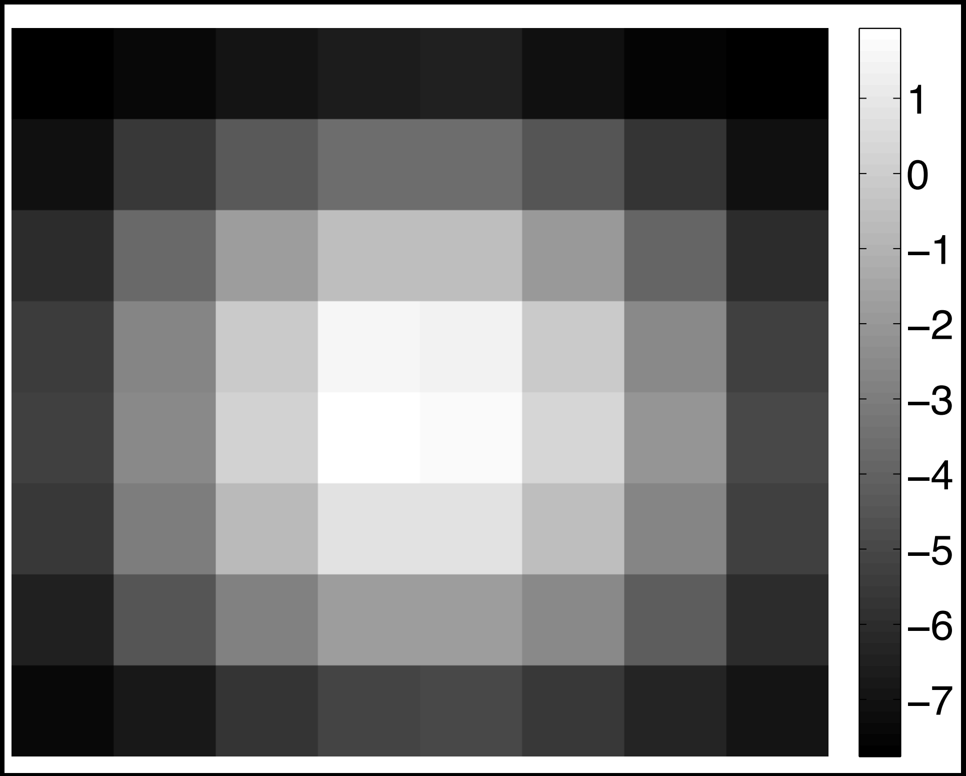
- CIELAB histogram 15x15x15 bins.
- z values indicate significance of a keyword w.r.t. to a characteristic.

A. Lindner, A. Shaji, N. Bonnier, and S. Süsstrunk, Joint Statistical Analysis of Images and Keywords with Applications in Semantic Image Enhancement, *ACM Multimedia*, 2012.

Other Characteristics

Spatial lightness layout.

light

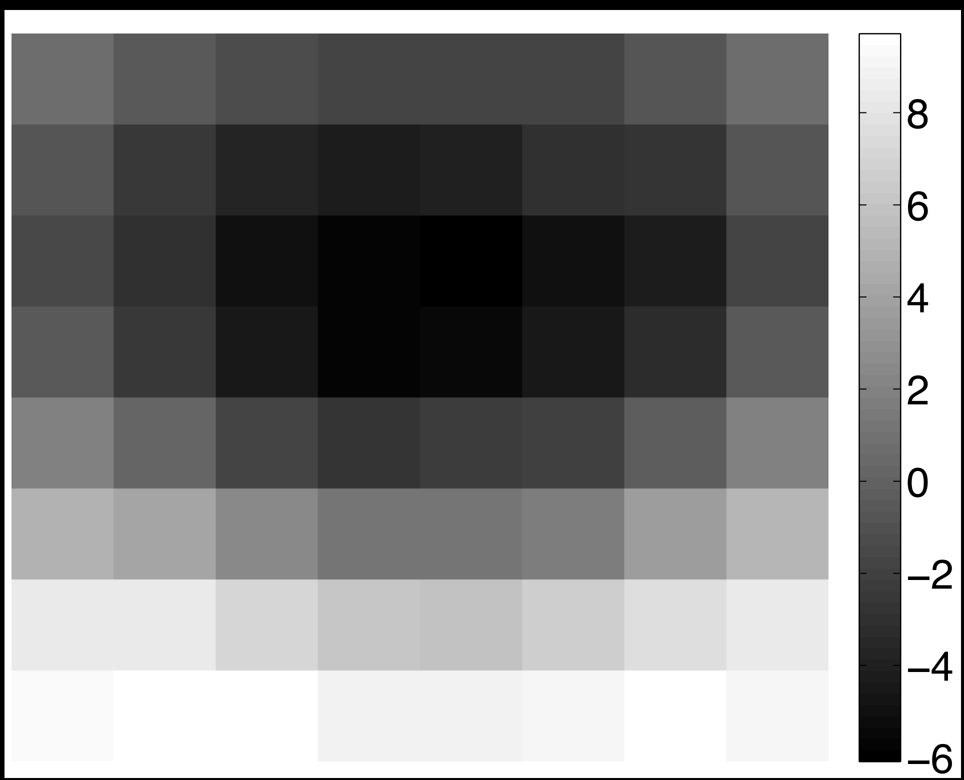


A. Lindner, A. Shaji, N. Bonnier, and S. Süsstrunk, Joint Statistical Analysis of Images and Keywords with Applications in Semantic Image Enhancement, *ACM Multimedia*, 2012.

Other Characteristics

Spatial chroma layout.

barn

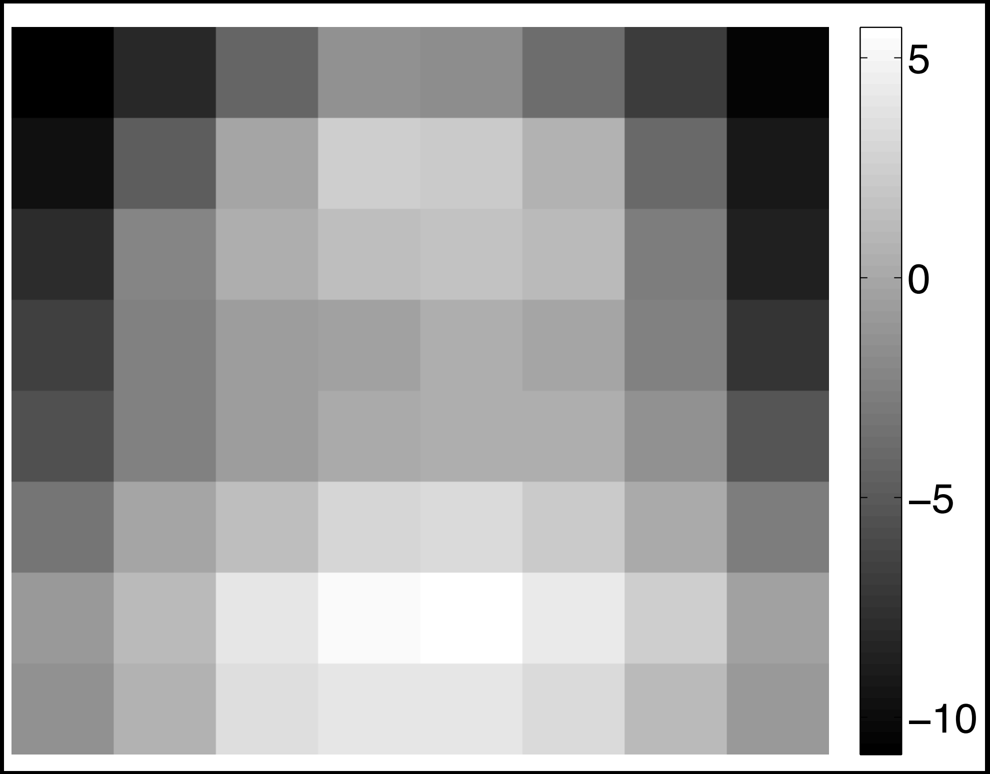


A. Lindner, A. Shaji, N. Bonnier, and S. Süsstrunk, Joint Statistical Analysis of Images and Keywords with Applications in Semantic Image Enhancement, *ACM Multimedia*, 2012.

Other Characteristics

Spatial Gabor filter layout.

fireworks



A. Lindner, A. Shaji, N. Bonnier, and S. Süsstrunk, Joint Statistical Analysis of Images and Keywords with Applications in Semantic Image Enhancement, *ACM Multimedia*, 2012.

Semantic Image Enhancement

Gray scale tone mapping

snow

Color enhancement

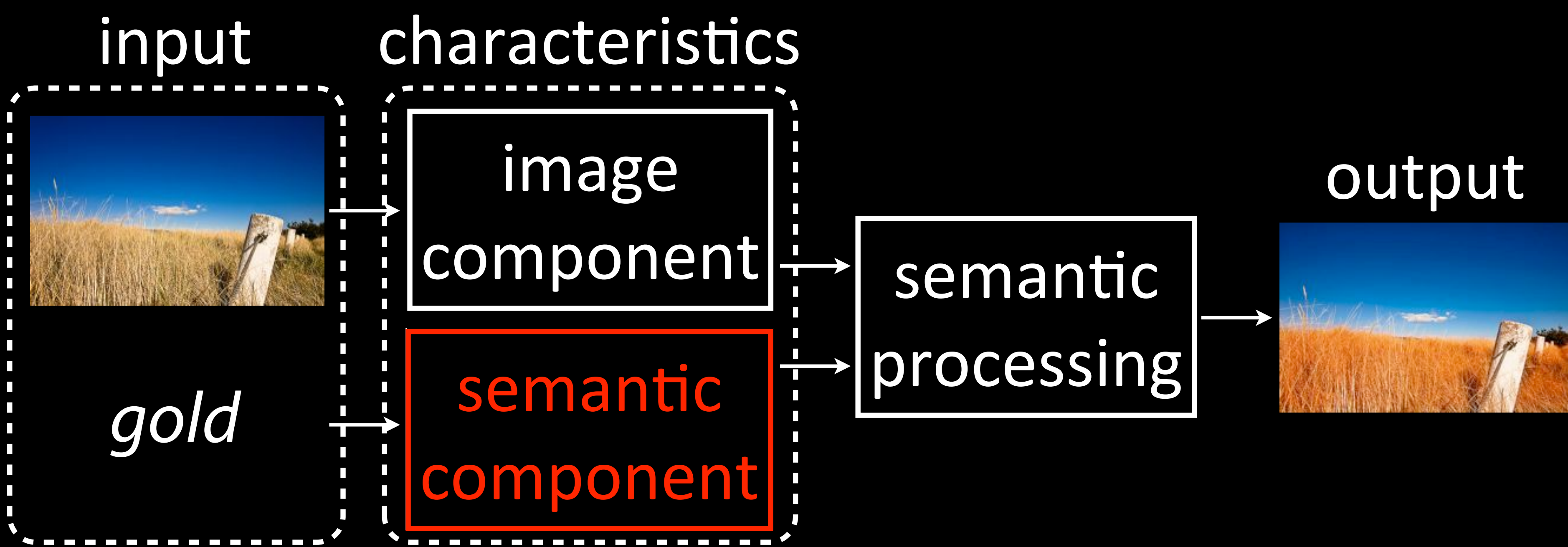
gold

Change depth-of-field

macro

[Zhuo and Sim, 2011]

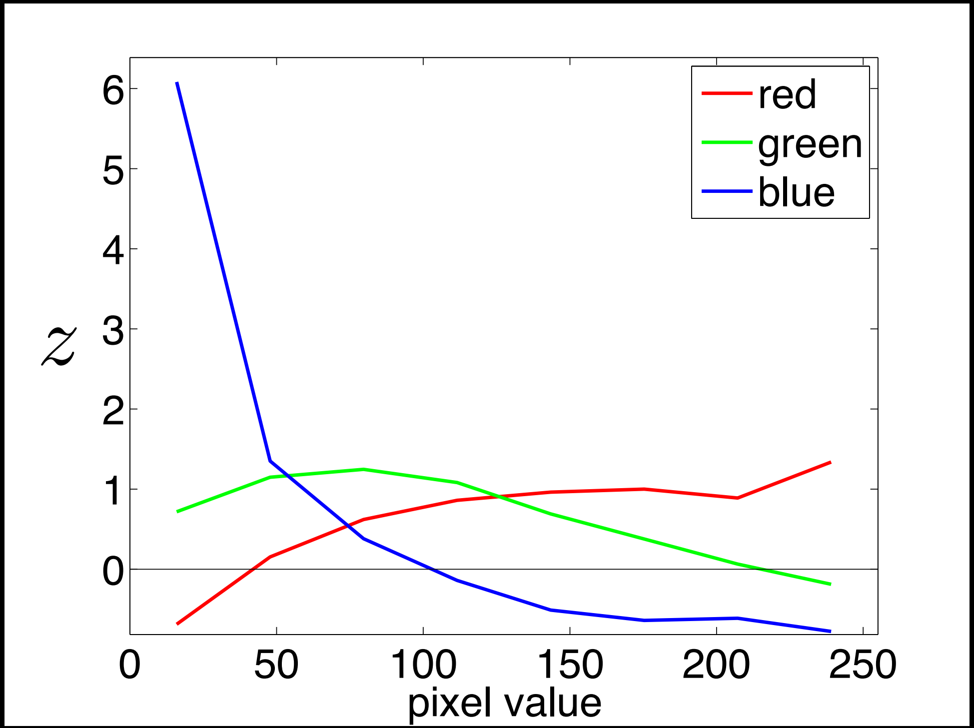
Semantic Enhancement



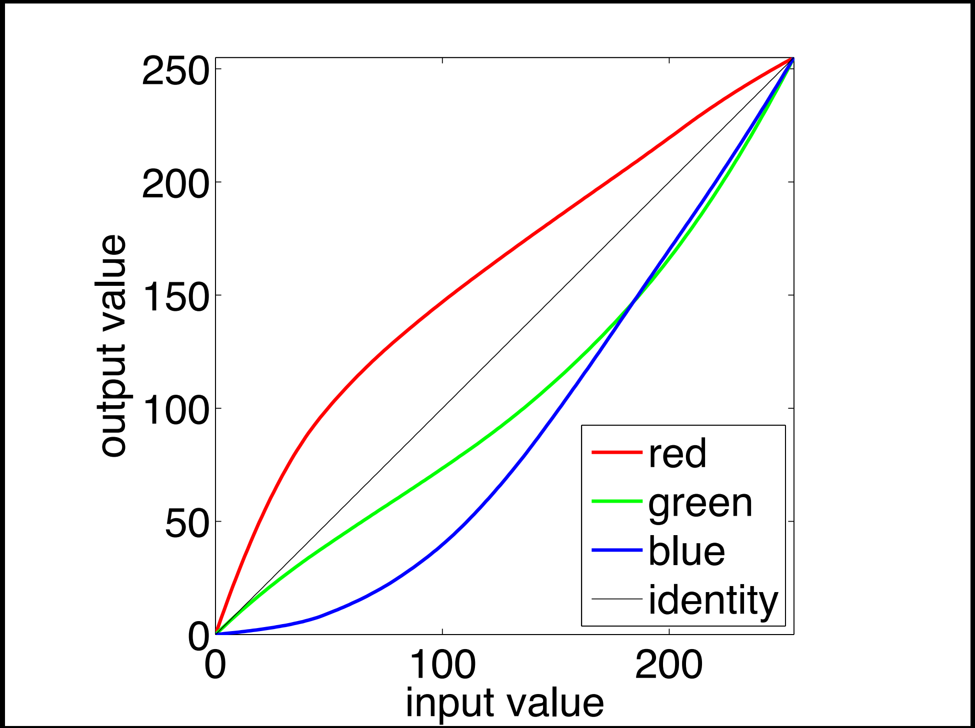
A. Lindner, A. Shaji, N. Bonnier, and S. Süsstrunk, Joint Statistical Analysis of Images and Keywords with Applications in Semantic Image Enhancement, *ACM Multimedia*, 2012.

Semantic Component

significance values for *gold*



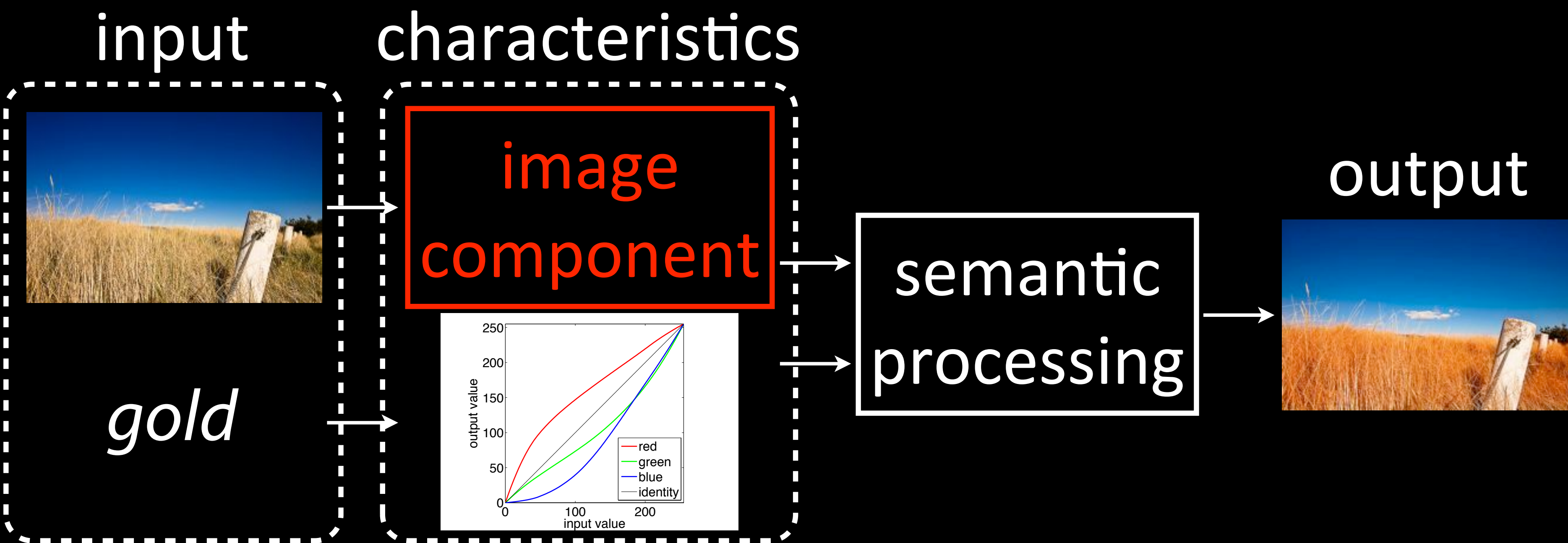
Tone mapping function f



$$f' = \begin{cases} 1/\left(1 + Sz\right) & \text{if } z \geq 0 \\ 1 + S|z| & \text{if } z < 0 \end{cases}$$

S global scale parameter

Semantic Enhancement



A. Lindner, A. Shaji, N. Bonnier, and S. Süsstrunk, Joint Statistical Analysis of Images and Keywords with Applications in Semantic Image Enhancement, *ACM Multimedia*, 2012.

Image Component

gold

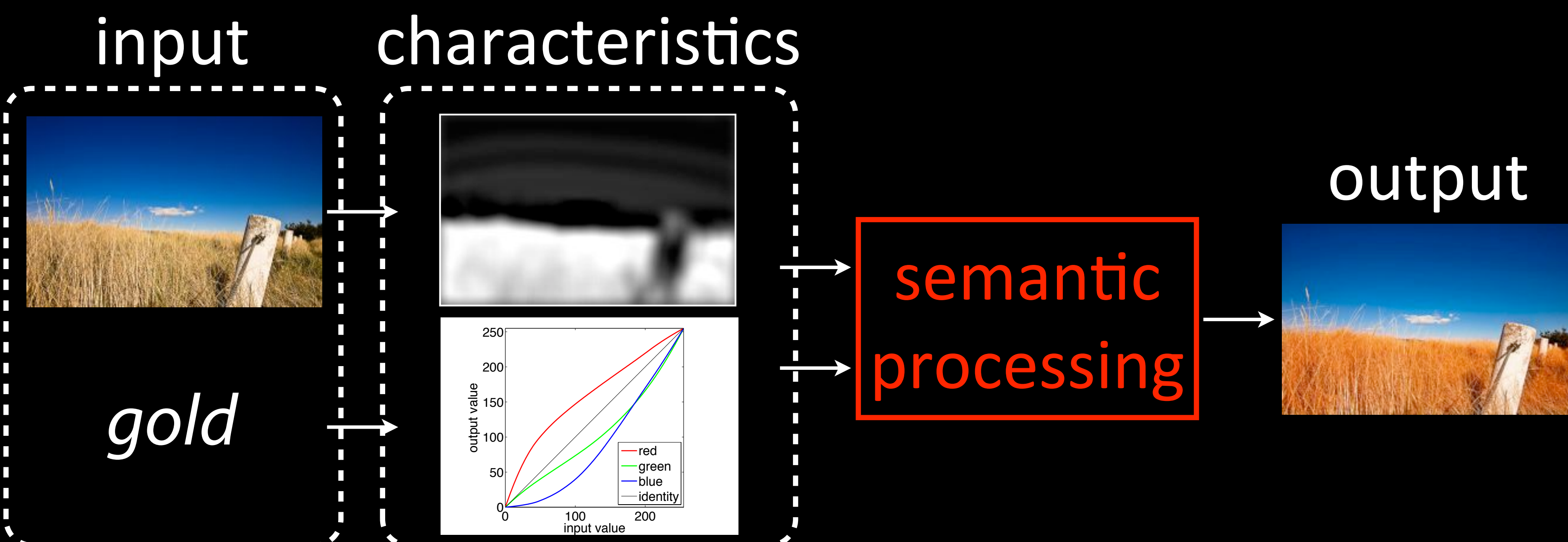
weight map ω

$$\omega = \left[g_\sigma * z_w\big(\mathrm{col}(p)\big)\right]_0^1$$

g_σ Gaussian blurring kernel (1% of image diagonal)

$[\,\cdot\,]_0^1$ normalization operator

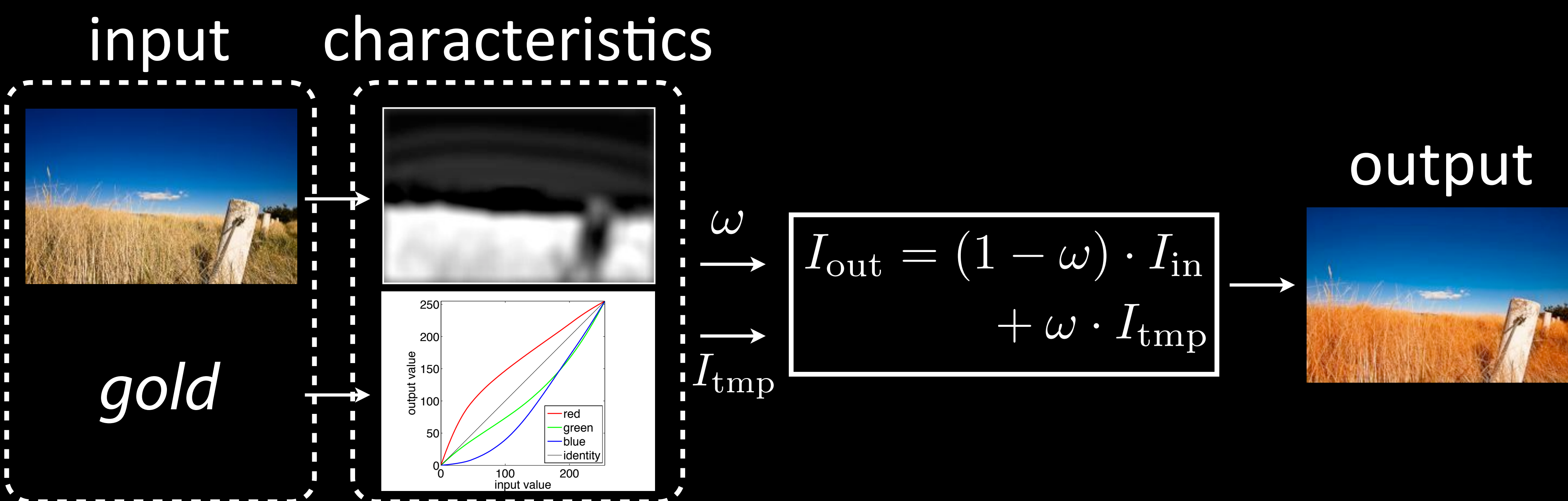
Semantic Enhancement



A. Lindner, A. Shaji, N. Bonnier, and S. Süsstrunk, Joint Statistical Analysis of Images and Keywords with Applications in Semantic Image Enhancement, *ACM Multimedia*, 2012.

Semantic Enhancement

Enhance relevant characteristics in relevant regions.



A. Lindner, A. Shaji, N. Bonnier, and S. Süsstrunk, Joint Statistical Analysis of Images and Keywords with Applications in Semantic Image Enhancement, *ACM Multimedia*, 2012.

sand

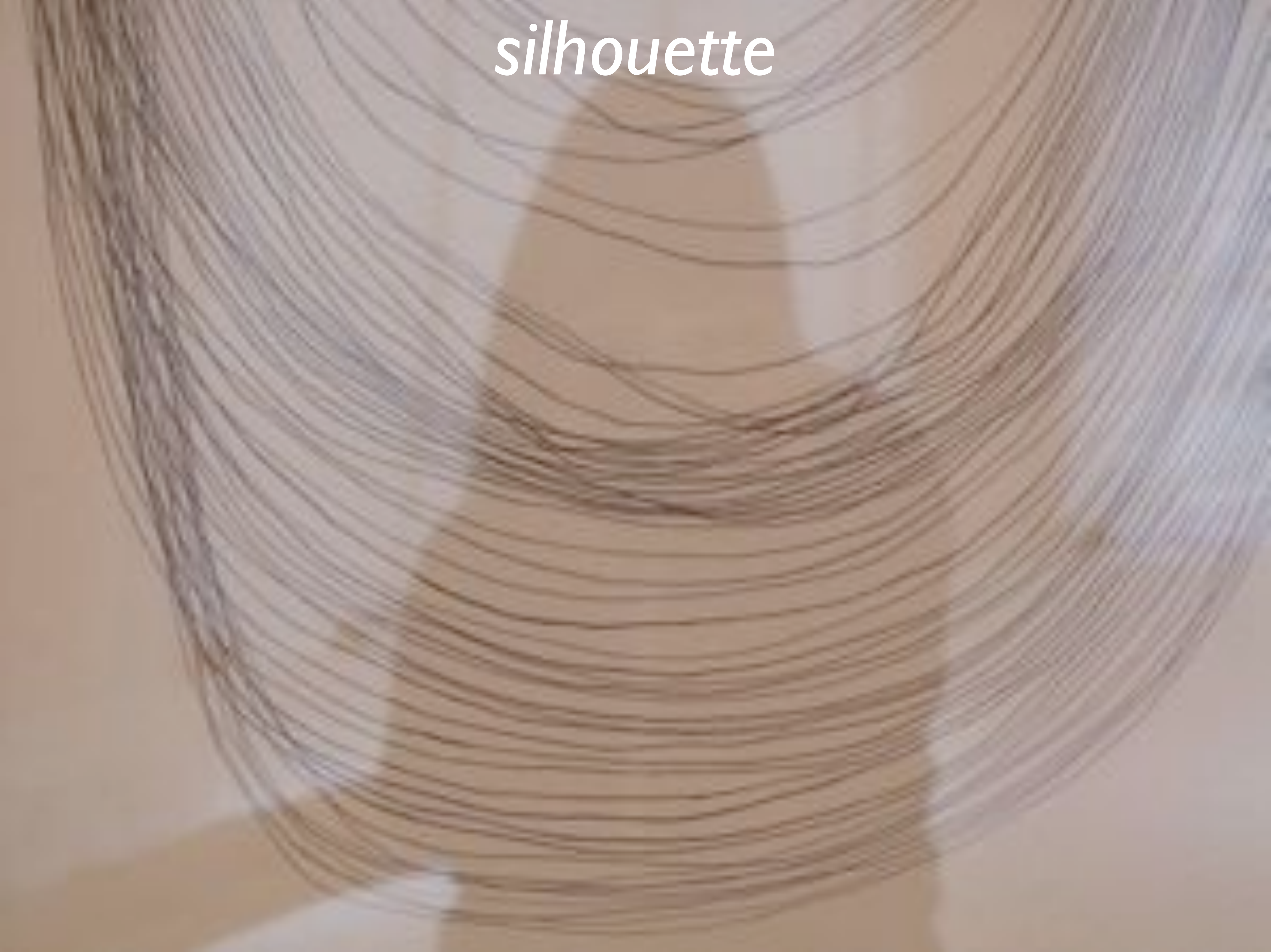
sand

snow

snow


dark


dark


silhouette


silhouette

sunset

sunset

grass

grass


autumn


autumn

strawberry

strawberry


sky


sky

banana


banana


macro


macro


flower

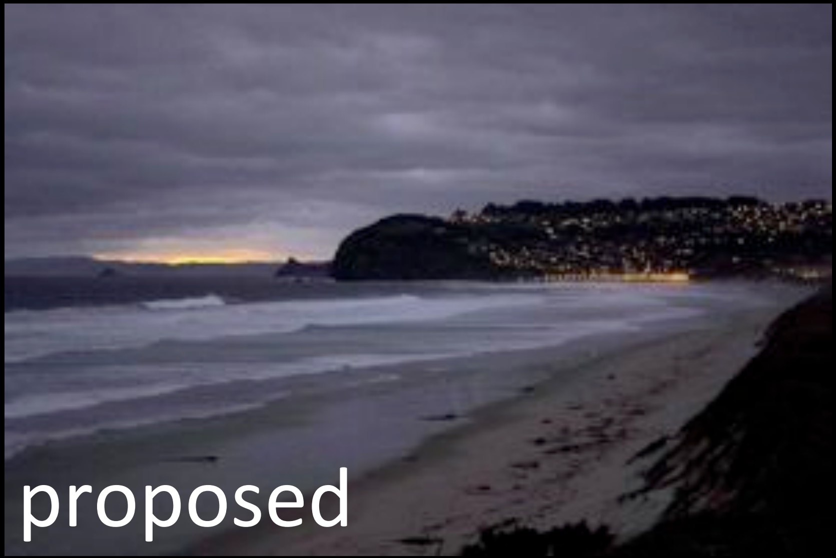
flower


macro


macro

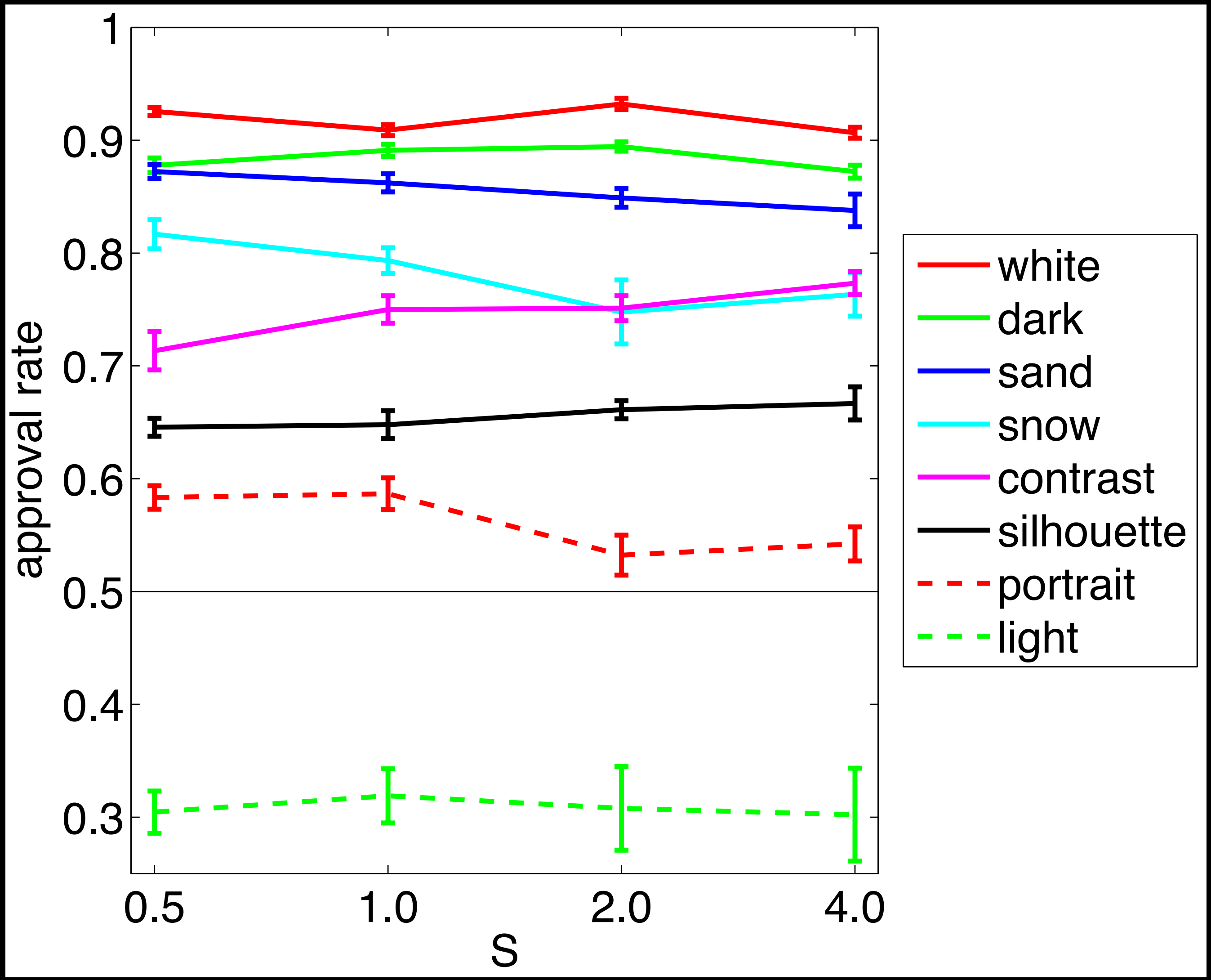
System Performance?

sand



8 keywords, 30 images, $S = \{0.5,\ 1,\ 2,\ 4\}$, 30 observers =28'800 image comparisons.

Amazon Mechanical Turk



A. Lindner, A. Shaji, N. Bonnier, and S. Süsstrunk, Joint Statistical Analysis of Images and Keywords with Applications in Semantic Image Enhancement, *ACM Multimedia*, 2012.

Conclusion

- Adding ***side information*** to the encoding, decoding or both can improve ***system performance,***
  - and that can be “like” or “dislike” of photos.

Conclusion

- Adding side information to the encoding, decoding or both can improve system performance
  - and that can be like or dislike of photos
- The given examples are in ***image enhancement*** (computational photography), but the same concept is also applicable to ***computer vision*** applications.

Conclusion

- Considering todays capture devices (cell phones, android cameras) with ***multiple sensors*** and ***connected to the internet***, there is more side information available than ever!

http://gizmoreport.com/huawei-ascend-mate-info/

Conclusion

- Considering todays capture devices (cell phones, android cameras) with ***multiple sensors*** and ***connected to the internet***, there is more side information available than ever!

http://www.newgadgets.de/46122/nikon-coolpix-s800c-16-megapixel-kamera-mit-android/

Conclusion

- Side Information needs to be ***correlated*** with the signal, and that might not always be so evident with non-image signals.
  - Let your imagination rule!

Acknowledgment

- Joint work with
  - *Nathalie Barbuscia, EPFL*
  - *Nicolas Bonnier, OCE now Canon CiSRA*
  - *Clément Fredembach, EPFL now Canon CiSRA*
  - *Albrecht Lindner, EPFL*
  - *Yue M. Lu, EPFL now Harvard*
  - *Neda Salamati, EPFL*
  - *Zahra Sadeghipoor, EPFL*
  - *Lex Schaul, EPFL now ETHZ*
  - *Appu Shaji, EPFL*
  - *Daniel Tamburrino, EPFL now ECAL*
- Funding
  - *Swiss National Science Foundation*
  - *OCE Print Logic, a Canon Group*
  - *Hasler Foundation*
  - *Xerox Foundation*

Thank you!

ivrg.epfl.ch