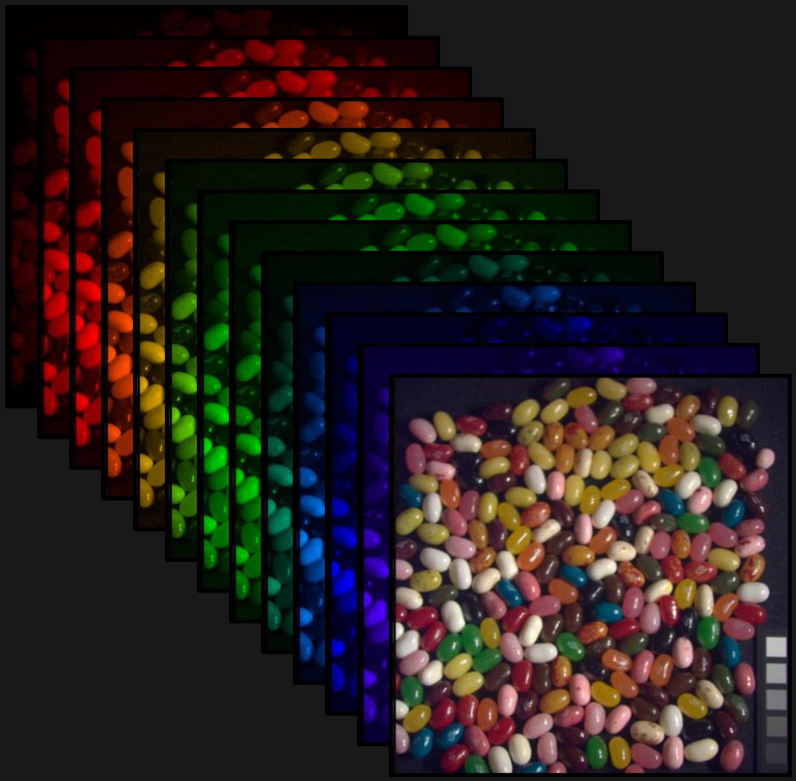
Generalized Assorted Camera Arrays: Robust Cross-channel Registration and Applications

Jason Holloway, Kaushik Mitra,
Sanjeev Koppal, Ashok Veeraraghavan

RICE®

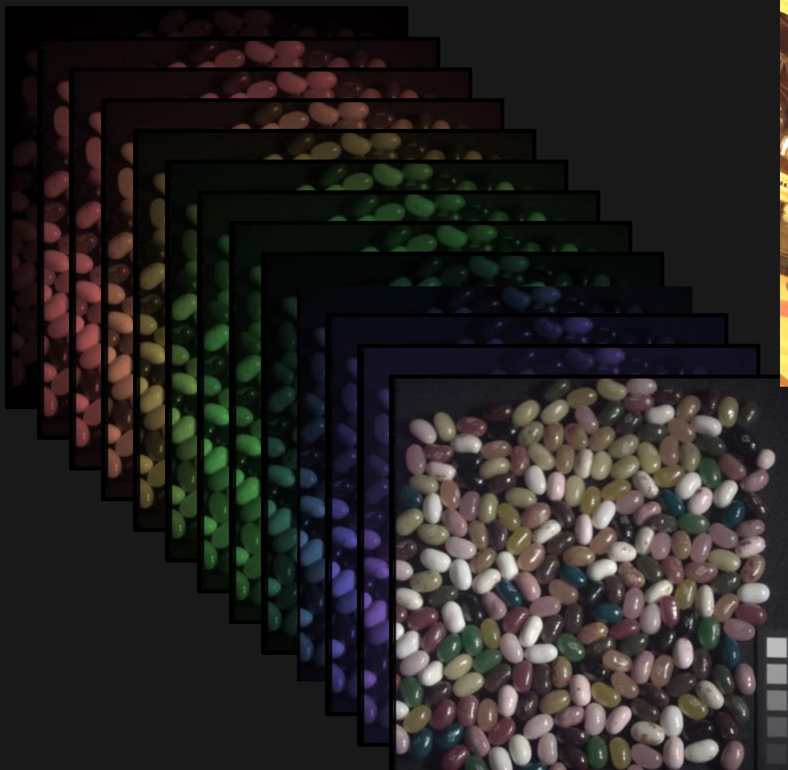
DSP

Cross-modal Imaging

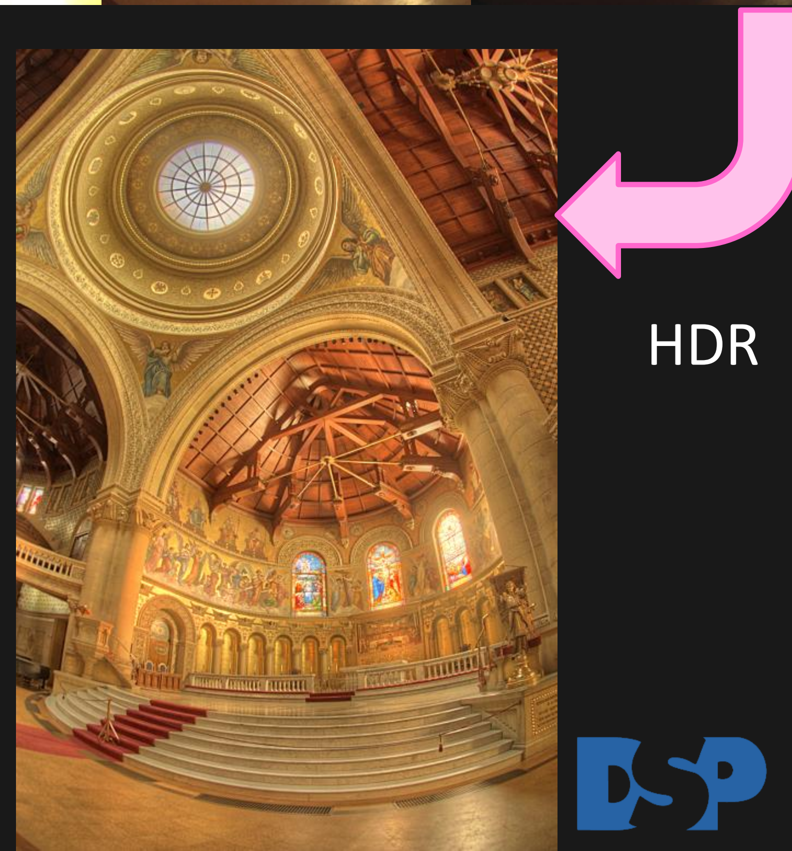


Hyperspectral

Cross-modal Imag

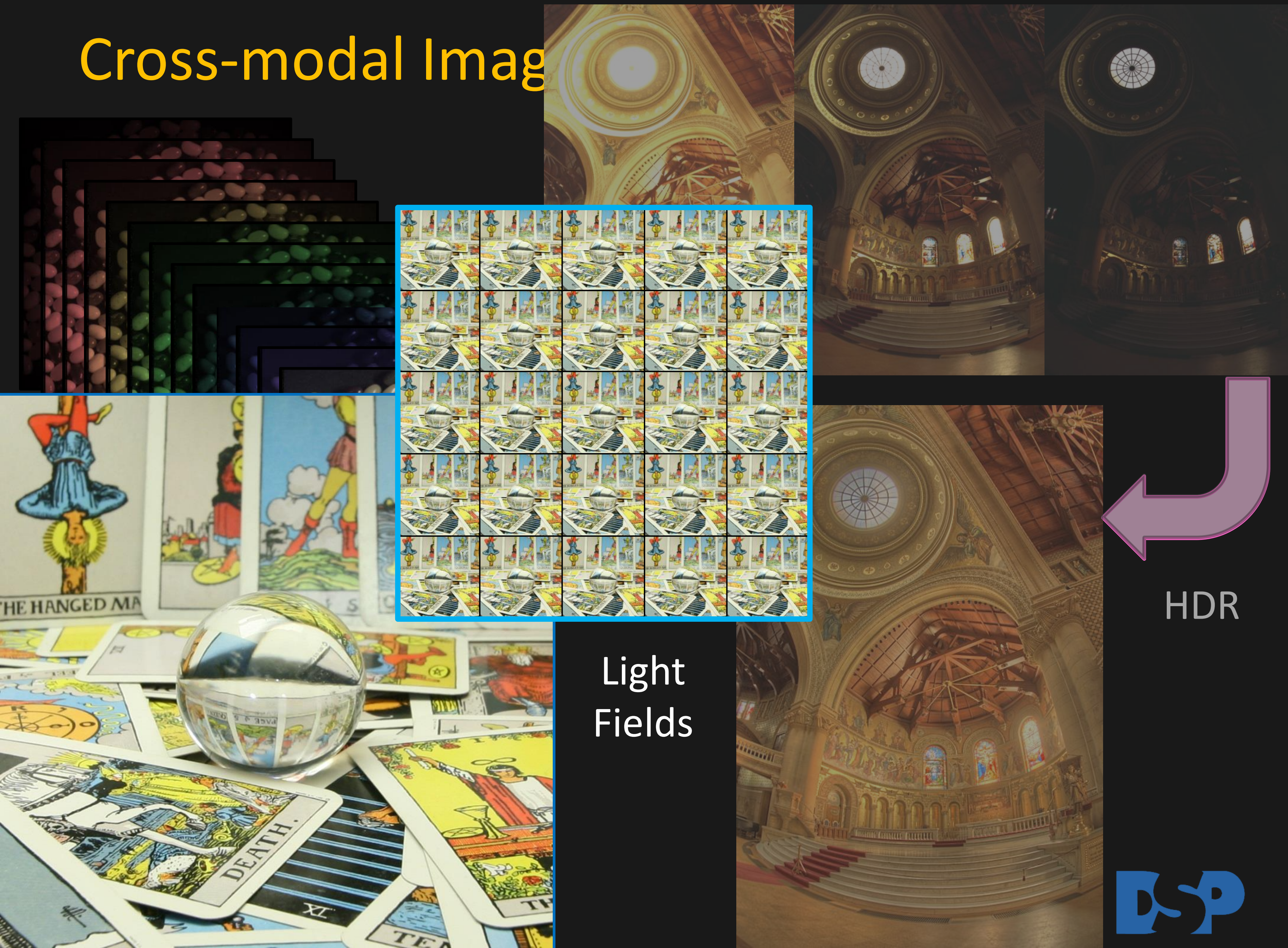


Hyperspectral



HDR

Cross-modal Imag

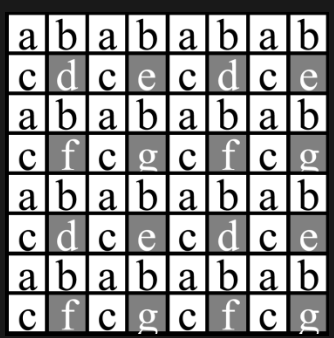


Limitations in Cross-modal

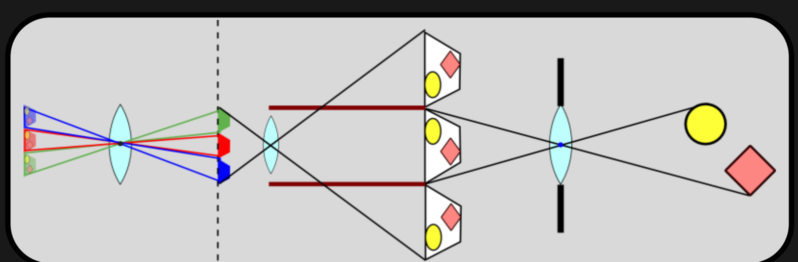
- Two common methods for cross-modal image acquisition

- Sequential capture
  - Filter wheels, liquid tunable filters

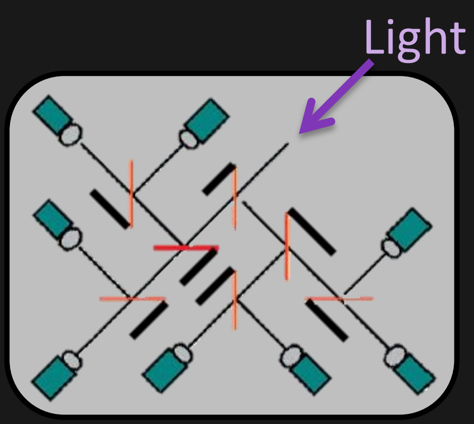
- Precise optical alignment
  - Beam splitting or Filter Array



Yasuma et al.



Manakov et al.



McGuire et al.

Solution: Camera arrays

- Simultaneous capture (dynamic scenes)
- Each view can be high resolution, different channels
- Provides angular information

Pelican Imaging

Stanford Camera Array

ProFUSION

Set back: Parallax

- Aligning images: scene-dependent registration
- Computing stereo correspondence requires redundant cameras
- Pelican Imaging – 16 cameras record only 3 unique channels
- There is a need for <u>cross-channel image registration</u>
  - Remove redundancy
  - Shrink array size

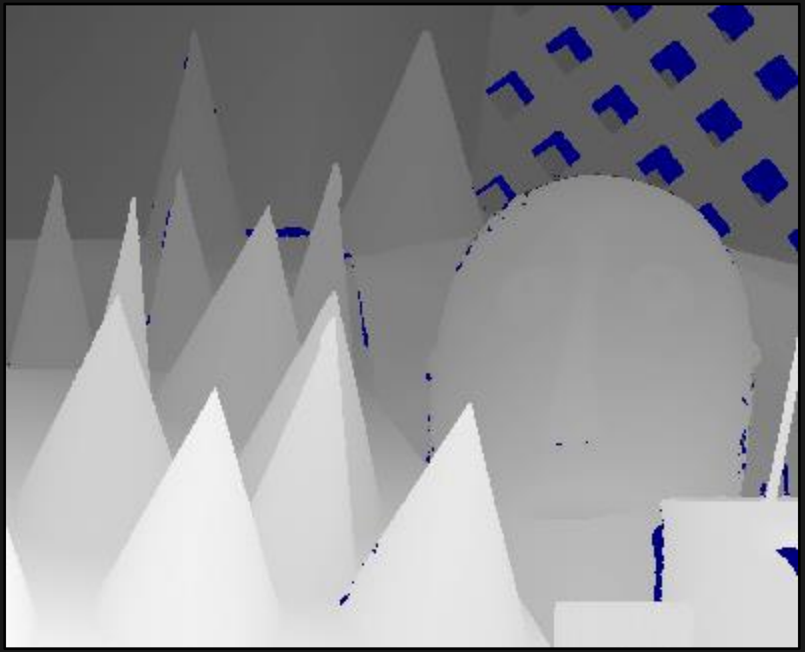
Contributions

1. We develop a novel cost metric for cross-channel registration
2. Reduce camera-to-channel ratio of camera arrays without sacrificing resolution or light throughput
3. Demonstrate GAC for consumer imaging
4. Enable flexible application-specific imaging applications
5. Capture hyperspectral video with high SNR

Cross-channel Image Registration

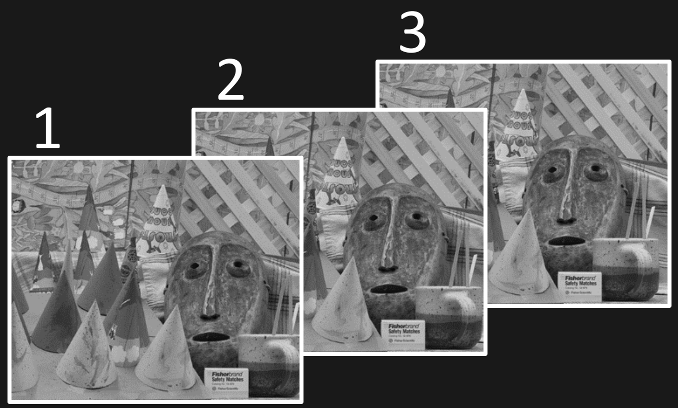
- Simulated cross-channel matching using Middlebury dataset
  - Multi-view stereo with 3 viewpoints

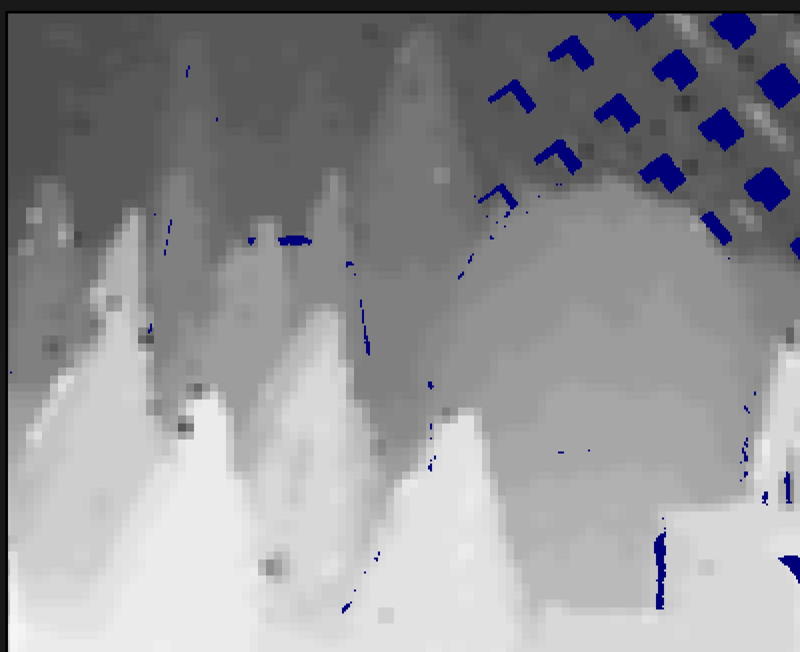
Reference view



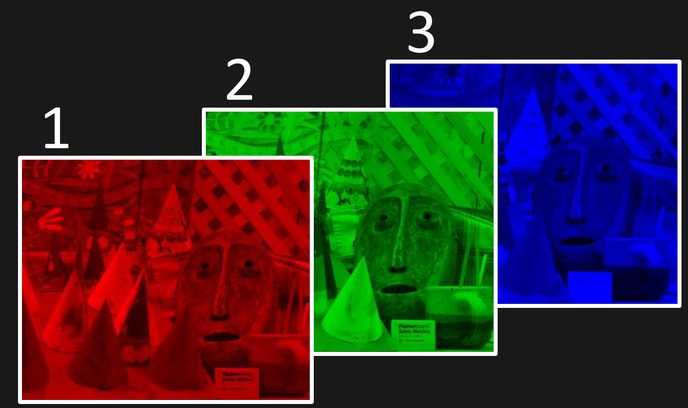
Ground truth disparity

SSD Intra- and Inter-channel Performance



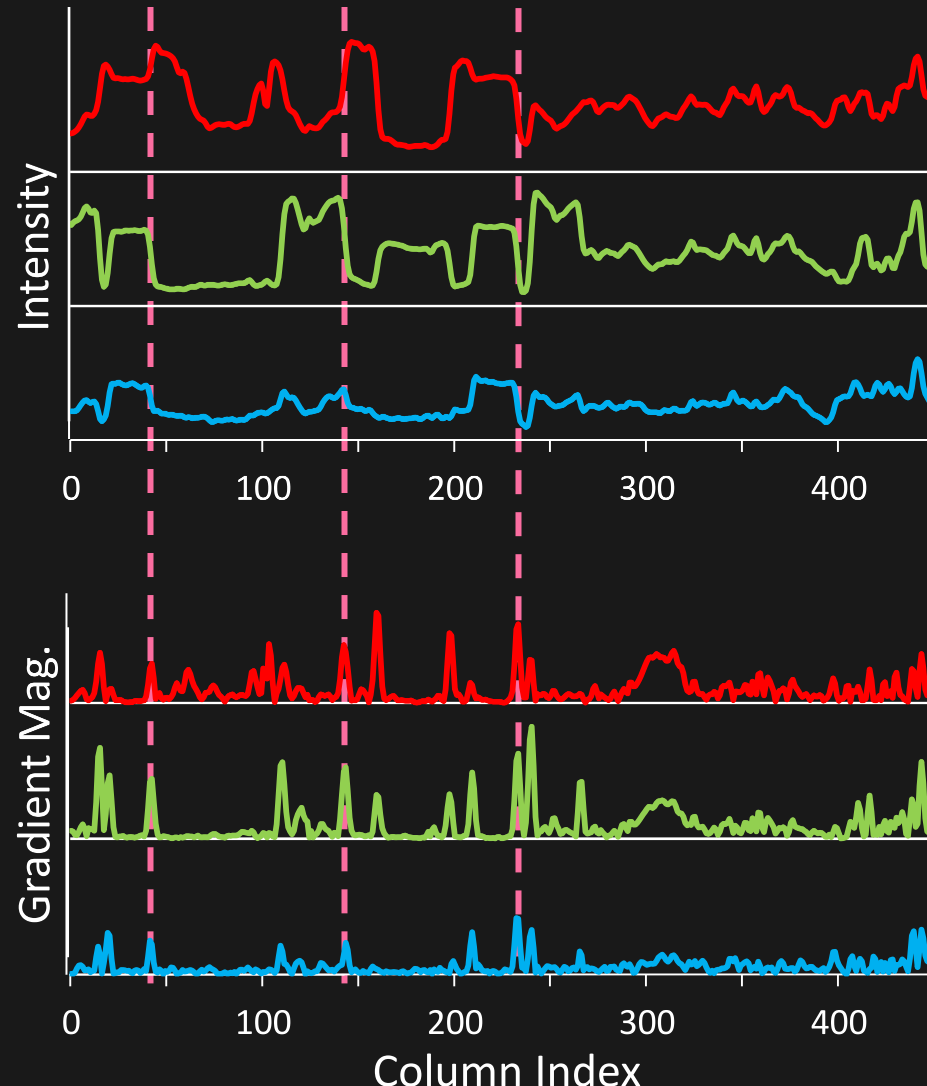


Intra-channel: 88% ± 1 pixel



Inter-channel: 39% ± 1 pixel

Edge alignment across color channels



Improving cross-channel correspondence

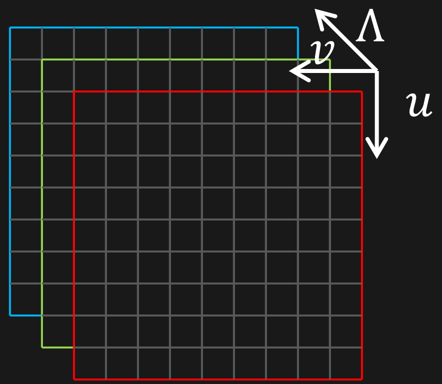
- Pixel intensities differ in each color channel
  - Traditional methods (SAD, SSD, cross-correlation, census) fail

- Edge locations correspond, but gradient magnitudes differ

- Solution: Use normalized gradient magnitudes to find correspondence

Correspondence Via Normalized Gradients

- We employ a window-based cost metric to compute correspondence likelihood at each disparity d



$$I_{\{p,d\}}(u, v, \Lambda)$$

- Compute gradients in u and v directions for each patch

$$G_{u,\{p,d\}}(u, v, \Lambda),\ \ G_{v,\{p,d\}}(u, v, \Lambda)$$

Correspondence Via Normalized Gradients

- Gradients are normalized in each channel independently

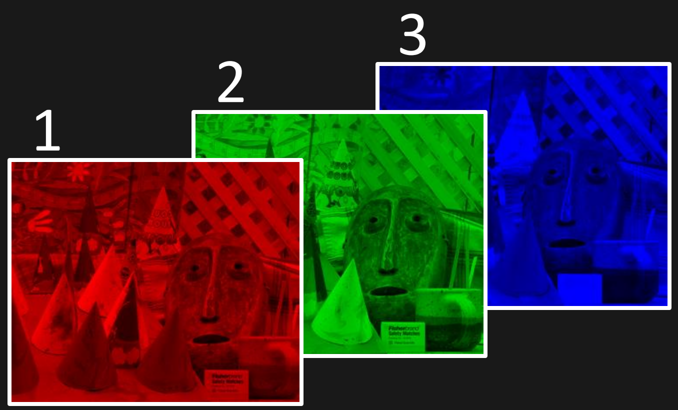
$$\hat{G}_{u,\{p,d\}}(u,v,\Lambda) = \frac{G_{u,\{p,d\}}(u,v,\Lambda)}{\left\| G_{u,\{p,d\}(\cdot,\cdot,\Lambda)} \right\|}$$

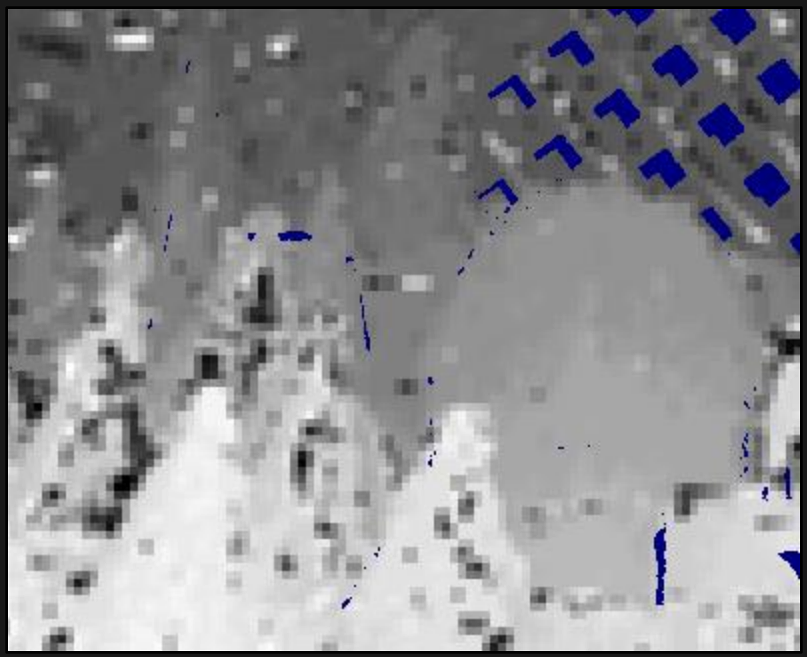
- u and v gradients are concatenated to give $\hat{G}_{\{p,d\}}(u,v,\Lambda)$

Edges must be aligned across the M color channels, giving our cost $C(p,d)$:

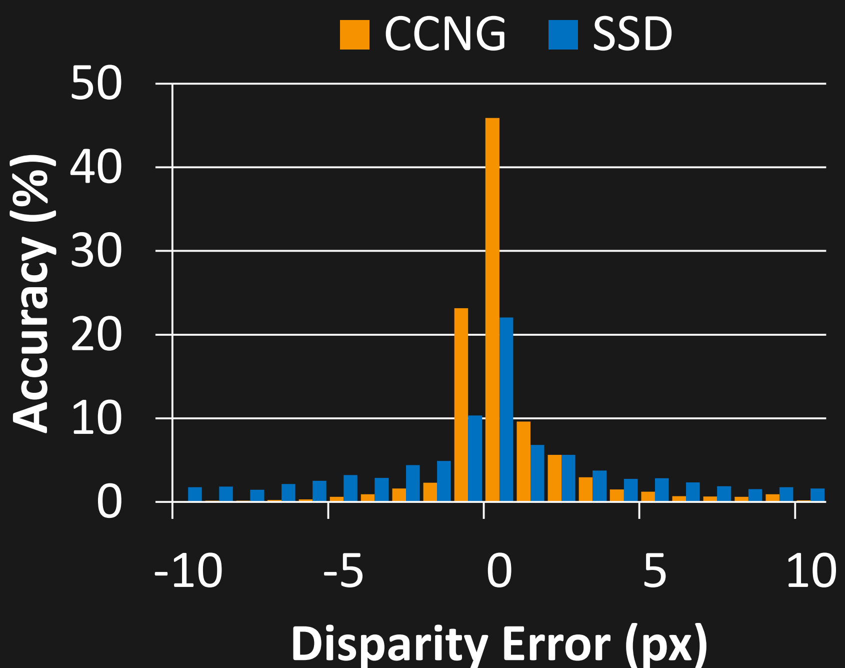
$$C(p,d) = -\sqrt[M]{\sum_{u,v}\prod_{\Lambda=1}^{M}\hat{G}_{\{p,d\}}(u,v,\Lambda)}$$

Cross-channel Normalized Gradients (CCNG)





CCNG Inter-channel



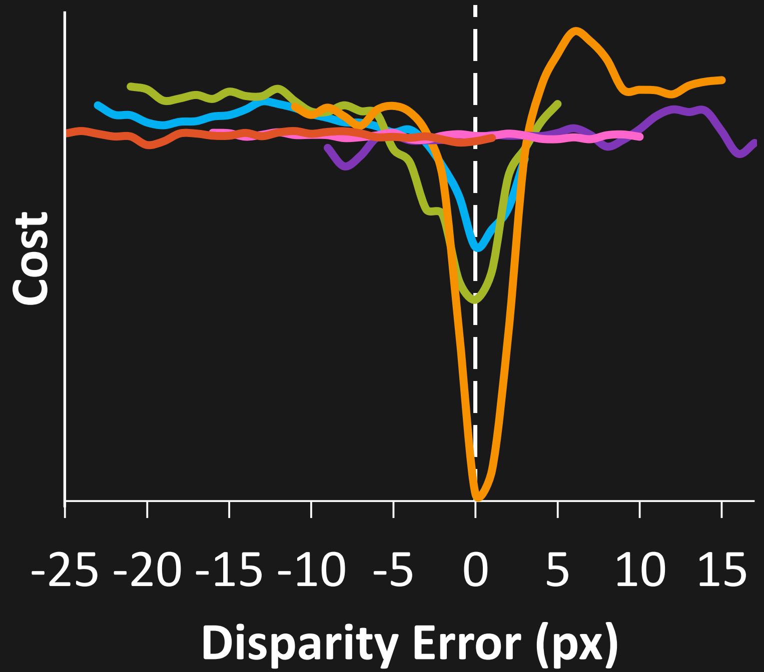
<u>Cross-channel accuracy</u>

SSD: 39% ±1 pixel

CCNG: 79% ±1 pixel

Confidence in Disparity Assignment

- CCNG shows a strong preference for the correct disparity in textured regions



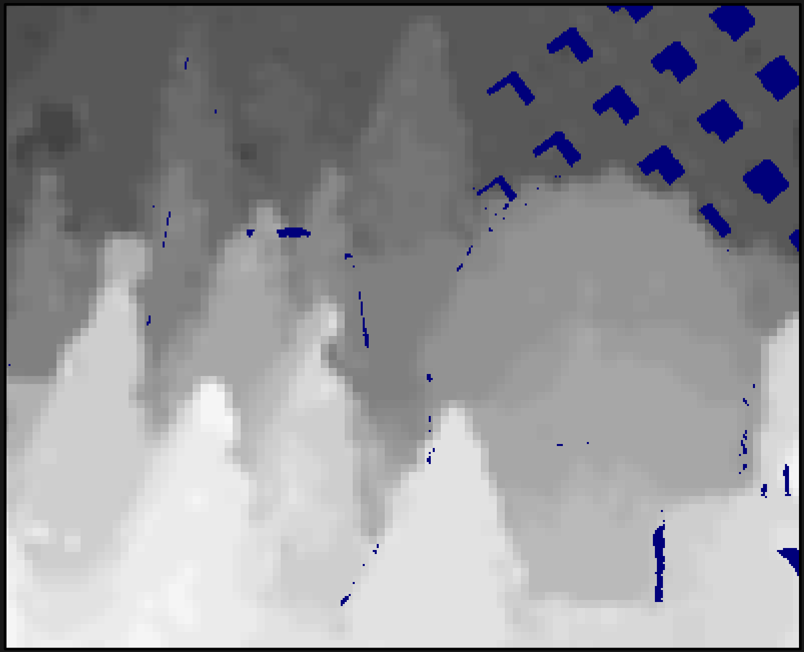
Correspondence in Textureless Regions

- CCNG cost performs well in textured regions
- Textureless regions are ambiguous, require priors to solve
  - Use larger patch sizes in smooth regions
  - Impose a smoothing penalty when computing disparities
- We use bilateral graph cuts to find a disparity map D:

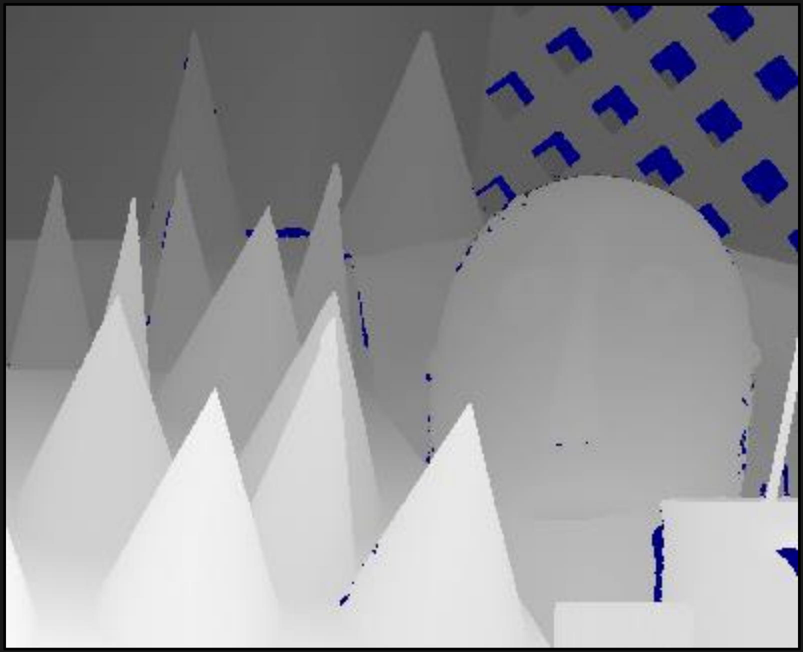
$$D(p) = \arg\min_{d} C(p,d) + \mu S(p,d)$$

Full CCNG Disparity Estimate

- 88% Accurate
  - The same accuracy as SSD *within* color channels
  - Accuracy improves with more channels



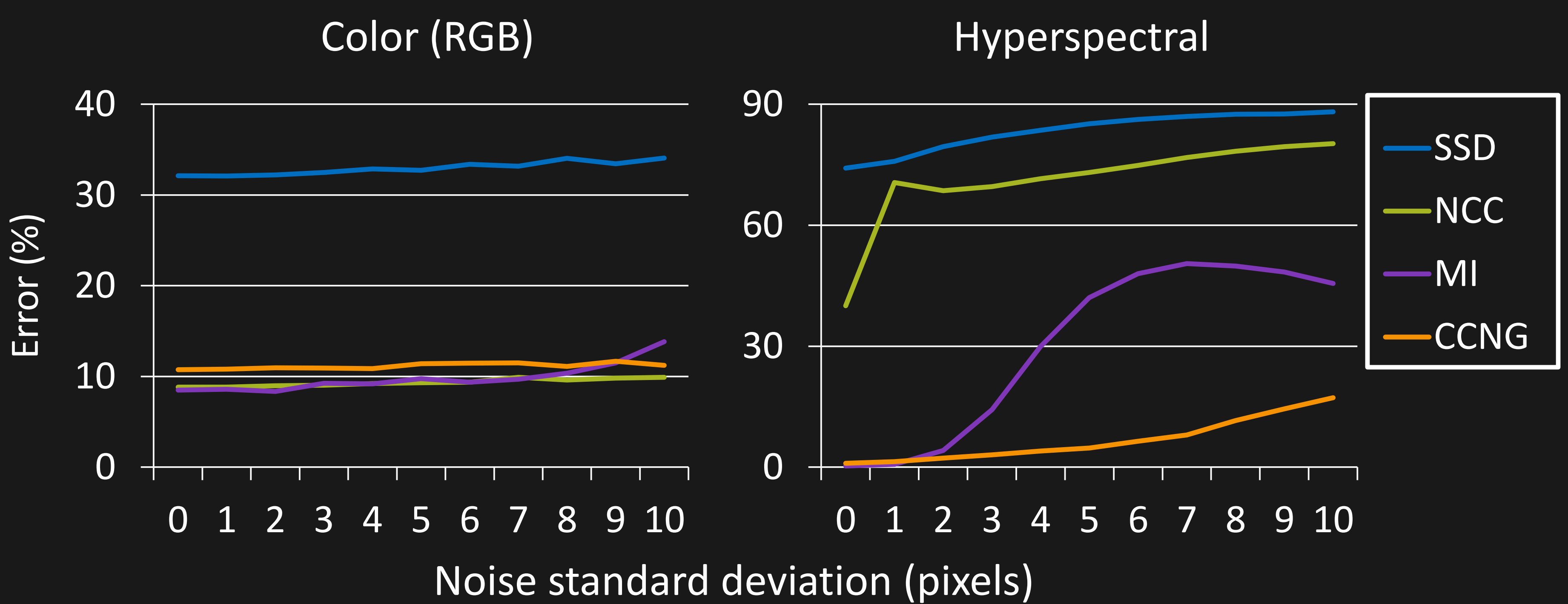
CCNG cross-channel disparity



Ground truth disparity

Robustness to noise

- AWGN is added to the three input channels, accuracy is the average of 10 trials per noise level
  - CCNG cost degrades gracefully with increasing noise



GAC Correspondence

- Assume camera array is calibrated such that internal and external camera parameters are known
- Sweep a virtual plane through the scene to hypothesize depths
- Given the hypothesized depths, the algorithm proceeds as described

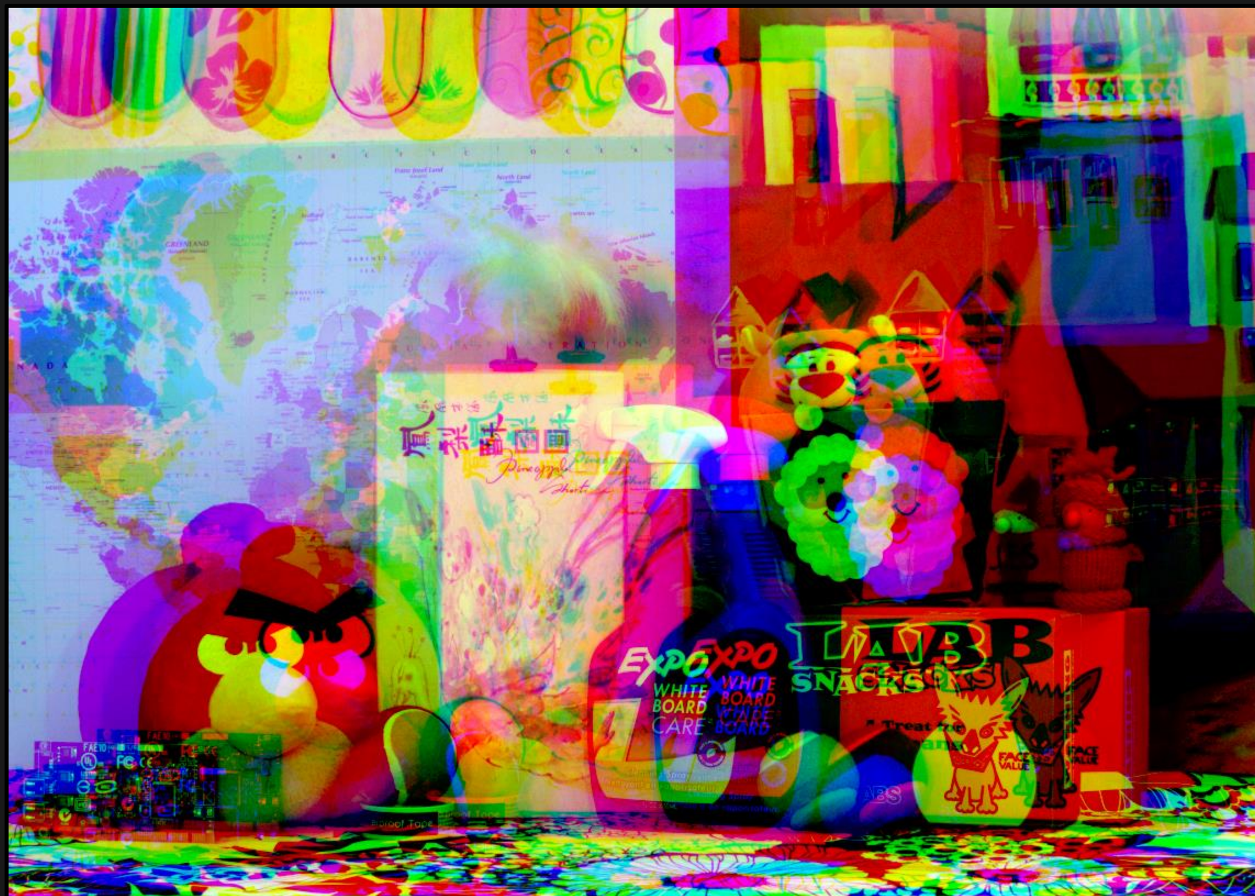
APPLICATION I: CONSUMER IMAGING

Input RGBY Images

- A 2 × 2 array of cameras capture 4 color channels
  - Red, Green, Blue, and Panchromatic (Y). All have IR cut filters

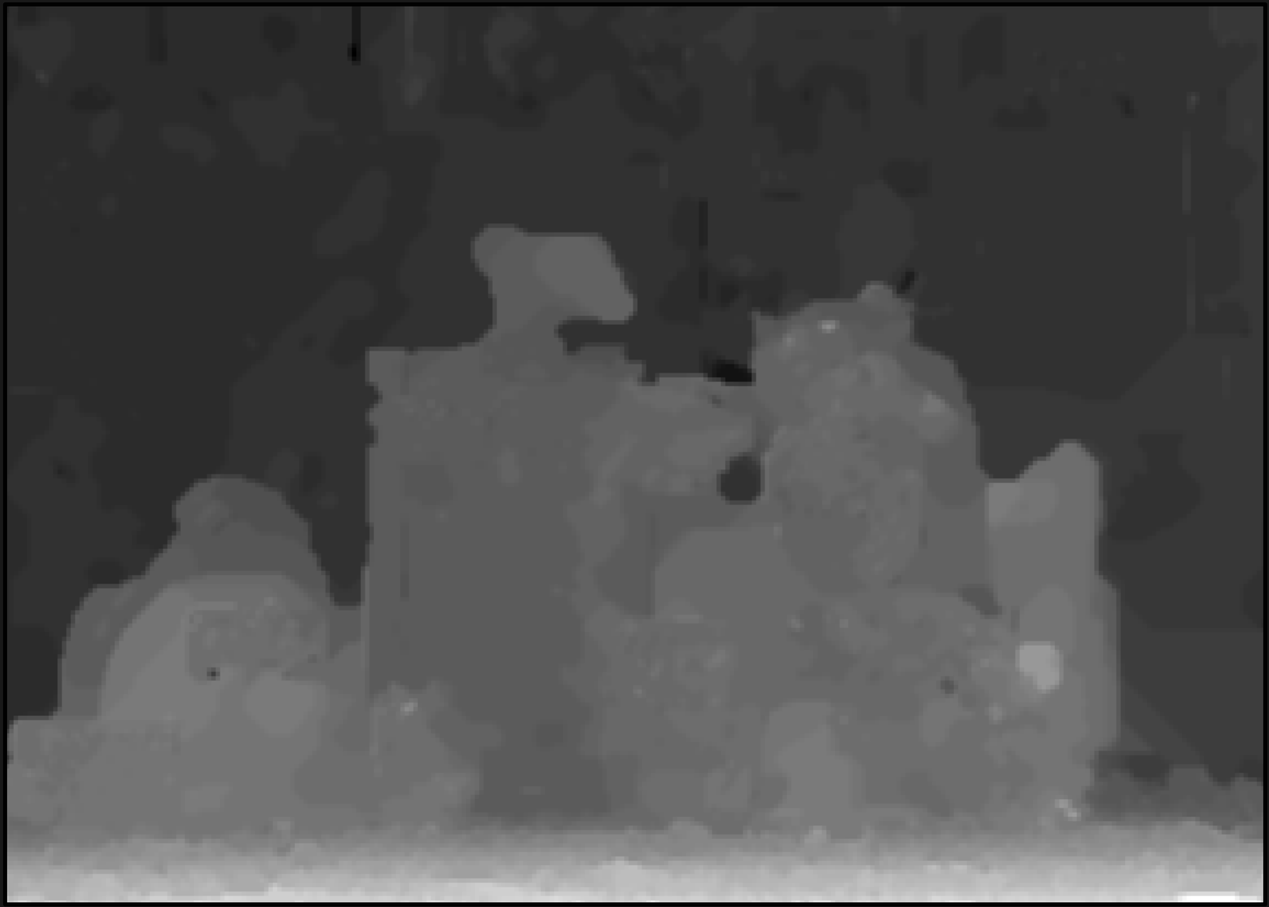
Direct Overlap Fails to Recover Color Images

- The cameras have a wide baseline (30mm)
- Direct image fusion is not possible



Computing Depth with CCNG

- Using our CCNG cost we recover a depth map
- The Y channel is used as reference



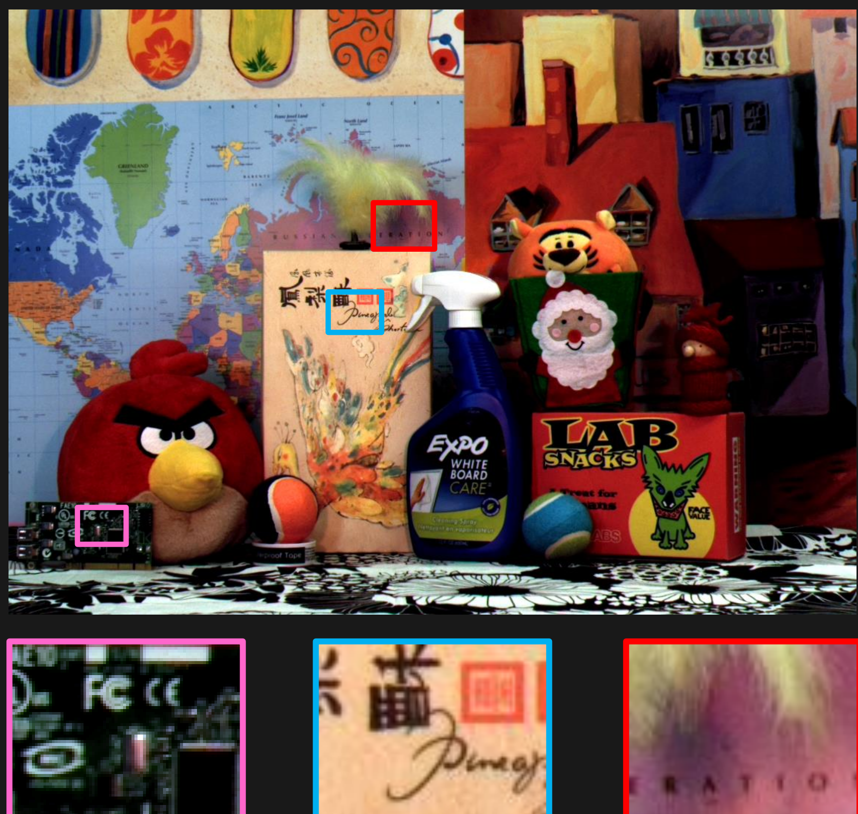
RGB Fusion

- R, G, B images are aligned using the depth map
- Chrominance from the RGB channels is added to the Y image

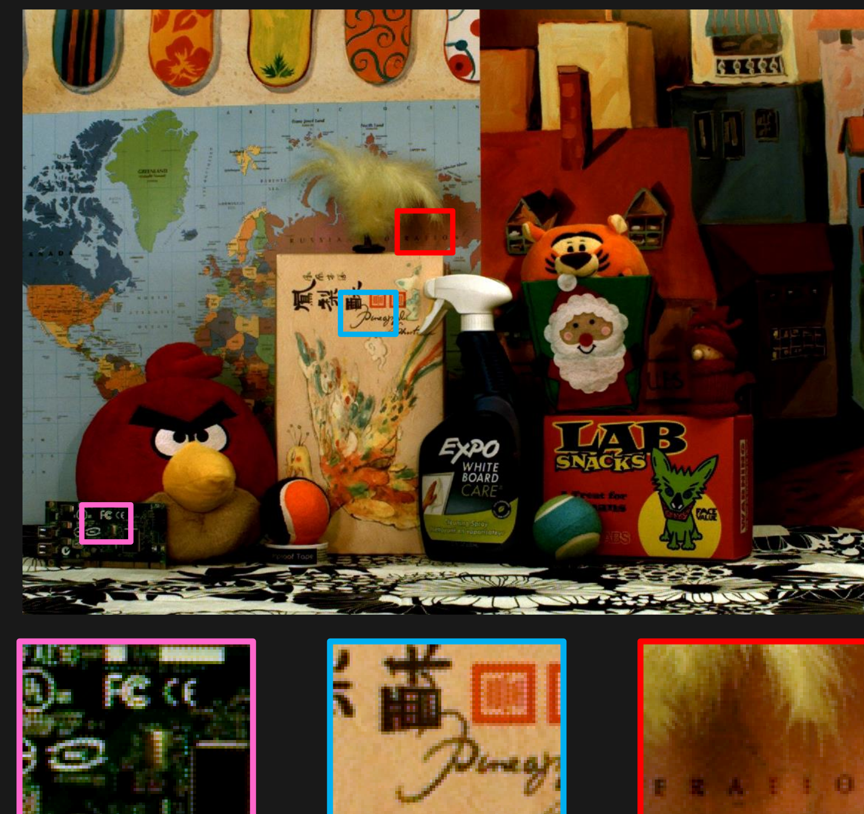
Color Image Comparison

- Quality of GAC image is comparable to a Bayer Sensor

GAC RGB Image



Bayer Color Image



Color Image Comparison

GAC RGB Image

Bayer RGB Image

GAC for Low-light Imaging

- Panchromatic camera in the GAC increases light throughput
  - Higher SNR in low light environments

GAC RGB Image

Noisy Bayer Image

Post-capture Refocusing

- GAC arrays provide finer angular resolution than single sensor cameras
- The depth map computed when using GACs enables post-capture refocusing
- Users may specify an aperture size and focal plane, affording greater artistic freedom

Post-capture Refocusing

Near Focus

In focus

Post-capture Refocusing

Mid Focus

Out of focus

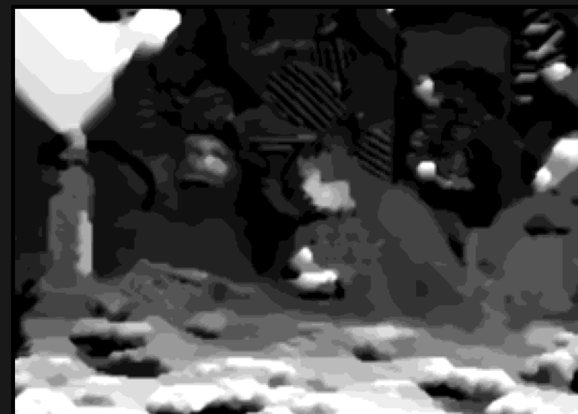
Post-capture Refocusing

Far Focus

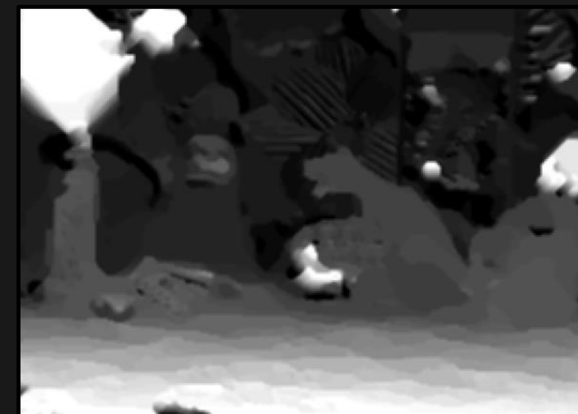
Out of focus

Depth Comparison

Recovered Scene

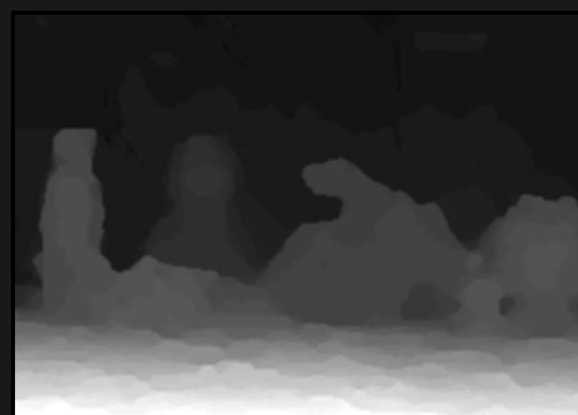


SAD

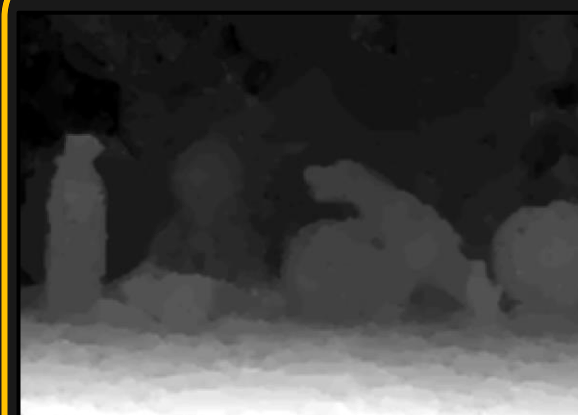


SSD

Mutual Information



Generalized NCC



CCNG (ours)

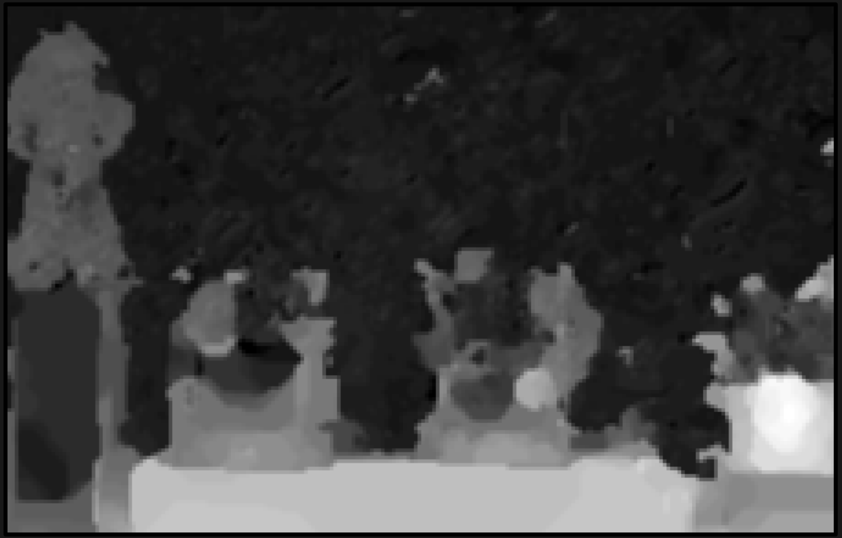
Additional GAC Color Images

Additional GAC Color Images

GAC—Limitations

- As with other stereo matching algorithms, specular surfaces are not faithfully recovered

Color image from Bayer sensor



Recovered depth map

GAC—Limitations

- As with other stereo matching algorithms, specular surfaces are not faithfully recovered

Color image from Bayer sensor

GAC color image

GAC—Limitations

- As with other stereo matching algorithms, specular surfaces

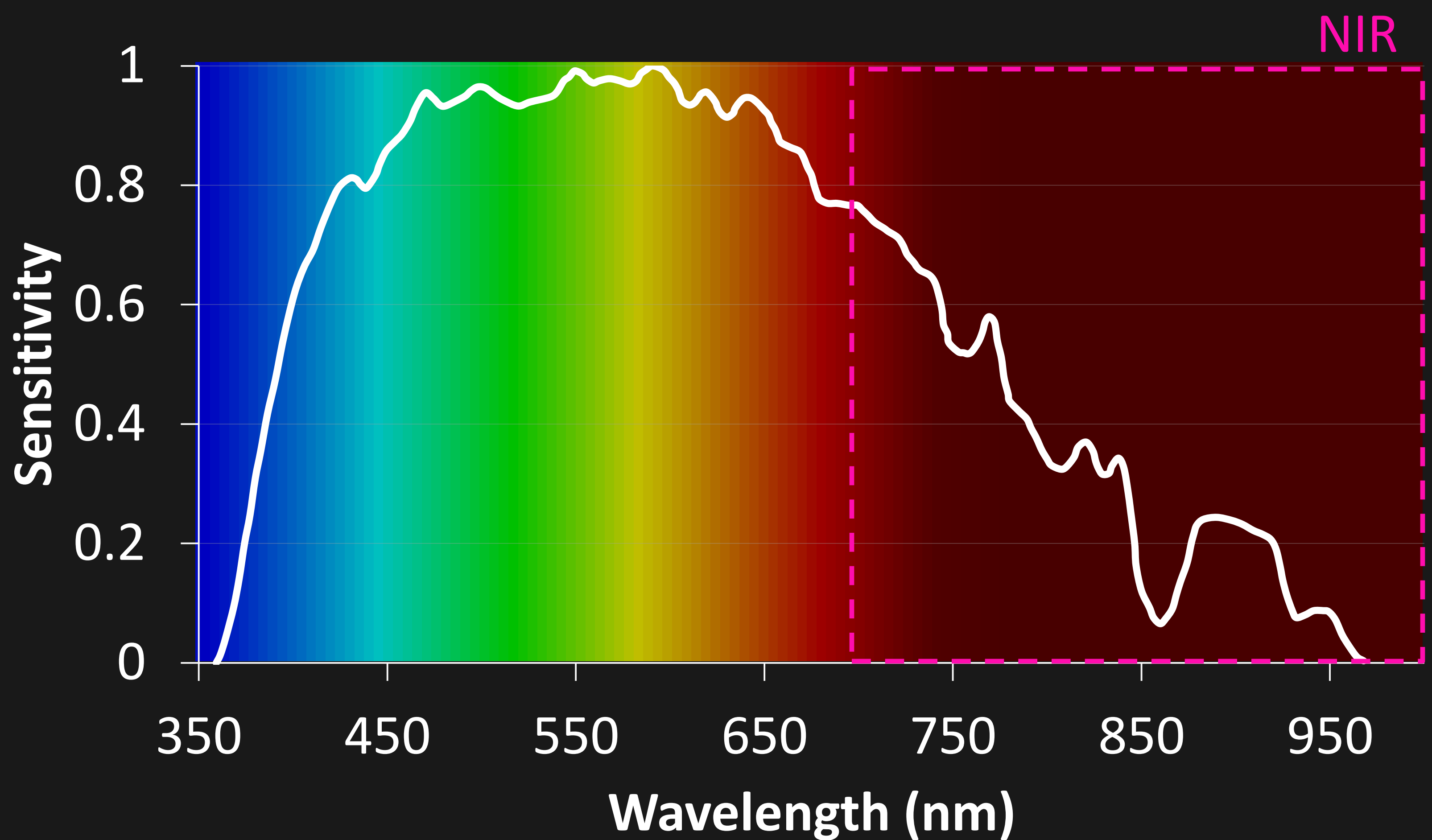
APPLICATION II: SKIN PERFUSION IMAGING

GAC: Flexible Application Driven Imaging

- Cameras and filters can be easily added or exchanged
- Appropriate tools can be designed for the task at hand
- Information in disparate modalities can be easily integrated
  - E.g. Near infrared, Narrowband, Polarized
- We demonstrate two applications for RGB+NIR imaging
- By simply adding an additional camera to our color imaging GAC, we obtain RGB+NIR images

Silicon Spectral Sensitivity

- Camera sensors are sensitive to near infrared light



Near Infrared Imaging Applications

- Dehazing (Feng et al.)

Input RGB image

+

NIR image

=

Dehazed image

- Shadow Detection (Rüfenacht et al.)

+

=

RICE

DSP

Skin Perfusion Imaging

- IR light penetrates skin to $\sim 100\mu$m
- Bypasses surface blemishes in the face (Süsstrunk et al.)
  - Using a co-axial camera setup
- Improves visibility of subsurface veins (Paquit et al.)
- Same reconstruction as before, but substitute high frequencies in NIR for high frequencies in luminance

$$Y_{\text{fused}} = Y_{\text{low freq.}} + \left((1-\alpha)Y_{\text{high freq.}} + \alpha NIR_{\text{high freq.}}\right)$$

“Natural” Image Retouching

- NIR images reduce the appearance of facial blemishes
  - Wrinkles, freckles, light facial hair, etc.

α=0.75

Color Image

NIR Image

RGB+NIR

Enhanced Vein Viewing

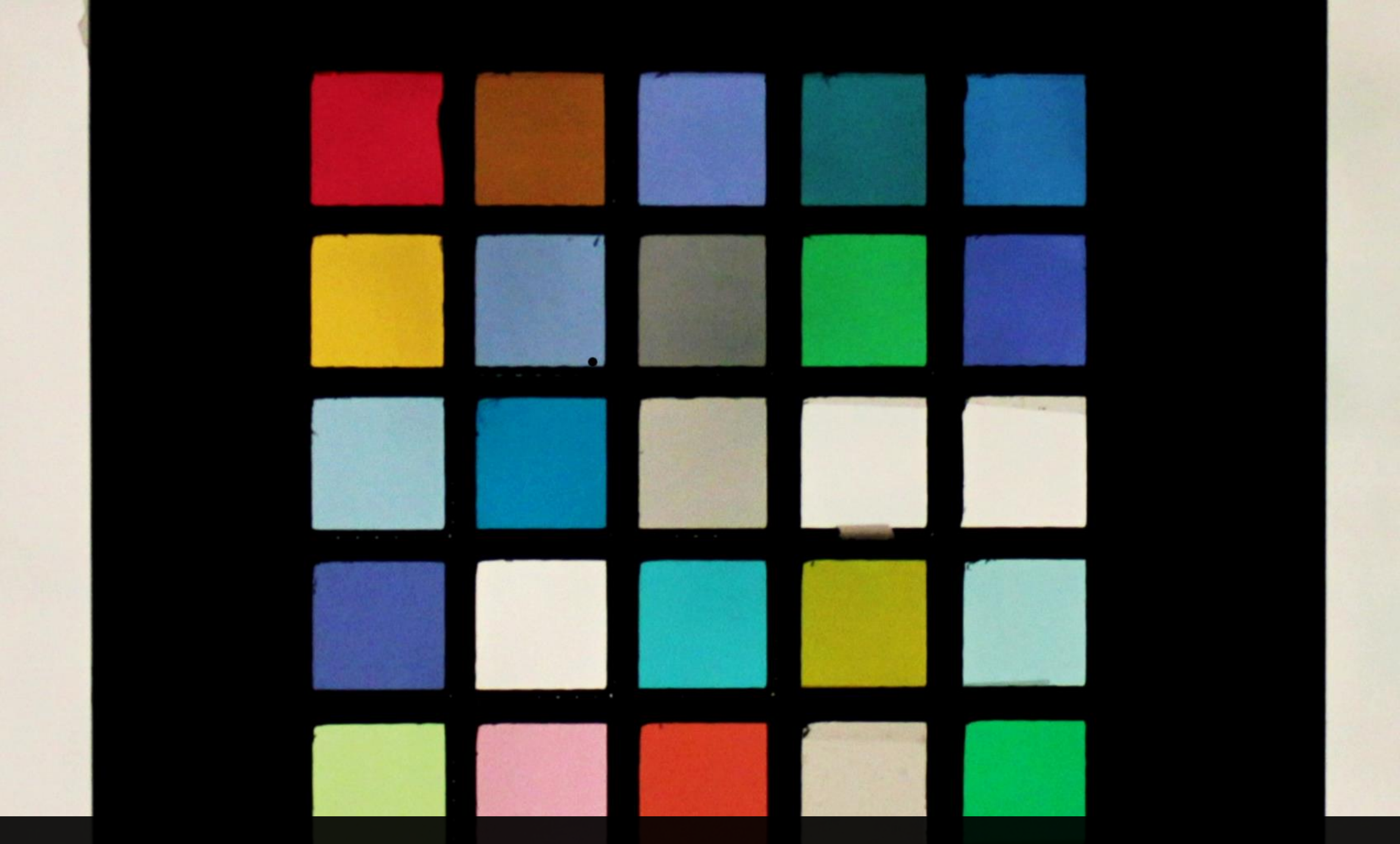
- Veins are prominent in NIR, helpful in medical environments

Color Image

NIR Image

RGB + NIR

α=1



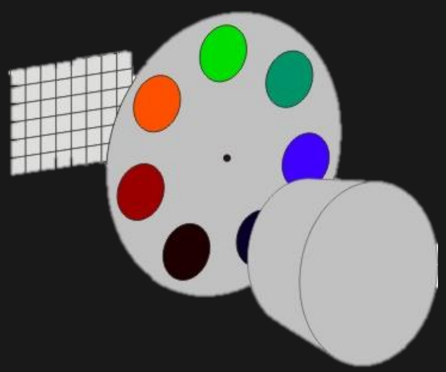
APPLICATION III: HYPERSPECTRAL IMAGING FOR DYNAMIC SCENES

Hyperspectral Image Acquisition ($$$)

- Serial image acquisition with different bandpass filters

External filters

Filter wheels [Brauers et al.]

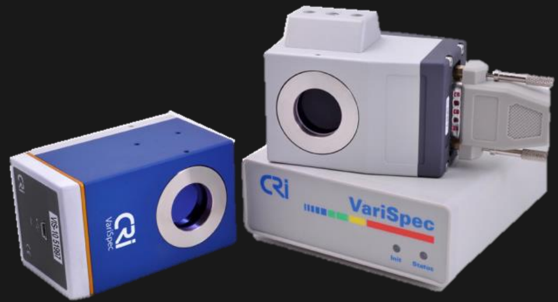


Remote sensing

Earth Observing-1

Tunable filters

Liquid crystal tunable filter [Harris and Wallace]

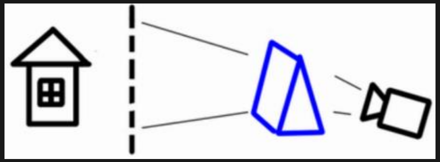


Snapshot Hyperspectral ($$$)

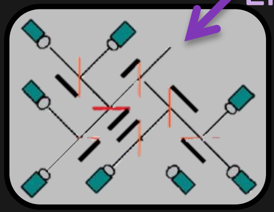
- Simultaneous image capture—low SNR and low resolution

Prism and Beam splitting

Dispersing prism
[Du et al.]

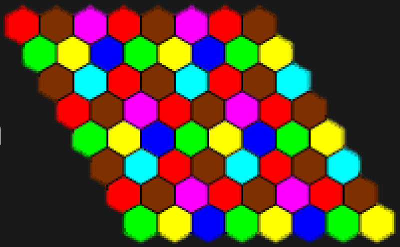


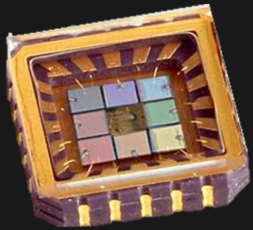
Optical Splitting
Trees [McGuire et al.]



Filter Array

Multi-spectral
filter array [Shrestha
et al., Miao et al.]





Monolithic sensor
[PIXELTEQ, IMEC]

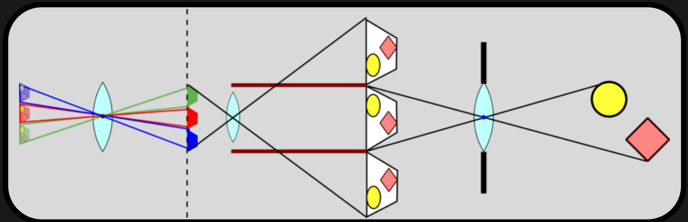
Rigid Camera Arrays

Wide band filters
[Frese and Gheta],

Planar scenes [Lau and Yang]

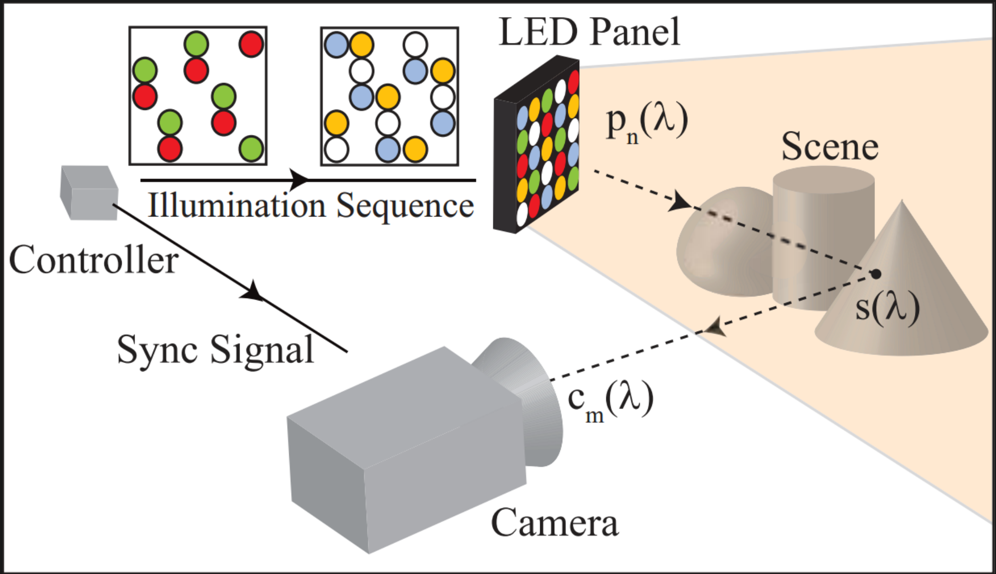
Coded aperture

Kaleidocam filtered aperture
[Manakov et al.]



Improving SNR

- Bandpass filters restrict light throughput in each channel
  - Resulting images are noisy
- Solution, multiplex light to improve SNR
- Park et al. use a multiplexed illumination scheme
  - Serial, static scenes only

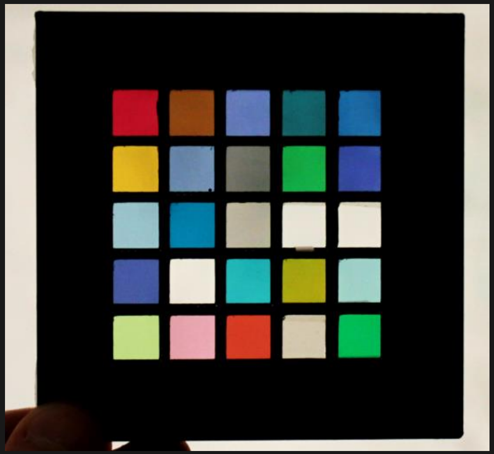


Multiplexed Image Capture

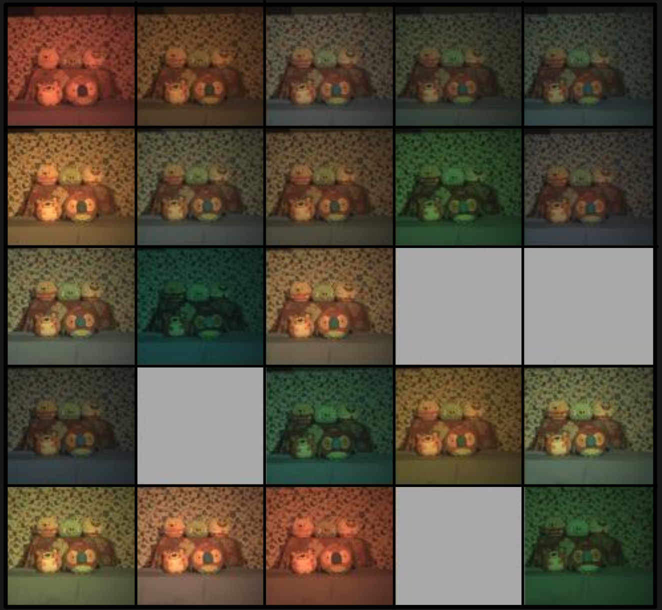
- Multiplexing increases light throughput and gives higher SNR

- Use a GAC with broadband filters with a single light source
  - 5×5 ProFUSION color camera array (21 of 25 cameras are used)
  - Commodity Roscolux filters ($\sim$ \$1 total cost)
  - 63 spectral measurements per scene point

+



=

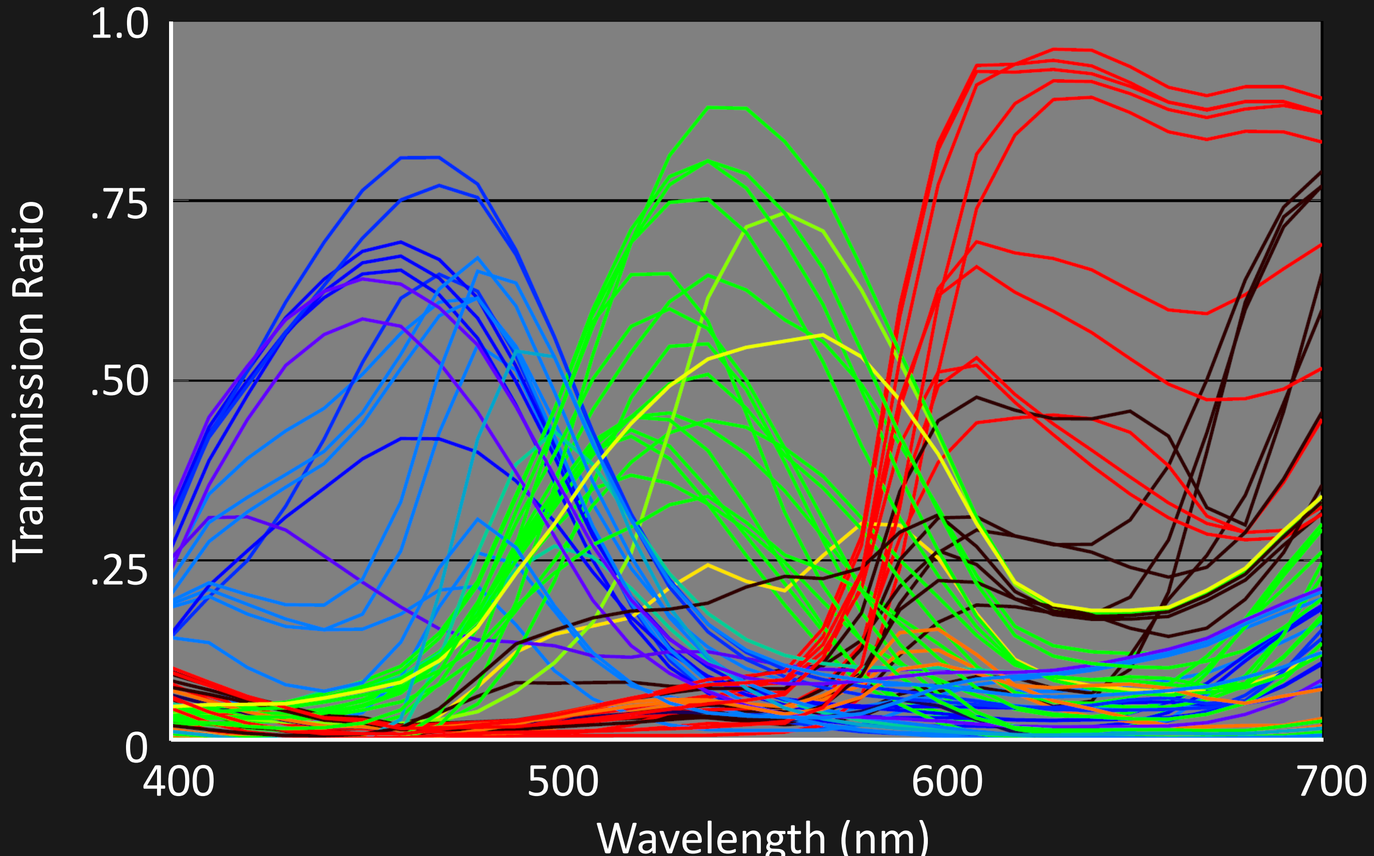


RICE

DSP

Commodity Broadband Filters

- Filters chosen using a greedy algorithm to minimize the condition number of the mixing matrix



Mulitplexing

- Images are aligned using our CCNG algorithm to compute a depth map
- Spectral profiles are recovered *without* needing to know the mixing matrix

$$I_m = \sum_{i=1}^{S} F_m(\lambda_i) R(\lambda_i), \qquad m = 1, \ldots, 63,$$

$$I = \boldsymbol{F} R$$

- I (63×1) — Image measurements for a given scene point
- $\boldsymbol{F}$ $(63 \times S)$ — Effective filter (broadband * Bayer response)
- R $(S \times 1)$ — Effective reflectance (Illumination * Reflectance)

Demultiplexing

- Given a dictionary of N known true/multiplexed spectral measurements, we demultiplex each scene point:

  $\boldsymbol{X}$ $(63 \times \mathrm{N})$ — Known multiplexed measurements

  $\boldsymbol{T}$ $(S \times \mathrm{N})$ — Known spectral profiles*

- Using $\boldsymbol{X}$ as a dictionary we find the K-sparse weights (ω) which recover the profile of I:

$$\arg\min_{\omega} \|I - \boldsymbol{X}\omega\|, \text{such that } \|\omega\|_0 < \mathrm{K}$$

- The same weights are used to recover $\hat{R}$

$$\hat{R} = \boldsymbol{T}\omega$$

* $\boldsymbol{T}$ is recorded using a Headwall hyperspectral imager

Color Checker Profiles

- We validate our method on a standard 24 square Color Checker
  - Dictionary learned from 140 square Digital SG Color Checker
  - Average reconstruction SNR: 23.7dB

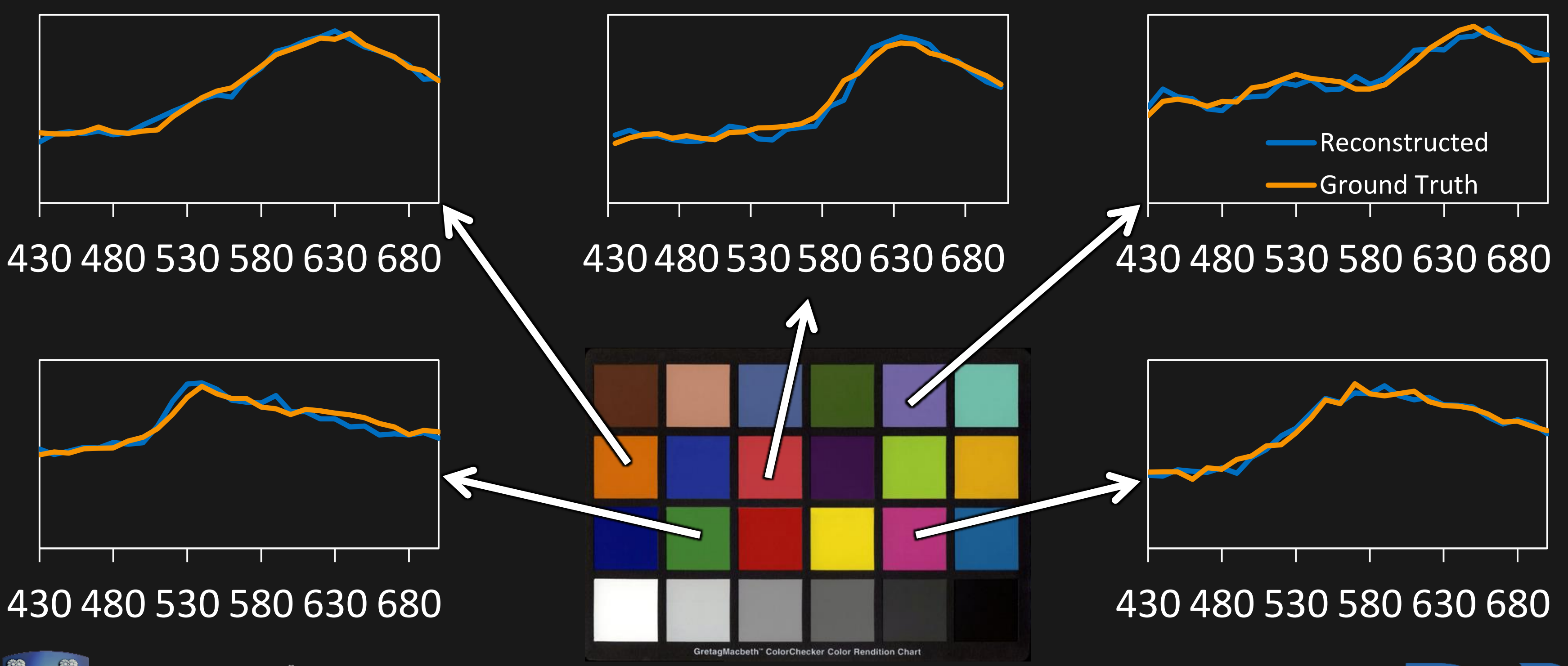
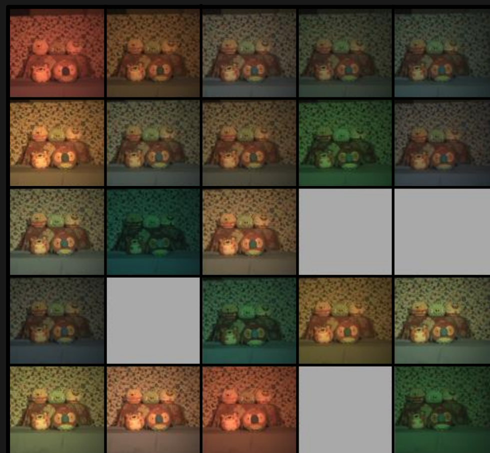
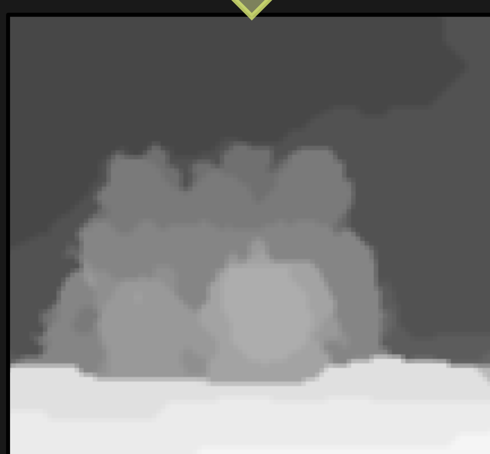


Image Alignment

Input Images



Direct overlap
(averaged)



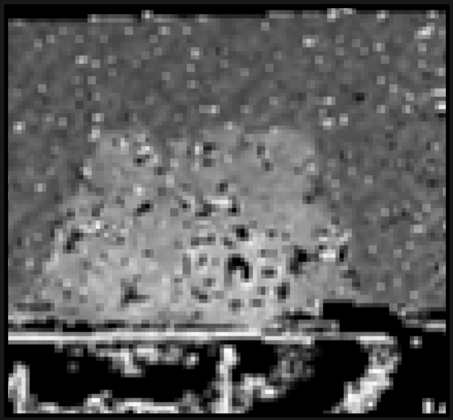
Aligned images
(averaged)

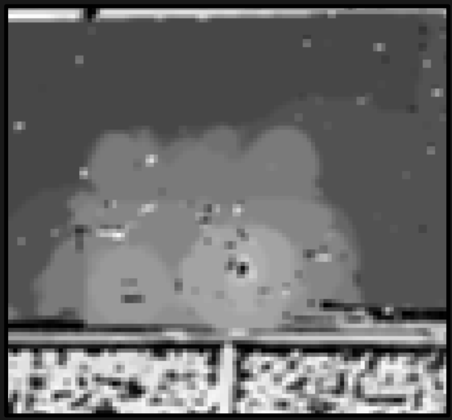
Depth Comparison


Recovered Scene


SAD


SSD

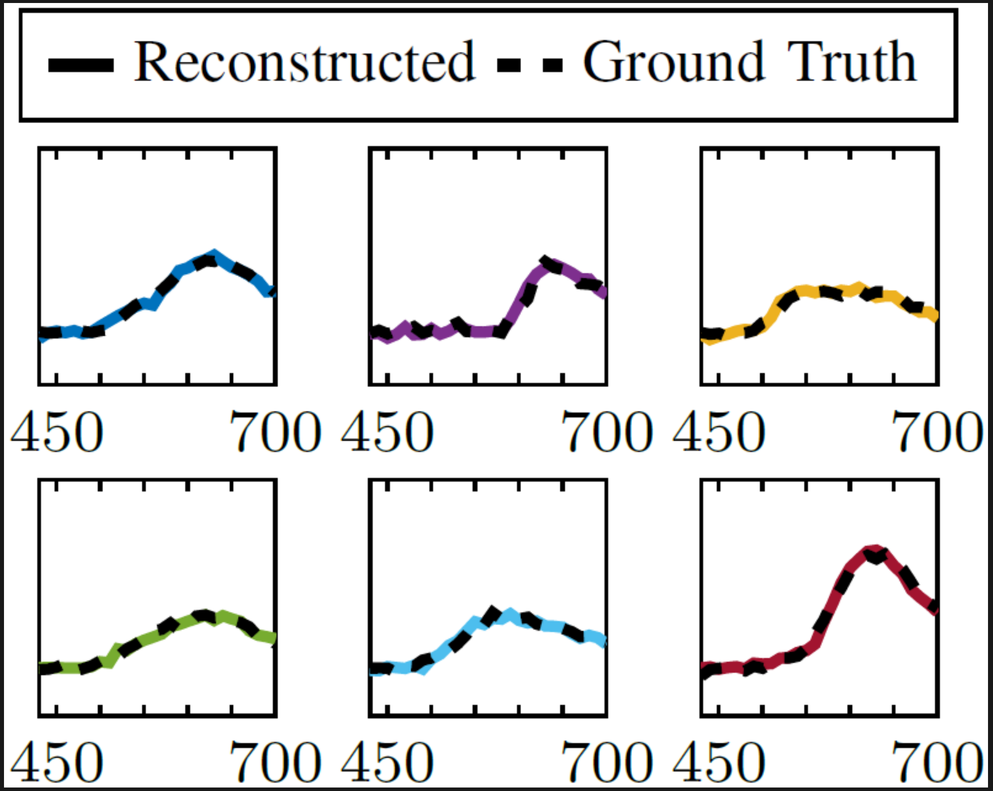

Mutual Information


Generalized NCC


CCNG (ours)

Static Scene Profile

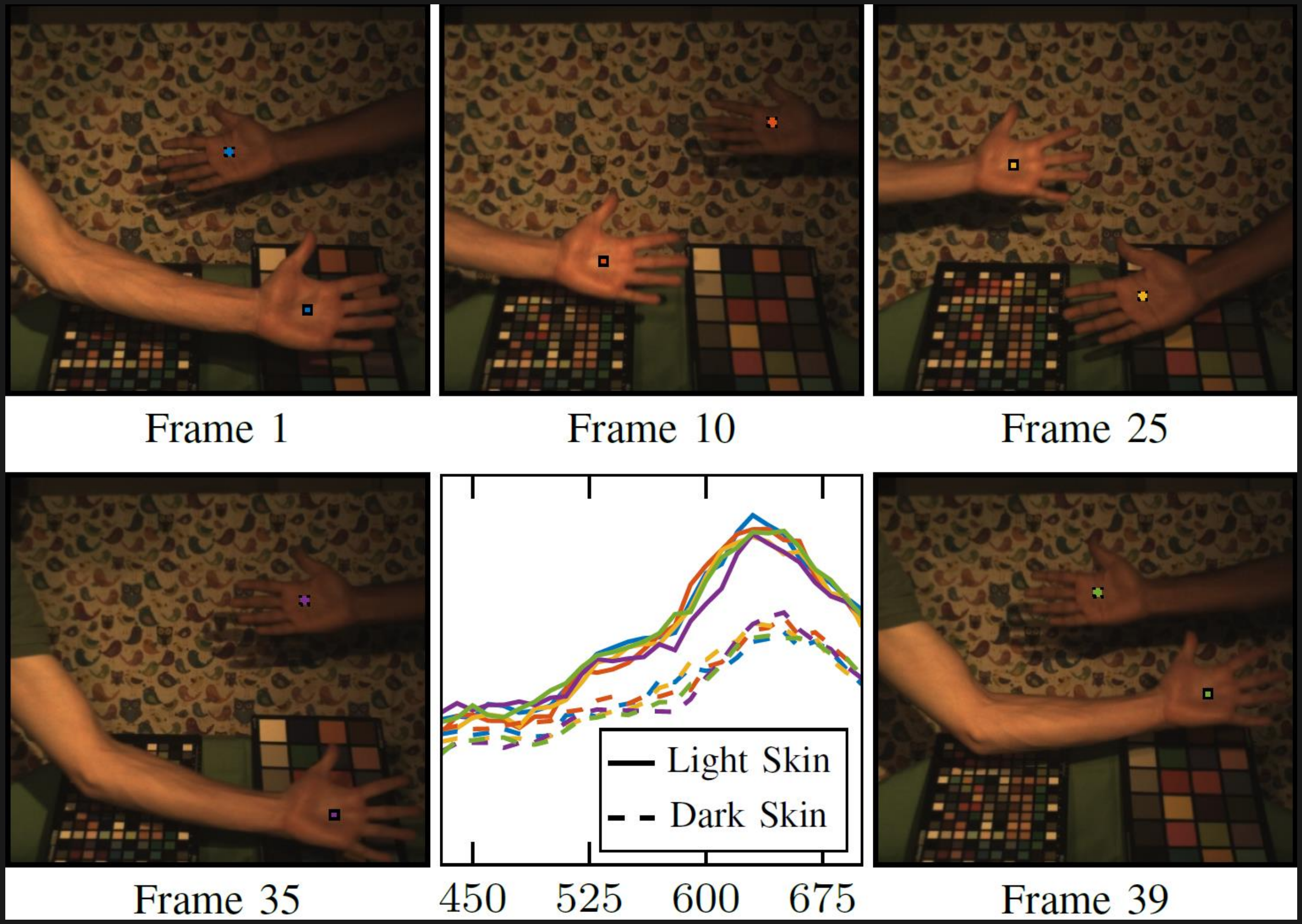
- Average spectral recovery SNR: 26.7dB



Hyperspectral Video

Hyperspectral Video

- Recovered spectral profiles of the hands (marked manually)



Average SNR: 27.8dB
(ground truth taken with arms resting on the table)

Limitations

- Currently need hyperspectral calibration
- Illumination dependent calibration
- Can remove HS calibration by assuming a known profile for the calibration target
  - Fold illumination into the unknown mixing matrix $\boldsymbol{F}$
  - Recover true reflectance of the material

Conclusion

- Generalized Assorted Cameras are well-suited for a wide range of imaging tasks
- Flexible architecture allows for rapid prototyping
- Scalable platform permits any combination of cameras
  - Can increase performance by using additional cameras
- Our cross-channel stereo algorithm accurately estimates depth *without* the need for redundant channels