Unifying Image Fusion and Compression: A Variational Approach

Daniele Picone, Laurent Condat, Mauro Dalla Mura

GIPSA-Lab, Grenoble Institute of Technology, France

GDR CNRS MaDICS: Rencontre Annuelle

November 20th, 2018

Pansharpening

HIGH SPATIAL RESOLUTION

Panchromatic (PAN)

HIGH SPECTRAL DIVERSITY

Multispectral (MS)

SPATIAL RESOLUTION: Minimum spatial distance required to distinguish two objects on the scene

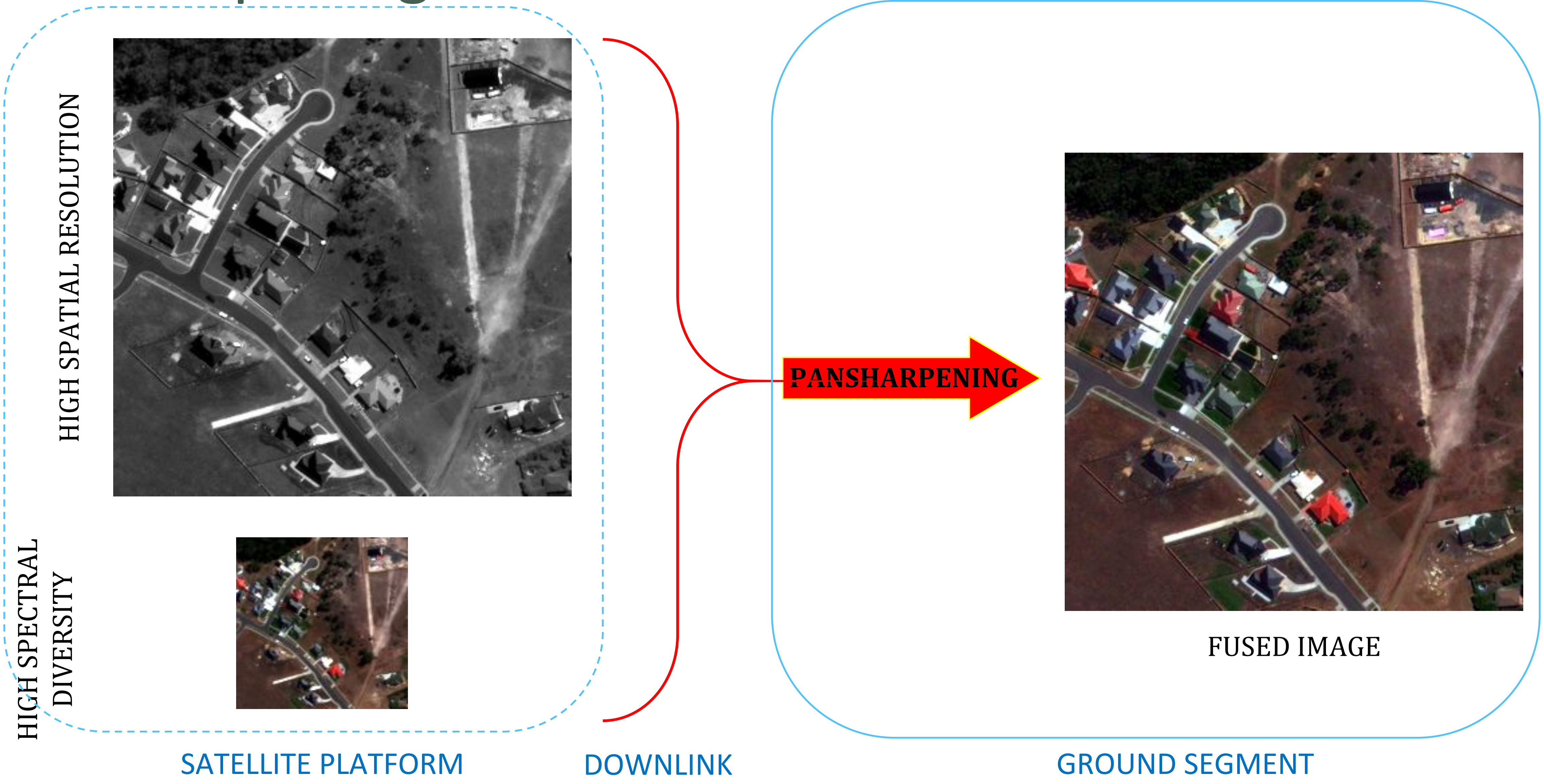
PANSHARPENING

SPECTRAL DIVERSITY: Minimum distance between two separable spectra

FUSED IMAGE

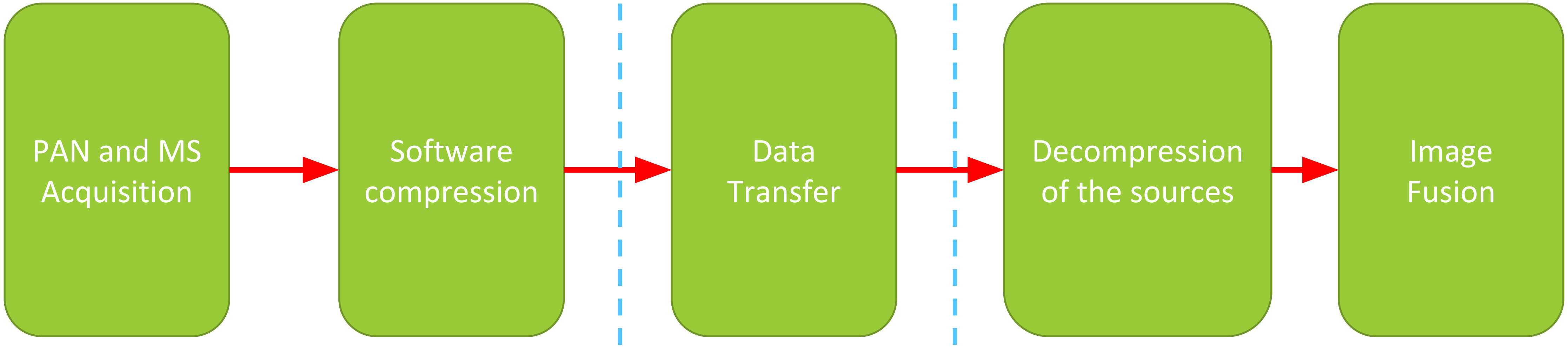
Definition: Sharpening (i/e: enhanching) a multispectral image with a panchromatic one [Vivone et al., 2015, Loncan et al., 2016]

Pansharpening

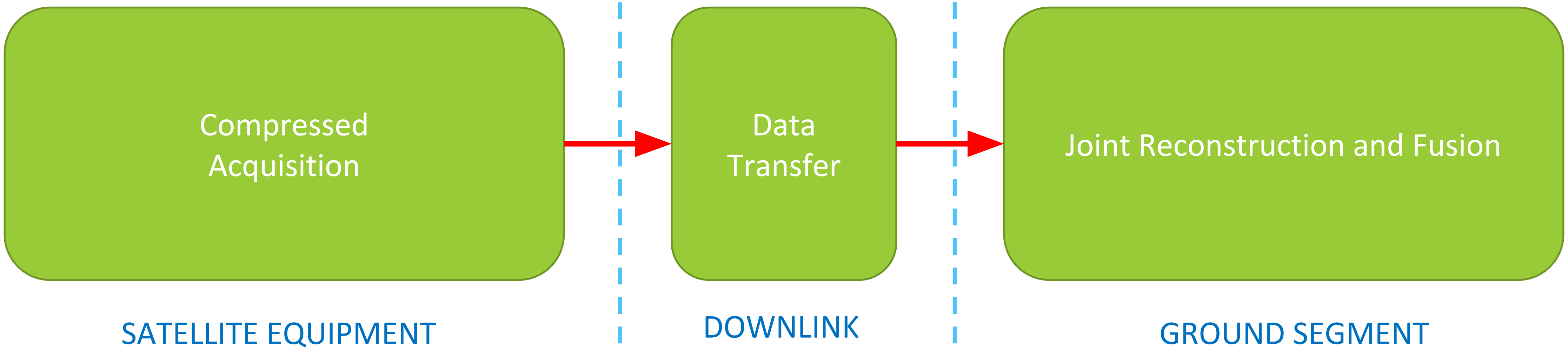


A joint model

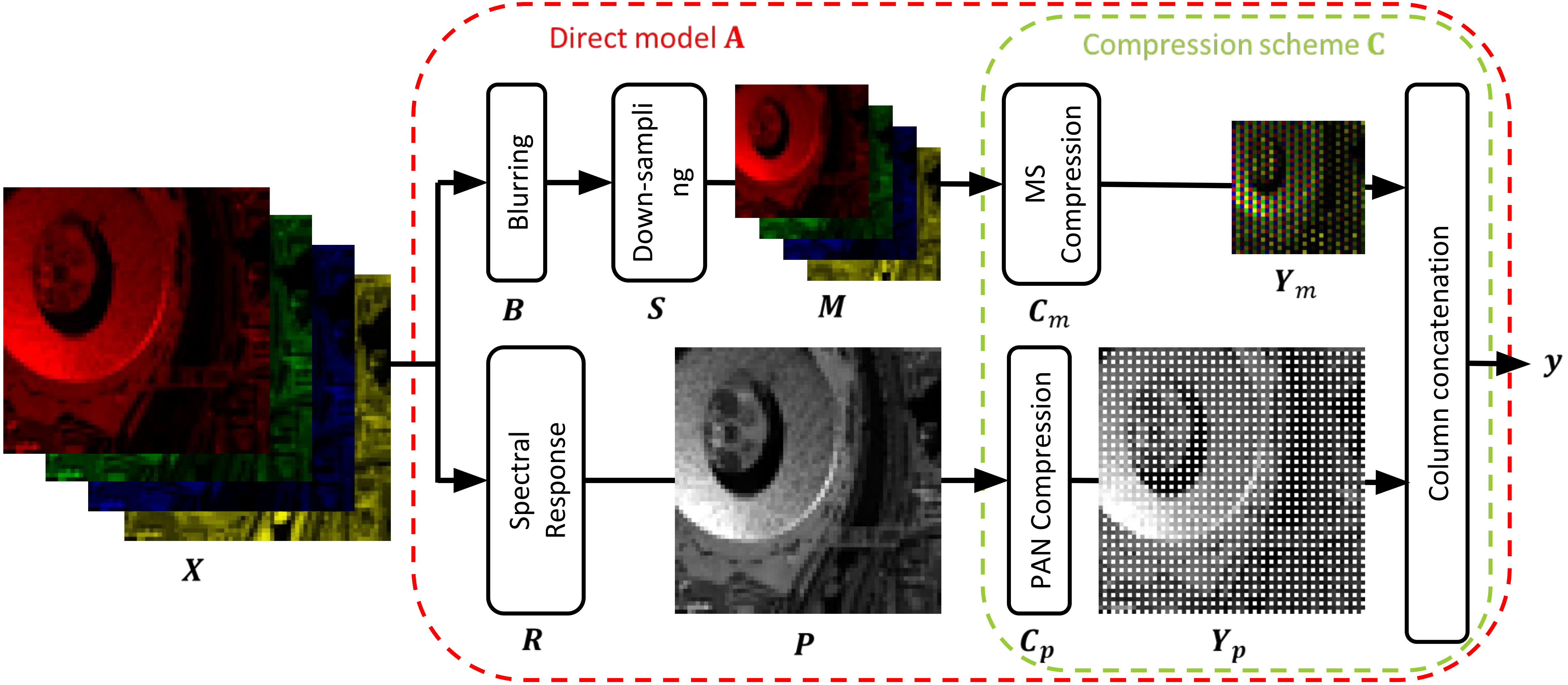
- Classical Approach



- Proposed Model



Reconstruction scheme: Direct Model

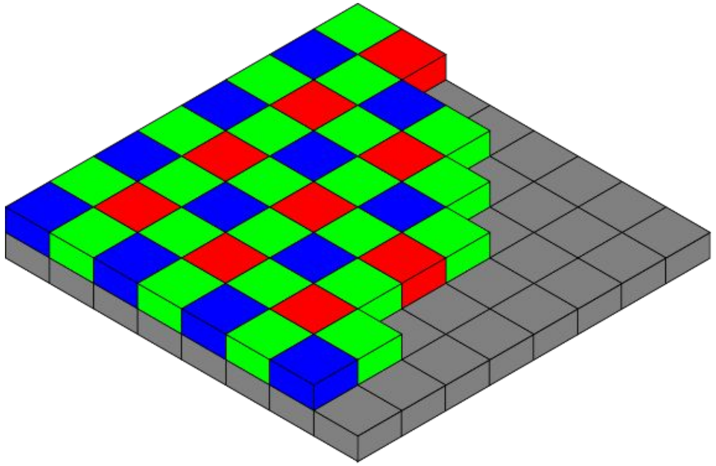


See you at the poster

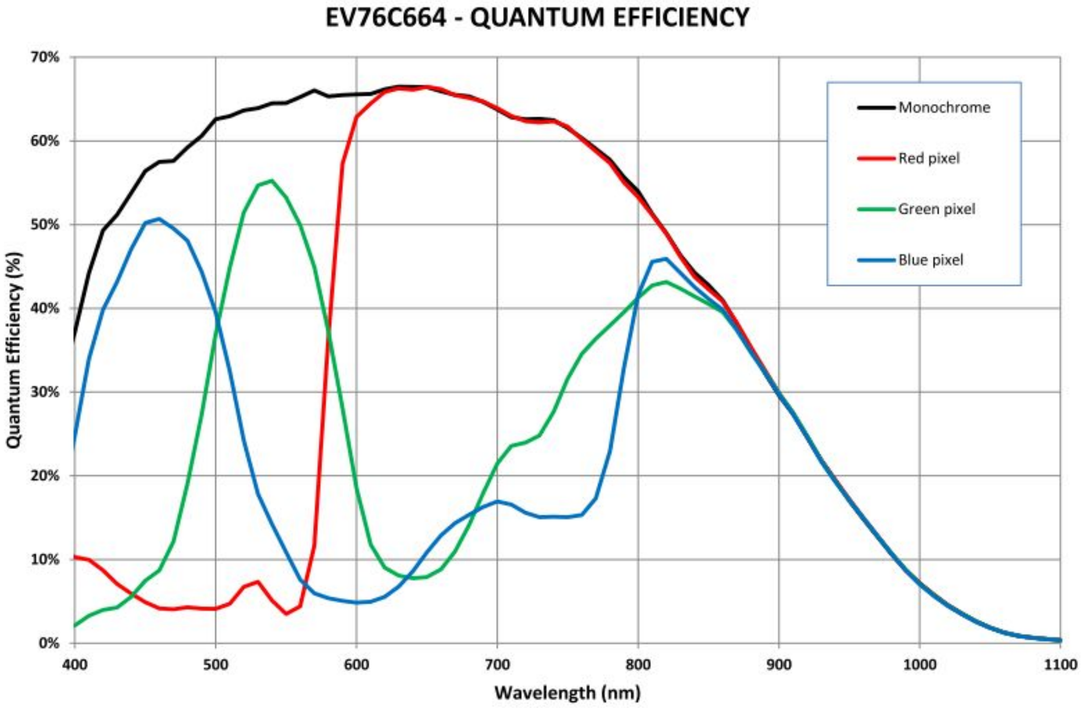
Summary

- Background
  - Pansharpening
  - Compressed Acquisitions
- Contribution
  - Joint model of compression and fusion
  - Employed Regularizers
- Experimental results
- Conclusion

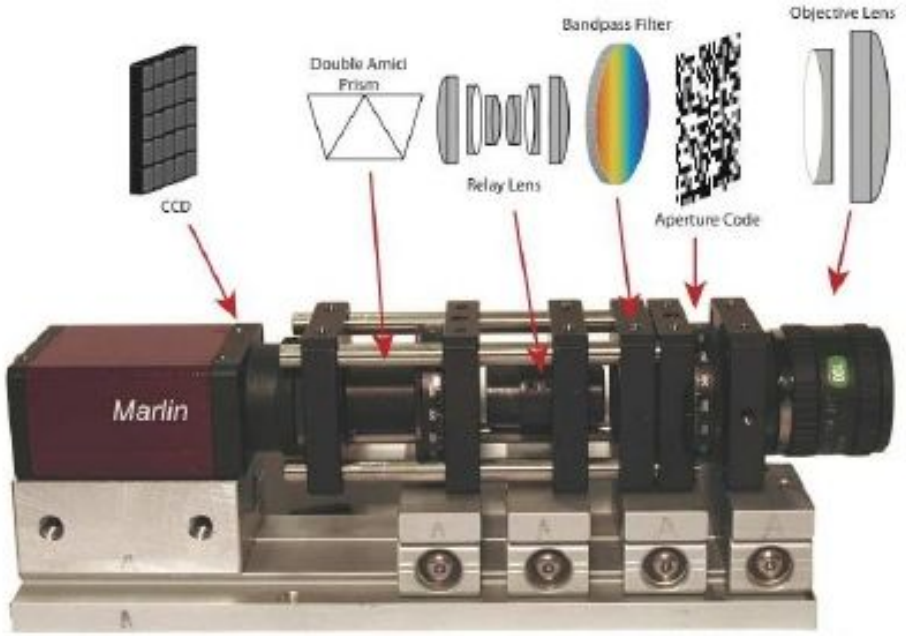
Compressed Acquisitions: Color Filter Arrays



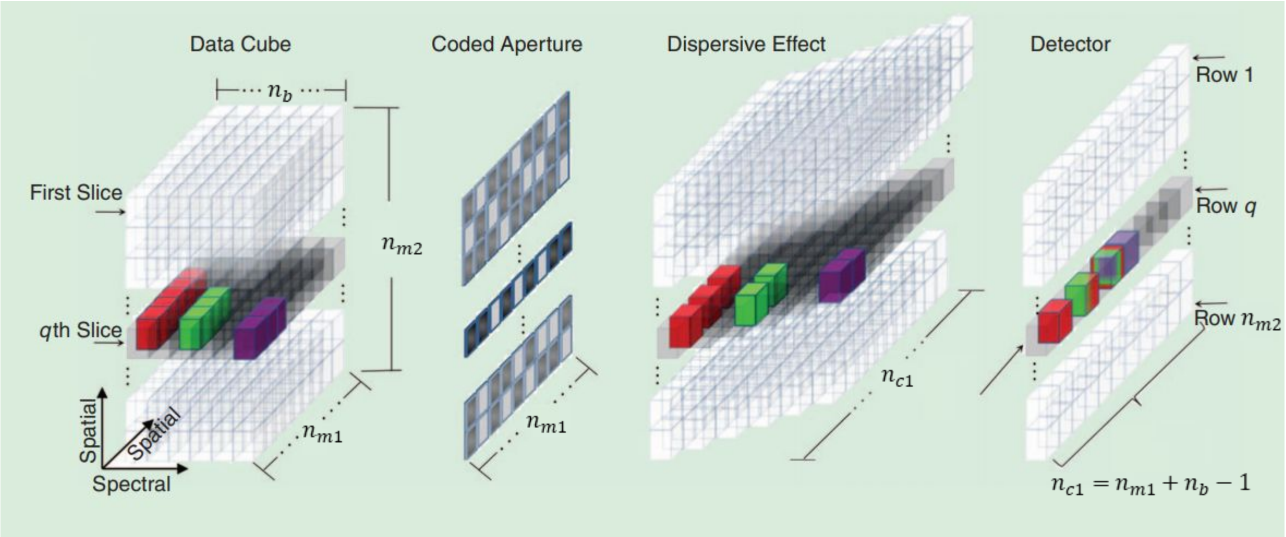
CFA (Bayer Pattern)



Compressed Acquisitions: CASSI



CASSI [Arce et al., 2014]



Compressed Acquisition: Model

- Given:
  - A PAN signal $\boldsymbol{p} \in \mathbb{R}^{n_p}$
  - A MS signal $\boldsymbol{m} \in \mathbb{R}^{n_m n_b}$
- Target: Generate an easy model for the optical compressed acquisition $\boldsymbol{y} \in \mathbb{R}^{n_c}$ of multimodal sources
  - Compression ratio: $\rho = n_c / (n_m n_b + n_p)$

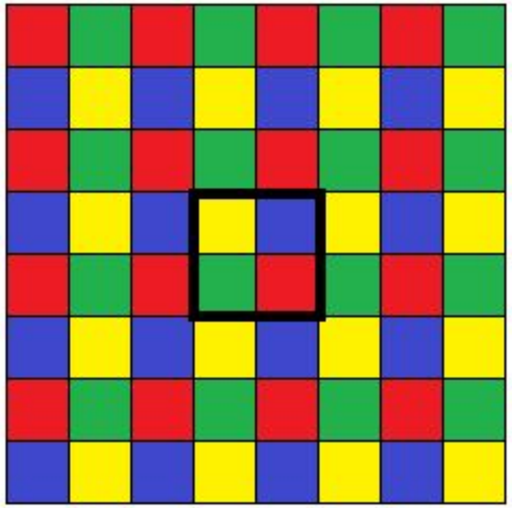
Column Concatenation

n_m=#pixels MS
n_b=#bands MS
n_p=#pixels PAN
n_c=#output samples

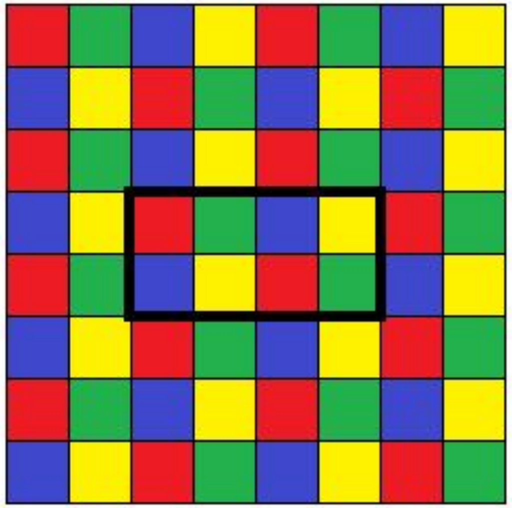
Property	Description	Mathematical model
Linearity	Optical devices are linear systems	
Separability	Each source is compressed indipendently	
Boolean Matrix	Each output sample is a sum of input samples	
Sub-sampling	Each output sample is equal to a single input pixel	

Compressed acquisitions: Binary masks

RGB/NIR CAMERAS

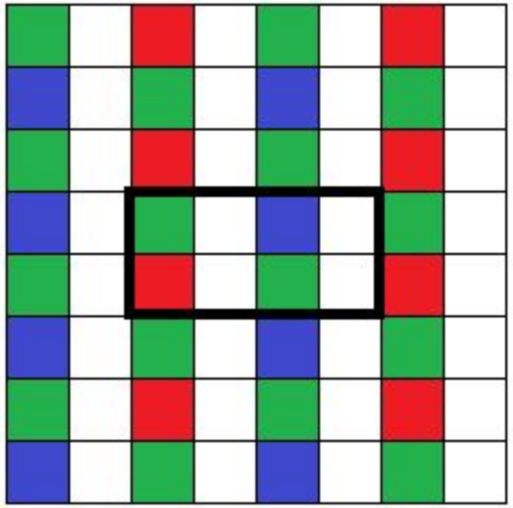


Uniform

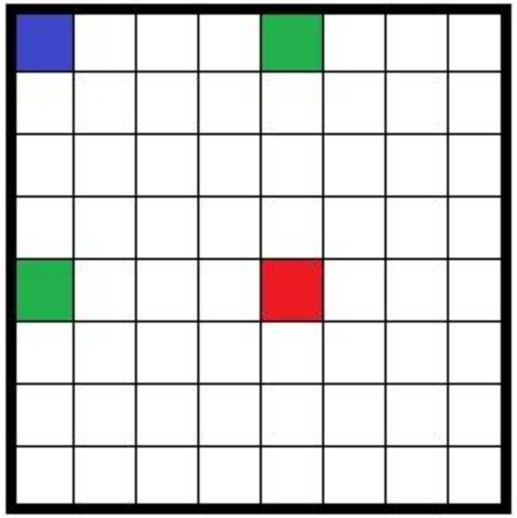


Maximum Distance
[Condat, 2009]

CAMERAS WITH DOMINANT WIDE BAND



Kodak v.3



Teledyne Onyx

Compression: Test environment

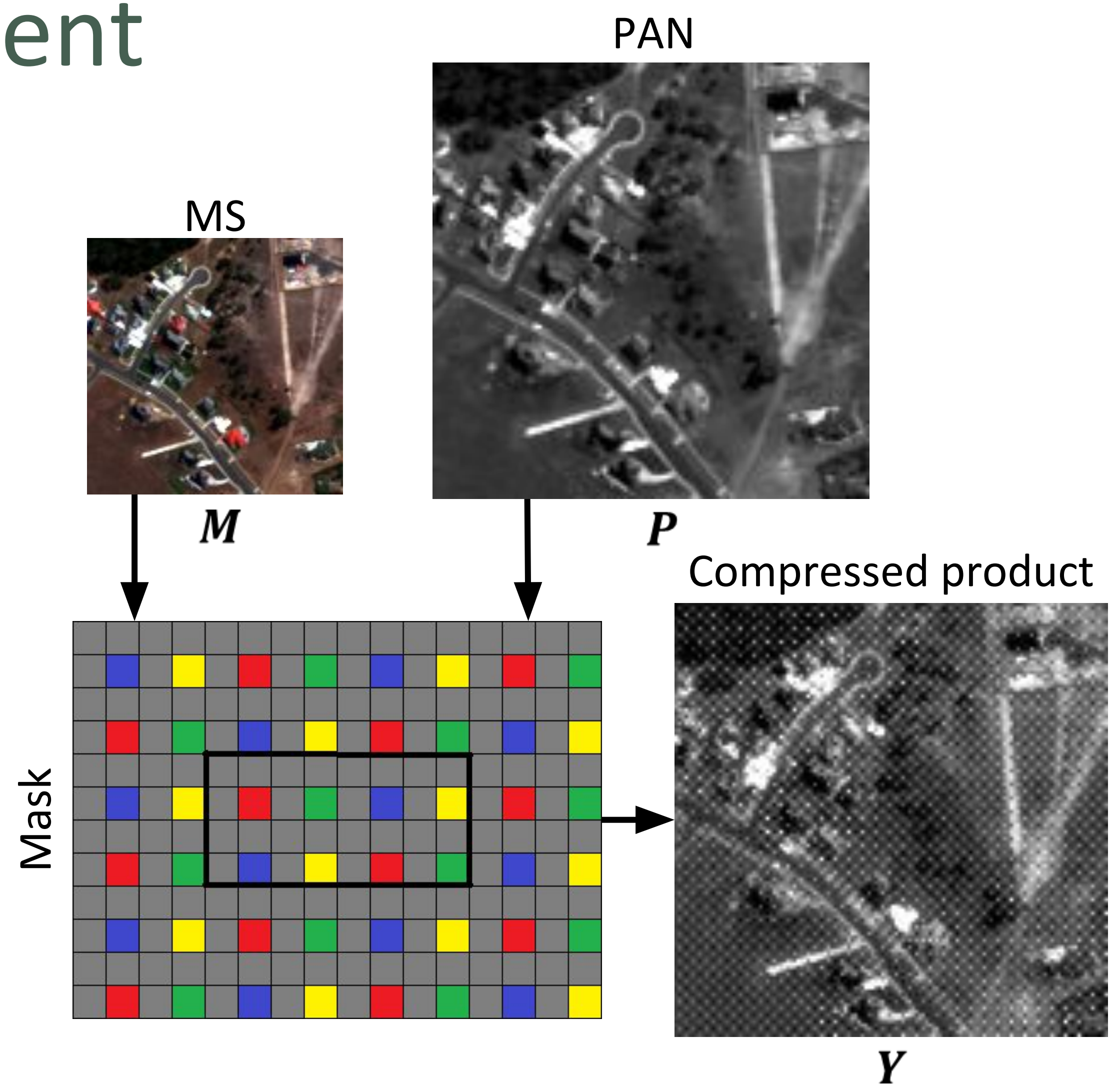
- In our experimental framework, we choose $\boldsymbol{y}$ such that n_c=n_p
- For CFA-style compression we describe each mask $\boldsymbol{H}_0$ (for the PAN) and $\boldsymbol{H}_{1,}, \dots, \boldsymbol{H}_{n_b}$ (for each band of the MS) as binary subsampling matrices
- Final compressed product is hence:

$$\boldsymbol{Y} = \boldsymbol{P} \otimes \boldsymbol{H}_0 + \boldsymbol{U}\left(\sum_{k=1}^{n_b} \boldsymbol{M}_k \otimes \boldsymbol{H}_k\right)$$

Where

- $\otimes$ stands for element-wise product
- $\boldsymbol{M}_k$ is the k-th band of the MS source
- $\boldsymbol{U}$ is a zero-padding (upsampling) oparator

- For CASSI-style compression, the equation can be easily modified by introducing a shift within parenthesis and taking random masks.



Reconstruction scheme: Direct Model

- We propose to solve this problem with a variational approach:
  - We suppose $\boldsymbol{x} \in \mathbb{R}^{n_p n_b}$ is the unknown ideal vector image to reconstruct (written in lexicographic order) and we want to find an estimation $\widehat{\boldsymbol{x}} \in \mathbb{R}^{n_p n_b}$ of such signal
- The PAN and MS sources are supposed to be generated according to this model:

$$\begin{cases} \boldsymbol{p} = \boldsymbol{R}\boldsymbol{x} + \boldsymbol{e}_P \\ \boldsymbol{m} = \boldsymbol{S}\boldsymbol{B}\boldsymbol{x} + \boldsymbol{e}_M \end{cases}$$

- Where:
  - $\boldsymbol{R} \in \mathbb{R}^{n_p \times n_p n_b}$ is a matrix related to the how the spectral response of the MS covers the one of the PAN
  - $\boldsymbol{B} \in \mathbb{R}^{n_p n_b \times n_p n_b}$ is a blurring matrix
  - $\boldsymbol{S} \in \mathbb{R}^{n_m n_b \times n_p n_b}$ is a subsampling matrix
  - $\boldsymbol{e}_P$ and $\boldsymbol{e}_M$ are instances of i.i.d. AWGN with zero mean and an unknown variance

Reconstruction scheme: Inverse Model

- The inversion is achieved by minimizing a cost function, for which we consider two approaches:

Regularization	Cost function	Solver
Vector Total Variation (VTV)		Primal-dual PDFP2O [Chen et al., 2013]
LASSO		SPARSA [Wright et al., 2009]

- Where:
  - $\boldsymbol{A}$ is the linear direct model which includes compression and degradation
  - $||\boldsymbol{Ax}-\boldsymbol{y}||_2^2$ is the maximum likelihood estimator
  - The remaining term is a regularization function
  - λ weights the two contributes
  - Δ_i and Δ_j indicate discrete gradient in the horizontal and vertical direction

Reconstruction scheme: Iterations

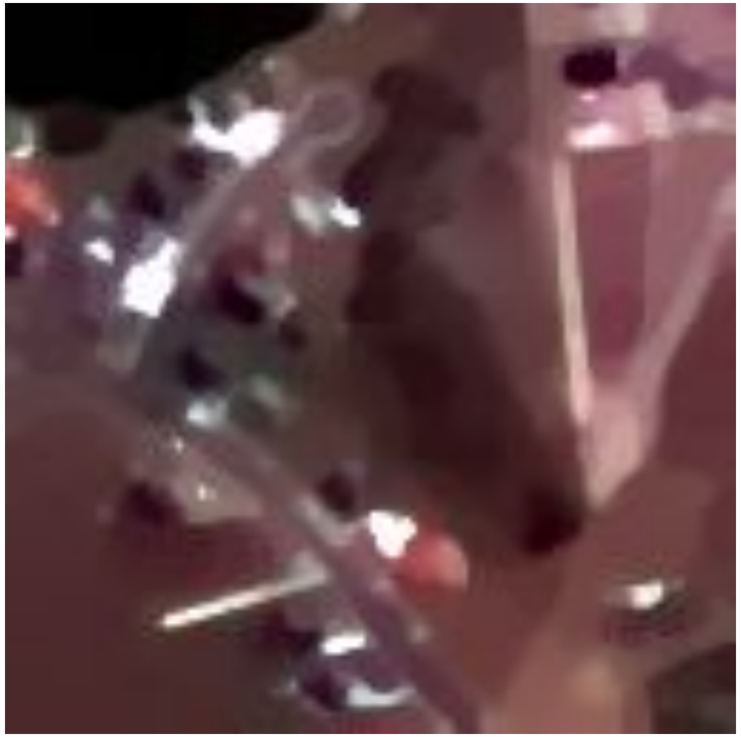
- Iteration: 0
- Iteration: 1
- Iteration: 2
- Iteration: 5
- Iteration: 10
- Iteration: 50
- Iteration: 100
- Iteration: 150
- Iteration: 250

- Dataset specifics:
  - Region: Hobart, Canada
  - Acquisition platform: IKONOS
  - PAN GSD: 2m
  - PAN sizes: 512x512 px
  - Spatial ratio: 2
  - MS bands: 4

Reconstruction scheme: Effect of λ parameter

$\lambda = 0.0001$

$\lambda = 0.0015\ (optimal)$



- $\lambda = 0.0040$

Reduced Resolution validation

- Objective quality assessment was performed according to the Wald's protocol [Wald et al., 1997]

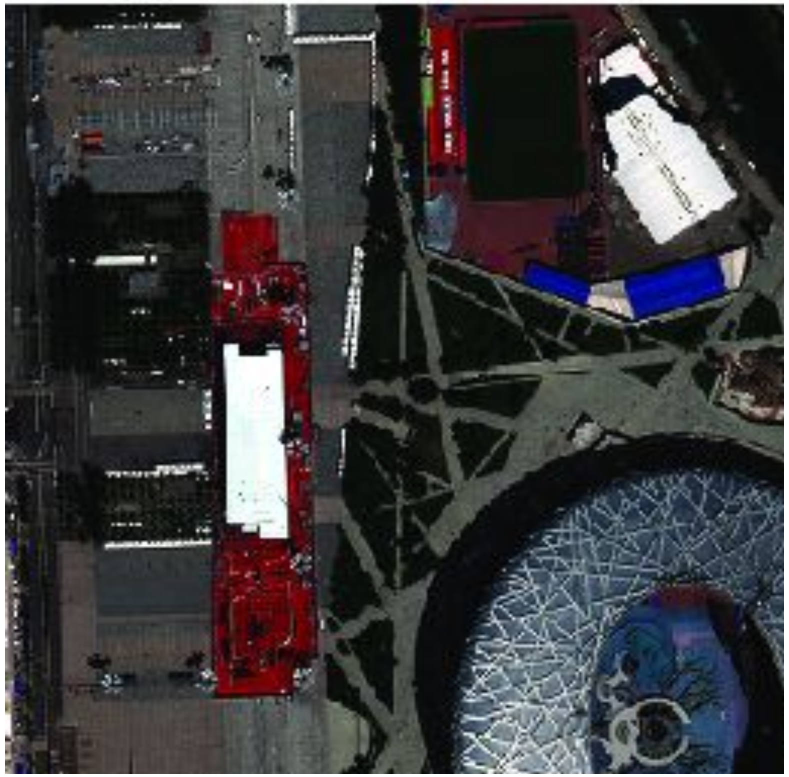
		ERGAS	SAM	Q4	sCC
	Ideal value	0	0	1	1
Hobart	Interpolated MS	6.5	3.0	0.88	0.52
Hobart	MTF-GLP-CBD	3.4	3.0	0.96	0.82
Hobart	CASSI+LASSO	8.2	6.5	0.82	0.53
Hobart	CASSI+VTV	7.0	5.3	0.88	0.62
Hobart	CFA+LASSO	6.3	4.8	0.89	0.57
Hobart	CFA+VTV	**5.2**	**4.0**	**0.93**	**0.65**
Beijing	Interpolated MS	12.5	4.4	0.78	0.30
Beijing	MTF-GLP-CBD	8.3	4.5	0.91	0.74
Beijing	CASSI+LASSO	13.2	9.5	0.77	0.53
Beijing	CASSI+VTV	11.5	6.5	0.82	0.59
Beijing	CFA+LASSO	11.4	6.9	0.83	0.56
Beijing	CFA+VTV	**10.5**	**5.6**	**0.85**	**0.60**

Reference fusion

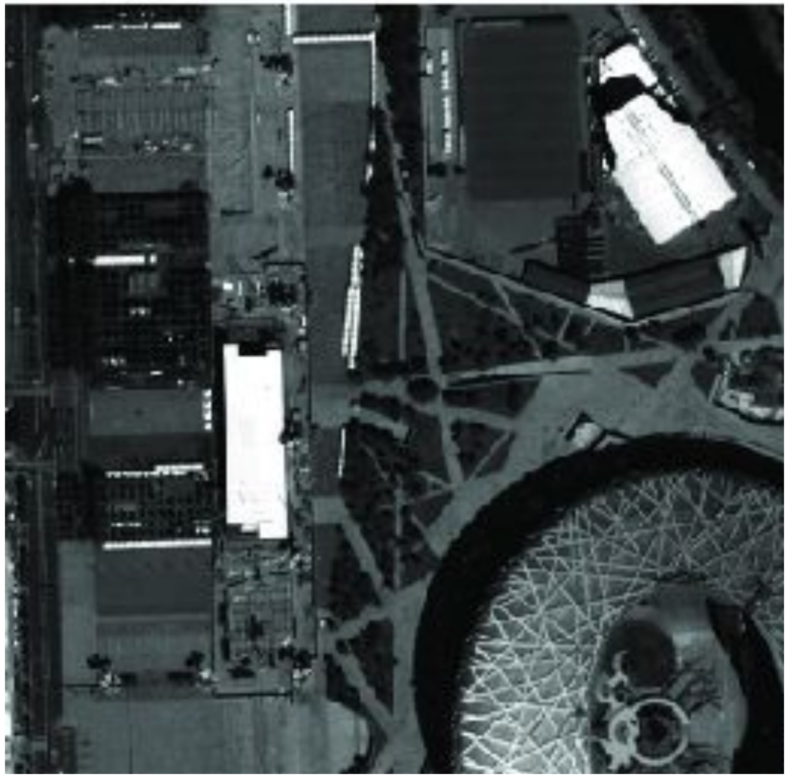
Compressed fusion

Dataset	Hobart	Beijing
Region	Canada	China
Aquisition Platform	IKONOS	Worldview-3
PAN GSD (red. res.)	2m	1,6m
PAN sizes (px)	512x512	512x512
Spatial ratio	2	2
MS bands	4	4

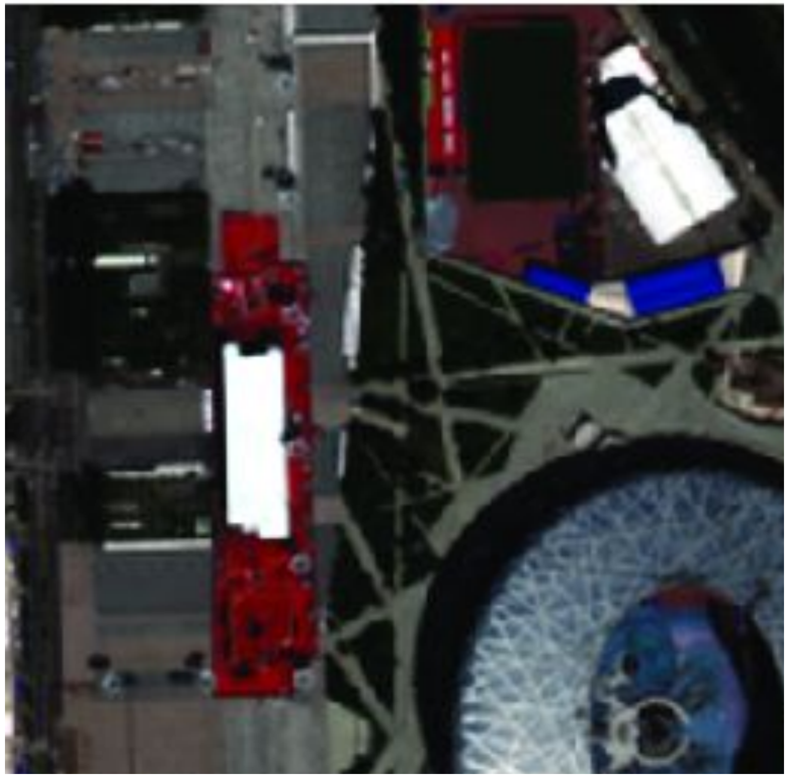
Visual analysis: Beijing dataset



GT (Ground Truth)

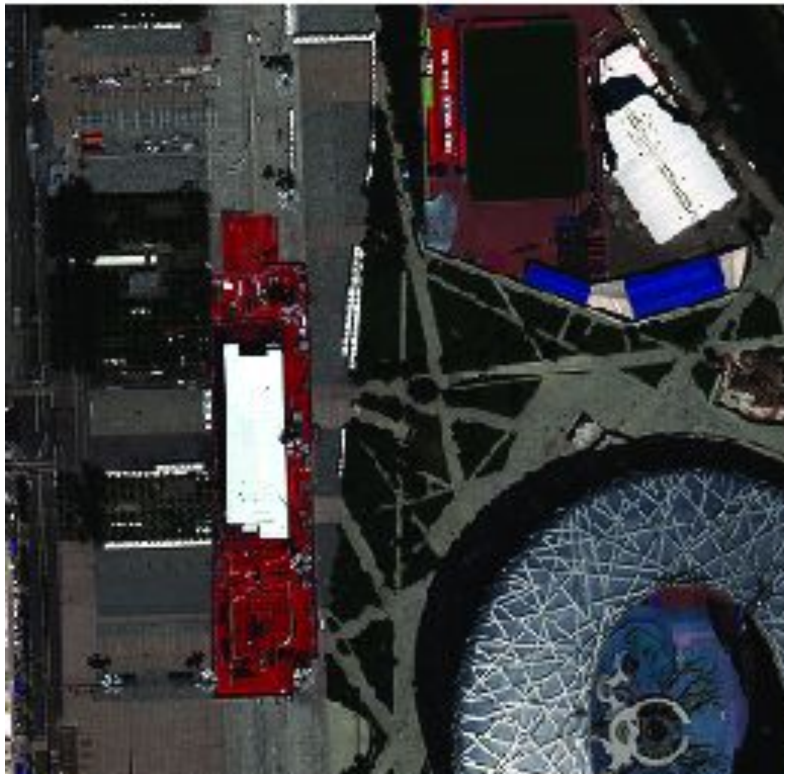


PAN



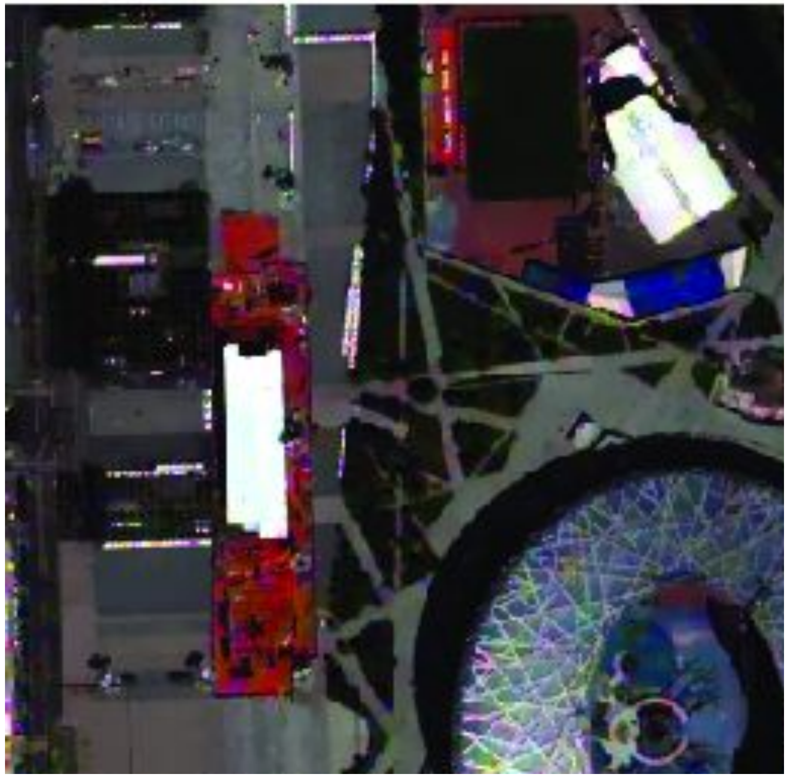
Interpolated MS

Visual analysis: Beijing dataset



GT (Ground Truth)

CFA+VTV



CASSI+VTV

Visual analysis: Hobart dataset

GT (Ground Truth)

PAN

Interpolated MS

Visual analysis: Hobart dataset

GT (Ground Truth)

CFA+VTV

CASSI+LASSO

Conclusions and future perspectives

- We presented a flexible model for joint approach of fusion and reconstruction of compressed images
- Compression can be tailored for optical hardware implementation
- Preliminary tests show potential for the reconstruction with total variation based regularization
- Future perspectives:
  - Comparison with software compression (e.g. JPEG2000)
  - Investigate mathematical conditions which link compression with loss of quality on the fused image
  - Expansion of the framework to hyperspectral images

References

- [Alparone et al., 2007] L. Alparone et al., “Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data fusion contest,” IEEE Trans. Geosci. Remote Sens., vol. 45, no. 10, pp. 3012–3021, Oct. 2007.
- [Arce et al., 2014] Arce, G. R., Brady, D. J., Carin, L., Arguello, H., and Kittle, D. S., “Compressive coded aperture spectral imaging: An introduction,” IEEE Signal Processing Magazine 31, 105–115 (jan 2014).
- [Chen et al., 2013] P. Chen, H. Jianguo, and Z. Xiaoqun, “A primaldual fixed point algorithm for convex separable minimization with applications to image restoration,” Inverse Problems, vol. 29, no. 2, p. 025011, 2013.
- [Condat, 2009] Condat, L., “A new random color filter array with good spectral properties,” in [2009 16th IEEE International Conference on Image Processing (ICIP)], IEEE (nov 2009).
- [Dalla Mura et al., 2015] Dalla Mura, M., Prasad, S., Pacifici, F., Gamba, P., Chanussot, J., and Benediktsson, J. A., “Challenges and opportunities of multimodality and data fusion in remote sensing,” Proc. IEEE 103, 1585–1601 (Sept 2015).
- [Loncan et al., 2015] Loncan, L., Fabre, S., Almeida, L. B., Bioucas-Dias, J. M., Wenzhi, L., Briottet, X., Licciardi, G. A., Chanussot, J., Simoes, M., Dobigeon, N., Tourneret, J. Y., Veganzones, M. A., Qi, W., Vivone, G., and Yokoya, N., “Hyperspectral pansharpening: A review,” IEEE Geosci. Remote Sens. Mag. 3, 27–46 (Sep. 2015).
- [Vivone et al., 2015] Vivone, G., Alparone, L., Chanussot, J., Dalla Mura, M., Garzelli, A., Licciardi, G., Restaino, R., and Wald, L., “A critical comparison among pansharpening algorithms,” IEEE Trans. Geosci. Remote Sens. 53, 2565–2586 (May 2015).
- [Wald et al., 1997] Wald, L., Ranchin, T., and Mangolini, M., “Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images,” Photogramm. Eng. Remote Sens. 63, 691–699 (Jun. 1997).
- [Wright et al., 2009] Wright, Stephen J., Robert D. Nowak, and Mário AT Figueiredo. "Sparse reconstruction by separable approximation." IEEE Transactions on Signal Processing 57.7 (2009): 2479-2493.

Pansharpening (1/2)

- Particular instance of data fusion between satellite acquired images
- Sources:
  - An image with high spatial resolution, such as a panchromatic (PAN)
    - Spectral resolution: Minimum spatial distance required to distinguish two objects on the scene
  - An image with high spectral diversity, such as a multispectral (MS)
    - Spectral diversity: Minimum distance between two separable spectra
- Target: Generate a fused image that features both characteristics at the same time

Pansharpening (2/2)

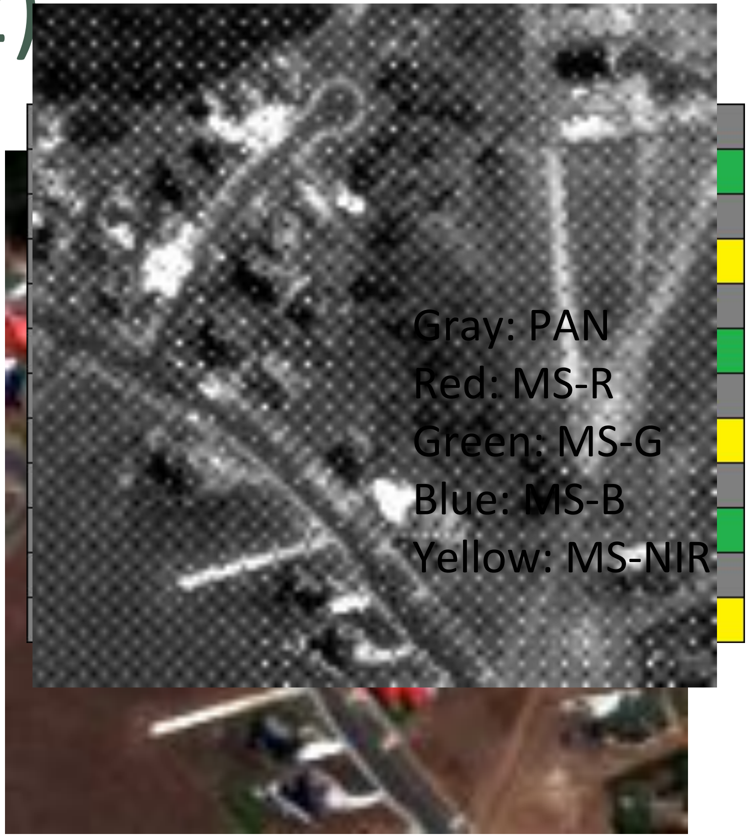
- PAN and MS are acquired by many commercial satellites (Worldview, Geo-eye, QuickBird, IKONOS, etc.)
- Physical and technology constraints don’t allow sensors to acquire an image which jointly features the highest available spatial resolution and spectral diversity
- There is a high commercial demand of fused products (eg: Google Earth)
- Various scientific application: change detection, object recognition, scene interpretation [Dalla Mura et al., 2015]

Compressed Acquisition: Introduction

- Compression allows to save storage resources on board and trasmission bandwidth
- Large amount of image data acquired by satellite platforms
  - Increase of image sizes are expected due to technological improvements (better spatial resolution and larger swaths)
- Allows a plan for a constellation of low-cost satellites
- Example of TOPSAT (a low-cost data satellite [Yu et al., 2009])
  - PAN images of 8000 pixels at 2.5m GSD
  - MS images of three bands of 2100 pixels each at 5m GSD
  - Each snapshot acquisition is around 70 Mbits and the satellite can’t store more than four of those packets

Compression: Proposed approach (1/2)

- Assumption: we want to generate a compression by generating a synthetic P which embeds information from both so
- Sources:
  - A PAN image $\boldsymbol{P} \in \mathbb{R}^{n_p}$
  - A MS image $\boldsymbol{M}$; it will be cal $\widetilde{\boldsymbol{M}}$ when scale e GSD of the (for simplicity
- Method: For each pixel coordinate, each source is masked and then overlayed
- Target: A compressed image $\boldsymbol{Y} \in \mathbb{R}^{n_p}$, reaching a compression rate $n_p/(n_p + n_m n_b)$

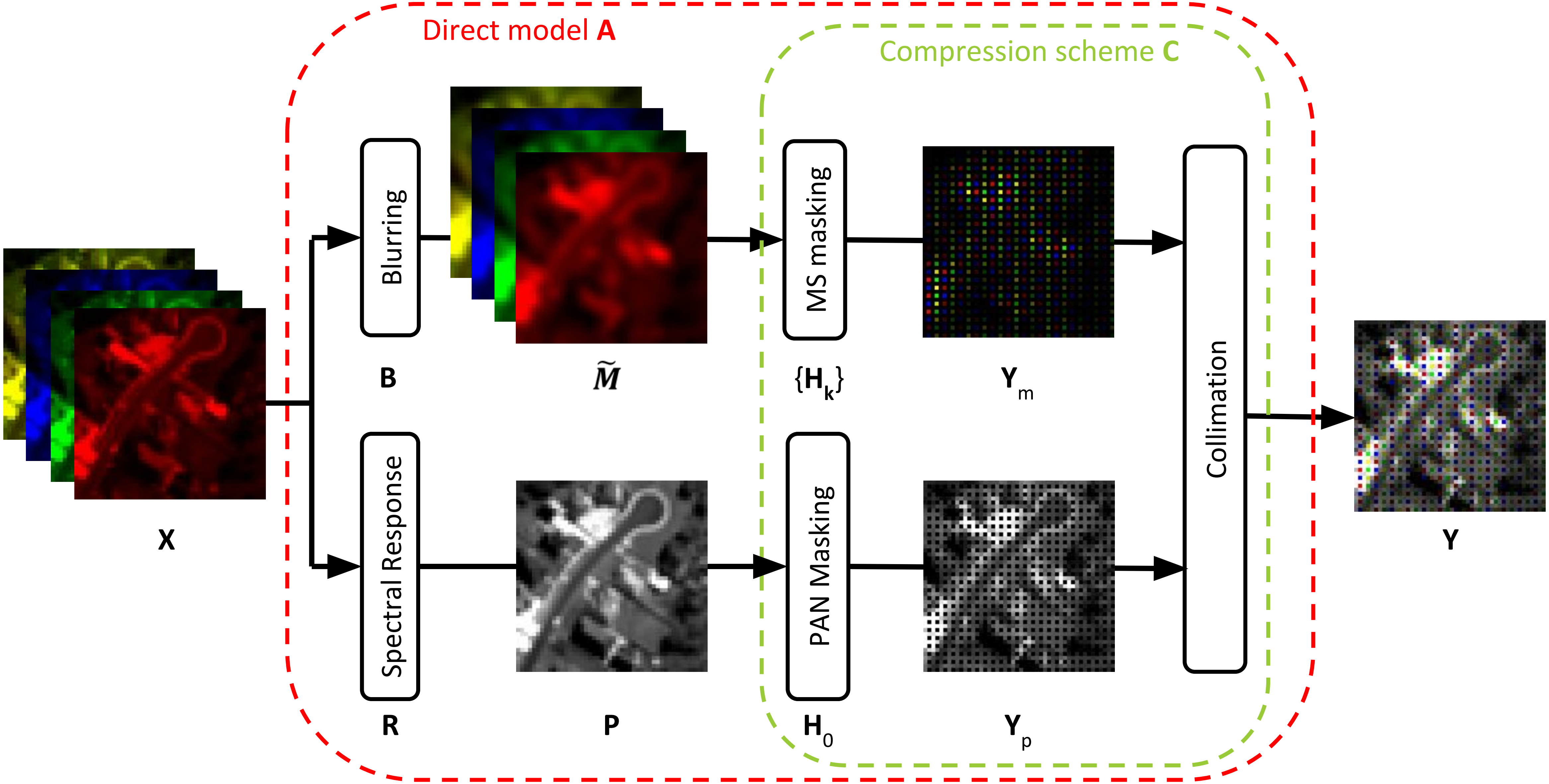


Compressed product
Interpolated MS

Reconstruction scheme

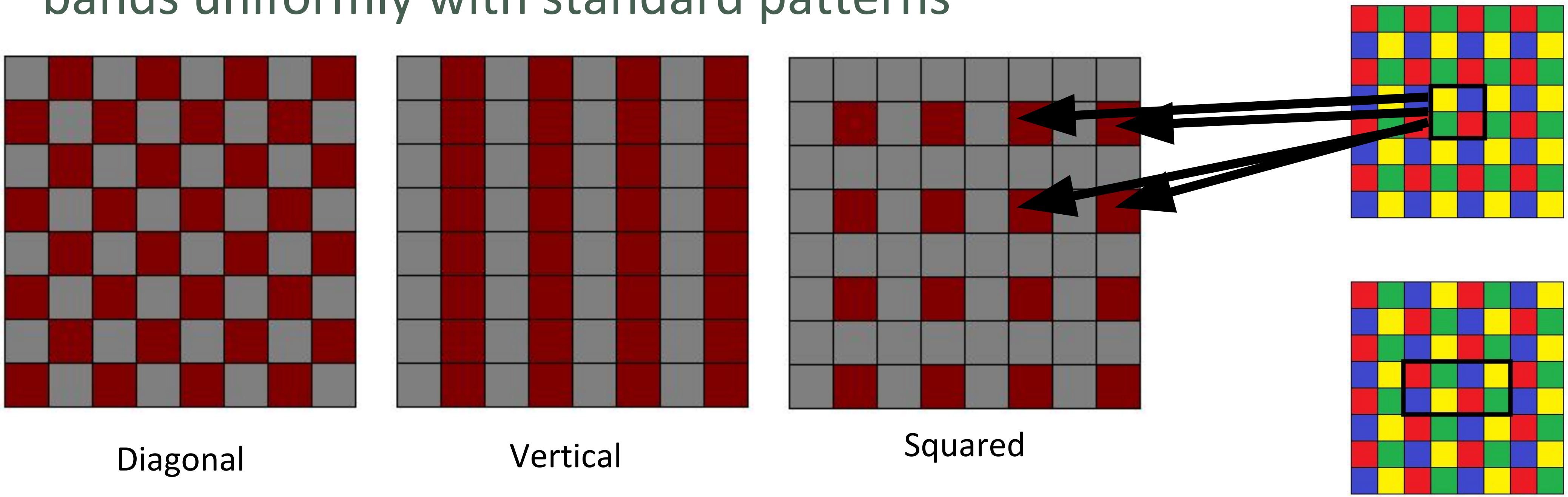
- Deconstruction/demosaicing is done by the ground segment
- It's a joint approach:
  - Not simply a fusion, as we just have a compressed signal
  - Not simply a decompression, as the target is to recover the fused image and not the original sources
- We propose to solve this problem with a variational approach:
  - We suppose $\boldsymbol{X} \in \mathbb{R}^{n_p \times n_b}$ is the unknown ideal image to reconstruct
  - We want to find an estimation $\widehat{\boldsymbol{X}} \in \mathbb{R}^{n_p \times n_b}$ of such signal

RECOSTRUCTION SCHEME: DIRECT MODEL



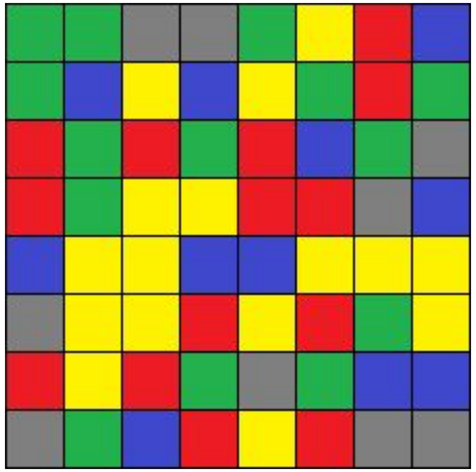
Testbed: Uniform masks

- We use the PAN as dominant pattern and we fill the holes with the MS bands uniformly with standard patterns

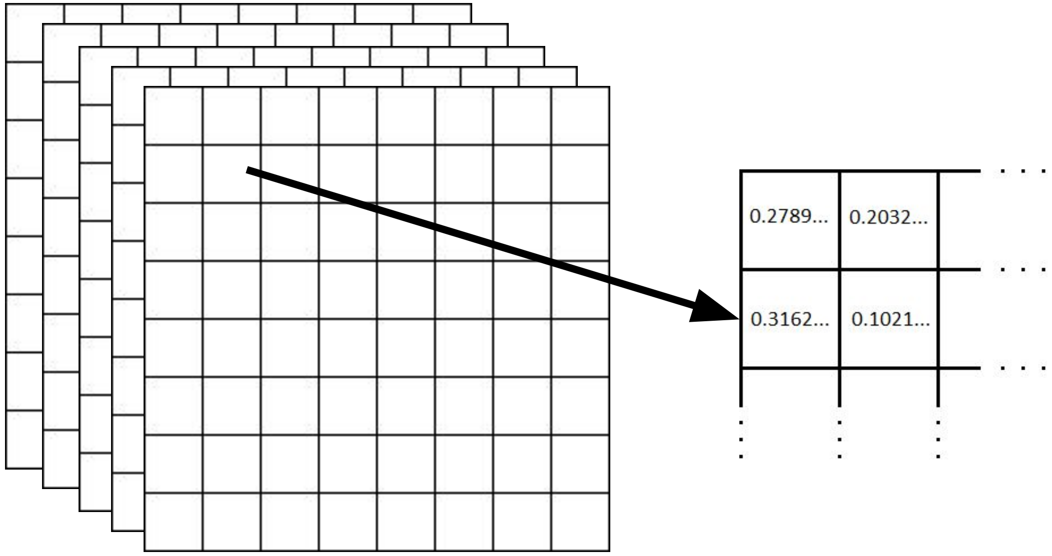


Testbed: Random masks

- We propose a new approach for masks based on matrices with random coefficients

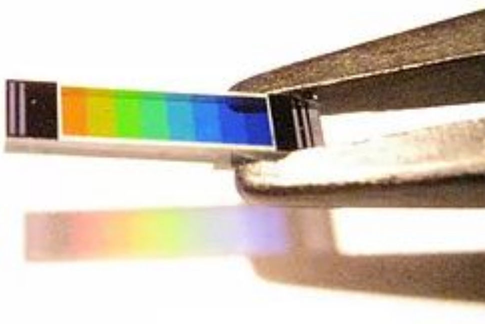


Random pattern
[Amba et al.,2016]



Proposed

Color Shades™



A comparison

- We have made 2 comparison to test the viabililty of the approach

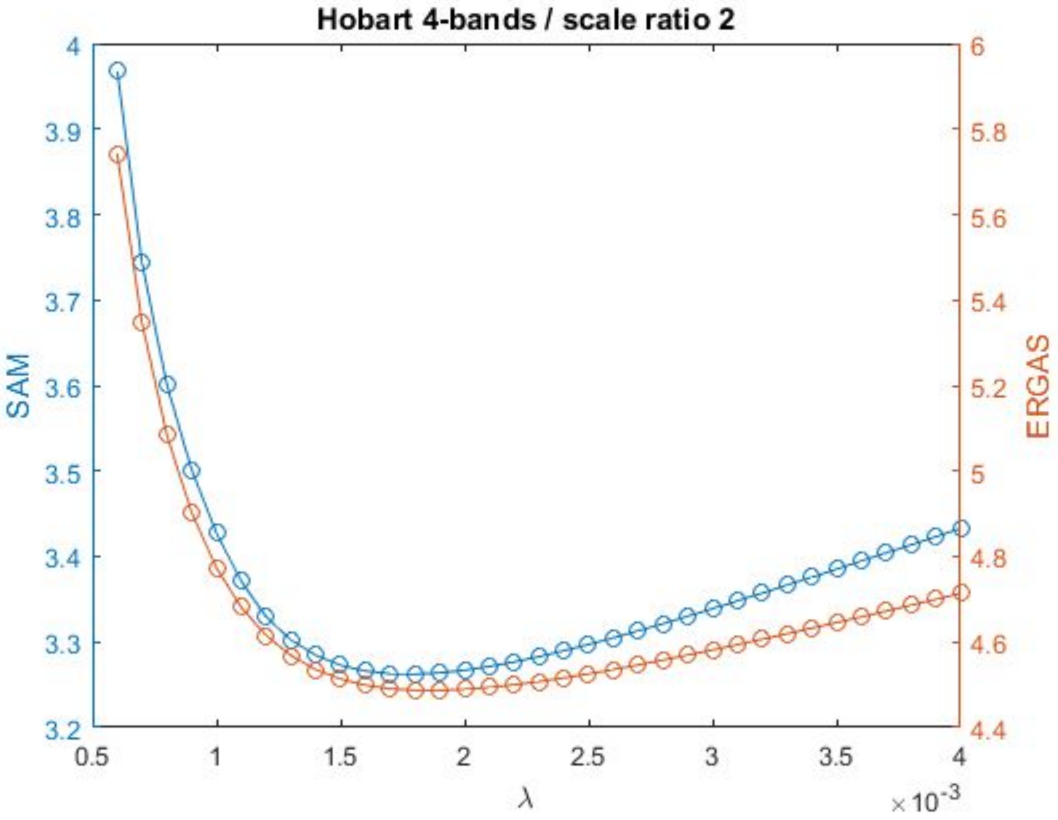
1. We check the most performing classical approach to pansharpening, without any compression, to assess the ideal performance of the fusion; it turns out to be the MTF-GLP-CBD [Alparone et al., 2007]
2. We test all the proposed combinations of masks with our proposed reconstruction framework, including our proposed mask

Quality assessment is performed with the Wald's protocol [Wald et al., 1997]

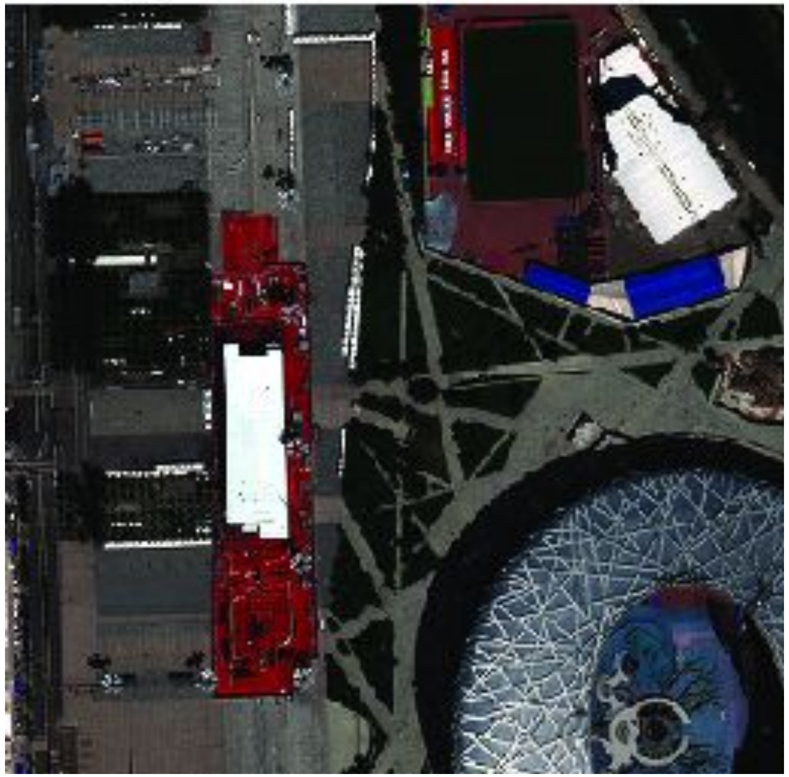
- Dataset specifics:
  - Region: San Francisco, USA
  - Acquisition platform: QuickBird
  - PAN GSD: 2.4m
  - PAN sizes: 128x128 px
  - Spatial ratio: 2
  - MS bands: 4

Reconstruction scheme: Effect of lambda parameter (2/2)

- Tradeoff between imposing a smoothness on the final product and remotion of the texturing effect (overfitting)
- Results are consistent between different quality indices used for the Wald protocol's test [Wald et al., 1997]

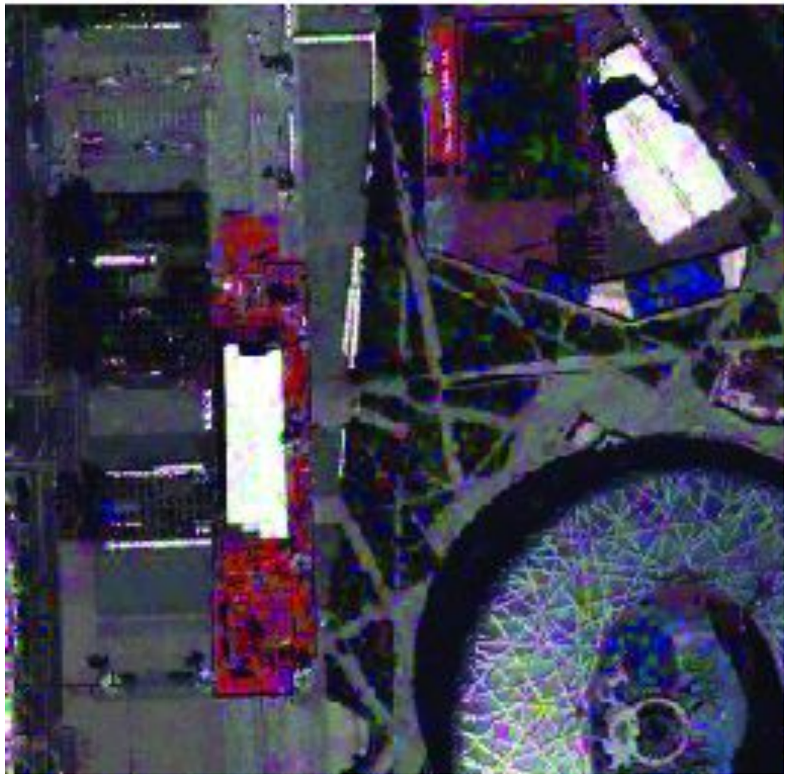


Mask comparison: Visual analysis



GT (Ground Truth)

CFA+TVT



CASSI+LASSO

A comparison: Visual analysis

GT (Ground Truth)

Proposed

PAN

Diagonal+Uniform

MS

References

- [Alparone et al., 2007] L. Alparone et al., “Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data fusion contest,” IEEE Trans. Geosci. Remote Sens., vol. 45, no. 10, pp. 3012–3021, Oct. 2007.
- [Amba et al.,2016] P. Amba, J. Dias, and D. Alleysson, “Random color filter arrays are better than regular ones,” Journal of Imaging Science and Technology, vol. 60, no. 5, pp. 504 061–504 066, sep 2016.
- [Arce et al., 2014] Arce, G. R., Brady, D. J., Carin, L., Arguello, H., and Kittle, D. S., “Compressive coded aperture spectral imaging: An introduction,” IEEE Signal Processing Magazine 31, 105–115 (jan 2014).
- [Chen et al., 2013] P. Chen, H. Jianguo, and Z. Xiaoqun, “A primaldual fixed point algorithm for convex separable minimization with applications to image restoration,” Inverse Problems, vol. 29, no. 2, p. 025011, 2013.
- [Condat, 2009] Condat, L., “A new random color filter array with good spectral properties,” in [2009 16th IEEE International Conference on Image Processing (ICIP)], IEEE (nov 2009).
- [Dalla Mura et al., 2015] Dalla Mura, M., Prasad, S., Pacifici, F., Gamba, P., Chanussot, J., and Benediktsson, J. A., “Challenges and opportunities of multimodality and data fusion in remote sensing,” Proc. IEEE 103, 1585–1601 (Sept 2015).
- [Loncan et al., 2015] Loncan, L., Fabre, S., Almeida, L. B., Bioucas-Dias, J. M., Wenzhi, L., Briottet, X., Licciardi, G. A., Chanussot, J., Simoes, M., Dobigeon, N., Tourneret, J. Y., Veganzones, M. A., Qi, W., Vivone, G., and Yokoya, N., “Hyperspectral pansharpening: A review,” IEEE Geosci. Remote Sens. Mag. 3, 27–46 (Sep. 2015).
- [Vivone et al., 2015] Vivone, G., Alparone, L., Chanussot, J., Dalla Mura, M., Garzelli, A., Licciardi, G., Restaino, R., and Wald, L., “A critical comparison among pansharpening algorithms,” IEEE Trans. Geosci. Remote Sens. 53, 2565–2586 (May 2015).
- [Wald et al., 1997] Wald, L., Ranchin, T., and Mangolini, M., “Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images,” Photogramm. Eng. Remote Sens. 63, 691–699 (Jun. 1997).