CREATIS

INSTITUT NATIONAL DES SCIENCES APPLIQUÉES LYON

ISBI 2020

Recurrent Neural Networks for Compressive Video Reconstruction

March 30, 2020

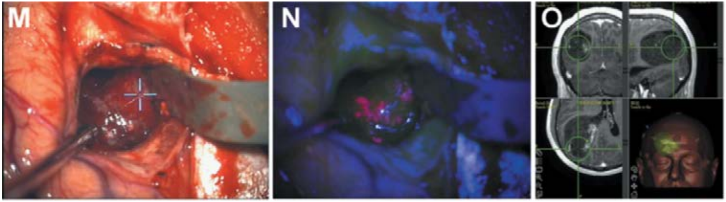
Antonio Tomas Lorente Mur [1] Françoise Peyrin [1] Nicolas Ducros [1]

[1] Univ Lyon, INSA-Lyon, UCB Lyon 1,CNRS, Inserm, CREATIS UMR 5220, U1206, Lyon, France

This work was supported by the French National Research Agency (ANR), under Grant ANR-17-CE19-0003 (ARMONI Project). It was performed within the framework of the LABEX PRIMES (ANR-11-LABX-0063) of Université de Lyon, within the programme "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the ANR.

Context : Medical Applications

- Fluorescence-guided surgery[Valdés *et al.*, *J. of neurosurgery*, 2015].



- Point detection of tumours [Alston *et al.*, *J. of Biomedical Optics*, 2018] - full emission spectrum needed.

Pros

- High spectral resolution

Cons

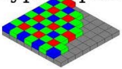
- No spatial resolution
- Surgical gesture is perturbated.

CREATIS

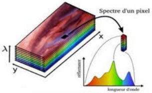
Array

Colour

Hyper-spectral



Multi-spectral



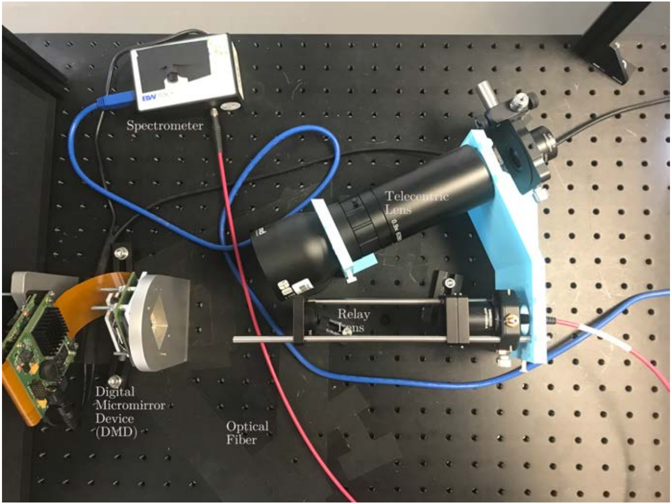
Point

Spectrometer

	Colour	Multi-spectral	Hyper-spectral	Spectrometer
Spatial resolution	yes	yes	yes	no
Number of spectral channels / *Spectral channels*	3	2—10	10—100	100—500
Cost	~€1k	~€10k	~€100k	~€1k

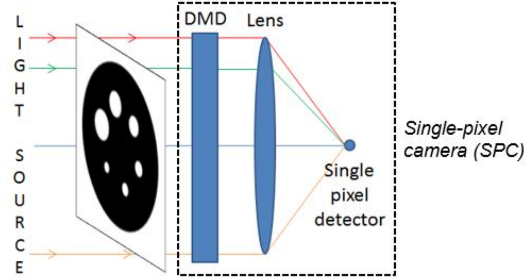
Need for **low cost array** with **high spectral** resolution

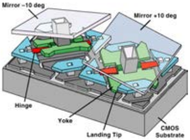
Context : The single-pixel Imaging

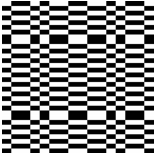


Single-pixel camera experimental setup.

CREATIS







- Single point detector [Duarte *et al.*, *IEEE, SPM*, . 2008].
- **Observe the image through** a spatial light modulator (DMD)
- **Acquire a sequence of measurements for** different patterns ...

... and post-process them!

- Acquisition model.

$$\boldsymbol{m}_t = \boldsymbol{P}\boldsymbol{f}_t, \tag{1}$$

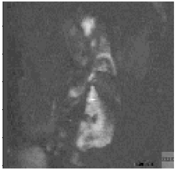
Goal

- Recover $\boldsymbol{f}_t \in \mathbb{R}^N$ from $\boldsymbol{m}_t \in \mathbb{R}^M$ ($N >> M$).
- **Real time.**

Proposed

- Use spatiotemporal redundancy in natural videos.

scene at time t



scene at time $t + \delta t$

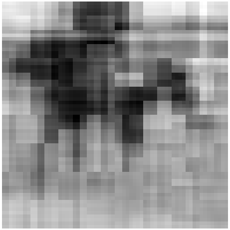
- The image was recovered by solving :

$$\min_{\boldsymbol{f}_t} \mathcal{R}(\boldsymbol{f}_t) \quad \text{such that} \quad \boldsymbol{m}_t = \boldsymbol{P}\boldsymbol{f}_t. \tag{2}$$

Problem :

Ground Truth

$\mathcal{R} = \|\cdot\|_2$



Lack of resolution

$\mathcal{R} = \|\cdot\|_{\mathrm{TV}}$

Iterative scheme

Reconstruction algorithms - neural network approach

- The neural network reconstruction

$$\mathcal{H}_{\boldsymbol{\theta}} = \mathcal{H}_{\boldsymbol{\theta}}^{L} \circ \ldots \circ \mathcal{H}_{\boldsymbol{\theta}}^{1} \tag{3}$$

- The network is then trained :

$$\min_{\boldsymbol{\theta}} \frac{1}{K} \sum_{k=1}^{K} \|\mathcal{H}_{\boldsymbol{\theta}}(\boldsymbol{m}_t^{(k)}) - \boldsymbol{f}_t^{(k)}\|^2 \tag{4}$$

where $\{\boldsymbol{f}_t^{(k)}\}_{k=1}^{K}$ is an image database and $\{\boldsymbol{m}_t^{(k)} = \boldsymbol{P}\boldsymbol{f}_t^{(k)}\}_{k=1}^{K}$.

Exploiting the spatio-temporal redundancy of natural videos

CREATIS

- Online reconstructor with a memory state as a Recurrent Neural Network

$$(\boldsymbol{f}_t^*, \boldsymbol{h}_t) = \Psi_{\theta^*}(\boldsymbol{m}_t, \boldsymbol{h}_{t-1}) \tag{5}$$

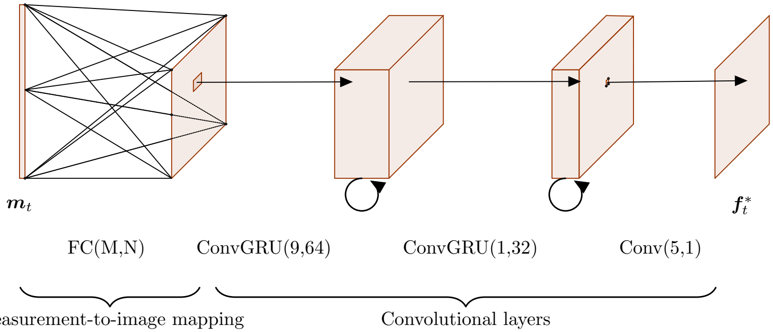
- The network is trained to minimize :

$$\boldsymbol{\theta}^* \in \arg\min \sum_{q=1}^{Q} \sum_{t=1}^{T} \frac{\|\boldsymbol{f}_t^q - \Psi_\theta(\boldsymbol{m}_t^q, \boldsymbol{h}_{t-1}^q)\|_2^2}{2QT} + \lambda \|\boldsymbol{\theta}\|_2^2 \tag{6}$$

For a given training set $\{(\boldsymbol{f}^q_{\{1,\dots,T\}}, \boldsymbol{m}^q_{\{1,\dots,T\}})\}_{1\leq q\leq Q}$

Proposed Network

CREATIS



Proposed recurrent neural network for single-pixel video reconstruction.

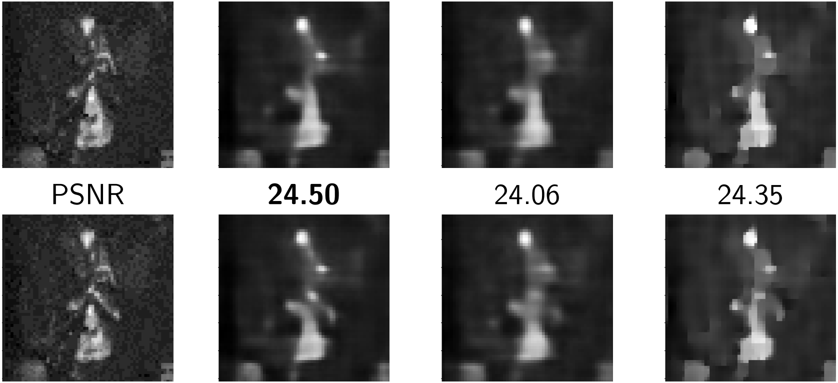
Results on simulated images

- Trained using the UFC-101 [Soomro *et al.*, *CoRR*, . N.d.] with 13 320 videos.
- 1,033,601 learned parameters [Ducros *et al.*, *IEEE, ISBI*, . 2020]
- $M = 333$ Hadamard patterns of size $N = 64 \times 64$ used for the simulation

Method	PSNR	SSIM
Least squares solution (or inverse transform)	20.81	0.9013
Completion method [Ducros *et al.*, *IEEE, ISBI*, . 2020]	21.77	0.9205
Static network [Higham *et al.*, *Sci. Rep.*, 2018]	22.17	0.9255
Proposed recurrent network	**22.25**	**0.9263**

Results on simulated images

	a) Ground truth	**b)** Proposed RNN	**c)** Static network	**d)** Total Variation
PSNR		**24.64**	24.35	24.16
PSNR		**24.50**	24.06	24.35



Conclusion :

- Recurrent neural network to solve the single-pixel video inverse problem
- Nearly instantaneous ($\approx 10ms/$frame) reconstruction

Perspectives :

- Take into account noisy acquisition in the model.

Thanks !!

References I

- Alston, L., Rousseau, D., Hébert, M., Mahieu-Williame, L. & Montcel, B. *J. of Biomedical Optics* **23,** 1–7 (2018).
- Duarte, M. *et al. IEEE, SPM* **25,** 83–91 (2008).
- Ducros, N. *et al. IEEE, ISBI* (2020).
- Higham, C., Murray-Smith, R., Padgett, M. & Edgar, M. *Sci. Rep.* (2018).
- Soomro, K. *et al. CoRR,* 2012.
- Valdés, P. A. *et al. J. of neurosurgery* **123,** 771–780 (2015).