A Completion Network for Reconstruction from Compressed Acquisition

Nicolas Ducros, A Lorente Mur, F. Peyrin

▶ To cite this version:

A COMPLETION NETWORK FOR RECONSTRUCTION FROM COMPRESSED ACQUISITION

N. Ducros, A Lorente Mur, F Peyrin

Univ. Lyon, INSA-Lyon, Université Claude Bernard Lyon 1, UJM Saint-Etienne,
CREATIS CNRS UMR 5220, Inserm U1206, F-69621, Lyon, France

ABSTRACT

We consider here the problem of reconstructing an image from a few linear measurements. This problem has many biomedical applications, such as computerized tomography, magnetic resonance imaging and optical microscopy.

While this problem has long been solved by compressed sensing methods, these are now outperformed by deep-learning approaches. However, understanding why a given network architecture works well is still an open question.

In this study, we proposed to interpret the reconstruction problem as a Bayesian completion problem where the missing measurements are estimated from those acquired. From this point of view, a network emerges that includes a fully connected layer that provides the best linear completion scheme. This network has a lot fewer parameters to learn than direct networks, and it trains more rapidly than image-domain networks that correct pseudo inverse solutions. Although, this study focuses on computational optics, it might provide some insight for inverse problems that have similar formulations.

Index Terms— Image reconstruction, deep learning, Bayesian completion, fully connected layer

1. INTRODUCTION

Single-pixel imaging is an extreme configuration of computational optics, where a single point detector is used to recover an image [1]. This has been applied to fluorescence microscopy [2], hyperspectral imaging [3, 4], diffuse tomography [5] and image-guided surgery [6]. Single-pixel measurements can be modelled as dot products between an image and some two-dimensional functions that are implemented through a spatial light modulator [7]. To limit acquisition times, it is highly desirable to reduce the number of light patterns, which leads to an undetermined inverse problem.

Image reconstruction from compressed acquisitions, where the number of unknowns is larger than the number of measurements, is a generic problem that has several applications in the field of biomedical imaging (e.g., limited angle computerized tomography, compressed magnetic resonance imaging, computational optics). While such problems have benefited from compressed sensing theory, which exploits sparsity priors, recent studies based on deep learning have shown promising results [8]. Much effort is currently devoted to the interpretation and understanding of why such inverse models perform well. This study is a small step in this direction. Although this study focuses on computational optics, it might provide some insight for problems that have similar formulations.

In [9], the authors proposed a convolutional auto-encoder for single-pixel imaging that outperforms compressed sensing approaches. This network directly maps the measurement vector to the desired image, using a fully connected layer (FCL) followed by convolutional layers. This architecture is similar to the automated transform by manifold approximation (AUTOMAP) network [10], which has been successfully applied to magnetic resonance imaging. Here, we focus on understanding the behaviour of the FCL and its connection to image space learning strategies where a network post-processes an approximate solution, usually obtained by an existing algorithm [11].

1.1. Contribution

Inspired by [12], we adopted a Bayesian framework. Reconstruction from compressed acquisition is interpreted as a completion problem where the missing measurements have to be estimated. In particular, we propose to estimate the missing measurements from their correlation to the acquired measurements. For a given database, we derive the best linear solution for the completion problem. Then, we interpret this linear solution as the FCL of an AUTOMAP-like network, where only the convolution layers are trained. Freezing the the FCL significantly decreases the number of parameters to learn.

1.2. Organization of the paper

In Section 2, we describe compressed acquisition and the associated reconstruction problem. In Section 3, we first reinter-

pret reconstruction as a completion problem. Next, we derive the best linear solution to the completion problem. Finally, we introduce a completion network that relies on setting the FCL as the best linear solution. In Section 4, we compare our approach to setting the FCL as the adjoint of the forward model (as in [11]), and learning the FCL (as in [9, 10]).

2. COMPRESSIVE IMAGING

2.1. Compressive image acquisition

Let $\boldsymbol{f} \in \mathbb{R}^N$ (in photons/s) be the image under acquisition. The main idea of compressive optics is to measure a compressed version $\boldsymbol{f}$ using hardware, and to recover it using software. The acquisition can be modeled by

$$\boldsymbol{m} = \boldsymbol{P}_1 \boldsymbol{f}, \qquad (1)$$

where $\boldsymbol{m} \in \mathbb{R}^M$ is the measurement vector and $\boldsymbol{P}_1 \in \mathbb{R}^{M \times N}$, with $M < N$, is the acquisition matrix. The matrix $\boldsymbol{P}_1$ collects the patterns that are sequentially uploaded on a spatial light modulator to get $\boldsymbol{m}$. The patterns in $\boldsymbol{P}_1$ are traditionally chosen on the basis $\boldsymbol{P} \in \mathbb{R}^{N \times N}$. Classical choices include Fourier, discrete cosine transform, wavelets, and the Hadamard approaches. The relevant patterns $\boldsymbol{P}_1$ can be chosen before acquisition (e.g., low frequency patterns) or adaptively during acquisition [13].

2.2. Compressive image reconstruction

Over the last 10 years, image reconstruction has typically proceeded by solving a minimization problem of the form

$$\min_{\boldsymbol{f}} \mathcal{R}(\boldsymbol{f}) \quad \text{such that} \quad \boldsymbol{m} = \boldsymbol{P}_1 \boldsymbol{f}. \qquad (2)$$

A popular choice for $\mathcal{R}$ is ℓ_2-norm $\mathcal{R} = \| \cdot \|_2$, which leads to the least squares solution $\boldsymbol{f}^* = \boldsymbol{P}_1^\top [\boldsymbol{P}_1 \boldsymbol{P}_1^\top]^{-1} \boldsymbol{m}$. An alternative choice is total variation $\mathcal{R} = \sum_n \|\boldsymbol{\nabla}_n \cdot \|_2$, where $\boldsymbol{\nabla}_n$ computes the gradient at pixel n, or variants that promote piecewise constant solutions. In this case, (2) has to be solved iteratively.

More recently, it was proposed to learn to reconstruct $\boldsymbol{f}$ using the non-linear model

$$\boldsymbol{f}^* = \mathcal{H}(\boldsymbol{m}; \boldsymbol{\theta}), \qquad (3)$$

where $\mathcal{H}$ is a neural network parameterized by $\boldsymbol{\theta}$. Given an image database, the parameters $\boldsymbol{\theta}$ are determined during a training stage by minimizing the reconstruction error. In particular, we consider the convolutional architecture depicted in figure 1.
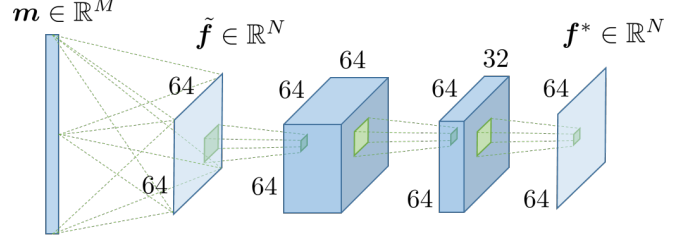

Fig. 1. Bayesian completion network. The first layer is a fully connected layer that outputs an image corrupted by artifacts that are corrected by convolution layers

3. COMPLETION-BASED APPROACH

3.1. Reinterpretation of least-squares

Let $\boldsymbol{y} \in \mathbb{R}^N$ be the full measurement vector such that

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix} = \boldsymbol{P} \boldsymbol{f} = \begin{bmatrix} \boldsymbol{P}_1 \\ \boldsymbol{P}_2 \end{bmatrix} \boldsymbol{f} \qquad (4)$$

where $\boldsymbol{y}_1 \in \mathbb{R}^M$ are the coefficients that are actually acquired in (3), and $\boldsymbol{y}_2 \in \mathbb{R}^L$, $L = N - M$ are the missing coefficients. Assuming $\boldsymbol{P}$ is an orthogonal matrix, the least-squares solution simplifies to

$$\boldsymbol{f}^* = \boldsymbol{P}^\top \boldsymbol{y}^*, \quad \text{with } \boldsymbol{y}^* = \begin{bmatrix} \boldsymbol{m} \\ \boldsymbol{0} \end{bmatrix}, \qquad (5)$$

where $\boldsymbol{y}^*$ is obtained by zero-padding the acquired coefficients. However, neglecting the missing coefficients leads to reduced image quality, as will be shown in Section 5.

3.2. Bayesian completion scheme

We propose to complete the measurement vector with non-zero coefficients before reconstruction; i.e., we consider

$$\boldsymbol{f}^* = \boldsymbol{P}^\top \boldsymbol{y}^*, \quad \text{with } \boldsymbol{y}^* = \begin{bmatrix} \boldsymbol{m} \\ \boldsymbol{y}_2^* \end{bmatrix}, \qquad (6)$$

Our idea is to estimate the missing coefficients $\boldsymbol{y}_2^*$ through their correlation with the acquired coefficients $\boldsymbol{m}$, exploiting a database.

To do so, we adopt a Bayesian point of view and assume that the measurement is a sample of a random vector. In this framework, we can estimate the missing coefficients $\boldsymbol{y}_2^*$ as the conditional expectation

$$\boldsymbol{y}_2^*(\boldsymbol{m}) = \mathbb{E}\left(\mathbf{y}_2 \mid \mathbf{y}_1 = \boldsymbol{m}\right) \qquad (7)$$

where $\mathbf{y}_1$ and $\mathbf{y}_2$ are the random vectors associated to the acquired coefficients and the missing coefficients, respectively, and $\boldsymbol{m}$ is the (deterministic) measured vector. For finite means and variance, the conditional expectation (7) can be computed as the minimum mean-square error solution

$$\boldsymbol{y}_2^* = \mathcal{G}^*(\boldsymbol{m}) \in \arg\min_{\mathcal{G}} \mathbb{E}(\|\mathcal{G}(\mathbf{y}_1) - \mathbf{y}_2\|^2) \qquad (8)$$

3.3. Best linear completion

We first assume that the data random vector follows a multivariate normal distribution with mean $\boldsymbol{\mu} \in \mathbb{R}^N$ and covariance matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{N \times N}$. In notations

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}). \tag{9}$$

where the covariance matrix takes the following form

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_{21}^\top \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_2 \end{bmatrix} \tag{10}$$

with $\boldsymbol{\Sigma}_1 \in \mathbb{R}^{M \times M}$ and $\boldsymbol{\Sigma}_{21} \in \mathbb{R}^{L \times M}$. Under this hypothesis, (8) allows the following closed-form solution

$$\boldsymbol{y}_2^*(\boldsymbol{m}) = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_1^{-1}(\boldsymbol{m} - \boldsymbol{\mu}_1) \tag{11}$$

where $\boldsymbol{\mu}_1 = \mathbb{E}(\mathbf{y}_1)$, $\boldsymbol{\mu}_2 = \mathbb{E}(\mathbf{y}_2)$. Plugging (11) into (6), we have

$$\boldsymbol{f}^* = \boldsymbol{P}^\top \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_1^{-1} \end{bmatrix} \boldsymbol{m} + \boldsymbol{P}^\top \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{\mu}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1 \end{bmatrix}. \tag{12}$$

When the condition (9) is unmet, the estimate provided by (12) achieves minimum squared error among all linear estimators [14]. Therefore, this will be referred to as the best linear completion in the following.

3.4. Proposed completion network

For the general probability distribution of $\mathbf{y}$, the solution of (8) is non-linear mapping. We propose to learn it through a family of non-linear mapping $\mathcal{H}_{\boldsymbol{\theta}}$ parameterized by $\boldsymbol{\theta}$. We consider a neural network model of the form

$$\mathcal{H}_{\boldsymbol{\theta}} = \mathcal{H}_{\boldsymbol{\theta}}^L \circ \ldots \circ \mathcal{H}_{\boldsymbol{\theta}}^1 \tag{13}$$

where $\mathcal{H}^\ell$, $1 \leq \ell \leq L$ is the ℓ-th (non-linear) layer of the network, and $\circ$ is the function composition.

The first layer is traditionally a fully connected layer that maps the measurement $\boldsymbol{m} \in \mathbb{R}^M$ to a raw solution in $\tilde{\boldsymbol{f}} \in \mathbb{R}^N$ (see figure 1). Here, we propose to parameterize the fully connected layer, such that it outputs the best linear solution, i.e.,

$$\tilde{\boldsymbol{f}} = \mathcal{H}^1(\boldsymbol{m}) = \boldsymbol{W}\boldsymbol{m} + \boldsymbol{b} \tag{14}$$

where the weight $\boldsymbol{W}$ and bias $\boldsymbol{b}$ are computed using (12).

Solving the original minimum mean square error problem (8) is computationally intractable in general. Therefore, we minimize the empirical mean square error over an image database

$$\min_{\boldsymbol{\theta}} \frac{1}{K} \sum_{k=1}^{K} \|\mathcal{H}_{\boldsymbol{\theta}}(\boldsymbol{m}^{(k)}) - \boldsymbol{f}^{(k)}\|^2 \tag{15}$$

where $\{\boldsymbol{f}^{(k)}\}_{k=1}^K$ represents an image database and $\{\boldsymbol{m}^{(k)} = \boldsymbol{P}_1\boldsymbol{f}^{(k)}\}_{k=1}^K$ are the associated measurements. Note that when the parameters of the fully connected layers are fixed, the optimization variable $\boldsymbol{\theta}$ has reduced dimension.
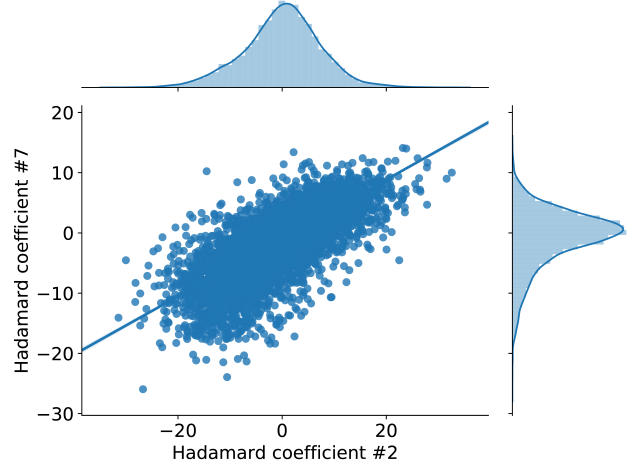

Fig. 2. Measurement scatter plot for two coefficients of the measurement vector. Each dot represents an image from the STL10 test set. The two histograms represents the marginal distribution of each variable while the full line on top of the scatter plot indicates the correlation between the two variables.

4. NUMERICAL EXPERIMENTS

4.1. Training Details

We consider $M = 333$ Hadamard patterns. We compare different choices for FCL of the convolutional neural network depicted in figure 1. In particular, we consider the following three variants:

- *Pseudo inverse network (pinvNET)*. The weights of the FCL are fixed during training such that it behaves as the pseudo inverse of the forward operator. The estimate $\tilde{\boldsymbol{f}}$ corresponds to the least squares solution (5). In this case, there are $L = 8\,129$ trainable parameters.

- *Completion (compNET)*. The weights of the FCL are fixed during training such that the estimate $\tilde{\boldsymbol{f}}$ corresponds to the best linear solution (12). In this case, there are $L = 8\,129$ trainable parameters.

- *Free network (freeNET)*. The weights of the FCL are optimized during training. In this case, there are $L = 1\,376\,193$ trainable parameters.

We implemented all networks using Pytorch [15]. The training was done using the ADAM optimizer, with an initial learning rate of 10^{-3} that is devided by 5 every 20 epochs.

4.2. Dataset

We consider images from the STL10 database [16]. All of the networks were trained using 105k images (i.e., the 'unlabeled' and 'train' subsets); 8k images were used for the test
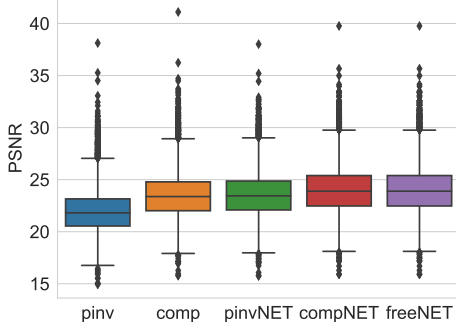
Fig. 3. Box plot of the distribution of the peak signal-to-noise ratios of the reconstructed test images for five different methods. From left to right: pseudo inverse (pinv); Bayesian completion (comp); pseudo inverse network (pinvNET); completion network (compNET); free network (freeNET).

(i.e., the 'test' subset). The original 96×96 images were resized to 64×64 using bicubic transform, and were normalized between -1 and 1.

Figure 2 illustrates two coefficients of the measurement vector for all of the training images. It can be seen that the marginal distributions for both of the coefficients are nearly Gaussian, which is representative of all coefficients, and motivates the hypothesis (9).

5. RESULTS AND DISCUSSION

Figure 3 illustrates the peak signal-to-noise ratios for the reconstructions of all of the test images. We consider the three networks described in Section 4.1 together with the zero padding solution (5) and the best linear completion (12). The highest peak signal-to-noise ratio is achieved by the proposed network (24.1 ± 2.3 dB), which slightly outperforms the free network (24.0 ± 2.2 dB) and the pseudo inverse network (23.6 ± 2.2 dB). Compared to the free network, our proposed network has no FCL to train, and hence it has a lot fewer parameters to train ($8\,129$ vs. $1\,376\,193$). We also observe that the pseudo inverse network particularly improves over the pseudo inverse solution (22.0 ± 2.2 dB) it is built from. However, it only just outperforms the best linear completion scheme (23.5 ± 2.2 dB).

Figure 4 shows the evolution of the empirical mean squared error of the test images during training, for the three network variants. The proposed completion network has the lowest MSE at all epochs. As it corrects the best linear solution, it has much lower MSE than the other two networks at the first epoch. The pseudo inverse network test error decreases at the same speed as the completion network; however, as the initial error is much higher, it fails to reach the same error over 60 epochs. The behaviour of the free network is quite different. Although the initial error is the highest, it decreases rapidly and gets close to the completion
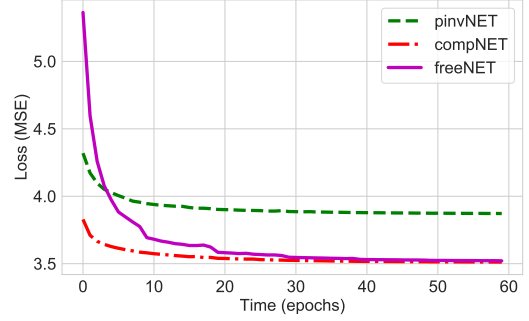

Fig. 4. Evolution of the mean squared errors of the 8k test images during training.
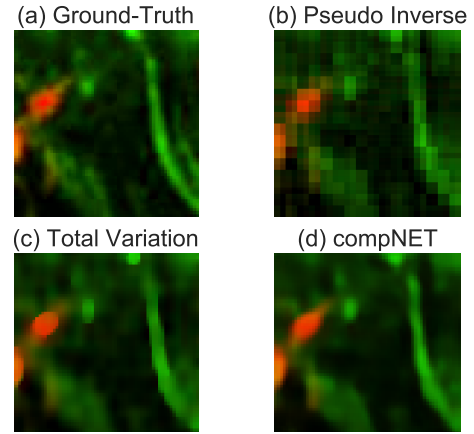

Fig. 5. Reconstrution of a fluorescence microscopy image. (a) ground-truth, (b) pseudo inverse solution given by (5), (c) total variation solution given by (2), and (d) proposed completion network.

network from 40 epochs.

In figure 5, we finally consider a fluorescence microscopy image of melanoma cells (green) migrating on blood vessel surfaces (red). We obtain an improved peak signal-to-noise ratio with the proposed completion network (31.47 dB and 27.84 dB for the red and green channel, respectively) compared to the pseudo inverse solution (27.15 dB and 24.27 dB) and the total variation solution (30.67 dB and 26.68 dB).

6. CONCLUSION

We propose a convolutional network for image reconstruction in single-pixel imaging. Our method is generic and applies to any underdetermined inverse problem. The fully connected layer is chosen to provide the best linear solution, and is frozen during training. This network trains more rapidly and outperforms variants that rely on adjoint operators, in terms of peak signal-to-noise ratio.

7. REFERENCES

[1] M. P. Edgar *et al.*, "Principles and prospects for single-pixel imaging," *Nature Photonics*, vol. 13, no. 1, pp. 13–20, Jan. 2019.

[2] V. Studer *et al.*, "Compressive fluorescence microscopy for biological and hyperspectral imaging," *Proceedings of the National Academy of Sciences*, vol. 109, no. 26, pp. E1679–E1687, 2012.

[3] F. Rousset *et al.*, "Time-resolved multispectral imaging based on an adaptive single-pixel camera," *Opt. Express*, vol. 26, no. 8, pp. 10 550–10 558, Apr 2018.

[4] G. R. Arce *et al.*, "Compressive coded aperture spectral imaging: An introduction," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 105–115, Jan 2014.

[5] Q. Pian *et al.*, "Compressive hyperspectral time-resolved wide-field fluorescence lifetime imaging." *Nature photonics*, vol. 11, pp. 411–414, 2017.

[6] E. Agunounon *et al.*, "Single snapshot imaging of optical properties using a single-pixel camera: a simulation study," *Journal of Biomedical Optics*, vol. 24, no. 7, pp. 1 – 6, 2019.

[7] M. P. Edgar *et al.*, "Principles and prospects for single-pixel imaging," *Nature Photonics*, vol. 13, no. 1, pp. 13–20, Jan. 2019.

[8] S. Ravishankar *et al.*, "Image reconstruction: From sparsity to data-adaptive methods and machine learning," *Proceedings of the IEEE*, pp. 1–24, 2019.

[9] C. F. Higham *et al.*, "Deep learning for real-time single-pixel video," *Scientific Reports*, vol. 8, no. 1, p. 2369, Feb. 2018.

[10] B. Zhu *et al.*, "Image reconstruction by domain-transform manifold learning," *Nature*, vol. 555, p. 487, Mar. 2018.

[11] K. H. Jin *et al.*, "Deep convolutional neural network for inverse problems in imaging," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509–4522, Sep. 2017.

[12] J. Adler *et al.*, "Deep bayesian inversion," 2018.

[13] F. Rousset *et al.*, "Adaptive basis scan by wavelet prediction for single-pixel imaging," *IEEE Transactions on Computational Imaging*, vol. 3, no. 1, pp. 36–46, March 2017.

[14] C. R. Henderson, "Best linear unbiased estimation and prediction under a selection model," *Biometrics*, vol. 31, no. 2, pp. 423–447, 1975.

[15] A. Paszke *et al.*, "Automatic differentiation in pytorch," 2017.

[16] A. Coates *et al.*, "An analysis of single-layer networks in unsupervised feature learning," in *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, 2011, pp. 215–223.