Paper Code: MLSP-P4.2
Paper Number: 1130
Session: Matrix Factorization, Deep Learning and Source Separation

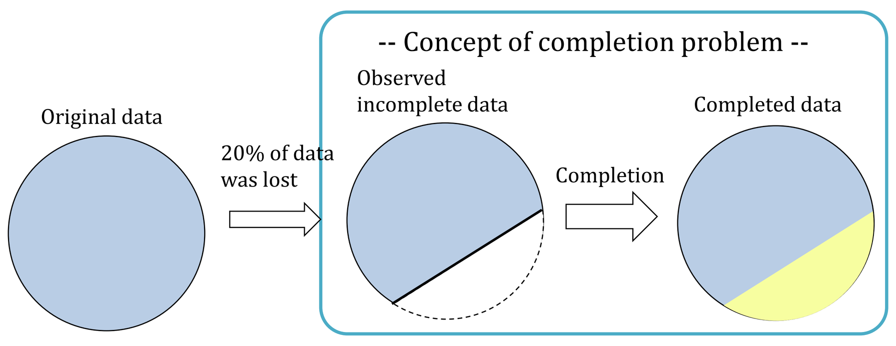
ICASSP 2016, Shanghai, China

# Tensor Completion via Functional Smooth Component Deflation


Tatsuya Yokota and Andrzej Cichocki
Nagoya Institute of Technology (NITECH), Nagoya, Japan
Laboratory for Advanced Brain Signal Processing, RIKEN BSI, Saitama, Japan

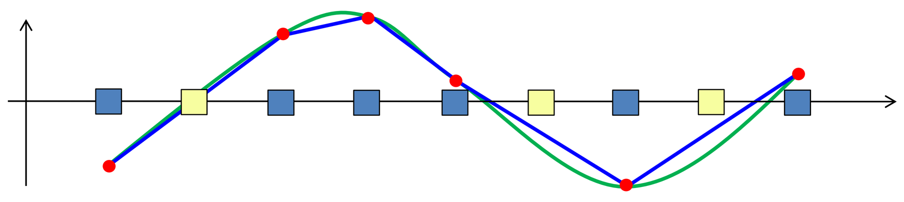

## Introduction

**Completion** is a procedure to estimate values for recovering the data by using
- **available parts of data**
- **structural assumptions**




### Vector Completion (interpolation or regression)

- linear interpolation
- polynomial interpolation



### Matrix Completion

Low rank based matrix completion(IALM) [Lin et al, 2010]

$$\min_{X} ||X||_* \ \text{ s.t. } X_\Omega = T_\Omega$$

$$||X||_* := \sum_i \sigma_i(X)$$

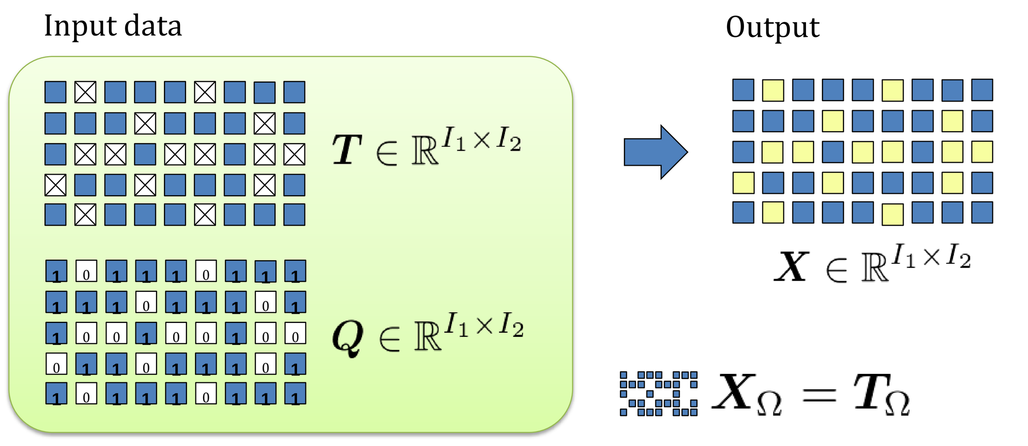
Smooth based matrix completion

$$\min_{X} ||X||_{TV} \ \text{ s.t. } X_\Omega = T_\Omega$$

$$||X||_{TV} = \sum_{i,j} \sqrt{\nabla_v X(i,j)^2 + \nabla_h X(i,j)^2}$$

Combination of both concepts (LTVNN) [Han et al, 2014]

$$\min_{X} ||X||_* + \gamma ||X||_{TV} \ \text{ s.t. } X_\Omega = T_\Omega$$




$T \in \mathbb{R}^{I_1 \times I_2}$

$Q \in \mathbb{R}^{I_1 \times I_2}$

$X \in \mathbb{R}^{I_1 \times I_2}$

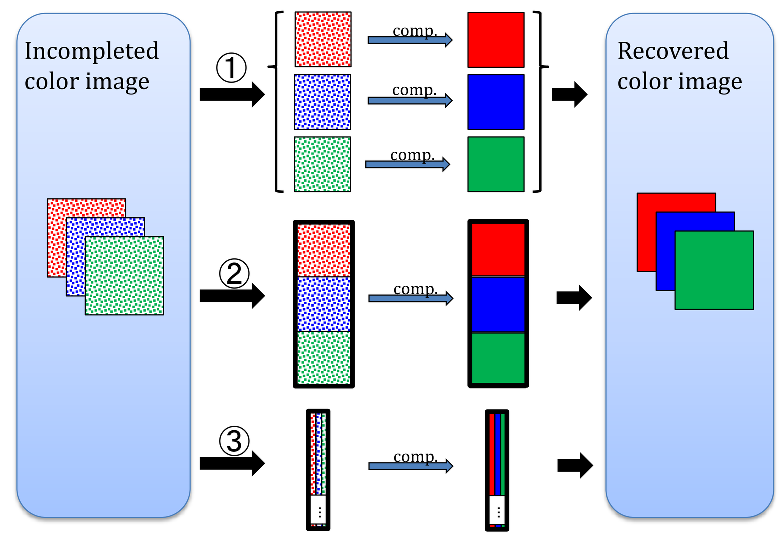
$X_\Omega = T_\Omega$

$X_{\bar{\Omega}} = f(T, Q)$

$$\Omega = \{(i,j) | Q(i,j) = 1, \forall (i,j)\}$$

## Existing Methods for Tensor

When data is given as a "higher order tensor", matrix completion with unfolding can be applied, however it is not good way for the tensor completion




① ignores similarity between color fames

② ignores column similarity between color frames

③ ignores low rank matrix structure of each frame

### Tensor-nuclear norm based algorithm (HaLRTC) [Liu et al, 2013]

$$\min_{\mathcal{X}} \sum_{n=1}^{N} \alpha_n ||X_{(n)}||_* \ \text{ s.t. } \mathcal{X}_\Omega = \mathcal{T}_\Omega$$

### Low-rank Tucker decomposition based algorithm (STDC) [Chen et al, 2014]

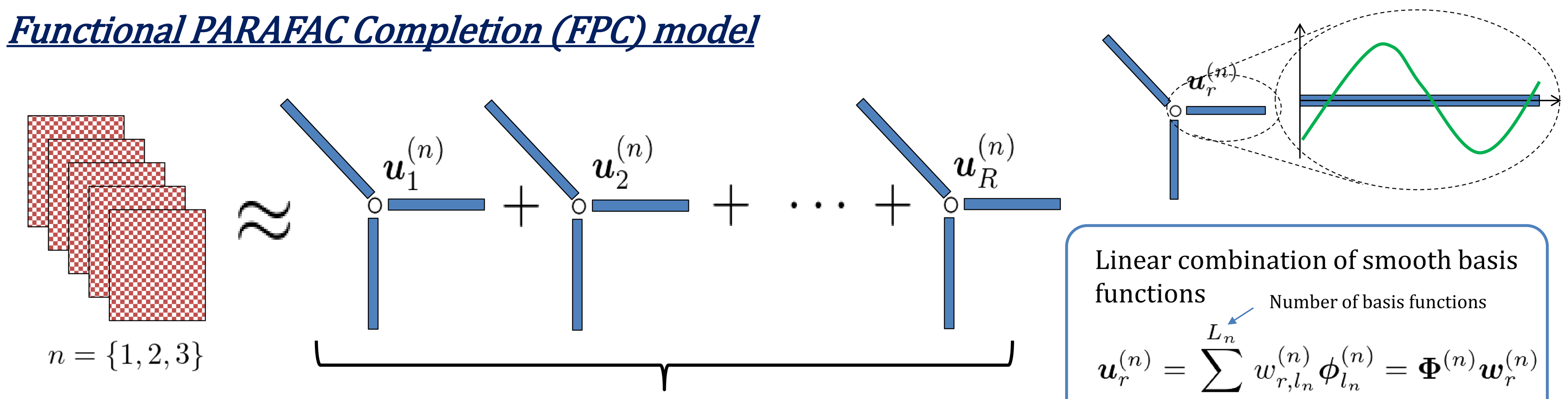
$$\min_{\mathcal{G}, U^{(1)},...,U^{(N)}} \sum_{n=1}^{N} \alpha_n ||U_{(n)}||_* + \delta \mathrm{tr}(\Phi L \Phi^T) + \gamma ||\mathcal{G}||_F^2$$
$$\text{s.t. } \mathcal{X} = \mathcal{G} \times_1 U^{(1)} \times_2 \cdots \times_N U^{(N)}, \ \mathcal{X}_\Omega = \mathcal{T}_\Omega$$

### Low-rank CP decomposition based algorithm (FBCP-MP) [Zhao et al, 2015]

- Low-rank CP decomposition
- Solved by using Bayesian framework
- Additional smoothness constraint (mixture prior)
- Remove redundant component one by one

## Proposed Method

### Functional PARAFAC Completion (FPC) model




Linear combination of smooth basis functions

Number of basis functions

$$u_r^{(n)} = \sum_{l_n=1}^{L_n} w_{r,l_n}^{(n)} \phi_{l_n}^{(n)} = \Phi^{(n)} w_r^{(n)}$$

### Algorithm of Smooth Component Deflation

$\mathcal{E}_1 :=$

For $r = 1, 2, ..., R_{\max}$
- (a) Smooth rank-1 tensor approximation

  $\mathcal{E}_r \approx \mathcal{U}_r :=$

- (b) Deflation

  $\mathcal{E}_{r+1} = \mathcal{E}_r - \mathcal{U}_r$

- Break if stopping criteria are satisfied;

End for

Output: $\mathcal{Z} = \sum_{r=1}^{R} \mathcal{U}_r$

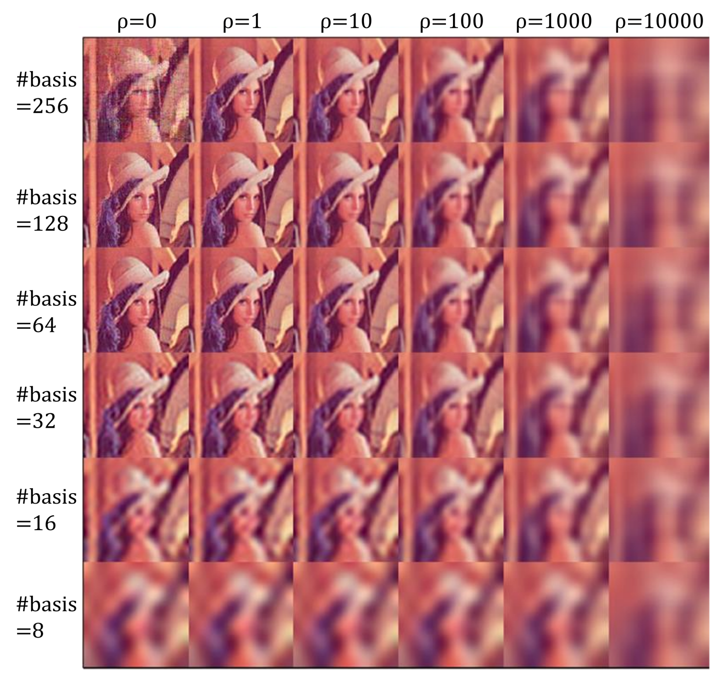
Minimize quadratic variation

$$\text{QV: } \sum_i \left( u_r^{(n)}(i+1) - u_r^{(n)}(i) \right)^2 = ||L^{(n)} u_r^{(n)}||^2$$
$$= w_r^{(n)T} \Phi^{(n)T} L^{(n)T} L^{(n)} \Phi^{(n)} w_r^{(n)}$$
$$\Lambda^{(n)} := \rho \Phi^{(n)T} L^{(n)T} L^{(n)} \Phi^{(n)}$$

#### Smooth rank-1 tensor approximation

$$\min \ \frac{1}{2} ||[\mathcal{E}_r - \mathcal{U}_r]_\Omega||_F^2 + \frac{g_r^2}{2} \sum_{n=1}^{N} w_r^{(n)T} \Lambda w_r^{(n)},$$
$$\text{s.t. } \mathcal{U}_r = g_r \Phi^{(1)} w_r^{(1)} \circ \Phi^{(2)} w_r^{(2)} \circ \cdots \circ \Phi^{(N)} w_r^{(N)}$$
$$||\Phi^{(n)} w_r^{(n)}|| = 1$$

### Selection of Smoothness and Robust FPC algorithm



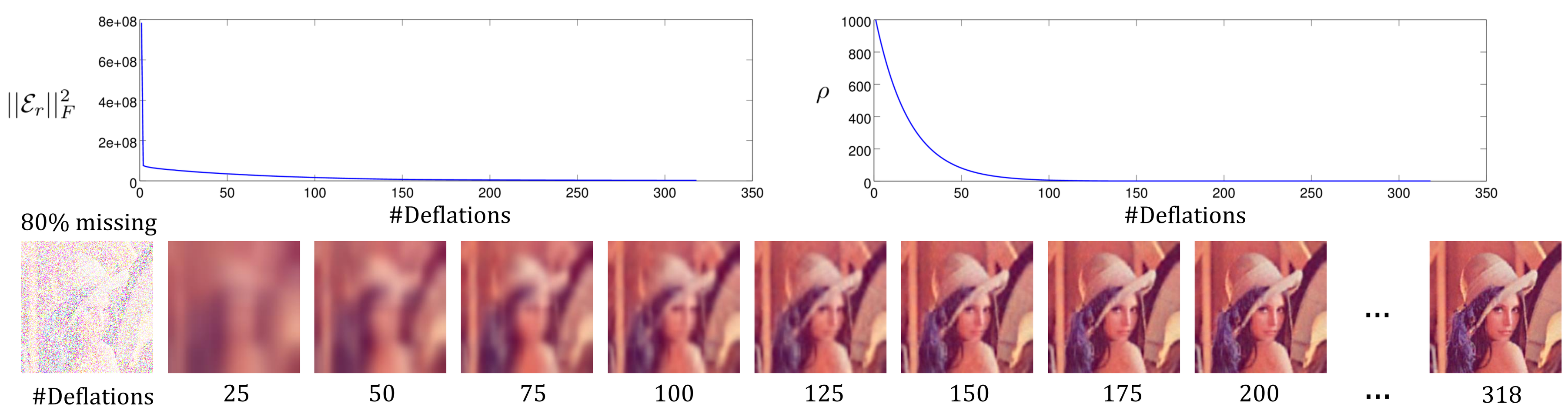

- Completion results are sensitive for smoothness parameter ρ, and number of basis functions.
- Few number of basis functions will reduce computational time of optimization, however too few number of basis functions can not represent complex original images.

#### Idea of Robust FPC (RFPC)

To get robustness for ρ, we start from large ρ and decay ρ in each deflation:

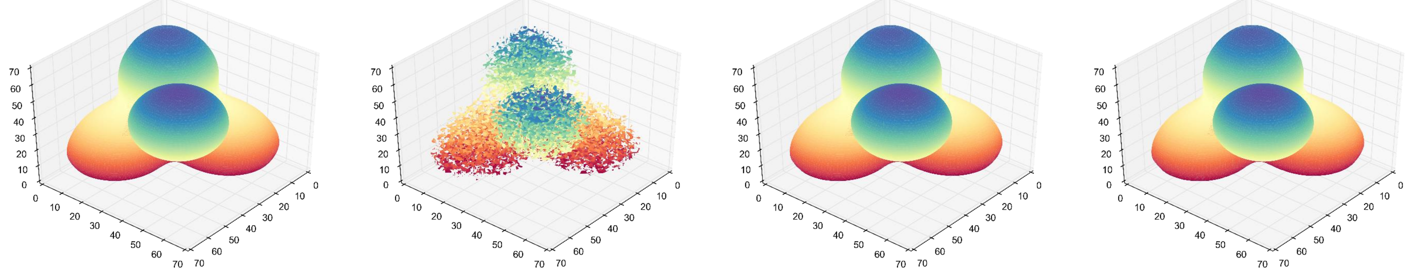
$$\rho \leftarrow \nu \rho \qquad 0 < \nu < 1$$

### Convergence Behavior (RFPC algorithm with 128 cosine basis functions)

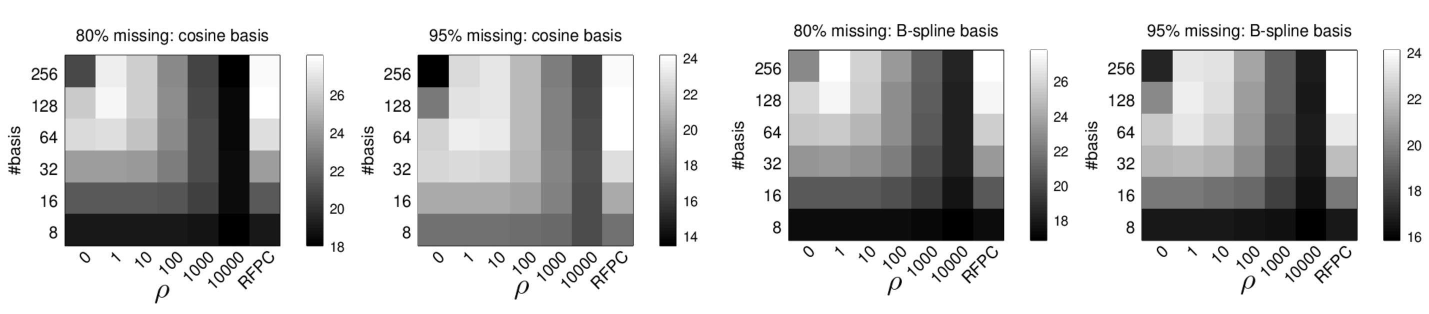



## Experimental results

### Toy problem (signal to inference ratio)

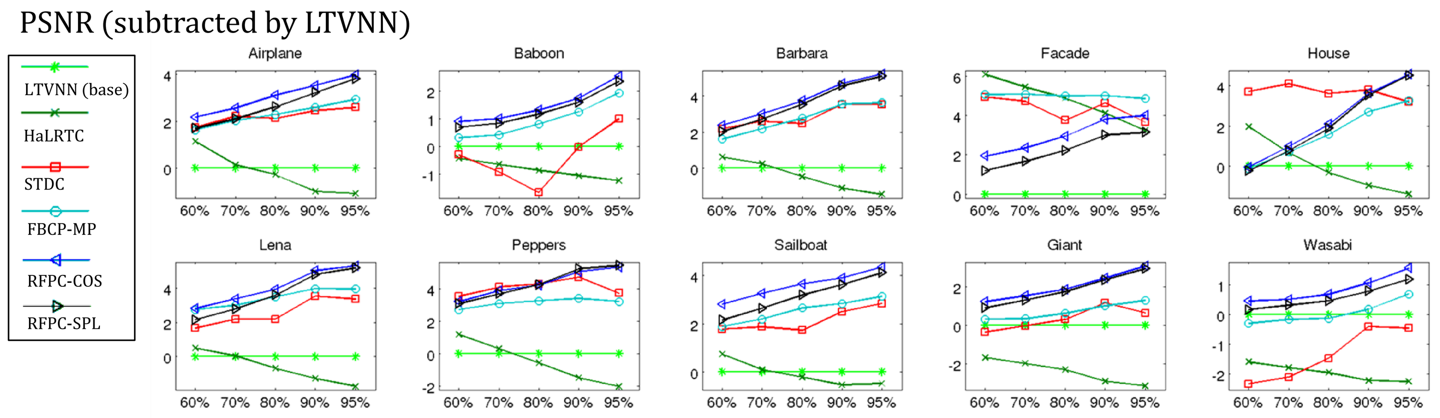



### Results of PSNR [dB] for various settings of FPC and Robust FPC (Lena)




### Results of PSNR [dB] for 10 benchmark images


PSNR (subtracted by LTVNN)




### Selected Images