Fully-Connected Tensor Network Decomposition and Its Application to Higher-Order Tensor Completion

Yu-Bang Zheng[1]

Ting-Zhu Huang[1], Xi-Le Zhao[1], Qibin Zhao[2], Tai-Xiang Jiang[3]

[1]University of Electronic Science and Technology of China, China

[2]Tensor Learning Team, RIKEN AIP, Japan

[3]Southwestern University of Finance and Economics, China

AAAI 2021

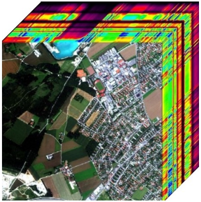
Outline

Outline

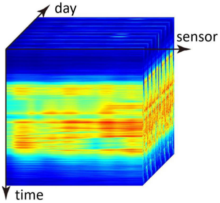
Higher-Order Tensors

Many real-world data are higher-order tensors: e.g., color video, hyperspectral image, and traffic data.

color video



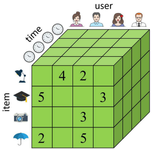
hyperspectral image



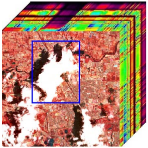
traffic data

Tensor Completion

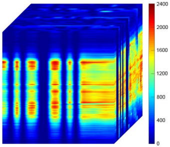
Missing Values Problems: recommender system design, image/video inpainting, and traffic data completion.



recommender system



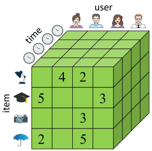
hyperspectral image



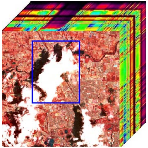
traffic data

Tensor Completion

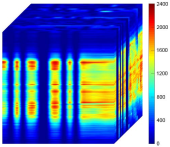
Missing Values Problems: recommender system design, image/video inpainting, and traffic data completion.



recommender system

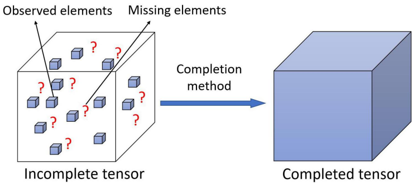


hyperspectral image



traffic data

Tensor Completion (TC): complete a tensor from its partial observation.



Ill-Posed Inverse Problem

Ill-posed inverse problem

Ill-Posed Inverse Problem

Ill-posed inverse problem

$\Uparrow$

Prior/Intrinsic property

- Piecewise smoothness
- Nonlocal self-similarity
- Low-rankness

Ill-Posed Inverse Problem

Ill-posed inverse problem

$\Uparrow$

Prior/Intrinsic property

- Piecewise smoothness
- Nonlocal self-similarity
- Low-rankness

$\Rightarrow$

Low-Rank Tensor Decomposition (Φ)

$$\min_{\mathcal{X},\mathcal{G}} \ \frac{1}{2}\|\mathcal{X} - \Phi(\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_N)\|_F^2, \quad \text{s.t. } \mathcal{P}_\Omega(\mathcal{X}) = \mathcal{P}_\Omega(\mathcal{F}).$$

Minimizing Tensor Rank

$$\min_{\mathcal{X}} \ \mathrm{Rank}(\mathcal{X}), \quad \text{s.t. } \mathcal{P}_\Omega(\mathcal{X}) = \mathcal{P}_\Omega(\mathcal{F}).$$

Here $\mathcal{F} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is an incomplete observation of $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, Ω is the index of the known elements, and $\mathcal{P}_\Omega(\mathcal{X})$ is a projection operator which projects the elements in Ω to themselves and all others to zeros.

Tensor Decomposition

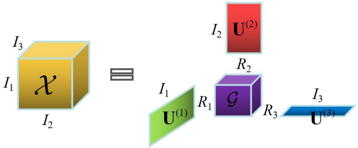
Tensor Decomposition

- decomposes a higher-order tensor to a set of **low-dimensional** factors;
- has powerful capability to **capture the global correlations** of tensors.

Tensor Decomposition

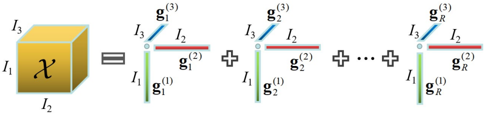
Tensor Decomposition

- decomposes a higher-order tensor to a set of **low-dimensional** factors;
- has powerful capability to **capture the global correlations** of tensors.



$$\mathcal{X} = \mathcal{G} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \cdots \times_N \mathbf{U}^{(N)}$$

Tucker decomposition



$$\mathcal{X} = \sum_{r=1}^{R} \lambda_r \, \mathbf{g}_r^{(1)} \circ \mathbf{g}_r^{(2)} \circ \cdots \circ \mathbf{g}_r^{(N)}$$

CANDECOMP/PARAFAC (CP) decomposition

Tensor Decomposition

Limitations of Tucker Decomposition

- only characterizes correlations among one mode and all the rest of modes, rather than between any two modes;
- needs high storage cost.

Tensor Decomposition

Limitations of Tucker Decomposition

- only characterizes correlations among one mode and all the rest of modes, rather than between any two modes;
- needs high storage cost.

Limitations of CP Decomposition

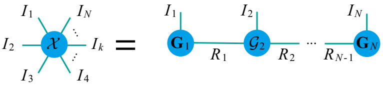
- difficulty in flexibly characterizing different correlations among different modes;
- difficulty in finding the optimal solution.

Tensor Decompositions

Recently, the popular **tensor train (TT) and tensor ring (TR) decompositions** have emerged and shown great ability to deal with **higher-order, especially beyond third-order tensors**.

Tensor Decompositions

Recently, the popular **tensor train (TT) and tensor ring (TR) decompositions** have emerged and shown great ability to deal with **higher-order, especially beyond third-order tensors**.

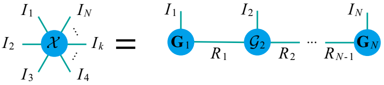


$$\mathcal{X}(i_1, i_2, \cdots, i_N) = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \cdots \sum_{r_{N-1}=1}^{R_{N-1}} \left\{ \mathbf{G}_1(i_1, r_1) \mathcal{G}_2(r_1, i_2, r_2) \cdots \mathbf{G}_N(r_{N-1}, i_N) \right\}$$

TT decomposition

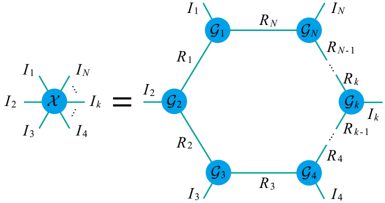
Tensor Decompositions

Recently, the popular **tensor train (TT) and tensor ring (TR) decompositions** have emerged and shown great ability to deal with **higher-order, especially beyond third-order tensors**.



$$\mathcal{X}(i_1, i_2, \cdots, i_N) = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \cdots \sum_{r_{N-1}=1}^{R_{N-1}} \left\{ \mathbf{G}_1(i_1, r_1) \mathcal{G}_2(r_1, i_2, r_2) \cdots \mathbf{G}_N(r_{N-1}, i_N) \right\}$$

TT decomposition



$$\mathcal{X}(i_1, i_2, \cdots, i_N) = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \cdots \sum_{r_N=1}^{R_N} \left\{ \mathcal{G}_1(r_N, i_1, r_1) \mathcal{G}_2(r_1, i_2, r_2) \cdots \mathcal{G}_N(r_{N-1}, i_N, r_N) \right\}$$

TR decomposition

Motivations

Limitations of TT and TR Decomposition

- **A limited correlation characterization**: only establish a connection (operation) between adjacent two factors, rather than any two factors;

Motivations

Limitations of TT and TR Decomposition

- **A limited correlation characterization**: only establish a connection (operation) between adjacent two factors, rather than any two factors;
- **Without transpositional invariance**: keep the invariance only when the tensor modes make a reverse permuting (TT and TR) or a circular shifting (only TR), rather than any permuting.

Examples:

$\triangleright$ reverse permuting: $[1, 2, 3, 4] \rightarrow [4, 3, 2, 1]$;

$\triangleright$ circular shifting: $[1, 2, 3, 4] \rightarrow [2, 3, 4, 1], [3, 4, 1, 2], [4, 1, 2, 3]$.

Motivations

Limitations of TT and TR Decomposition

- **A limited correlation characterization**: only establish a connection (operation) between adjacent two factors, rather than any two factors;
- **Without transpositional invariance**: keep the invariance only when the tensor modes make a reverse permuting (TT and TR) or a circular shifting (only TR), rather than any permuting.

Examples:

▷ reverse permuting: $[1, 2, 3, 4] \rightarrow [4, 3, 2, 1]$;

▷ circular shifting: $[1, 2, 3, 4] \rightarrow [2, 3, 4, 1], [3, 4, 1, 2], [4, 1, 2, 3]$.

How to break through?

Outline

FCTN Decomposition

Definition 1 (FCTN Decomposition)

The FCTN decomposition aims to decompose an Nth-order tensor $\mathcal{X}$ into a set of ***low-dimensional*** *Nth-order factor tensors $\mathcal{G}_k$ ($k = 1, 2, \cdots, N$).The element-wise form of the FCTN decomposition can be expressed as*

$$\begin{aligned}\mathcal{X}(i_1, i_2, \cdots, i_N) = \sum_{r_{1,2}=1}^{R_{1,2}} \sum_{r_{1,3}=1}^{R_{1,3}} \cdots \sum_{r_{1,N}=1}^{R_{1,N}} \sum_{r_{2,3}=1}^{R_{2,3}} \cdots \sum_{r_{2,N}=1}^{R_{2,N}} \cdots \sum_{r_{N-1,N}=1}^{R_{N-1,N}} \\ \{\mathcal{G}_1(i_1, r_{1,2}, r_{1,3}, \cdots, r_{1,N}) \\ \mathcal{G}_2(r_{1,2}, i_2, r_{2,3}, \cdots, r_{2,N}) \cdots \\ \mathcal{G}_k(r_{1,k}, r_{2,k}, \cdots, r_{k-1,k}, i_k, r_{k,k+1}, \cdots, r_{k,N}) \cdots \\ \mathcal{G}_N(r_{1,N}, r_{2,N}, \cdots, r_{N-1,N}, i_N)\}.\end{aligned} \tag{1}$$

Note: Here $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ and $\mathcal{G}_k \in \mathbb{R}^{R_{1,k} \times R_{2,k} \times \cdots \times R_{k-1,k} \times I_k \times R_{k,k+1} \times \cdots \times R_{k,N}}$.

FCTN Decomposition

Definition 1 (FCTN Decomposition)

The FCTN decomposition aims to decompose an Nth-order tensor $\mathcal{X}$ into a set of ***low-dimensional*** *Nth-order factor tensors $\mathcal{G}_k$ ($k = 1, 2, \cdots, N$). The element-wise form of the FCTN decomposition can be expressed as*

$$\begin{aligned}\mathcal{X}(i_1, i_2, \cdots, i_N) = \sum_{r_{1,2}=1}^{R_{1,2}} \sum_{r_{1,3}=1}^{R_{1,3}} \cdots \sum_{r_{1,N}=1}^{R_{1,N}} \sum_{r_{2,3}=1}^{R_{2,3}} \cdots \sum_{r_{2,N}=1}^{R_{2,N}} \cdots \sum_{r_{N-1,N}=1}^{R_{N-1,N}} & \{\mathcal{G}_1(i_1, r_{1,2}, r_{1,3}, \cdots, r_{1,N}) \\ & \mathcal{G}_2(r_{1,2}, i_2, r_{2,3}, \cdots, r_{2,N}) \cdots \\ & \mathcal{G}_k(r_{1,k}, r_{2,k}, \cdots, r_{k-1,k}, i_k, r_{k,k+1}, \cdots, r_{k,N}) \cdots \\ & \mathcal{G}_N(r_{1,N}, r_{2,N}, \cdots, r_{N-1,N}, i_N)\}.\end{aligned} \tag{1}$$

Note: Here $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ and $\mathcal{G}_k \in \mathbb{R}^{R_{1,k} \times R_{2,k} \times \cdots \times R_{k-1,k} \times I_k \times R_{k,k+1} \times \cdots \times R_{k,N}}$.

FCTN-ranks: the vector (length: $N(N-1)/2$) collected by R_{k_1,k_2} ($1 \leq k_1 < k_2 \leq N$ and $k_1, k_2 \in \mathbb{N}^+$).

FCTN Decomposition

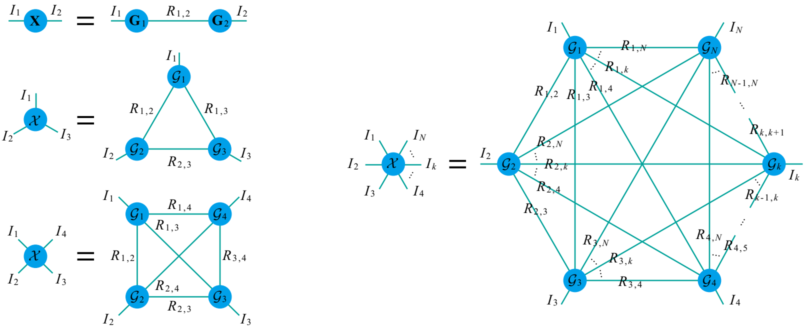


Figure 1: The Fully-Connected Tensor Network Decomposition.

FCTN Decomposition

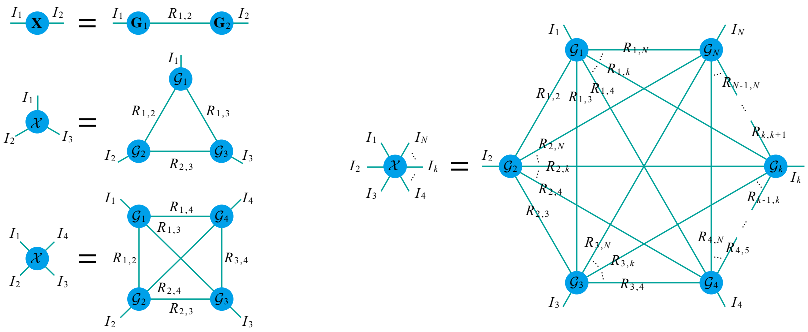


Figure 1: The Fully-Connected Tensor Network Decomposition.

R_{k_1,k_2}: characterizes the intrinsic correlations between the k_1th and k_2th modes of $\mathcal{X}$.

FCTN Decomposition: characterizes the correlations between any two modes.

FCTN Decomposition

Matrices/Second-Order Tensors

$$\mathbf{X} = \mathbf{G}_1\mathbf{G}_2 \Leftrightarrow \mathbf{X}^{\mathsf{T}} = \mathbf{G}_2^{\mathsf{T}}\mathbf{G}_1^{\mathsf{T}}$$

$\Rightarrow$

Higher-Order Tensors

? ? ?

FCTN Decomposition

Matrices/Second-Order Tensors

$$\mathbf{X} = \mathbf{G}_1\mathbf{G}_2 \Leftrightarrow \mathbf{X}^{\mathsf{T}} = \mathbf{G}_2^{\mathsf{T}}\mathbf{G}_1^{\mathsf{T}}$$

$\Rightarrow$

Higher-Order Tensors

? ? ?

Theorem 1 (Transpositional Invariance)

Supposing that an Nth-order tensor $\mathcal{X}$ has the following FCTN decomposition: $\mathcal{X} = \text{FCTN}(\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_N)$. Then, its vector $\mathbf{n}$-based generalized tensor transposition $\vec{\mathcal{X}}^{\mathbf{n}}$ can be expressed as $\vec{\mathcal{X}}^{\mathbf{n}} = \text{FCTN}(\vec{\mathcal{G}}^{\mathbf{n}}_{n_1}, \vec{\mathcal{G}}^{\mathbf{n}}_{n_2}, \cdots, \vec{\mathcal{G}}^{\mathbf{n}}_{n_N})$, where $\mathbf{n} = (n_1, n_2, \cdots, n_N)$ is a reordering of the vector $(1, 2, \cdots, N)$.

Note: $\vec{\mathcal{X}}^{\mathbf{n}} \in \mathbb{R}^{I_{n_1} \times I_{n_2} \times \cdots \times I_{n_N}}$ is generated by rearranging the modes of $\mathcal{X}$ in the order specified by the vector $\mathbf{n}$.

FCTN Decomposition: has transpositional invariance.

FCTN Decomposition

Theorem 2 (The FCTN Rank and the Unfolding Matrix Rank)

Supposing that an Nth-order tensor $\mathcal{X}$ can be represented by Equation (1), the following inequality holds:

$$\text{Rank}\left(\mathbf{X}_{[n_{1:d};n_{d+1:N}]}\right) \leq \prod_{i=1}^{d} \prod_{j=d+1}^{N} R_{n_i,n_j},$$

where $R_{n_i,n_j} = R_{n_j,n_i}$ if $n_i > n_j$ and $(n_1, n_2, \cdots, n_N)$ is a reordering of the vector $(1, 2, \cdots, N)$.

Note: $\mathbf{X}_{[n_{1:d};n_{d+1:N}]} = \texttt{reshape}\left(\vec{\mathcal{X}}^{\mathbf{n}}, \prod_{i=1}^{d} I_{n_i}, \prod_{i=d+1}^{N} I_{n_i}\right)$.

Comparison:

- ▷ TT-rank: $\text{Rank}\left(\mathbf{X}_{[1:d;d+1:N]}\right) \leq R_d$;
- ▷ TR-rank: $\text{Rank}\left(\mathbf{X}_{[1:d;d+1:N]}\right) \leq R_d R_N$;
- ▷ FCTN-ank: $\text{Rank}\left(\mathbf{X}_{[1:d;d+1:N]}\right) \leq \prod_{i=1}^{d} \prod_{j=d+1}^{N} R_{i,j}$.

FCTN Decomposition

Theorem 2 (The FCTN Rank and the Unfolding Matrix Rank)

Supposing that an Nth-order tensor $\mathcal{X}$ can be represented by Equation (1), the following inequality holds:

$$\operatorname{Rank}\left(\mathbf{X}_{[n_{1:d};n_{d+1:N}]}\right) \leq \prod_{i=1}^{d} \prod_{j=d+1}^{N} R_{n_i,n_j},$$

where $R_{n_i,n_j} = R_{n_j,n_i}$ if $n_i > n_j$ and $(n_1, n_2, \cdots, n_N)$ is a reordering of the vector $(1, 2, \cdots, N)$.

Note: $\mathbf{X}_{[n_{1:d};n_{d+1:N}]} = \mathtt{reshape}\left(\vec{\mathcal{X}}^{\mathbf{n}}, \prod_{i=1}^{d} I_{n_i}, \prod_{i=d+1}^{N} I_{n_i}\right)$.

Comparison:

$\triangleright$ TT-rank: $\operatorname{Rank}\left(\mathbf{X}_{[1:d;d+1:N]}\right) \leq R_d$;

$\triangleright$ TR-rank: $\operatorname{Rank}\left(\mathbf{X}_{[1:d;d+1:N]}\right) \leq R_d R_N$;

$\triangleright$ FCTN-ank: $\operatorname{Rank}\left(\mathbf{X}_{[1:d;d+1:N]}\right) \leq \prod_{i=1}^{d} \prod_{j=d+1}^{N} R_{i,j}$.

- the FCTN-rank can bound the rank of all generalized tensor unfolding;
- can capture more informations than TT-rank and TR-rank;

A Discussion of the Storage Cost

CP Decomposition

$\mathcal{O}(NR_1I)$

TT/TR Decomposition

$\mathcal{O}(NR_2^2I)$

Tucker Decomposition

$\mathcal{O}(NIR_3 + R_3^N)$

FCTN Decomposition

$\mathcal{O}(NR_4^{N-1}I)$

A Discussion of the Storage Cost

CP Decomposition

$\mathcal{O}(NR_1 I)$

TT/TR Decomposition

$\mathcal{O}(NR_2^2 I)$

Tucker Decomposition

$\mathcal{O}(NIR_3 + R_3^N)$

FCTN Decomposition

$\mathcal{O}(NR_4^{N-1} I)$

The storage cost of the FCTN decomposition seems to theoretical high. But when we express real-world data, the required FCTN-rank **is usually less** than CP, TT, TR, and Tucker-ranks.

FCTN Composition

Definition 2 (FCTN Composition)

We call the process of generating $\mathcal{X}$ by its FCTN factors $\mathcal{G}_k$ $(k = 1, 2, \cdots N)$ as the FCTN composition, which is also denoted as FCTN$(\{\mathcal{G}_k\}_{k=1}^N)$. If one of the factors $\mathcal{G}_t$ $(t \in \{1, 2, \cdots, N\})$ does not participate in the composition, we denote it as FCTN$(\{\mathcal{G}_k\}_{k=1}^N, /\mathcal{G}_t)$

Theorem 3

Supposing that $\mathcal{X} =$ FCTN$(\{\mathcal{G}_k\}_{k=1}^N)$ and $\mathcal{M}_t =$ FCTN$(\{\mathcal{G}_k\}_{k=1}^N, /\mathcal{G}_t)$, we obtain that

$$\mathbf{X}_{(t)} = (\mathbf{G}_t)_{(t)} (\mathbf{M}_t)_{[m_{1:N-1}; n_{1:N-1}]},$$

where

$$m_i = \begin{cases} 2i, & \text{if } i < t, \\ 2i-1, & \text{if } i \geq t, \end{cases} \quad \text{and} \quad n_i = \begin{cases} 2i-1, & \text{if } i < t, \\ 2i, & \text{if } i \geq t. \end{cases}$$

Outline

FCTN-TC Model

Incomplete Observation		Relationship		Underlying Tensor
$\mathcal{F} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$	$\Leftarrow$	$\mathcal{P}_{\Omega}(\mathcal{X}) = \mathcal{P}_{\Omega}(\mathcal{F})$	$\Rightarrow$	$\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$

FCTN-TC Model

Incomplete Observation		Relationship		Underlying Tensor
$\mathcal{F} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$	$\Leftarrow$	$\mathcal{P}_\Omega(\mathcal{X}) = \mathcal{P}_\Omega(\mathcal{F})$	$\Rightarrow$	$\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$

$\Downarrow$

FCTN Decomposition-Based TC (FCTN-TC) Model

$$\min_{\mathcal{X},\mathcal{G}} \ \frac{1}{2}\|\mathcal{X} - \mathrm{FCTN}(\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_N)\|_F^2 + \iota_{\mathbb{S}}(\mathcal{X}), \tag{2}$$

where $\mathcal{G} = (\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_N)$,

$$\iota_{\mathbb{S}}(\mathcal{X}) := \begin{cases} 0, & \text{if } \mathcal{X} \in \mathbb{S}, \\ \infty, & \text{otherwise}, \end{cases} \quad \text{with } \mathbb{S} := \{\mathcal{X} : \mathcal{P}_\Omega(\mathcal{X} - \mathcal{F}) = 0\},$$

Ω is the index of the known elements, and $\mathcal{P}_\Omega(\mathcal{X})$ is a projection operator which projects the elements in Ω to themselves and all others to zeros.

PAM-Based Algorithm

Proximal Alternating Minimization (PAM)

$$\begin{cases} \mathcal{G}_k^{(s+1)} = \underset{\mathcal{G}_k}{\operatorname{argmin}} \left\{ f(\mathcal{G}_{1:k-1}^{(s+1)}, \mathcal{G}_k, \mathcal{G}_{k+1:N}^{(s)}, \mathcal{X}^{(s)}) + \frac{\rho}{2}\|\mathcal{G}_k - \mathcal{G}_k^{(s)}\|_F^2 \right\}, \ k = 1, 2, \cdots, N, \\ \mathcal{X}^{(s+1)} = \underset{\mathcal{X}}{\operatorname{argmin}} \left\{ f(\mathcal{G}^{(s+1)}, \mathcal{X}) + \frac{\rho}{2}\|\mathcal{X} - \mathcal{X}^{(s)}\|_F^2 \right\}, \end{cases} \tag{3}$$

where $f(\mathcal{G}, \mathcal{X})$ is the objective function of (2) and $\rho > 0$ is a proximal parameter.

PAM-Based Algorithm

Proximal Alternating Minimization (PAM)

$$\begin{cases} \mathcal{G}_k^{(s+1)} = \underset{\mathcal{G}_k}{\operatorname{argmin}} \left\{ f(\mathcal{G}_{1:k-1}^{(s+1)}, \mathcal{G}_k, \mathcal{G}_{k+1:N}^{(s)}, \mathcal{X}^{(s)}) + \frac{\rho}{2} \|\mathcal{G}_k - \mathcal{G}_k^{(s)}\|_F^2 \right\}, \ k=1,2,\cdots,N, \\ \mathcal{X}^{(s+1)} = \underset{\mathcal{X}}{\operatorname{argmin}} \left\{ f(\mathcal{G}^{(s+1)}, \mathcal{X}) + \frac{\rho}{2} \|\mathcal{X} - \mathcal{X}^{(s)}\|_F^2 \right\}, \end{cases} \tag{3}$$

where $f(\mathcal{G}, \mathcal{X})$ is the objective function of (2) and $\rho > 0$ is a proximal parameter.

$\mathcal{G}_k$-Subproblems $(k=1,2,\cdots,N)$

$$\begin{aligned} (\mathbf{G}_k^{(s+1)})_{(k)} &= [\mathbf{X}_{(k)}^{(s)} (\mathbf{M}_k^{(s)})_{[\mathbf{n}_{1:N-1};\mathbf{m}_{1:N-1}]} + \rho (\mathbf{G}_k^{(s)})_{(k)}] [(\mathbf{M}_k^{(s)})_{[\mathbf{m}_{1:N-1};\mathbf{n}_{1:N-1}]} (\mathbf{M}_k^{(s)})_{[\mathbf{n}_{1:N-1};\mathbf{m}_{1:N-1}]} + \rho \mathbf{I}]^{-1}, \\ \mathcal{G}_k^{(s+1)} &= \texttt{GenFold}((\mathbf{G}_k^{(s+1)})_{(k)}, k; 1, \cdots, k-1, k+1, \cdots, N), \end{aligned} \tag{4}$$

where $\mathcal{M}_k^{(s)} = \mathrm{FCTN}(\mathcal{G}_{1:k-1}^{(s+1)}, \mathcal{G}_k, \mathcal{G}_{k+1:N}^{(s)}, /\mathcal{G}_k)$, and vectors $\mathbf{m}$ and $\mathbf{n}$ have the same setting as that in Theorem 3.

$\mathcal{X}$-Subproblem

$$\mathcal{X}^{(s+1)} = \mathcal{P}_{\Omega^c}\left(\frac{\mathrm{FCTN}(\{\mathcal{G}_k^{(s+1)}\}_{k=1}^N) + \rho \mathcal{X}^{(s)}}{1+\rho} \right) + \mathcal{P}_{\Omega}(\mathcal{F}). \tag{5}$$

PAM-Based Algorithm

Algorithm 1 PAM-Based Solver for the FCTN-TC Model.

Input: $\mathcal{F} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, Ω, the maximal FCTN-rank $R^{\max}$, and $\rho = 0.1$.

Initialization: $s = 0$, $s^{\max} = 1000$, $\mathcal{X}^{(0)} = \mathcal{F}$, the initial FCTN-rank $R = \max\{\text{ones}(N(N-1)/2, 1), R^{\max}-5\}$, and $\mathcal{G}_k^{(0)} = \text{rand}(R_{1,k}, R_{2,k}, \cdots, R_{k-1,k}, I_k, R_{k,k+1}, \cdots, R_{k,N})$, where $k = 1, 2, \cdots, N$.

while not converged and $s < s^{\max}$ **do**

Update $\mathcal{G}_k^{(s+1)}$ via (4).

Update $\mathcal{X}^{(s+1)}$ via (5).

Let $R = \min\{R+1, R^{\max}\}$ and expand $\mathcal{G}_k^{(s+1)}$ if $\|\mathcal{X}^{(s+1)} - \mathcal{X}^{(s)}\|_F / \|\mathcal{X}^{(s)}\|_F < 10^{-2}$.

Check the convergence condition: $\|\mathcal{X}^{(s+1)} - \mathcal{X}^{(s)}\|_F / \|\mathcal{X}^{(s)}\|_F < 10^{-5}$.

Let $s = s + 1$.

end while

Output: The reconstructed tensor $\mathcal{X}$.

Theorem 4 (Convergence)

The sequence $\{\mathcal{G}^{(s)}, \mathcal{X}^{(s)}\}_{s\in\mathbb{N}}$ obtained by the Algorithm 1 globally converges to a critical point of (2).

Outline

Synthetic Data Experiments

- Compared Methods: TT-TC (PAM), TR-TC (PAM), and FCTN-TC (PAM);
- Quantitative Metric: the relative error (RSE) between the reconstructed tensor and the ground truth.

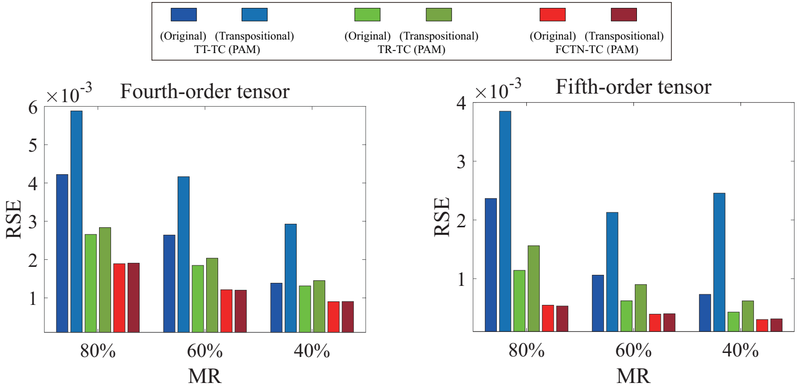


Figure 2: Reconstructed results on the synthetic dataset.

Real Data Experiments

Compared Methods:

- HaLRTC [*Liu et al. 2013; IEEE TPAMI*];
- TMac [*Xu et al. 2015; IPI*];
- t-SVD [*Zhang and Aeron 2017; IEEE TSP*];
- TMacTT [*Bengua et al. 2017; IEEE TIP*];
- TRLRF [*Yuan et al. 2019; AAAI*].

Quantitative Metric:

- PSNR;
- RSE.

Color Video Data

Table 1: The PSNR values and the running times of all utilized methods on the color video data.

Dataset	MR	95%	90%	80%	Mean time (s)	Dataset	MR	95%	90%	80%	Mean time (s)
	Observed	8.7149	8.9503	9.4607	—		Observed	4.5969	4.8315	5.3421	—
	HaLRTC	14.490	18.507	22.460	36.738		HaLRTC	18.617	21.556	25.191	34.528
	TMac	25.092	27.035	29.778	911.14		TMac	26.941	26.142	32.533	1224.4
news	t-SVD	25.070	28.130	31.402	74.807	*containe*	t-SVD	28.814	34.912	39.722	71.510
	TMacTT	24.699	27.492	31.546	465.75		TMacTT	28.139	31.282	37.088	450.70
	TRLRF	22.558	27.823	31.447	891.96		TRLRF	30.631	32.512	38.324	640.41
	FCTN-TC	**26.392**	**29.523**	**33.048**	473.50		FCTN-TC	**30.805**	**37.326**	**42.974**	412.72

Dataset	MR	95%	90%	80%	Mean time (s)	Dataset	MR	95%	90%	80%	Mean time (s)
	Observed	3.8499	4.0847	4.5946	—		Observed	6.4291	6.6638	7.1736	—
	HaLRTC	16.651	20.334	24.813	38.541		HaLRTC	14.561	19.128	23.396	32.882
	TMac	26.753	28.648	31.010	500.70		TMac	25.464	28.169	30.525	779.78
elephants	t-SVD	21.810	27.252	30.975	63.994	*bunny*	t-SVD	21.552	26.094	30.344	66.294
	TMacTT	25.918	28.880	32.232	204.64		TMacTT	26.252	29.512	33.096	264.15
	TRLRF	27.120	28.361	32.133	592.13		TRLRF	27.749	29.034	33.224	652.03
	FCTN-TC	**27.780**	**30.835**	**34.391**	455.71		FCTN-TC	**28.337**	**32.230**	**36.135**	468.25

The data is available at *http://trace.eas.asu.edu/yuv/*.

Color Video Data

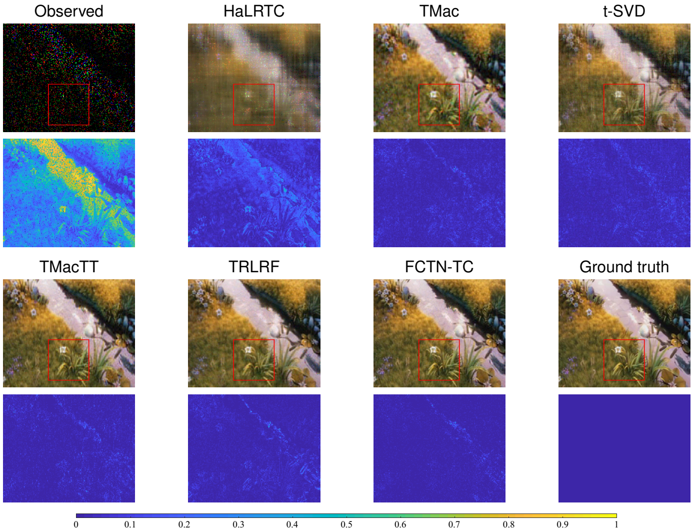


Figure 3: Reconstructed results on the 35th frame of the CV *bunny*.

Traffic Data

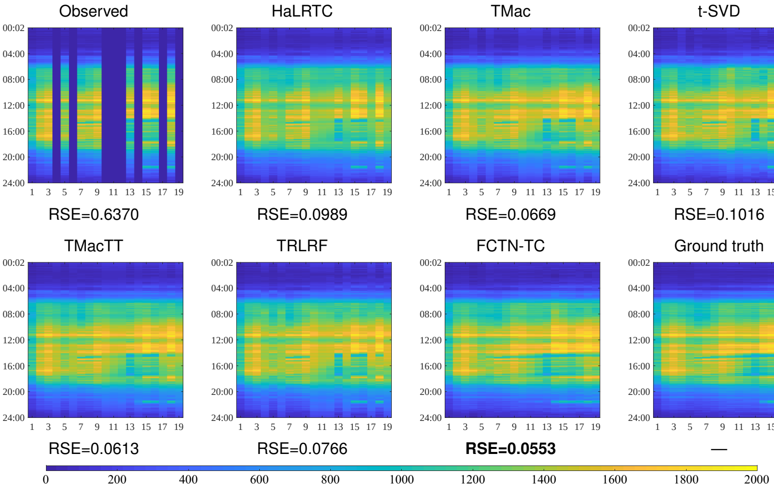


Figure 4: Reconstructed results on the traffic flow dataset with MR=40%. The first and the second rows are the results on the 2nd day and the corresponding residual results, respectively.

The data is available at *http://gtl.inrialpes.fr/*.

Conclusion

Contributions

1. Propose an FCTN decomposition, which breaks through the limitations of TT and TR decompositions;
2. Employ the FCTN decomposition to the TC problem and develop an efficient PAM-based algorithm to solve it;
3. Theoretically demonstrate the convergence of the developed algorithm.

Conclusion

Contributions

1. Propose an FCTN decomposition, which breaks through the limitations of TT and TR decompositions;
2. Employ the FCTN decomposition to the TC problem and develop an efficient PAM-based algorithm to solve it;
3. Theoretically demonstrate the convergence of the developed algorithm.

Challenges and Future Directions

1. Difficulty in finding the optimal FCTN-ranks $\Leftarrow$ Exploit prior knowledge of factors;
2. Storage cost seems to theoretical high $\Leftarrow$ Introduce probability graphical model.

Thank you very much for listening!

Wechat

Homepage: https://yubangzheng.github.io