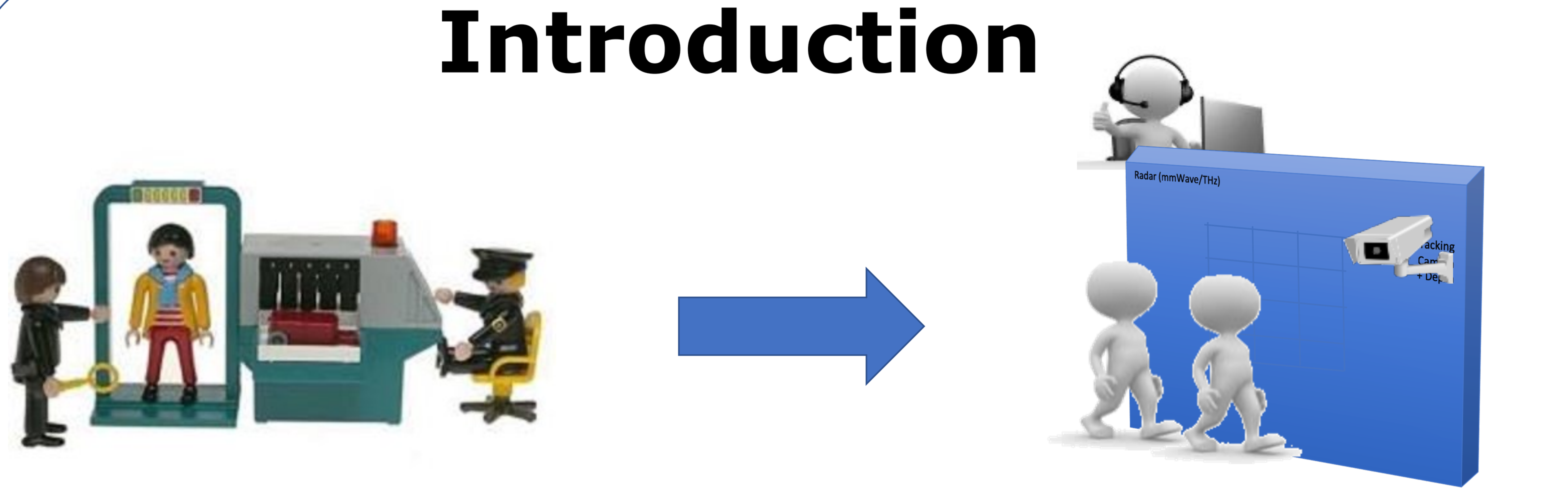
Multiview Sensing with Unknown Permutations: An Optimal Transport Approach

Yanting Ma*, Petros Boufounos*, Hassan Mansour*, Shuchin Aeron#

***Mitsubishi Electric Research Laboratories (MERL)**

#Department of Electrical and Computer Engineering, Tufts University

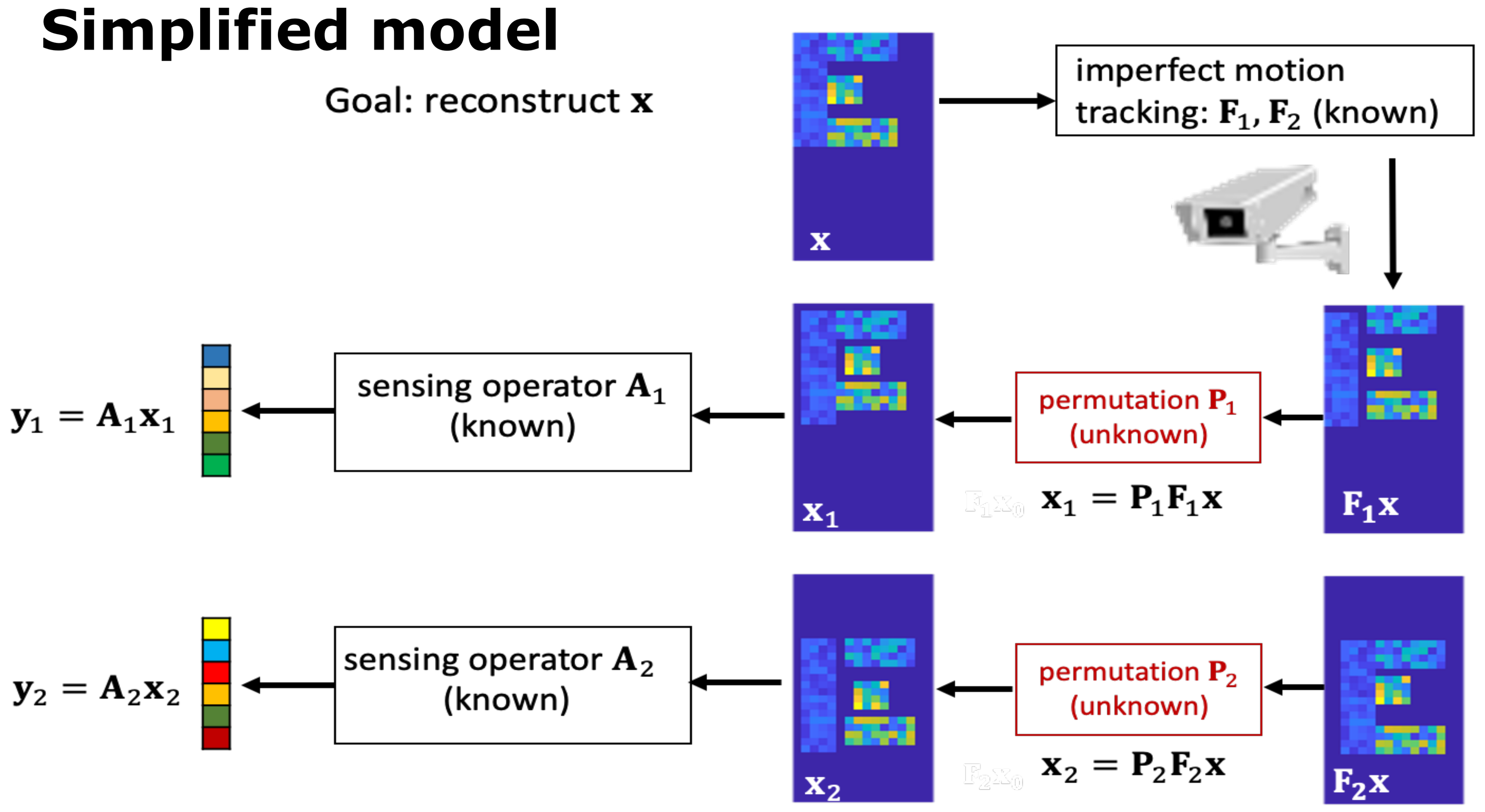
Introduction



traditional security checkpoint — walk-through security checkpoint

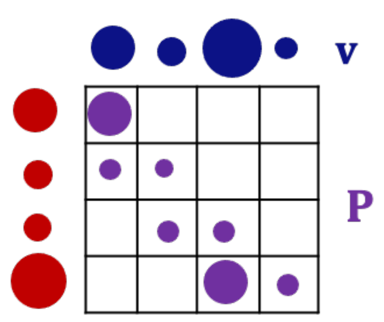
- **Motivation**
  - Traditional security checkpoint requires each person to stand still while being scanned -> slowdowns and long queues
  - By combining the latest developments in optical and depth sensing, tracking, and array processing, we wish to allow scanning of moving individuals with irregular motion as they pass the scanner
  - In this work, we assume an imperfect motion tracking system is available and tackle the image reconstruction task while simultaneously correcting the tracking error

- **Simplified model**



  - As the object $\mathbf{x}$ moves, it is measured in a sequence of frames. Due to the motion, it undergoes a series of deformations which determines its pose in each frame of the measurements.
  - Motion tracking error is assumed to be corrected by an unknown permutation matrix $\mathbf{P}_i$
  - Measurement model for the i^{th} view: $\mathbf{y}_i = \mathbf{A}_i\mathbf{P}_i\mathbf{F}_i\mathbf{x} + \mathbf{w}_i$

- **Optimal transport (OT)**

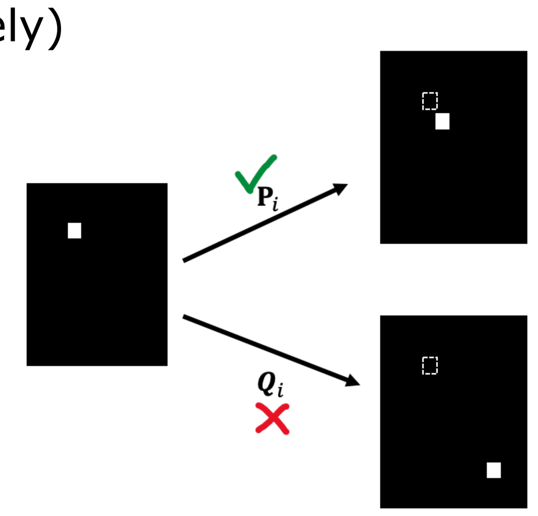


  - A Key component of the proposed method
  - Given two probability vectors $\mathbf{u}$ and $\mathbf{v}$ and a predefined ground cost matrix $\mathbf{C}$
  - Find the optimal coupling $\mathbf{P}$ between $\mathbf{u}$ and $\mathbf{v}$ by solving
    $$\min_{\mathbf{P}\in\Pi(\mathbf{u},\mathbf{v})} \langle \mathbf{C},\mathbf{P}\rangle \coloneqq \sum_{m,n}\mathbf{C}[m,n]\mathbf{P}[m,n]$$
    where $\Pi(\mathbf{u},\mathbf{v}) = \{\mathbf{P}\in\mathbb{R}_+^{M\times N} \mid \mathbf{P1}=\mathbf{u}, \mathbf{P}^T\mathbf{1}=\mathbf{v}\}$ is the set of all joint probability distributions whose marginals are $\mathbf{u}$ and $\mathbf{v}$
  - When $M=N$ and $\mathbf{u}=\mathbf{v}=\frac{1}{N}\mathbf{1}$, the optimal $\mathbf{P}$ is permutation matrix
  - Efficient algorithms available for solving OT problems

Proposed Method

- **Assumptions**
  - Support of $\mathbf{x}$ is known
  - Permutations moving pixels far away from its original location are less likely (i.e., big tracking error is less likely)

- **Optimization formulation**



(1) $$\min_{\mathbf{P}_i\in\mathcal{P},\mathbf{x},\mathbf{x}_i} \sum_i \|\mathbf{y}_i - \mathbf{A}_i\mathbf{x}_i\|_2^2 + \beta R(\mathbf{P}_i) \quad \text{subject to } \mathbf{x}_i = \mathbf{P}_i\mathbf{F}_i\mathbf{x} \ \ \forall i$$

  - $R(\mathbf{P}_i) = \sum_{n,n'} \|\mathbf{r}[n]-\mathbf{r}[n']\|_2^2\, \mathbf{P}_i[n,n']$
  - $\mathbf{r}[n]$: coordinate of n^{th} pixel on 2D grid
  - $\mathcal{P}$: set of all $N\times N$ permutation matrices
  - Hard to solve due to combinatorial search for permutation matrix

- **Relaxation of equality constraint**

(2) $$\min_{\mathbf{P}_i\in\mathcal{P},\mathbf{x},\mathbf{x}_i} \sum_i \|\mathbf{y}_i - \mathbf{A}_i\mathbf{x}_i\|_2^2 + \beta R(\mathbf{P}_i) + \lambda\|\mathbf{x}_i - \mathbf{P}_i\mathbf{F}_i\mathbf{x}\|_2^2$$

- **Connection to OT**

(3) $$\min_{\mathbf{x},\mathbf{x}_i} \sum_i \|\mathbf{y}_i - \mathbf{A}_i\mathbf{x}_i\|_2^2 + \beta \min_{\mathbf{P}_i\in\Pi(\mathbf{u}_i,\mathbf{v}_i)} \langle \mathbf{C}(\mathbf{x}_i,\mathbf{F}_i\mathbf{x}),\ \mathbf{P}_i\rangle$$ **OT Problem**

  - $\Pi(\mathbf{u}_i,\mathbf{v}_i) = \{\mathbf{P}\in\mathbb{R}_+^{N\times N} \mid \mathbf{P1}=\mathbf{u}_i, \mathbf{P}^T\mathbf{1}=\mathbf{v}_i\}$
  - $\mathbf{C}(\mathbf{x}_i,\mathbf{F}_i\mathbf{x})[\mathrm{n},\mathrm{n}'] = \|\mathbf{r}[\mathrm{n}]-\mathbf{r}[\mathrm{n}']\|_2^2 + \gamma(\mathbf{x}_i[\mathrm{n}] - (\mathbf{F}_i\mathbf{x})[\mathrm{n}'])^2$
  - Equivalent to the relaxed problem if $\mathbf{u}_i = \mathbf{v}_i = \frac{1}{\mathrm{N}}\mathbf{1}$
  - Efficient algorithms from OT literature (e.g., Sinkhorn iterations and its variants) can be applied to solve for $\mathbf{P}_i$
  - Other marginals $\mathbf{u}_i, \mathbf{v}_i$ can also be used to further relax the constraint that $\mathbf{P}_i$ is permutation

- **Proposed algorithm**
  - Alternating between estimation of $\mathbf{x}$ and $\mathbf{x}_i$ in (3)
  - Gradient descent for each subproblem
  - Envelope theorem can be used to compute the gradient of minimization function
  - Marginal $\mathbf{v}_i$ is uniform over support of $\mathbf{F}_i\mathbf{x}$ (known)
  - Marginal $\mathbf{u}_i$ is uniform over estimated support of $\mathbf{x}_i$

- **Extensions**
  - Other similarity measure between $\mathbf{r}[n]$ and $\mathbf{r}[n']$, $\mathbf{x}_i$ and $\mathbf{P}_i\mathbf{F}_i\mathbf{x}$, can be used depending on specific problems
  - Regularization for $\mathbf{x}$ and $\mathbf{x}_i$ can be easily incorporated

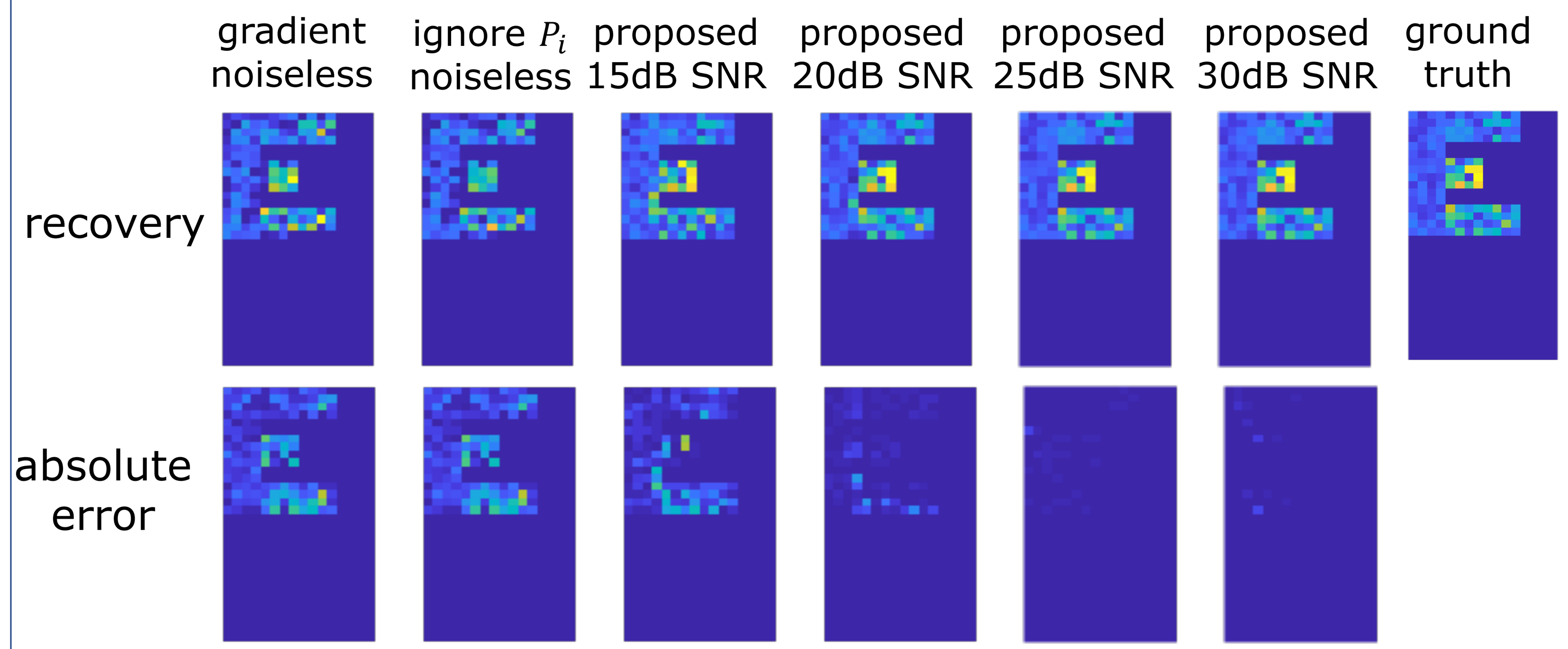
- **Baseline methods for comparison**
  - Since tracking error is small, a naïve method is to ignore $\mathbf{P}_i$

(ignore $\mathbf{P}_i$) $$\min_{x} \sum_i \|\mathbf{y}_i - \mathbf{A}_i\mathbf{F}_i\mathbf{x}\|_2^2$$

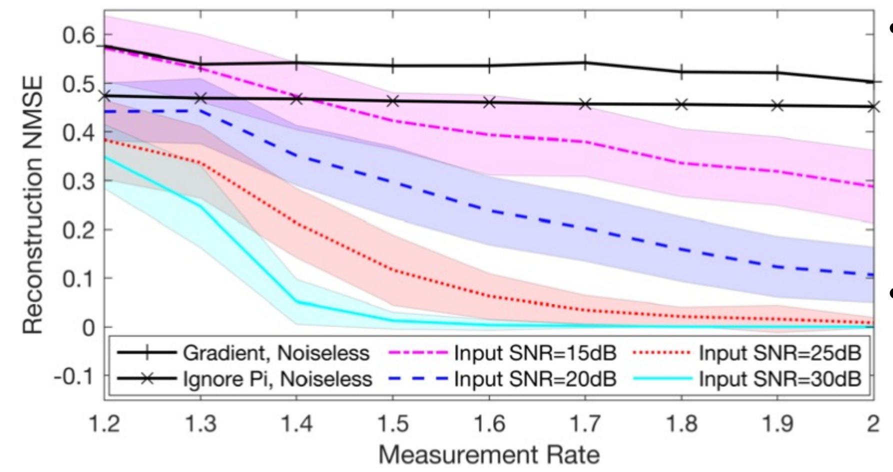
  - Alternatively, a more straightforward relaxation of the permutation constraint is to replace it with a differentiable penalty

(gradient) $$\min_{\mathbf{P}_i\in[0,1]^{N\times N},\mathbf{x},\mathbf{x}_i} \sum_i \|\mathbf{y}_i - \mathbf{A}_i\mathbf{x}_i\|_2^2 + \beta R(\mathbf{P}_i) + \lambda\|\mathbf{x}_i - \mathbf{P}_i\mathbf{F}_i\mathbf{x}\|_2^2 + \|\mathbf{P}_i\mathbf{1}-\mathbf{1}\|_2^2 + \|\mathbf{P}_i^T\mathbf{1}-\mathbf{1}\|_2^2$$

Simulation Results

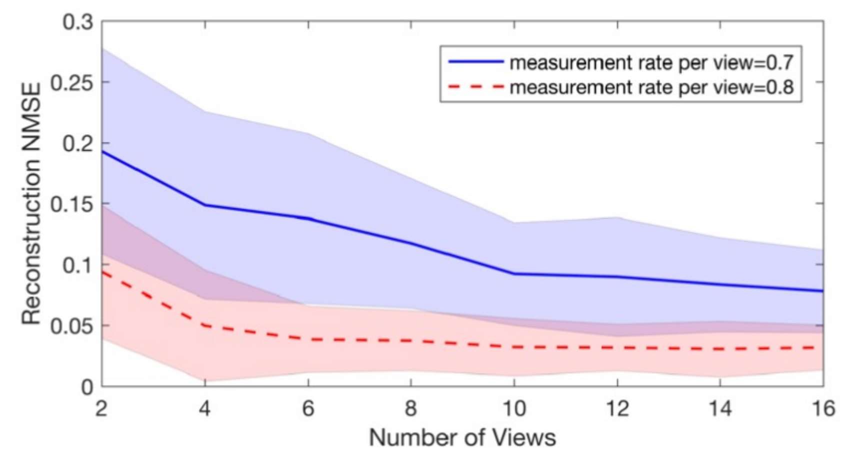


- **Simulation setup 1**
  - Ground truth pixel value is i.i.d. uniform within each bar
  - Sensing operator $\mathbf{A}$ has i.i.d. Gaussian entries
  - Number of views is 2
  - Measurement rate is defined as the ratio between total number of measurements (summing over all views) and number of pixels in $\mathbf{x}$
  - Increasing number of measurements also increases size of unknown permutation



  - With careful estimation of the permutation, proposed method improves reconstruction quality with increased measurements
  - Example visual results at measurement rate 1.6 is shown above

- **Simulation setup 2**
  - Input SNR is 20dB
  - Fixing number of measurements per view, reconstruction quality improves with increased number of views



  - This can be important for some applications, as the number of measurements per view can be limited by hardware

Summary

- Signal estimation with unknown permutations is challenging
- In practice, some permutations are more likely than others
- We introduced regularization to promote certain type of permutations
- Further relaxation allowed us making connection to optimal transport, which provides tractable algorithms
- Other regularization for permutations (depending on specific problem) may be translated to choosing certain OT ground cost