Feature Matching via Sparse Relaxation Models

江 波

jiangbo@ahu.edu.cn

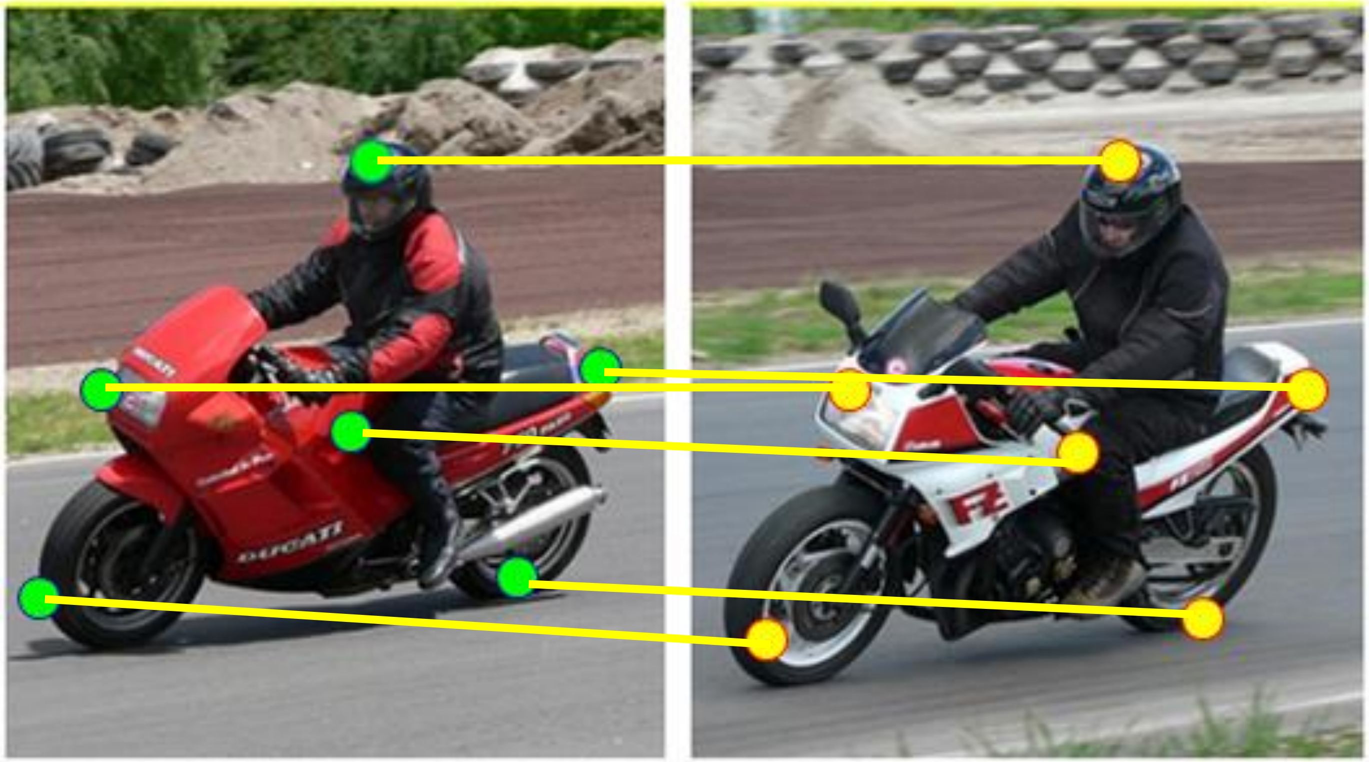
安徽大学 计算机科学与技术学院

2018-8-8

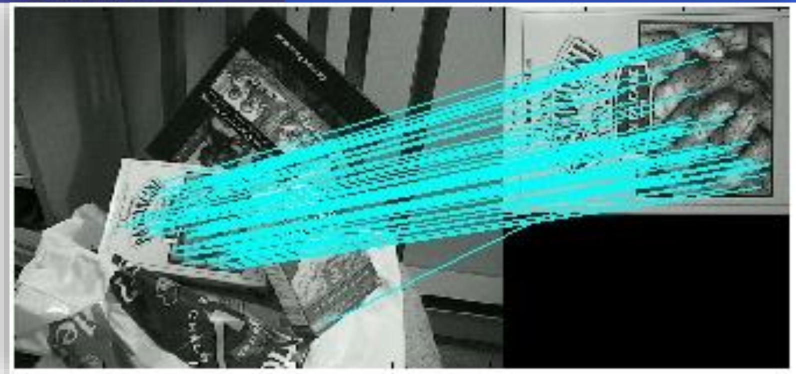
Content

Introduction

Introduction

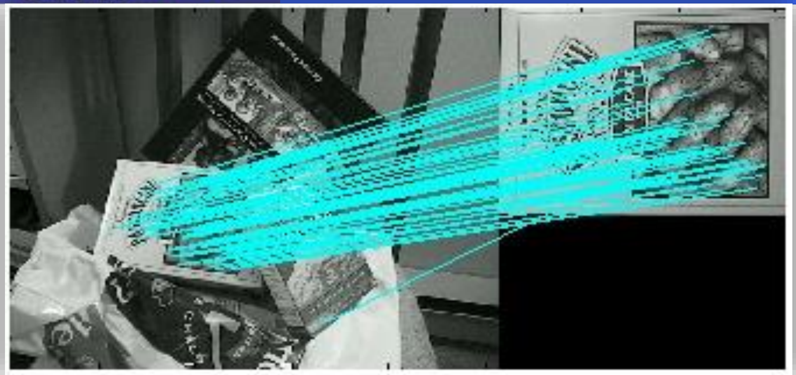


Introduction



Object detection

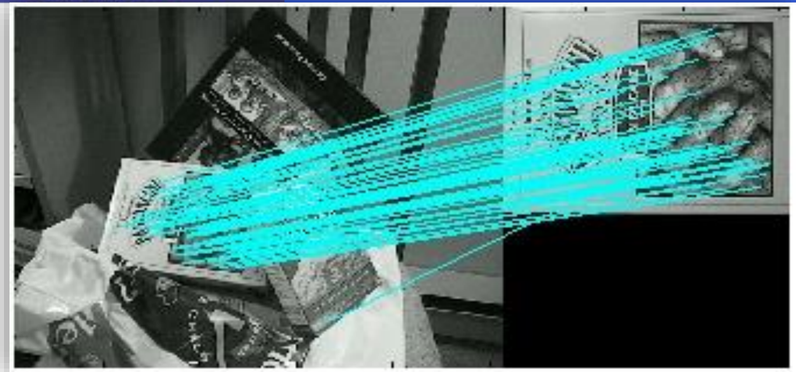
Introduction



Object detection

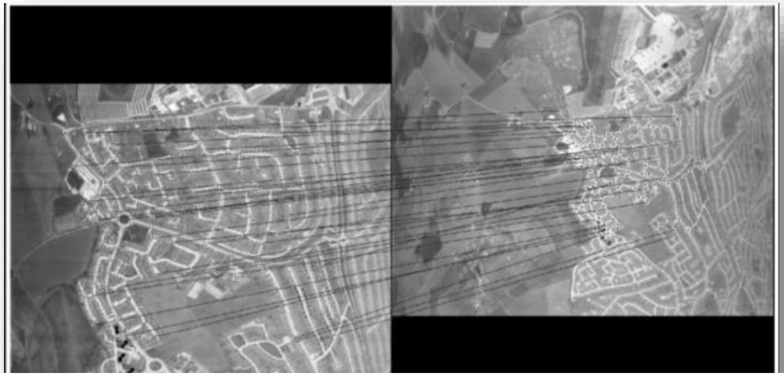
Person ReID, Zhou et al. AAAI 2018

Introduction



Object detection

Person ReID, Zhou et al. AAAI 2018



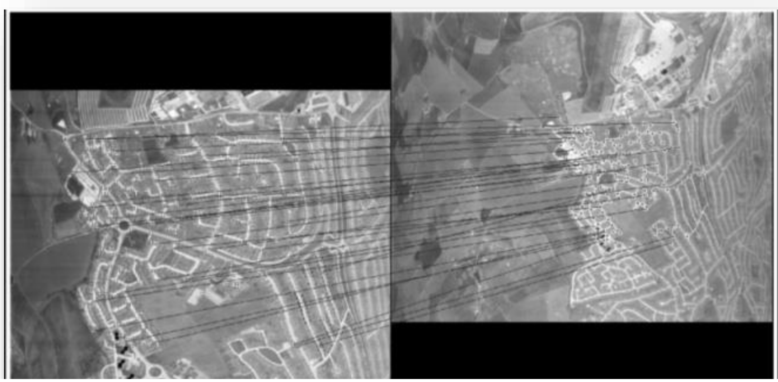
Luo et al. PAMI 2001

Introduction

Object detection

Object tracking, Nebehay et al. CVPR15

Person ReID, Zhou et al. AAAI 2018



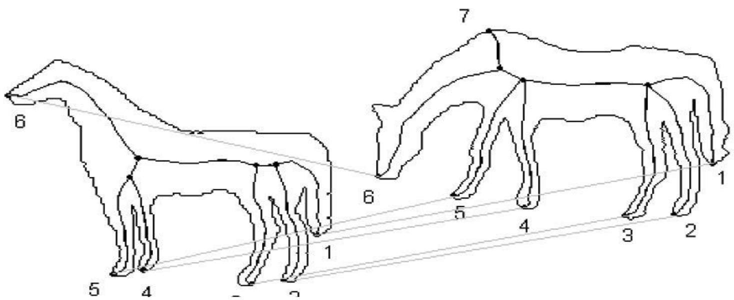
Luo et al. PAMI 2001

Introduction

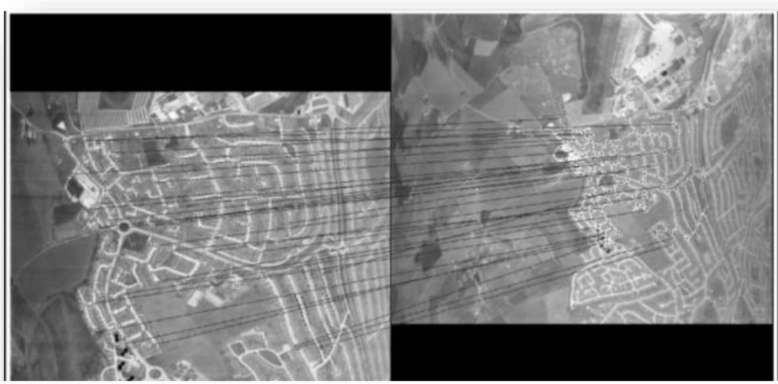
Object detection

Object tracking, Nebehay et al. CVPR15

Person ReID, Zhou et al. AAAI 2018

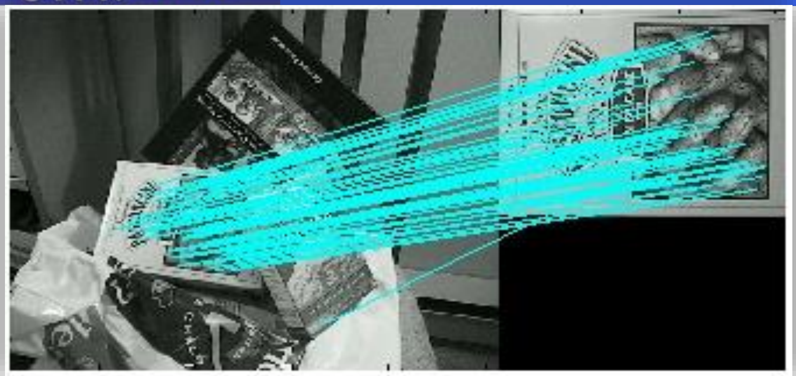


Shape matching, Bai et al. PAMI2008



Luo et al. PAMI 2001

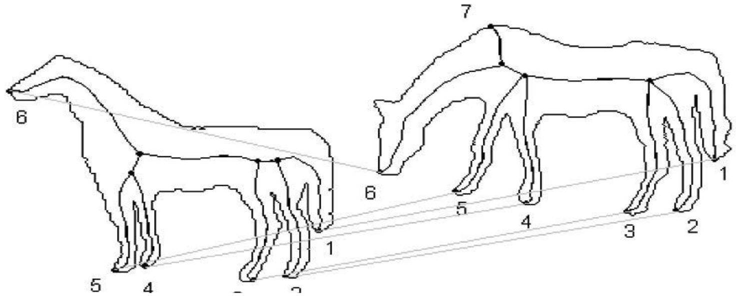
Introduction



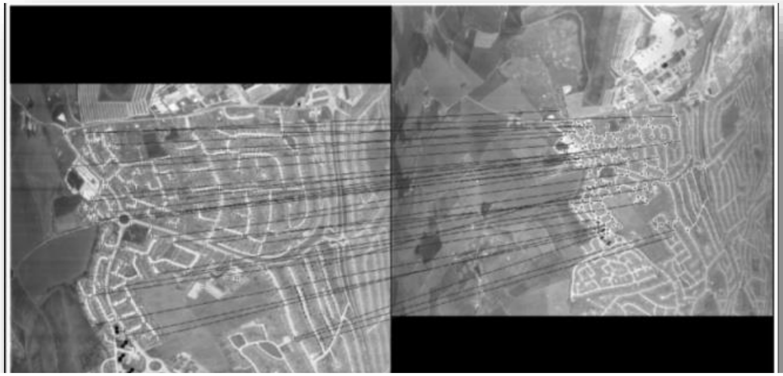
Object detection

Object tracking, Nebehay et al. CVPR15

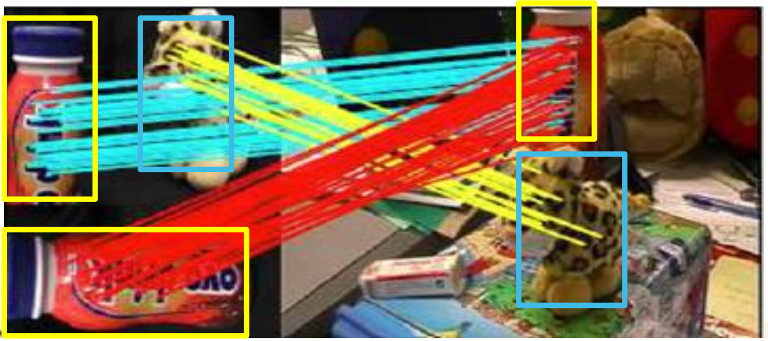
Person ReID, Zhou et al. AAAI 2018



Shape matching, Bai et al. PAMI2008

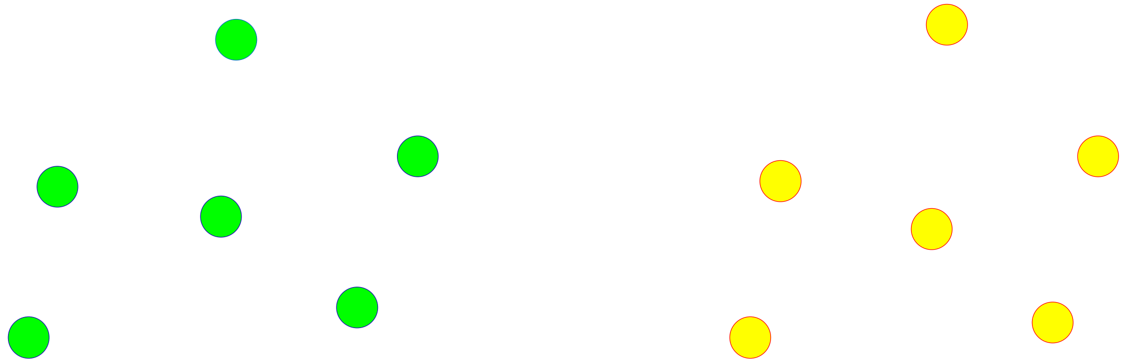


Luo et al. PAMI 2001

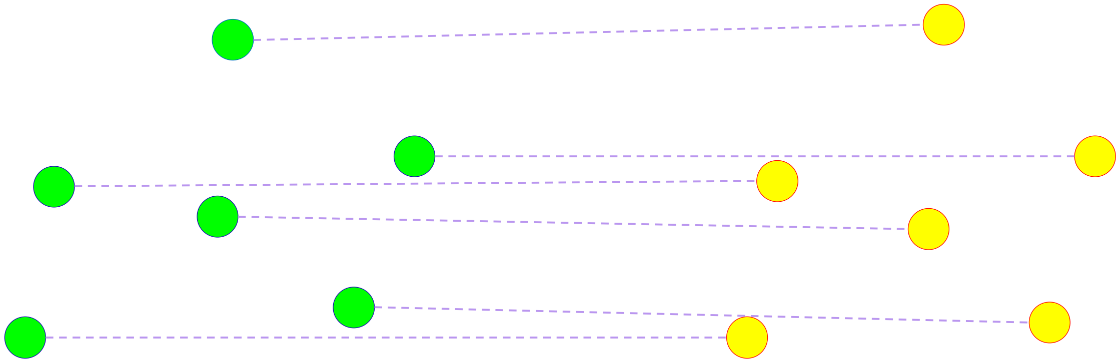


Common Visual Pattern Discovery

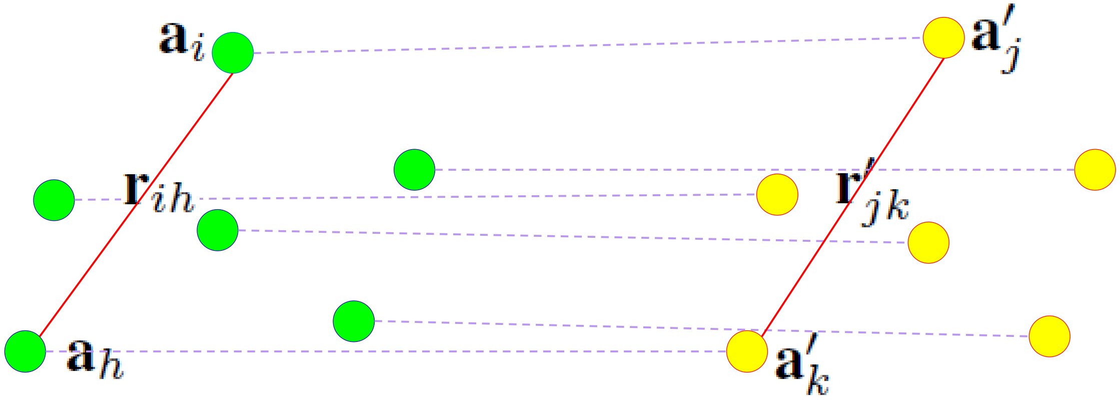
Problem Formulation



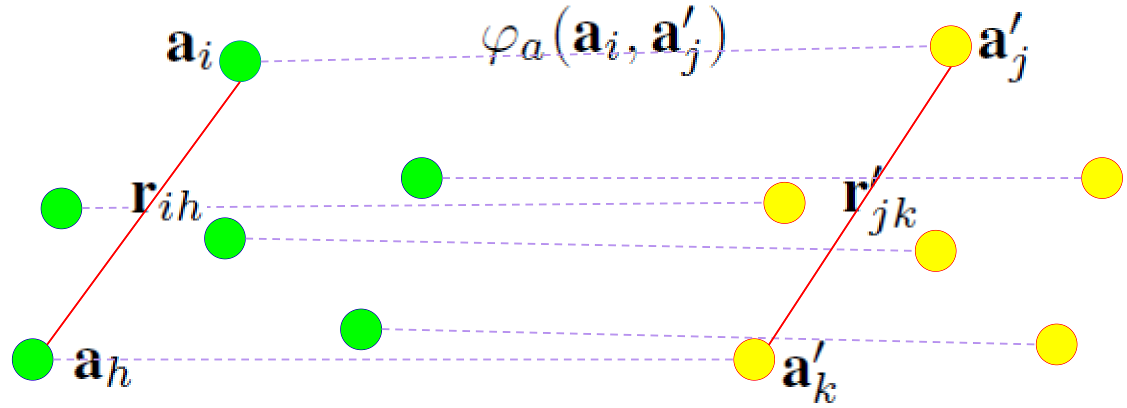
Problem Formulation



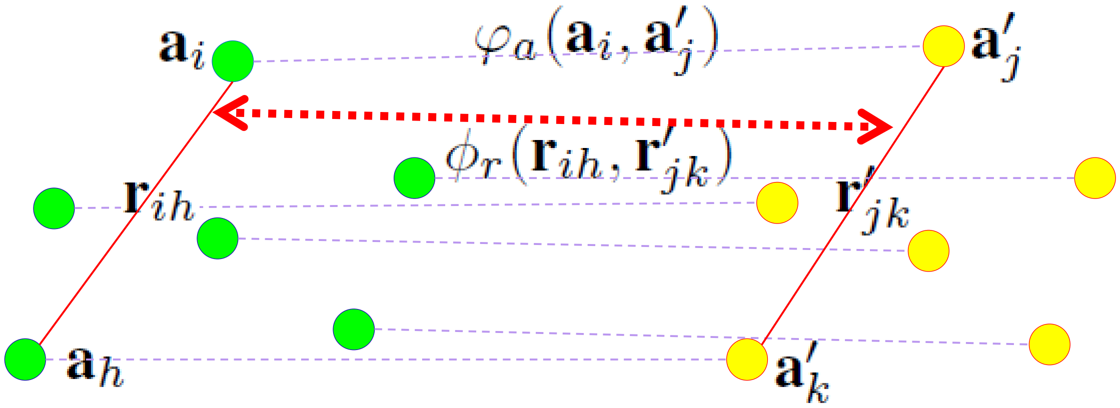
Problem Formulation



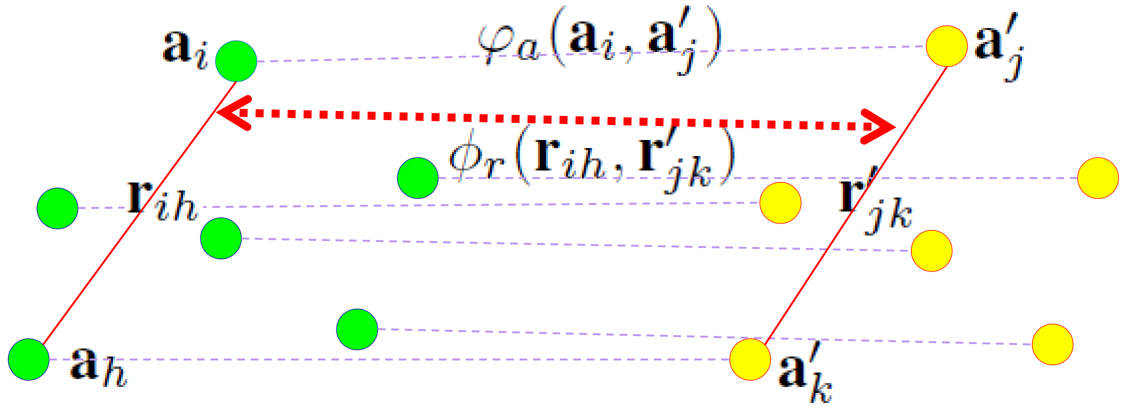
Problem Formulation



Problem Formulation



Problem Formulation

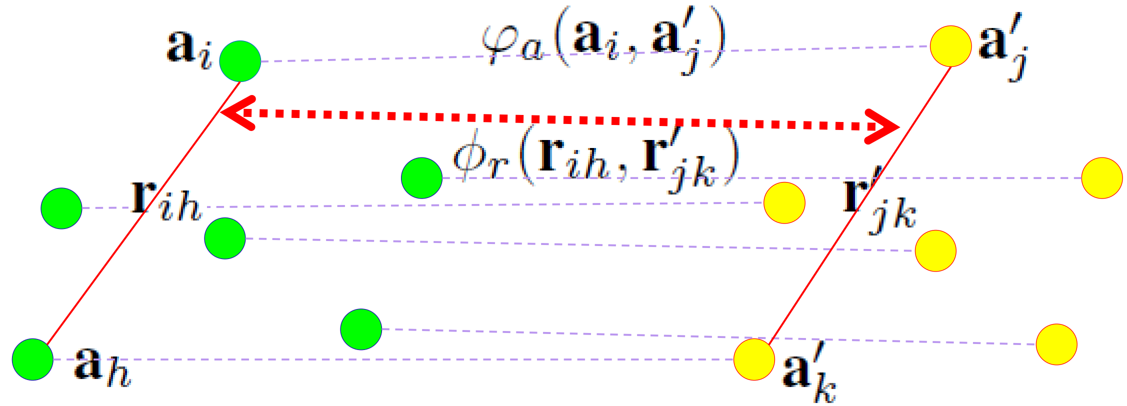


$$\mathbf{X}_{ij} \in \{0, 1\}$$

$$\mathbf{W}_{ij,ij} = \varphi_a(\mathbf{a}_i, \mathbf{a}'_j)$$

$$\mathbf{W}_{ij,hk} = \phi_r(\mathbf{r}_{ih}, \mathbf{r}'_{jk})$$

Problem Formulation



$$\mathbf{X}_{ij} \in \{0, 1\}$$

$$\mathbf{W}_{ij,ij} = \varphi_a(\mathbf{a}_i, \mathbf{a}'_j)$$

$$\mathbf{W}_{ij,hk} = \phi_r(\mathbf{r}_{ih}, \mathbf{r}'_{jk})$$

Integer Quadratic Programming (IQP) problem

$$\max_{\mathbf{X}} \quad \sum\nolimits_{ij} \sum\nolimits_{hk} \mathbf{W}_{ij,hk}\mathbf{X}_{ij}\mathbf{X}_{hk} = \mathrm{vec}(\mathbf{X})^{\mathrm{T}}\mathbf{W}\,\mathrm{vec}(\mathbf{X})$$

$$s.t. \quad \mathbf{X1} = \mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1} \leq \mathbf{1}, \mathbf{X}_{ij} \in \{0, 1\}$$

- NP-hard problem
- Approximate solution

Related works

Continuous Relaxation

$$\max_{\mathbf{X}} \ \mathrm{vec}(\mathbf{X})^{\mathrm{T}} \mathbf{W} \, \mathrm{vec}(\mathbf{X}) \quad s.t. \ \ \mathbf{X1} = \mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1} \leq \mathbf{1}, \mathbf{X}_{ij} \in \{0,1\}$$

Related works

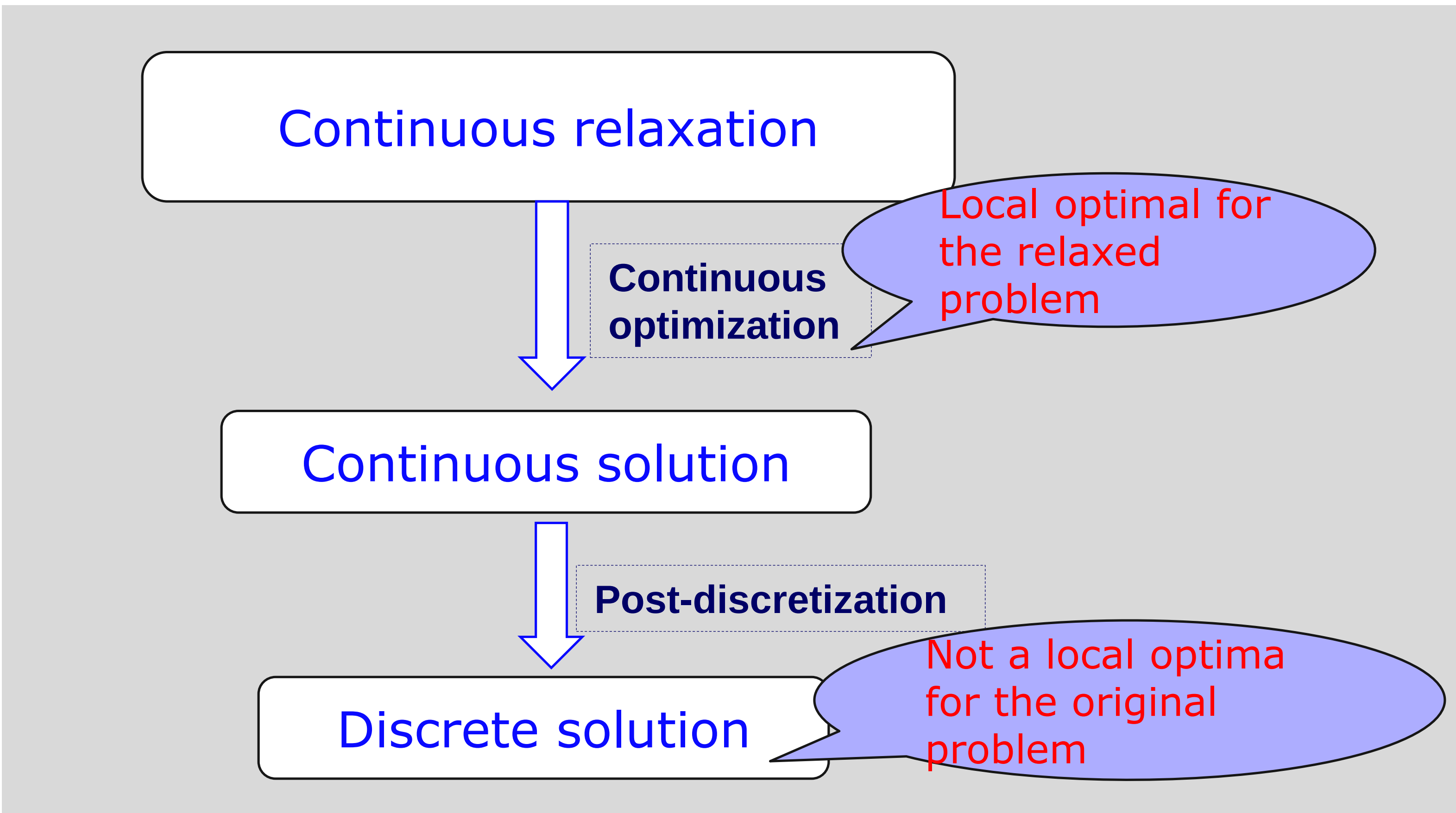
Continuous Relaxation

$$\max_{\mathbf{X}} \ \mathrm{vec}(\mathbf{X})^{\mathrm{T}} \mathbf{W} \, \mathrm{vec}(\mathbf{X}) \quad s.t. \ \mathbf{X}\mathbf{1} = \mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1} \leq \mathbf{1}, \mathbf{X}_{ij} \in \{0, 1\}$$

Related works

Continuous Relaxation

$$\max_{\mathbf{X}} \ \mathrm{vec}(\mathbf{X})^{\mathrm{T}}\mathbf{W}\,\mathrm{vec}(\mathbf{X}) \quad s.t. \ \ \mathbf{X}\mathbf{1}=\mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1}\le\mathbf{1}, \mathbf{X}_{ij}\in\{0,1\}$$



Related works

Continuous Relaxation

Spectral matching-ICCV 2005

$$\max \ \operatorname{vec}(\mathbf{X})^{\mathrm{T}} \mathbf{W} \operatorname{vec}(\mathbf{X}) \quad \text{s.t.} \quad \|\mathbf{X}\|_2 = 1$$

Spectral matching with affine constraint-NIPS 2006

$$\max_{\mathbf{X}} \ \operatorname{vec}(\mathbf{X})^{\mathrm{T}} \mathbf{M} \operatorname{vec}(\mathbf{X}) \quad s.t. \ \mathbf{A} \operatorname{vec}(\mathbf{X}) = \mathbf{1}, \ \|\mathbf{X}\|_2^2 = 1$$

Doubly stochastic relaxation

$$\max_{\mathbf{X}} \ \operatorname{vec}(\mathbf{X})^{\mathrm{T}} \mathbf{W} \operatorname{vec}(\mathbf{X})$$

$$s.t. \ \ \mathbf{X}\mathbf{1} = \mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1} \leq \mathbf{1}, \mathbf{X}_{ij} \geq 0$$

- GA-PAMI 1996
- POCS-PAMI 2004
- RRWM-ICCV 2010
- SCGA-ECCV 2012
- Probabilistic Models

……

Related works

Discrete Methods

$$\max_{\mathbf{X}} \ \mathrm{vec}(\mathbf{X})^{\mathrm{T}} \mathbf{W} \, \mathrm{vec}(\mathbf{X}) \quad s.t. \quad \mathbf{X1} = \mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1} \leq \mathbf{1}, \mathbf{X}_{ij} \in \{0, 1\}$$

Integer Projected Fixed Point (IPFP) -NIPS 2009

Factorized Graph Matching (FGM) - CVPR 2012

$$\max \ J_{\alpha}(\mathbf{X}) = (1-\alpha) J_{vex}(\mathbf{X}) + \alpha J_{cav}(\mathbf{X})$$

Discrete Tabu Search –ICCV 2015

Hungarian-BP-CVPR 2016

Related works

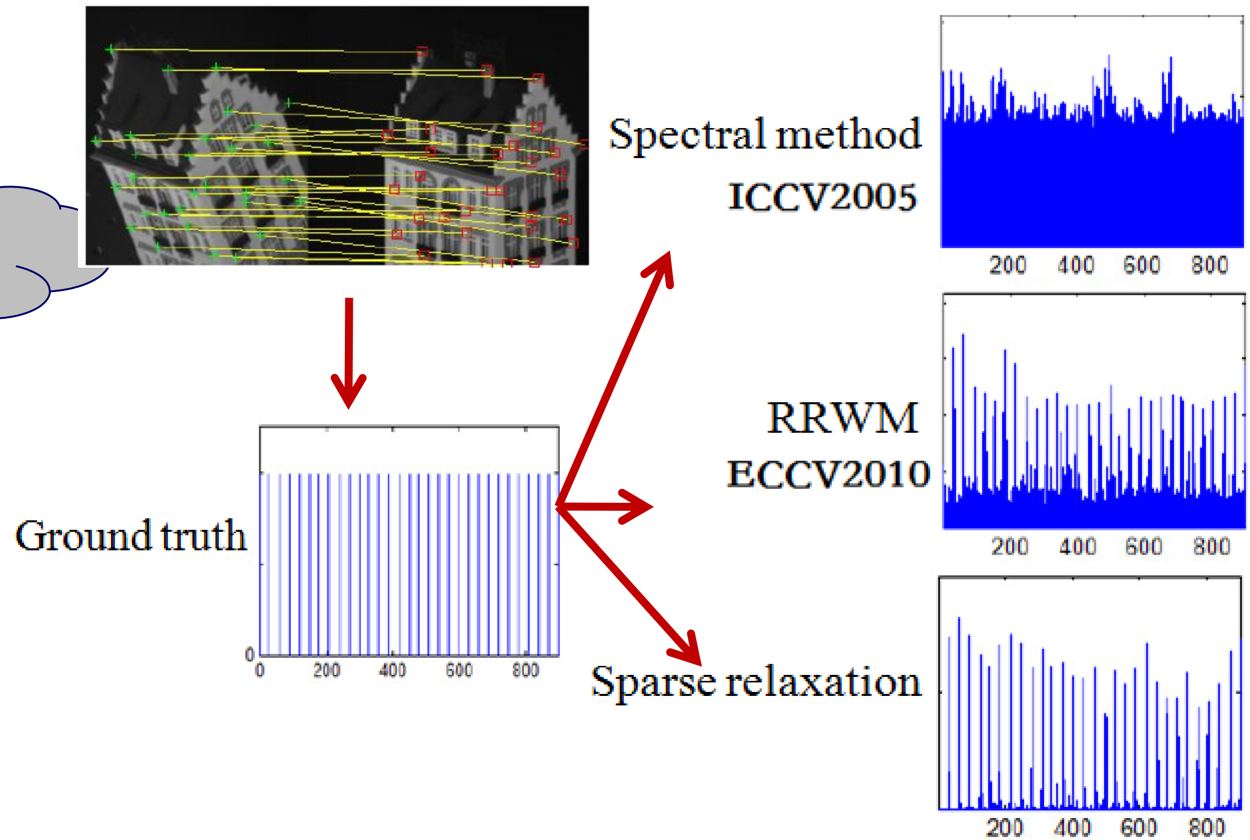
Sparse Relaxation

Discrete constraint

relaxation

Nonnegative sparse

Nonnegative sparse model



Related works

Sparse Relaxation

Spectral matching (SM)-ICCV 2005

$$\max \quad \text{vec}(\mathbf{X})^T \mathbf{W} \, \text{vec}(\mathbf{X}) \quad \text{s.t.} \quad \|\mathbf{X}\|_2 = 1$$

Game-theoretic matching (GameM)-ICCV 2009, IJCV 2011

$$\max \quad \text{vec}(\mathbf{X})^T \mathbf{W} \, \text{vec}(\mathbf{X}) \quad \text{s.t.} \quad \|\mathbf{X}\|_1 = 1, \mathbf{X} \geq 0$$

Elastic net matching (EnetM)-ICCV 2013

$$\max \quad \text{vec}(\mathbf{X})^T \mathbf{W} \, \text{vec}(\mathbf{X})$$

$$\text{s.t.} \quad (1-\alpha)\|\mathbf{X}\|_1 + \alpha\|\mathbf{X}\|_2^2 = 1, \mathbf{X} \geq 0$$

Sparse nonnegative matrix factorization (SNMF)-PR 2014

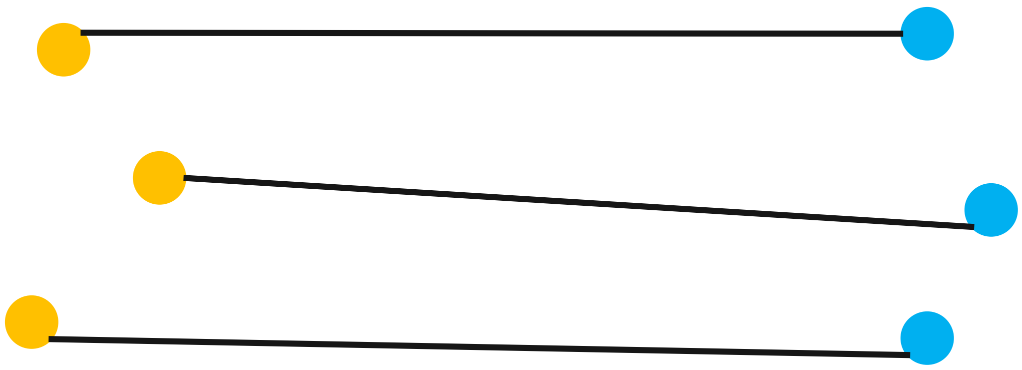
$$\max \quad \text{vec}(\mathbf{X})^T \mathbf{W} \, \text{vec}(\mathbf{X}) \quad \text{s.t.} \quad \|\mathbf{X}\|_p = 1, \mathbf{X} \geq 0$$

Local sparse model for matching

Motivation

Local sparse model for matching

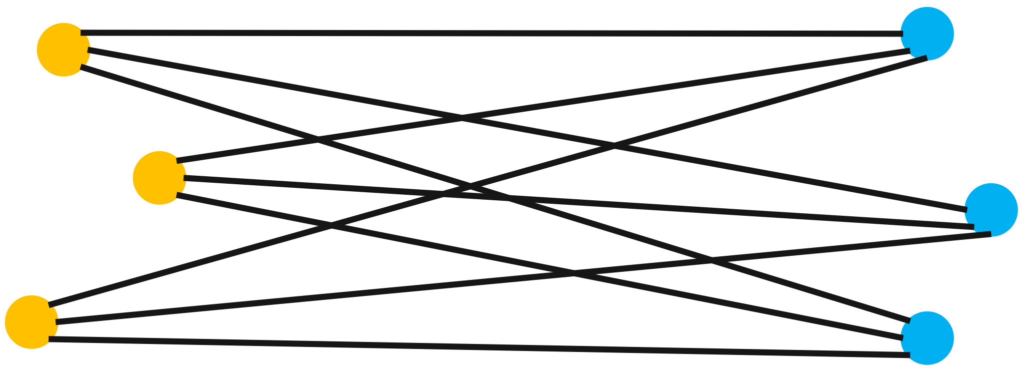
Motivation



$$\mathbf{X} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Local sparse model for matching

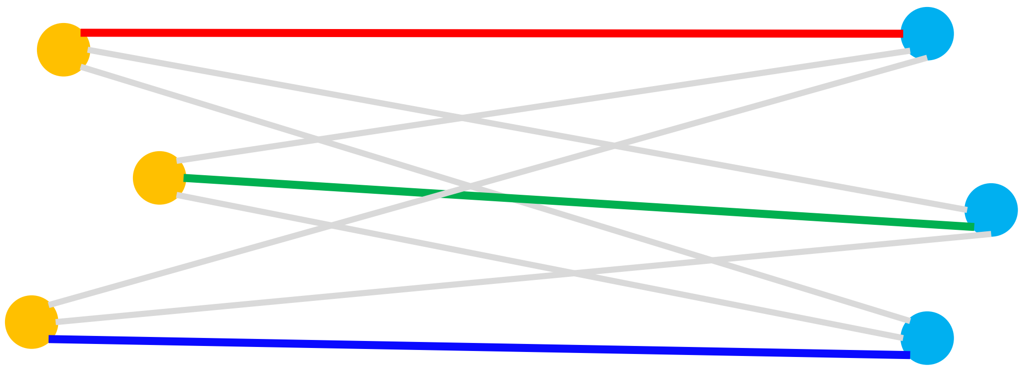
Motivation



$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Local sparse model for matching

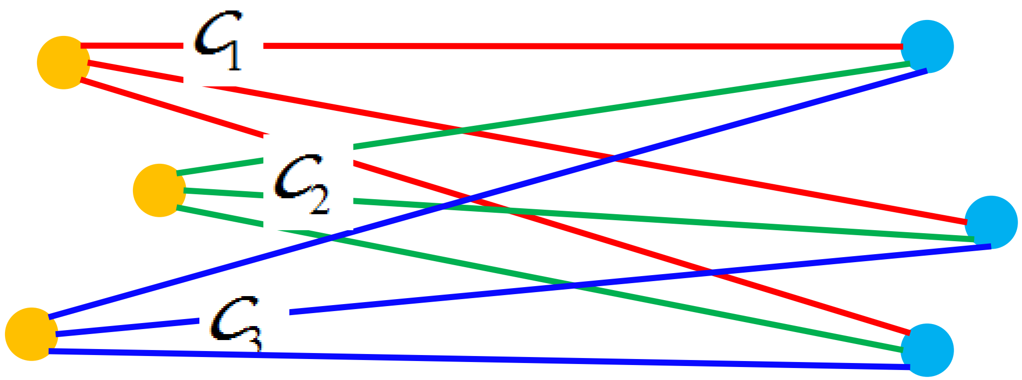
Motivation



$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Local sparse model for matching

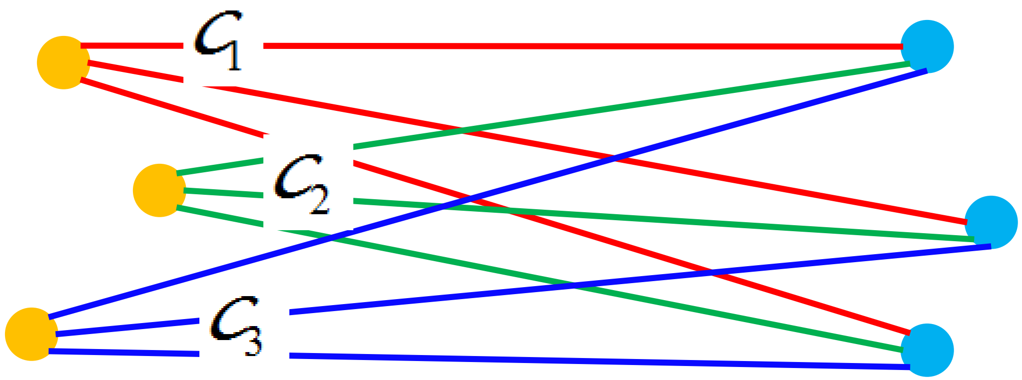
Motivation



$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{matrix} C_1 \\ C_2 \\ C_3 \end{matrix}$$

Local sparse model for matching

Motivation



$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{matrix} C_1 \\ C_2 \\ C_3 \end{matrix}$$

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} & \mathbf{X}_{13} \\ \mathbf{X}_{21} & \mathbf{X}_{22} & \mathbf{X}_{23} \\ \mathbf{X}_{31} & \mathbf{X}_{32} & \mathbf{X}_{33} \end{pmatrix}$$

➔ **Local Sparse Model**

Observations

- Each row of solution matrix **X** is sparse
- There is no zero row in solution matrix **X**

Local sparse model for matching

Local sparse matching

$$\max_{\mathbf{X}} \ \mathrm{vec}(\mathbf{X})^{\mathrm{T}} \mathbf{M} \,\mathrm{vec}(\mathbf{X}) \quad s.t. \ \ \|\mathbf{X}\|_{1,2} = 1, \mathbf{X}_{ij} \geq 0$$

L12 norm

$$\|\mathbf{X}\|_{1,2} = \left[\sum_i \left(\sum_j |\mathbf{X}_{ij}| \right)^2 \right]^{1/2}$$

Local sparse

- L1 norm on each row encourages sparsity
- L2 norm on rows encourages that there is no zero row

Bo Jiang, et al., A Local sparse model for matching problem, **AAAI 2015**

Local sparse model for matching

Algorithm

$$\mathbf{X}_{ij}^{(t+1)} = \mathbf{X}_{ij}^{(t)} \left(\frac{\mathbf{K}_{ij}^{(t)}}{\lambda \sum_i \mathbf{X}_{ij}^{(t)}} \right), \lambda = \mathrm{vec}(\mathbf{X}^{(t)})^T \mathbf{M}\,\mathrm{vec}(\mathbf{X}^{(t)})$$

$\mathbf{K}^{(t)} \in \mathbb{R}^{m\times n}$ is the matrix form of $[\mathbf{W}^{(t)}\,\mathrm{vec}\,(\mathbf{X}^{(t)})]$

Properties

Theorem 1 *Update rule satisfies the first-order Karush-Kuhn-Tucker (KKT) optimality condition.*

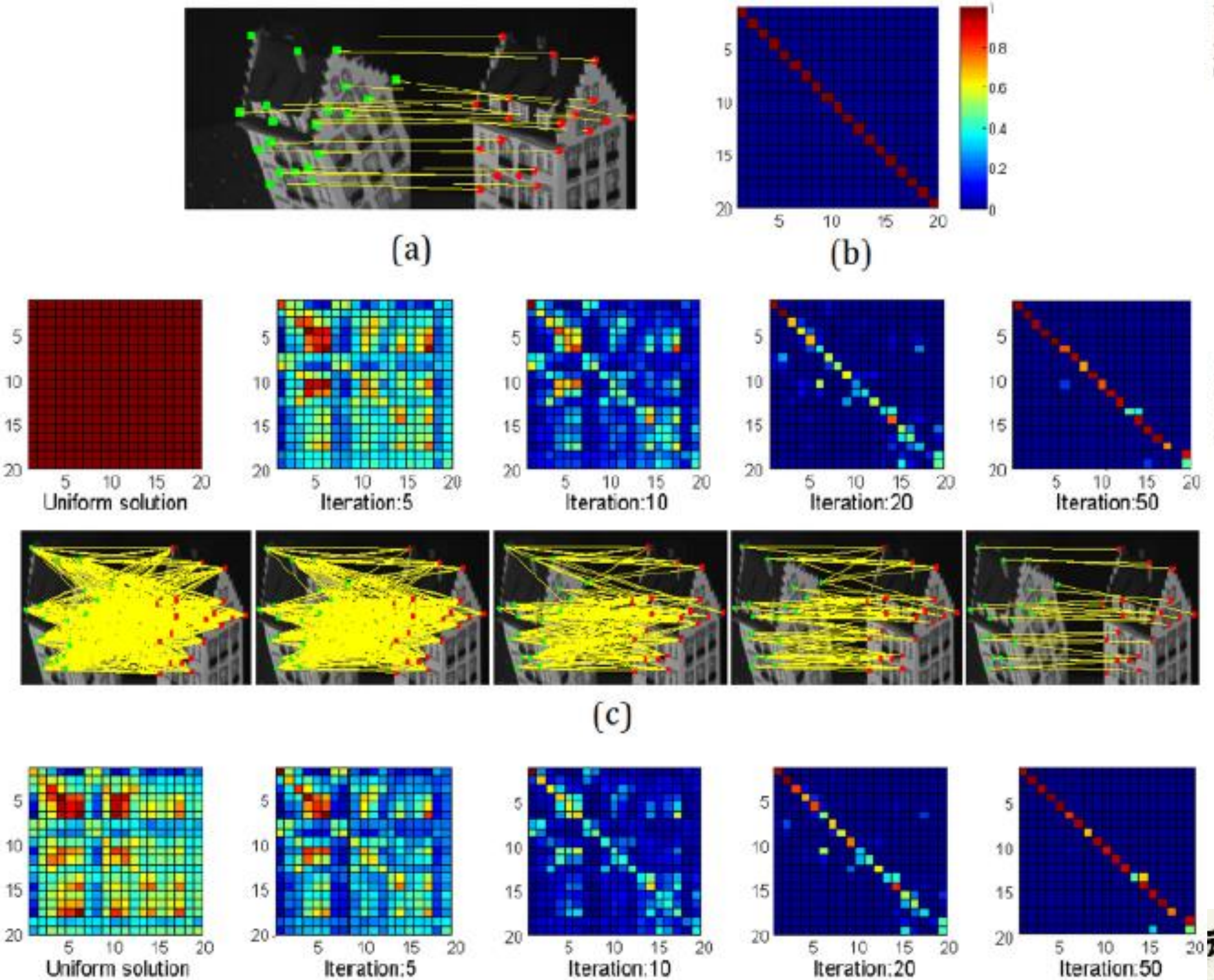
Theorem 2 *Under the update rule, the Lagrangian function $\mathcal{L}(X)$ is monotonically increasing.*

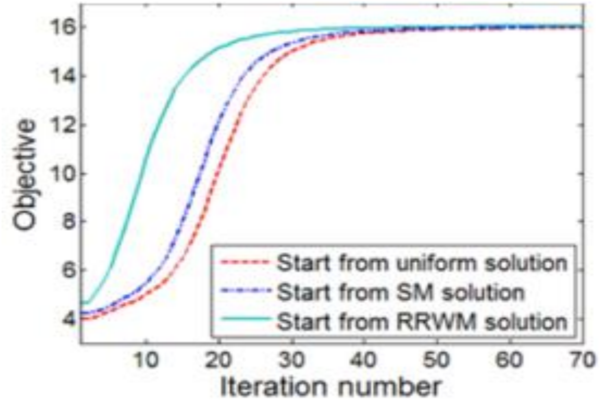
$$\mathcal{L}(\mathbf{X}) = \sum_{ij} \sum_{kl} \mathbf{W}_{ij,kl} \mathbf{X}_{ij} \mathbf{X}_{kl} - \lambda(\mathbf{e}^{\mathrm{T}} \mathbf{X}^{\mathrm{T}} \mathbf{X} \mathbf{e} - 1)$$

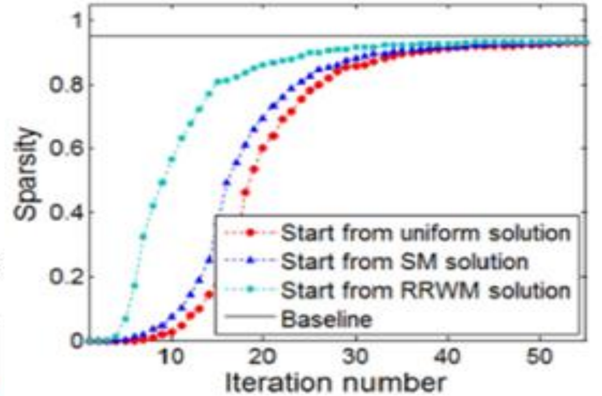
Bo Jiang, et al., A Local sparse model for matching problem, **AAAI 2015**

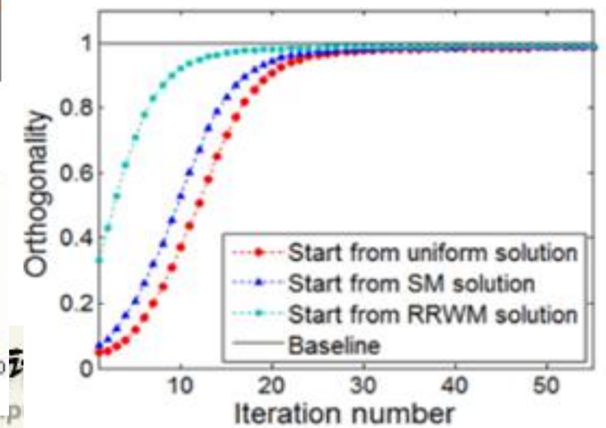
Local sparse model for matching

Illustration

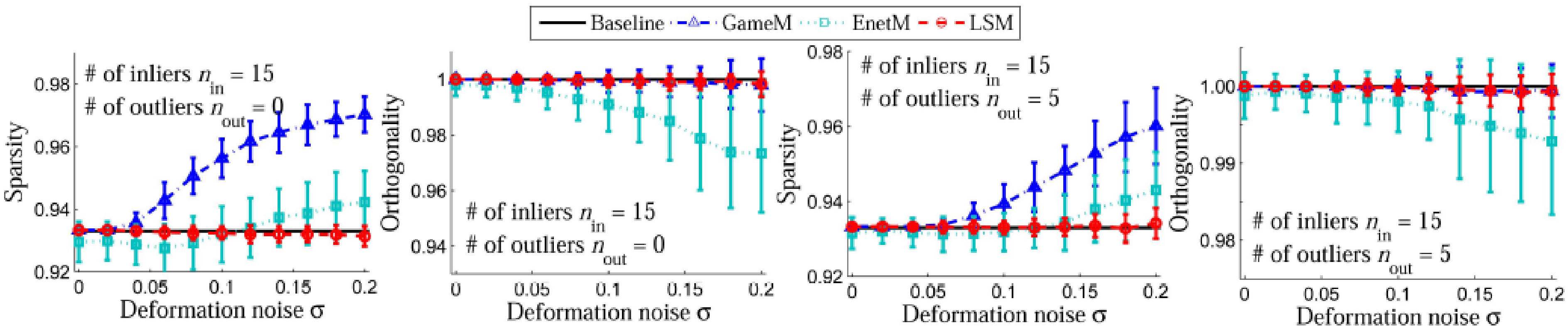


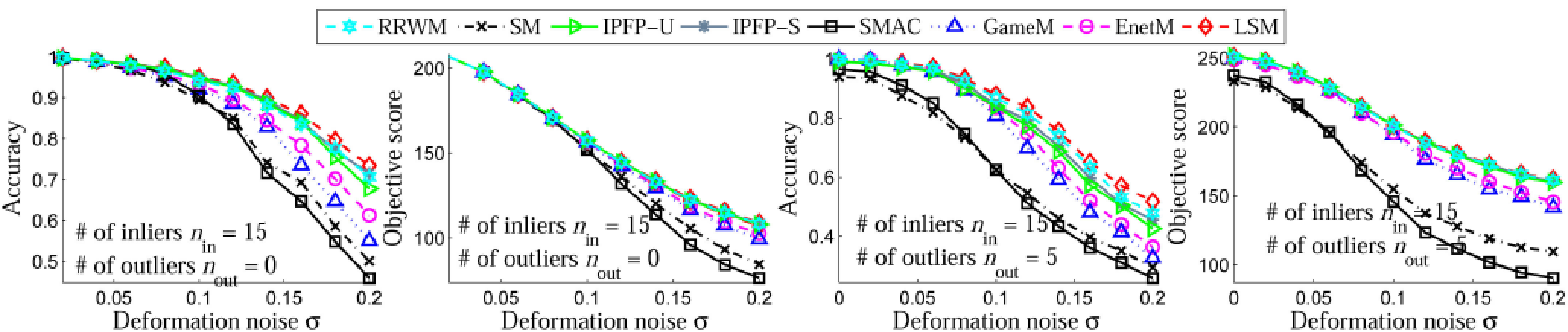






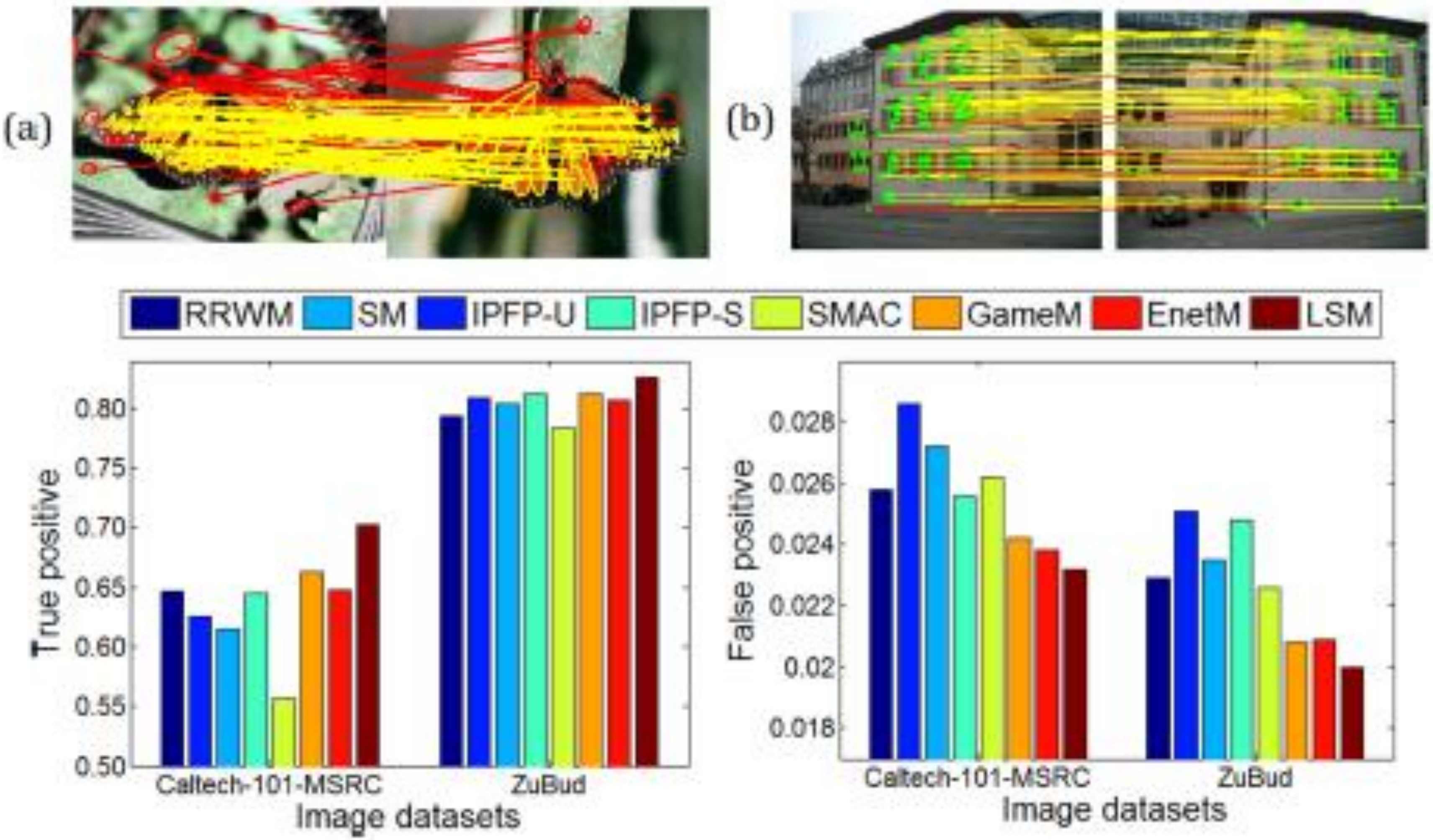
Local sparse model for matching





Bo Jiang, et al., A Local sparse model for matching problem, **AAAI 2015**

Local sparse model for matching



Bo Jiang, et al., A Local sparse model for matching problem, **AAAI 2015**

Binary constraint preserving matching

Motivation

Integer Quadratic Programming (IQP) problem

$$\max_{\mathbf{X}} \quad \sum_{ij} \sum_{hk} \mathbf{W}_{ij,hk} \mathbf{X}_{ij} \mathbf{X}_{hk} = \mathrm{vec}(\mathbf{X})^{\mathrm{T}} \mathbf{W} \, \mathrm{vec}(\mathbf{X})$$

$$s.t. \quad \mathbf{X1} = \mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1} \leq \mathbf{1}, \mathbf{X}_{ij} \in \{0, 1\}$$

$$\max_{\mathbf{X}} \quad \sum_{ij} \sum_{hk} \mathbf{W}_{ij,hk} \mathbf{X}_{ij} \mathbf{X}_{hk} = \mathrm{vec}(\mathbf{X})^{\mathrm{T}} \mathbf{W} \, \mathrm{vec}(\mathbf{X})$$

$$s.t. \quad \mathbf{X1} = \mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1} \leq \mathbf{1}, \|\mathbf{X}\|_2^2 = n$$

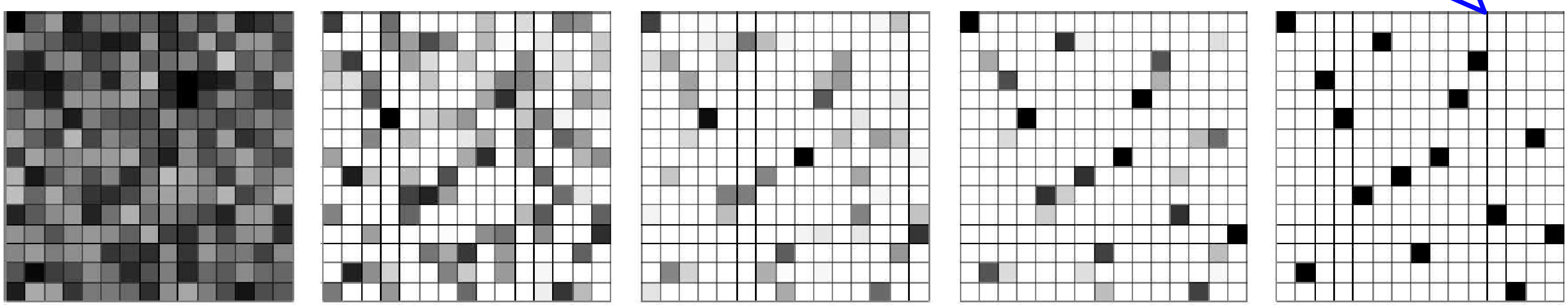
Binary constraint preserving matching

BPGM formulation

$$\max_{\mathbf{x}} \ \mathbf{x}^{\mathrm{T}}\mathbf{W}\mathbf{x} \quad s.t. \ \mathbf{A}\mathbf{x} = \mathbf{1}, \mathbf{x}_i \geq 0, \|\mathbf{x}\|_2^2 = \gamma.$$

$$\|\mathbf{h}_0\|_2^2 \leq \gamma \leq n, \text{ where } \mathbf{h}_0 = \mathbf{A}^{\mathrm{T}}(\mathbf{A}\,\mathbf{A}^{\mathrm{T}})^{-1}\,\mathbf{1}$$

Binary constraint preserving



$\gamma = 1$ $\gamma = 3$ $\gamma = 7$ $\gamma = 11$ $\gamma = 15$

- As ϒ becomes larger, the more closely **x** approximates to discrete
- It provides a series of relaxation models

Bo Jiang, et al., Binary constraint preserving graph matching, **CVPR 2017**

Binary constraint preserving matching

Theoretical analysis

$$\max_{\mathbf{x}} \ \mathbf{x}^{\mathrm{T}}\mathbf{W}\mathbf{x} \quad s.t. \ \mathbf{A}\mathbf{x} = \mathbf{1}, \mathbf{x}_i \geq 0, \|\mathbf{x}\|_2^2 = \gamma;$$

Property 1. When $\Upsilon = n$, where n is the number of features, BPGM model is equivalent to original matching problem

$$\max_{\mathbf{x}} \ \mathbf{x}^{\mathrm{T}}\mathbf{W}\mathbf{x} \quad s.t. \ \mathbf{A}\mathbf{x} = \mathbf{1}, \ \mathbf{x}_i \in \{0, 1\} \qquad (1)$$

Property 2. When $\Upsilon = ||\mathbf{x}^*||$, where $\mathbf{x}^*$ is the optimal solution of problem (2), BPGM model is equivalent to the matching problem (2)

$$\max_{\mathbf{x}} \ \mathbf{x}^{\mathrm{T}}\mathbf{W}\mathbf{x} \quad s.t. \ \mathbf{A}\mathbf{x} = \mathbf{1}, \ \mathbf{x}_i \geq 0. \qquad (2)$$

Balanced model between (1) and (2)

Bo Jiang, et al., Binary constraint preserving graph matching, **CVPR 2017**

CVPR AHU

Binary constraint preserving matching

Theoretical analysis

$$\max_{\mathbf{x}} \ \mathbf{x}^{\mathrm{T}}\mathbf{W}\mathbf{x} \quad s.t. \ \mathbf{A}\mathbf{x} = \mathbf{1}, \mathbf{x}_i \geq 0, \|\mathbf{x}\|_2^2 = \gamma.$$

Lemma 3. There exists a parameter Υ_0 such that BPGM with $\Upsilon = \Upsilon_0$ has a global optimal solution

Path-following strategy

$$\max_{\mathbf{x}} \ \mathbf{x}^{\mathrm{T}}\mathbf{W}\mathbf{x} \quad s.t. \ \mathbf{A}\mathbf{x} = \mathbf{1}, \mathbf{x}_i \geq 0, \|\mathbf{x}\|_2^2 = \gamma_k$$

Starting from global optimal solution and aims to obtain the discrete solution

Bo Jiang, et al., Binary constraint preserving graph matching, **CVPR 2017**

Binary constraint preserving matching

Algorithm

$$\mathbf{x}^{k+1} = \arg\min_{\mathbf{v}} \|\mathbf{v} - \widetilde{\mathbf{W}}\mathbf{x}^k\|_2^2$$

$$s.t.\ \ \mathbf{A}\mathbf{v} = \mathbf{1},\ \|\mathbf{v}\|_2^2 = \gamma, \mathbf{v}_i \geq 0$$

$$\widetilde{\mathbf{W}} = \mathbf{W} + \widetilde{\lambda}_m \mathbf{I}$$

$$\mathbf{u} = \widetilde{\mathbf{W}}\mathbf{x}^k$$

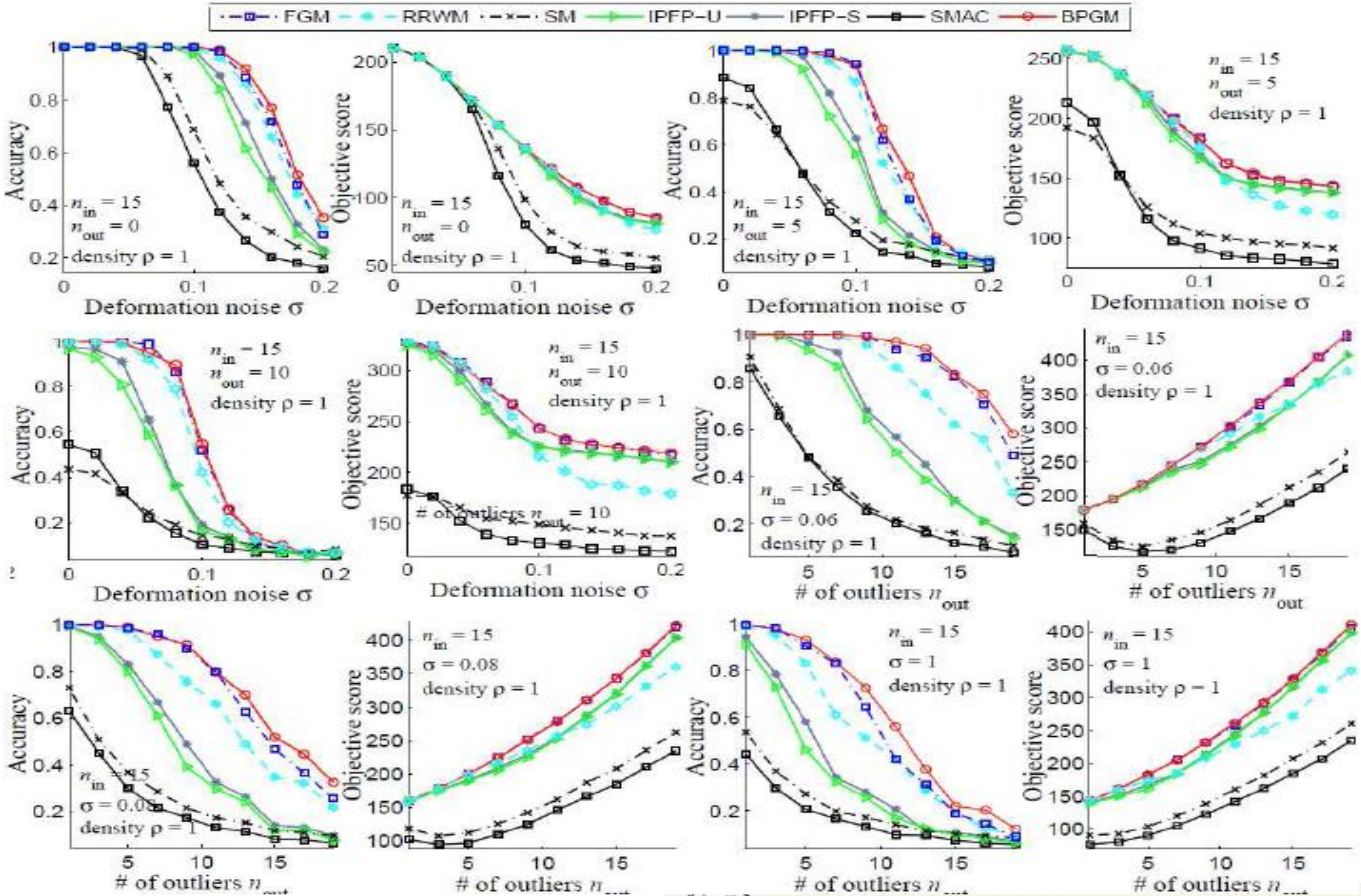
$$P_1(\mathbf{v}) = \arg\min_{\mathbf{v}} \|\mathbf{v} - \mathbf{u}\|_2^2 \quad s.t.\ \ \mathbf{A}\mathbf{v} = \mathbf{1},\ \|\mathbf{v}\|_2^2 = \gamma$$

$$P_2(\mathbf{v}) = \arg\min_{\mathbf{v}} \|\mathbf{v} - \mathbf{u}\|_2^2 \quad s.t.\ \mathbf{v}_i \geq 0$$

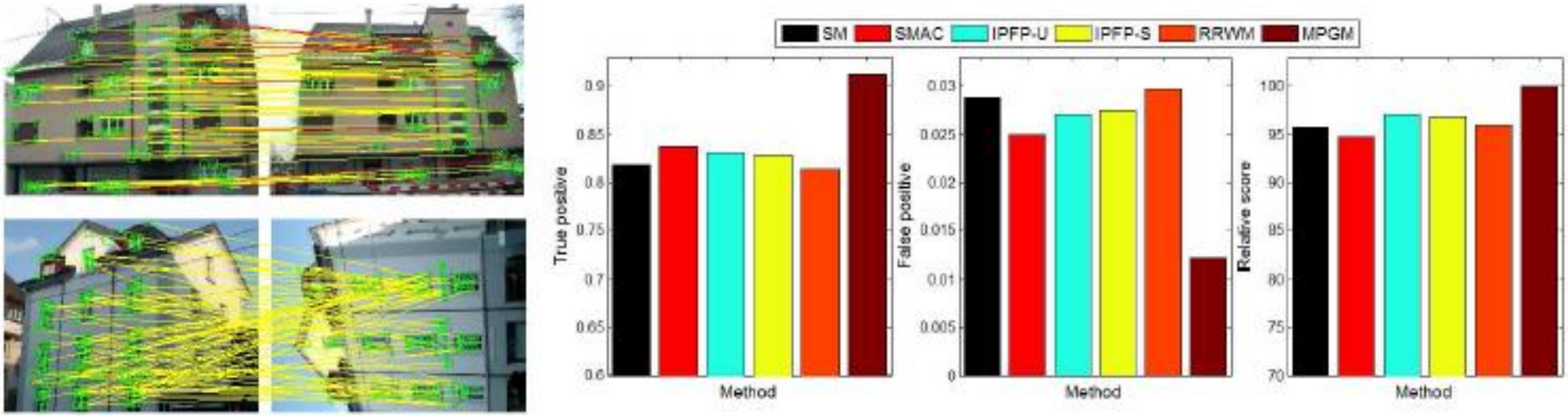
Theorem 1 *Under the update projection Eq.(7), the graph matching objective function* $J_{gm} = \mathbf{x}^{\mathrm{T}}\mathbf{W}\mathbf{x}$ *is monotonically increasing.*

Bo Jiang, et al., Binary constraint preserving graph matching, **CVPR 2017**

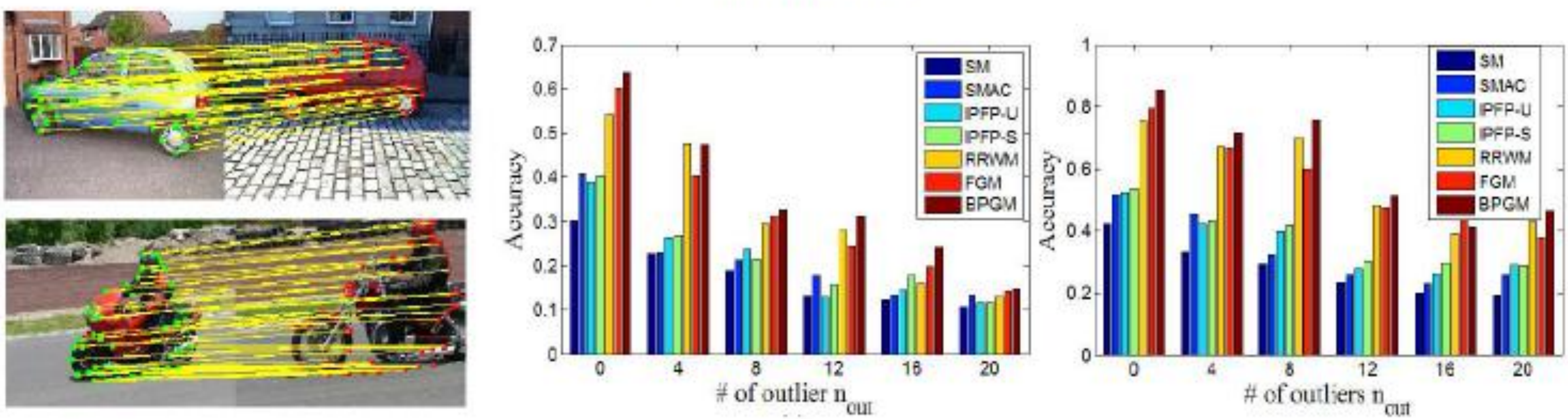
Binary constraint preserving matching



Binary constraint preserving matching



(a) ZuBud image dataset

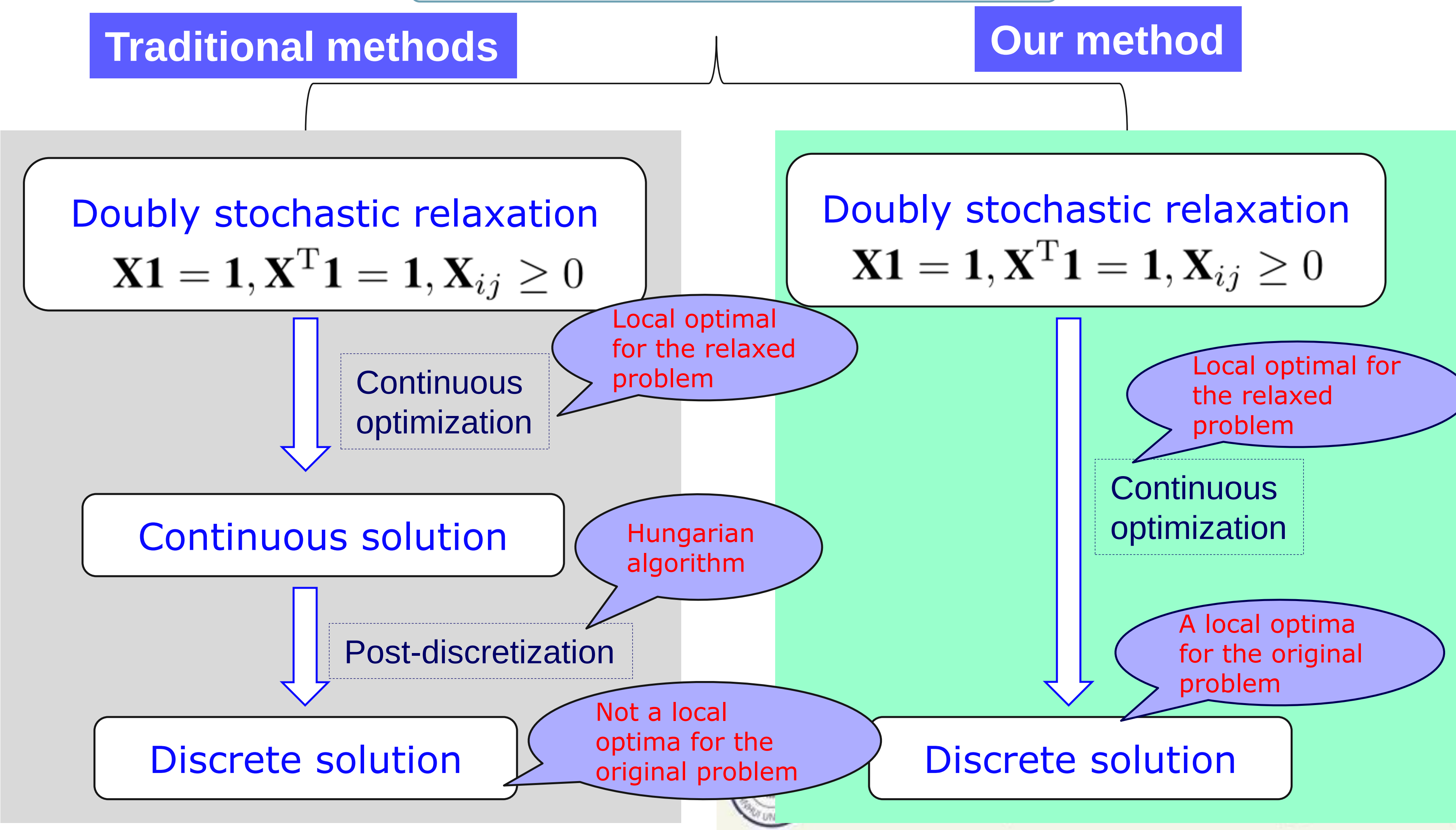


(b) Pascal dataset

Figure 5. Comparison results on real-world image datasets.

Multiplicative update matching

$$\max_{\mathbf{X}} \quad \mathrm{vec}(\mathbf{X})^{\mathrm{T}} \mathbf{W}\, \mathrm{vec}(\mathbf{X})$$
$$s.t. \quad \mathbf{X1} = \mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1} = \mathbf{1},\ \mathbf{X}_{ij} \in \{0, 1\}$$



Multiplicative update matching

Doubly-stochastic Relaxation

$$\max_{\mathbf{X}} \ \mathrm{vec}(\mathbf{X})^{\mathrm{T}} \mathbf{W} \, \mathrm{vec}(\mathbf{X})$$

$$s.t. \, \mathbf{X1} = \mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1} = \mathbf{1}, \mathbf{X}_{ij} \geq 0$$

Multiplicative update matching

Doubly-stochastic Relaxation

$$\max_{\mathbf{X}} \ \mathrm{vec}(\mathbf{X})^{\mathrm{T}}\mathbf{W}\,\mathrm{vec}(\mathbf{X})$$

$$s.t.\ \mathbf{X}\mathbf{1}=\mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1}=\mathbf{1}, \mathbf{X}_{ij}\geq 0$$

$$\mathbf{x} = \mathrm{vec}(\mathbf{X})$$

$$\mathcal{L}(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{W}\mathbf{x} - \sum_{i=1}^{n}\mathbf{\Lambda}_i(\sum_{j=1}^{n}\mathbf{x}_{ij}-1) - \sum_{j=1}^{n}\mathbf{\Gamma}_j(\sum_{i=1}^{n}\mathbf{x}_{ij}-1) - \sum_{i=1}^{n}\sum_{j=1}^{n}\mathbf{\Delta}_{ij}\mathbf{x}_{ij}$$

Multiplicative update matching

Doubly-stochastic Relaxation

$$\max_{\mathbf{X}} \ \mathrm{vec}(\mathbf{X})^{\mathrm{T}} \mathbf{W}\, \mathrm{vec}(\mathbf{X})$$
$$s.t. \ \mathbf{X}\mathbf{1} = \mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1} = \mathbf{1}, \mathbf{X}_{ij} \geq 0$$

$$\mathbf{x} = \mathrm{vec}(\mathbf{X})$$

$$\mathcal{L}(\mathbf{x}) = \mathbf{x}^{\mathrm{T}} \mathbf{W} \mathbf{x} - \sum_{i=1}^{n} \boldsymbol{\Lambda}_i (\sum_{j=1}^{n} \mathbf{x}_{ij} - 1) - \sum_{j=1}^{n} \boldsymbol{\Gamma}_j (\sum_{i=1}^{n} \mathbf{x}_{ij} - 1) - \sum_{i=1}^{n} \sum_{j=1}^{n} \boldsymbol{\Delta}_{ij} \mathbf{x}_{ij}$$

Multiplicative update matching

Doubly-stochastic Relaxation

$$\max_{\mathbf{X}} \ \operatorname{vec}(\mathbf{X})^{\mathrm{T}} \mathbf{W} \operatorname{vec}(\mathbf{X})$$
$$s.t. \mathbf{X1} = \mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1} = \mathbf{1}, \mathbf{X}_{ij} \geq 0$$

Multipliers

$$\frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}_{kl}} = 2(\mathbf{W}\mathbf{x})_{kl} - \mathbf{\Lambda}_k - \mathbf{\Gamma}_l - \mathbf{\Delta}_{kl} = 0$$
$$\frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{\Lambda}_k} = -(\sum_l \mathbf{x}_{kl} - 1) = 0$$
$$\frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{\Gamma}_l} = -(\sum_k \mathbf{x}_{kl} - 1) = 0$$
$$\mathbf{\Delta}_{kl}\mathbf{x}_{kl} = 0$$

Multiplicative update matching

Doubly-stochastic Relaxation

$$\max_{\mathbf{X}} \ \mathrm{vec}(\mathbf{X})^{\mathrm{T}} \mathbf{W} \, \mathrm{vec}(\mathbf{X})$$

$$s.t. \, \mathbf{X1} = \mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1} = \mathbf{1}, \mathbf{X}_{ij} \geq 0$$

Multipliers

$$\frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}_{kl}} = 2(\mathbf{Wx})_{kl} - \mathbf{\Lambda}_k - \mathbf{\Gamma}_l - \mathbf{\Delta}_{kl} = 0$$

$$\frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{\Lambda}_k} = -(\sum_l \mathbf{x}_{kl} - 1) = 0$$

$$\frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{\Gamma}_l} = -(\sum_k \mathbf{x}_{kl} - 1) = 0$$

$$\mathbf{\Delta}_{kl}\mathbf{x}_{kl} = 0$$

$$\left[2(\mathbf{Wx})_{kl} - \mathbf{\Lambda}_k - \mathbf{\Gamma}_l\right]\mathbf{x}_{kl} = 0$$

Multiplicative update matching

Doubly-stochastic Relaxation

$$\max_{\mathbf{X}} \ \mathrm{vec}(\mathbf{X})^{\mathrm{T}} \mathbf{W} \, \mathrm{vec}(\mathbf{X})$$

$$s.t. \ \mathbf{X}\mathbf{1} = \mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1} = \mathbf{1}, \mathbf{X}_{ij} \geq 0$$

Multipliers

$$\frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}_{kl}} = 2(\mathbf{W}\mathbf{x})_{kl} - \mathbf{\Lambda}_k - \mathbf{\Gamma}_l - \mathbf{\Delta}_{kl} = 0$$

$$\frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{\Lambda}_k} = -(\sum_l \mathbf{x}_{kl} - 1) = 0$$

$$\frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{\Gamma}_l} = -(\sum_k \mathbf{x}_{kl} - 1) = 0$$

$$\mathbf{\Delta}_{kl}\mathbf{x}_{kl} = 0$$

$$\left[2(\mathbf{W}\mathbf{x})_{kl} - \mathbf{\Lambda}_k - \mathbf{\Gamma}_l\right]\mathbf{x}_{kl} = 0$$

$$2\sum_{l=1}^{n} \mathbf{x}_{kl}(\mathbf{W}\mathbf{x})_{kl} - \sum_{l=1}^{n} \mathbf{\Gamma}_l \mathbf{x}_{kl} - \mathbf{\Lambda}_k = 0$$

$$2\sum_{k=1}^{n} \mathbf{x}_{kl}(\mathbf{W}\mathbf{x})_{kl} - \sum_{k=1}^{n} \mathbf{\Lambda}_k \mathbf{x}_{kl} - \mathbf{\Gamma}_l = 0$$

Multiplicative update matching

Doubly-stochastic Relaxation

$$\max_{\mathbf{X}} \ \mathrm{vec}(\mathbf{X})^{\mathrm{T}}\mathbf{W}\,\mathrm{vec}(\mathbf{X})$$
$$s.t. \ \mathbf{X1} = \mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1} = \mathbf{1}, \mathbf{X}_{ij} \geq 0$$

Multipliers

$$\frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}_{kl}} = 2(\mathbf{Wx})_{kl} - \mathbf{\Lambda}_k - \mathbf{\Gamma}_l - \mathbf{\Delta}_{kl} = 0$$
$$\frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{\Lambda}_k} = -(\sum_l \mathbf{x}_{kl} - 1) = 0$$
$$\frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{\Gamma}_l} = -(\sum_k \mathbf{x}_{kl} - 1) = 0$$
$$\mathbf{\Delta}_{kl}\mathbf{x}_{kl} = 0$$

$$\left[2(\mathbf{Wx})_{kl} - \mathbf{\Lambda}_k - \mathbf{\Gamma}_l\right]\mathbf{x}_{kl} = 0$$

$$2\sum_{l=1}^{n} \mathbf{x}_{kl}(\mathbf{Wx})_{kl} - \sum_{l=1}^{n} \mathbf{\Gamma}_l \mathbf{x}_{kl} - \mathbf{\Lambda}_k = 0$$
$$2\sum_{k=1}^{n} \mathbf{x}_{kl}(\mathbf{Wx})_{kl} - \sum_{k=1}^{n} \mathbf{\Lambda}_k \mathbf{x}_{kl} - \mathbf{\Gamma}_l = 0$$

$$2\,\mathrm{diag}(\mathbf{KX}^{\mathrm{T}}) - \mathbf{\Lambda} - \mathbf{X\Gamma} = \mathbf{0},$$
$$2\,\mathrm{diag}(\mathbf{K}^{\mathrm{T}}\mathbf{X}) - \mathbf{\Gamma} - \mathbf{X}^{\mathrm{T}}\mathbf{\Lambda} = \mathbf{0}$$

Multiplicative update matching

Doubly-stochastic Relaxation

$$\max_{\mathbf{X}} \ \ \mathrm{vec}(\mathbf{X})^{\mathrm{T}}\mathbf{W}\,\mathrm{vec}(\mathbf{X})$$
$$s.t. \mathbf{X1} = \mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1} = \mathbf{1}, \mathbf{X}_{ij} \geq 0$$

Multipliers

$$\frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}_{kl}} = 2(\mathbf{Wx})_{kl} - \mathbf{\Lambda}_k - \mathbf{\Gamma}_l - \mathbf{\Delta}_{kl} = 0$$
$$\frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{\Lambda}_k} = -(\sum_l \mathbf{x}_{kl} - 1) = 0$$
$$\frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{\Gamma}_l} = -(\sum_k \mathbf{x}_{kl} - 1) = 0$$
$$\mathbf{\Delta}_{kl}\mathbf{x}_{kl} = 0$$

$$\left[2(\mathbf{Wx})_{kl} - \mathbf{\Lambda}_k - \mathbf{\Gamma}_l\right]\mathbf{x}_{kl} = 0$$

$$2\sum_{l=1}^{n} \mathbf{x}_{kl}(\mathbf{Wx})_{kl} - \sum_{l=1}^{n} \mathbf{\Gamma}_l \mathbf{x}_{kl} - \mathbf{\Lambda}_k = 0$$
$$2\sum_{k=1}^{n} \mathbf{x}_{kl}(\mathbf{Wx})_{kl} - \sum_{k=1}^{n} \mathbf{\Lambda}_k \mathbf{x}_{kl} - \mathbf{\Gamma}_l = 0$$

$$2\,\mathrm{diag}(\mathbf{KX}^{\mathrm{T}}) - \mathbf{\Lambda} - \mathbf{X\Gamma} = \mathbf{0},$$
$$2\,\mathrm{diag}(\mathbf{K}^{\mathrm{T}}\mathbf{X}) - \mathbf{\Gamma} - \mathbf{X}^{\mathrm{T}}\mathbf{\Lambda} = \mathbf{0}$$

$$\mathbf{\Gamma} = 2(\mathbf{I} - \mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}(\mathrm{diag}(\mathbf{K}^{\mathrm{T}}\mathbf{X}) - \mathbf{X}^{\mathrm{T}}\mathrm{diag}(\mathbf{KX}^{\mathrm{T}}))$$
$$\mathbf{\Lambda} = 2\,\mathrm{diag}(\mathbf{KX}^{\mathrm{T}}) - \mathbf{X\Gamma}$$

Multiplicative update matching

Doubly-stochastic Relaxation

$$\max_{\mathbf{X}} \quad \mathrm{vec}\left(\mathbf{X}\right)^{\mathrm{T}} \mathbf{W} \, \mathrm{vec}\left(\mathbf{X}\right)$$

$$s.t. \, \mathbf{X1} = \mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1} = \mathbf{1}, \mathbf{X}_{ij} \geq 0$$

Solution update

$$\mathcal{L}(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{W}\mathbf{x} - \sum_{i=1}^{n} \mathbf{\Lambda}_i \Big(\sum_{j=1}^{n} \mathbf{x}_{ij} - 1\Big) - \sum_{j=1}^{n} \mathbf{\Gamma}_j \Big(\sum_{i=1}^{n} \mathbf{x}_{ij} - 1\Big)$$

Multiplicative update matching

Doubly-stochastic Relaxation

$$\max_{\mathbf{X}} \ \ \mathrm{vec}\,(\mathbf{X})^{\mathrm{T}}\mathbf{W}\,\mathrm{vec}\,(\mathbf{X})$$

$$s.t.\ \mathbf{X}\mathbf{1} = \mathbf{1}, \mathbf{X}^{\mathrm{T}}\mathbf{1} = \mathbf{1}, \mathbf{X}_{ij} \geq 0$$

Solution update

$$\mathcal{L}(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{W}\mathbf{x} - \sum_{i=1}^{n}\mathbf{\Lambda}_i\Big(\sum_{j=1}^{n}\mathbf{x}_{ij} - 1\Big) - \sum_{j=1}^{n}\mathbf{\Gamma}_j\Big(\sum_{i=1}^{n}\mathbf{x}_{ij} - 1\Big)$$

$$\Phi(\mathbf{x}, \mathbf{x}) = \mathcal{L}(\mathbf{x}), \Phi(\mathbf{x}, \tilde{\mathbf{x}}) \leq \mathcal{L}(\mathbf{x})$$

$$\mathbf{x}^{(t+1)} = \arg\max_{\mathbf{x}} \Phi(\mathbf{x}, \mathbf{x}^{(t)})$$

$$\mathcal{L}(\mathbf{x}^{(t)}) = \Phi(\mathbf{x}^{(t)}, \mathbf{x}^{(t)}) \leq \mathcal{L}(\mathbf{x}^{(t+1)})$$

Multiplicative update matching

$$\Phi(\mathbf{x}, \tilde{\mathbf{x}}) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} \mathbf{W}_{ij,kl}\tilde{\mathbf{x}}_{ij}\tilde{\mathbf{x}}_{kl}\Big(1 + \log\frac{\mathbf{x}_{ij}\mathbf{x}_{kl}}{\tilde{\mathbf{x}}_{ij}\tilde{\mathbf{x}}_{kl}}\Big)$$

$$-\sum_{i=1}^{n}\mathbf{\Lambda}_i^{+}\Big[\sum_{j=1}^{n}\frac{1}{2}(\frac{\mathbf{x}_{ij}^2}{\tilde{\mathbf{x}}_{ij}} + \tilde{\mathbf{x}}_{ij}) - 1\Big] + \sum_{i=1}^{n}\mathbf{\Lambda}_i^{-}\Big[\sum_{j=1}^{n}\tilde{\mathbf{x}}_{ij}(1 + \log\frac{\mathbf{x}_{ij}}{\tilde{\mathbf{x}}_{ij}}) - 1\Big]$$

$$-\sum_{j=1}^{n}\mathbf{\Gamma}_j^{+}\Big[\sum_{i=1}^{n}\frac{1}{2}(\frac{\mathbf{x}_{ij}^2}{\tilde{\mathbf{x}}_{ij}} + \tilde{\mathbf{x}}_{ij}) - 1\Big] + \sum_{j=1}^{n}\mathbf{\Gamma}_j^{-}\Big[\sum_{i=1}^{n}\tilde{\mathbf{x}}_{ij}(1 + \log\frac{\mathbf{x}_{ij}}{\tilde{\mathbf{x}}_{ij}}) - 1\Big]$$

Multiplicative update matching

$$\Phi(\mathbf{x},\tilde{\mathbf{x}}) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} \mathbf{W}_{ij,kl}\tilde{\mathbf{x}}_{ij}\tilde{\mathbf{x}}_{kl}\Big(1+\log\frac{\mathbf{x}_{ij}\mathbf{x}_{kl}}{\tilde{\mathbf{x}}_{ij}\tilde{\mathbf{x}}_{kl}}\Big)$$

$$-\sum_{i=1}^{n}\mathbf{\Lambda}_i^{+}\Big[\sum_{j=1}^{n}\frac{1}{2}\Big(\frac{\mathbf{x}_{ij}^2}{\tilde{\mathbf{x}}_{ij}}+\tilde{\mathbf{x}}_{ij}\Big)-1\Big]+\sum_{i=1}^{n}\mathbf{\Lambda}_i^{-}\Big[\sum_{j=1}^{n}\tilde{\mathbf{x}}_{ij}\Big(1+\log\frac{\mathbf{x}_{ij}}{\tilde{\mathbf{x}}_{ij}}\Big)-1\Big]$$

$$-\sum_{j=1}^{n}\mathbf{\Gamma}_j^{+}\Big[\sum_{i=1}^{n}\frac{1}{2}\Big(\frac{\mathbf{x}_{ij}^2}{\tilde{\mathbf{x}}_{ij}}+\tilde{\mathbf{x}}_{ij}\Big)-1\Big]+\sum_{j=1}^{n}\mathbf{\Gamma}_j^{-}\Big[\sum_{i=1}^{n}\tilde{\mathbf{x}}_{ij}\Big(1+\log\frac{\mathbf{x}_{ij}}{\tilde{\mathbf{x}}_{ij}}\Big)-1\Big]$$

Algorithm

$$\mathbf{x}^{(t+1)} = \arg\max_{\mathbf{x}} \Phi(\mathbf{x},\mathbf{x}^{(t)})$$

$$\mathbf{x}_{kl} = \tilde{\mathbf{x}}_{kl}\Big[\frac{2(\mathbf{W}\tilde{\mathbf{x}})_{kl}+\mathbf{\Lambda}_k^{-}+\mathbf{\Gamma}_l^{-}}{\mathbf{\Lambda}_k^{+}+\mathbf{\Gamma}_l^{+}}\Big]^{1/2}$$

Multiplicative update matching

$$\begin{aligned}\Phi(\mathbf{x},\tilde{\mathbf{x}}) = & \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\mathbf{W}_{ij,kl}\tilde{\mathbf{x}}_{ij}\tilde{\mathbf{x}}_{kl}\Big(1+\log\frac{\mathbf{x}_{ij}\mathbf{x}_{kl}}{\tilde{\mathbf{x}}_{ij}\tilde{\mathbf{x}}_{kl}}\Big)\\ & -\sum_{i=1}^{n}\mathbf{\Lambda}_i^{+}\Big[\sum_{j=1}^{n}\frac{1}{2}\big(\frac{\mathbf{x}_{ij}^2}{\tilde{\mathbf{x}}_{ij}}+\tilde{\mathbf{x}}_{ij}\big)-1\Big]+\sum_{i=1}^{n}\mathbf{\Lambda}_i^{-}\Big[\sum_{j=1}^{n}\tilde{\mathbf{x}}_{ij}\big(1+\log\frac{\mathbf{x}_{ij}}{\tilde{\mathbf{x}}_{ij}}\big)-1\Big]\\ & -\sum_{j=1}^{n}\mathbf{\Gamma}_j^{+}\Big[\sum_{i=1}^{n}\frac{1}{2}\big(\frac{\mathbf{x}_{ij}^2}{\tilde{\mathbf{x}}_{ij}}+\tilde{\mathbf{x}}_{ij}\big)-1\Big]+\sum_{j=1}^{n}\mathbf{\Gamma}_j^{-}\Big[\sum_{i=1}^{n}\tilde{\mathbf{x}}_{ij}\big(1+\log\frac{\mathbf{x}_{ij}}{\tilde{\mathbf{x}}_{ij}}\big)-1\Big]\end{aligned}$$

Algorithm

$$\mathbf{x}^{(t+1)} = \arg\max_{\mathbf{x}} \Phi(\mathbf{x},\mathbf{x}^{(t)})$$

$$\mathbf{x}_{kl} = \tilde{\mathbf{x}}_{kl}\Big[\frac{2(\mathbf{W}\tilde{\mathbf{x}})_{kl}+\mathbf{\Lambda}_k^{-}+\mathbf{\Gamma}_l^{-}}{\mathbf{\Lambda}_k^{+}+\mathbf{\Gamma}_l^{+}}\Big]^{1/2}$$

$$\begin{aligned}\mathbf{\Gamma} = & 2\big(\mathbf{I}-{\mathbf{X}^{(t)}}^{\mathrm{T}}\mathbf{X}^{(t)}\big)^{-1}\Big[\mathrm{diag}\,\big({\mathbf{K}^{(t)}}^{\mathrm{T}}\mathbf{X}^{(t)}\big)-{\mathbf{X}^{(t)}}^{\mathrm{T}}\,\mathrm{diag}\,\big(\mathbf{K}^{(t)}{\mathbf{X}^{(t)}}^{\mathrm{T}}\big)\Big]\\ \mathbf{\Lambda} = & 2\,\mathrm{diag}\,\big(\mathbf{K}^{(t)}{\mathbf{X}^{(t)}}^{\mathrm{T}}\big)-\mathbf{X}^{(t)}\mathbf{\Gamma}\end{aligned}$$

Multiplicative update matching

Convergence

Theorem 1. *Under update, the Lagrangian function $\mathcal{L}(\mathbf{x})$ is monotonically increasing,*

$$\mathcal{L}(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{W}\mathbf{x} - \sum_{i=1}^{n} \mathbf{\Lambda}_i \Big(\sum_{j=1}^{n} \mathbf{x}_{ij} - 1\Big) - \sum_{j=1}^{n} \mathbf{\Gamma}_j \Big(\sum_{i=1}^{n} \mathbf{x}_{ij} - 1\Big)$$

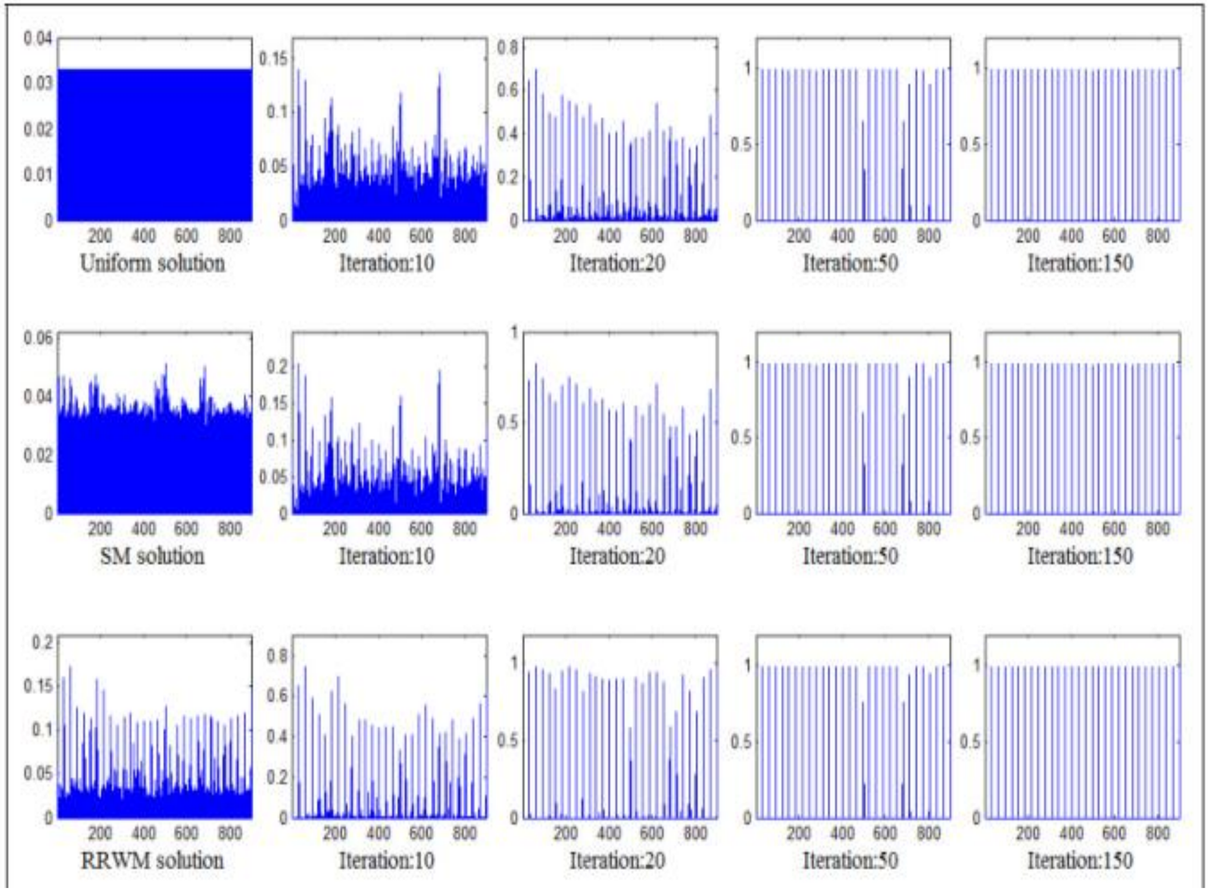
where $\mathbf{\Lambda}, \mathbf{\Gamma}$ are Lagrangian multipliers.

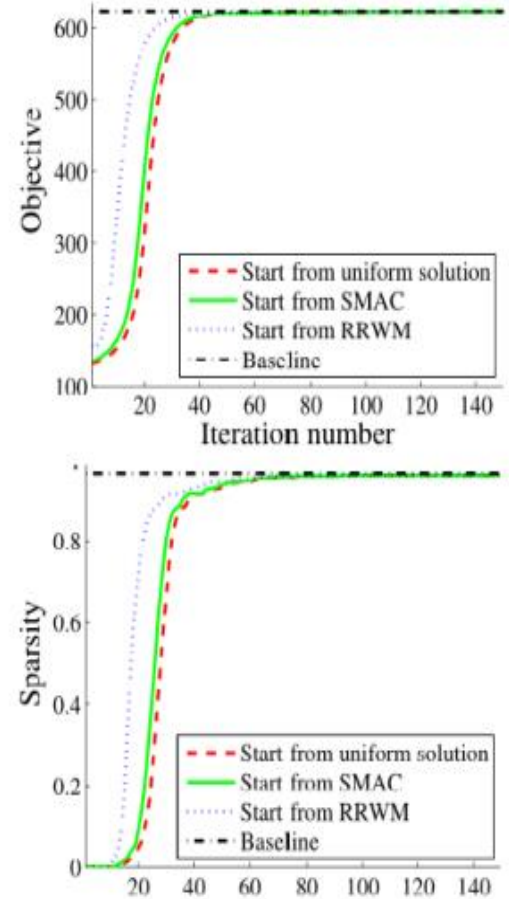
Optimality

Theorem 2. *Under update, the converged solution $\mathbf{x}^*$ is Karush-Kuhn Tucker (KKT) optimal.*

Bo Jiang, et al., Graph Matching via Multiplicative Update Algorithm, **NIPS** 2017

Multiplicative update matching





Top: start from uniform solution
Middle: start from Spectral Matching (SM) solution
Bottom: start from Random Walk (RRWM) solution

Bo Jiang, et al., Graph Matching via Multiplicative Update Algorithm, **NIPS** 2017

Multiplicative update matching

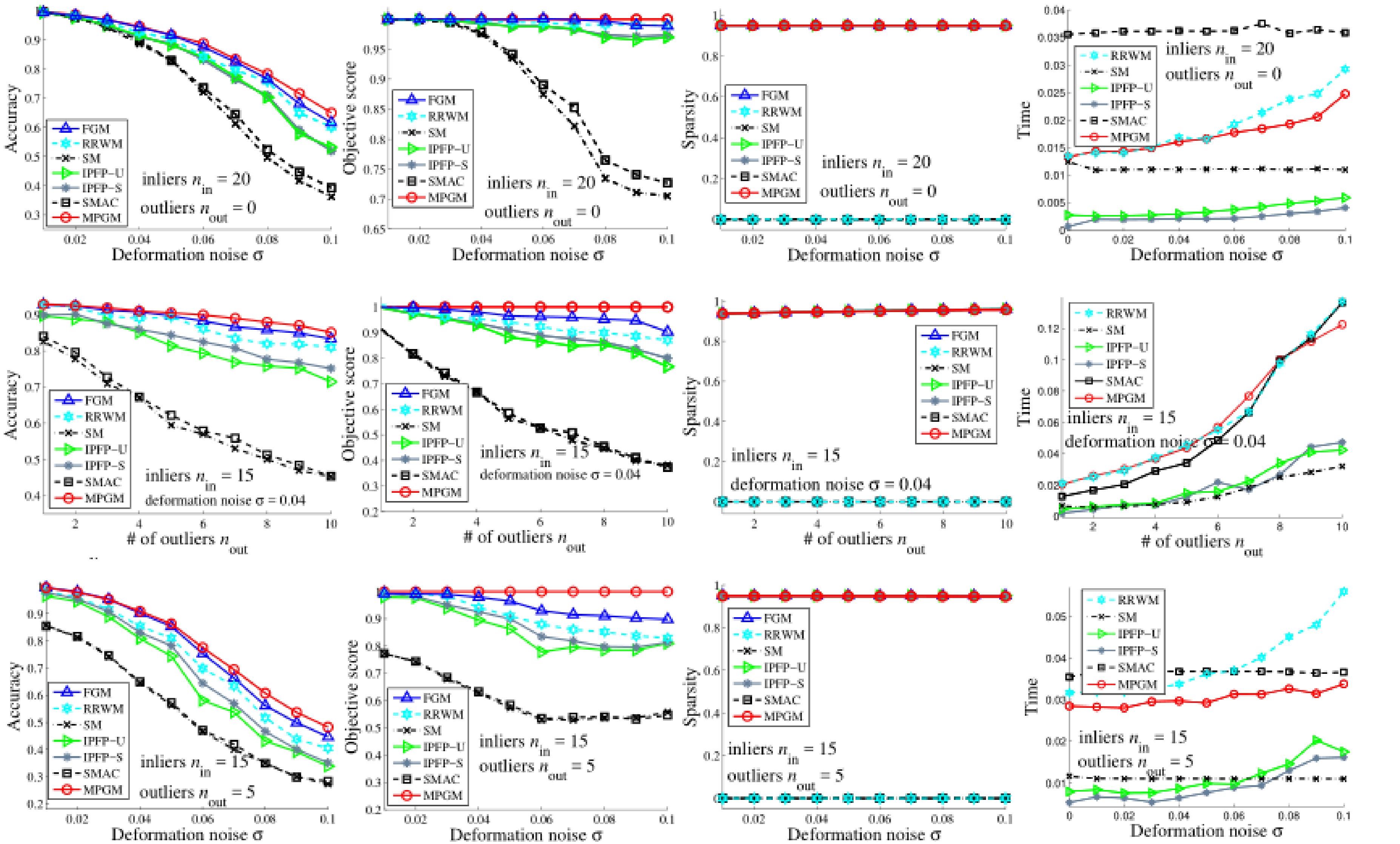


Figure 2: Comparison results of different methods on synthetic point sets matching

Multiplicative update matching

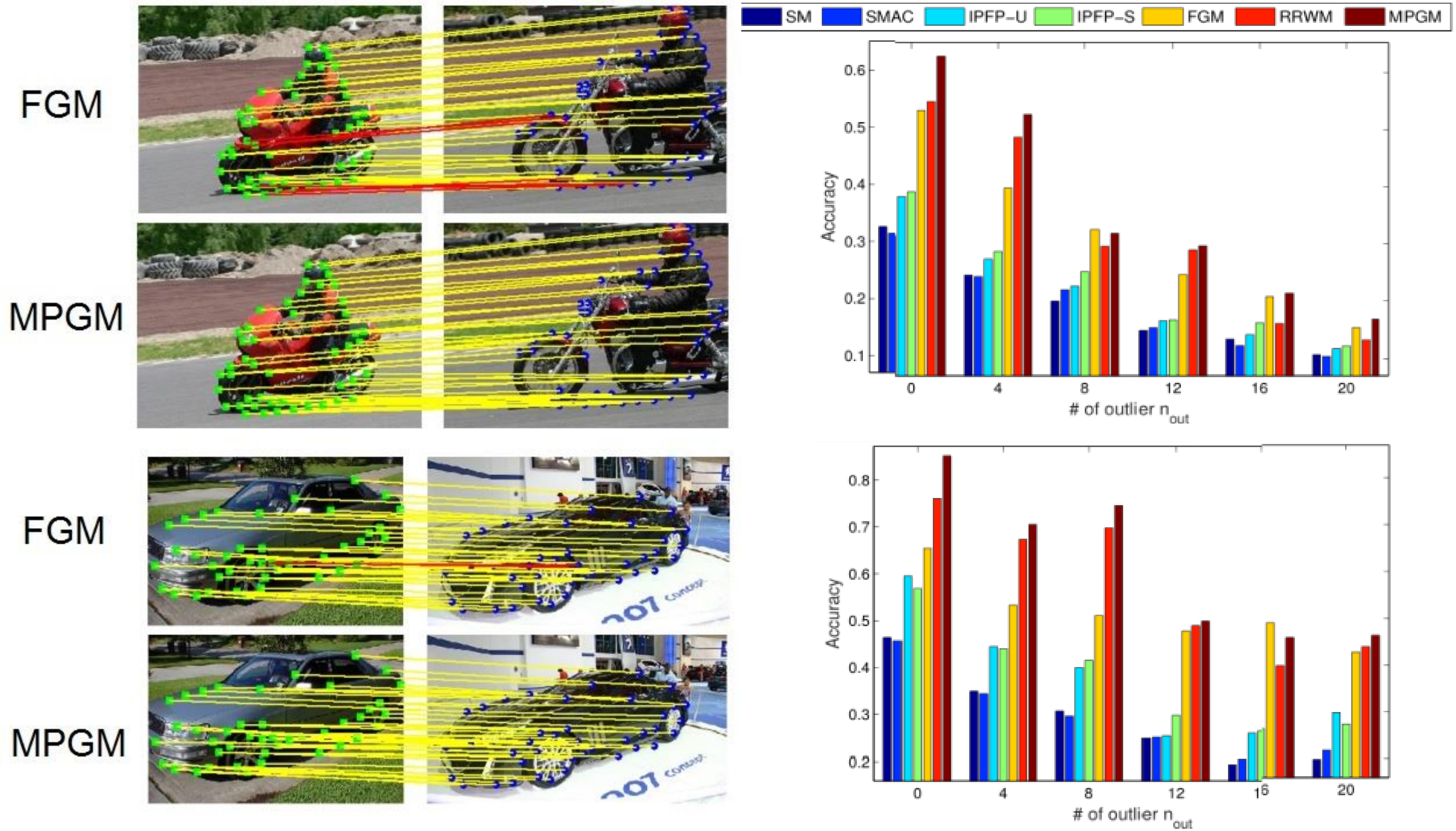


Figure 5: Comparison results of different graph matching methods on the Pascal 2007 dataset

Reference

- Bo Jiang, Jin Tang, Chris Ding, Yihong Gong and Bin Luo, Graph Matching via Multiplicative Update Algorithm, Neural Information Processing Systems (*NIPS*-2017)
- Bo Jiang, Jin Tang, Bin Luo and Chris Ding, Binary constraint preserving graph matching, IEEE Conference on Computer Vision and Pattern Recognition (*CVPR*), pp.4402-4409, 2017
- Bo Jiang, Jin Tang, Chris Ding and Bin Luo, Nonnegative Orthogonal Graph Matching, AAAI Conference on Artificial Intelligence (*AAAI*), pp.4089-4095, 2017
- Bo Jiang, Jin Tang, Xiaochun Cao, Bin Luo, Lagrangian relaxation graph matching, *Pattern Recognition*, 61: 255-265, 2017
- Bo Jiang, Jin Tang, Chris Ding and Bin Luo, A local sparse model for matching problem, AAAI Conference on Artificial Intelligence (*AAAI*), pp. 3790-3796, 2015
- Bo Jiang, Jin Tang, Bin Luo and Liang Lin, Robust feature point matching with sparse model, *IEEE Transactions on Image Processing*, 23(12):5175-5186, 2014

Conclusion and Future works

Conclusion

- Sparse relaxation model for matching problem
- Binary constraint preserving model for matching
- Multiplicative update algorithm for matching

Future works

- More theoretical analysis on Multiplicative matching
- More effective algorithm to solve sparse matching model
- Matching objective relaxation

Thank you !

jiangbo@ahu.edu.cn