# Robust Object Matching using Low-rank constraint and its Applications

**Kui Jia**

*University of Macau*

*ADSC*

*VALSE Webinar, May 6, 2015*

# References

- Kui Jia, Tsung-Han Chan, Zinan Zeng, Shenghua Gao, Gang Wang, Tianzhu Zhang, and Yi Ma, *"**ROML: A Robust Feature Correspondence Approach for Matching Objects in A Set of Images**"*, arXiv:1403.7877, 2014.

  *Data and code available at* [*https://sites.google.com/site/kuijia/research/roml*](https://sites.google.com/site/kuijia/research/roml)

- Tianzhu Zhang*, Kui Jia*, Changsheng Xu, Yi Ma, and N. Ahuja, *"**Partial Occlusion Handling for Visual Tracking via Robust Part Matching**"*, IEEE Conference on Computer Vision and Pattern Recognition, 2014. (* indicates equal contributions)

- Zinan Zeng, Shijie Xiao, Kui Jia, Tsung-Han Chan, Shenghua Gao, Dong Xu, and Yi Ma, "***Learning by Associating Ambiguously Labeled Images***", IEEE Conference on Computer Vision and Pattern Recognition, 2013.

- Zinan Zeng, Tsung-Han Chan, Kui Jia, and Dong Xu, "***Finding Correspondence from Multiple Images via Sparse and Low-rank Decomposition***", European Conference on Computer Vision, 2012.

# Outline

- Background knowledge and motivation
- ROML: Robust Object Matching using Low-rank constraint
  - *Formulation*
  - *Solving algorithm*
  - *Results*
- Applications of ROML to other data problems
  - *Tracking*
  - *Ambiguous learning*

# A job post

- *A few Research Assistant positions are available in my group at University of Macau, Macau SAR, China*

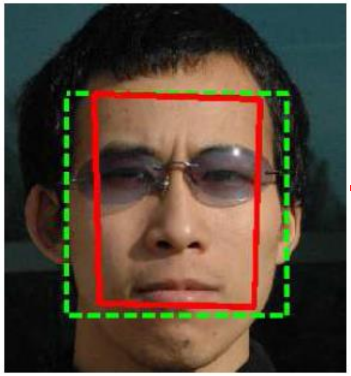
- *Payment is similar/identical to RA jobs in universities in Hong Kong (e.g., 1,5000 HKD per month)*

- *Interested students may send your CV to [kuijia@umac.mo](mailto:kuijia@umac.mo) or contact me via QQ (124401525) for a casual discussion of the potential research topics*

# Outline

- Background knowledge and motivation
- ROML: Robust Object Matching using Low-rank constraint
  - *Formulation*
  - *Solving algorithm*
  - *Results*
- Applications of ROML to other data problems
  - *Tracking*
  - *Ambiguous learning*

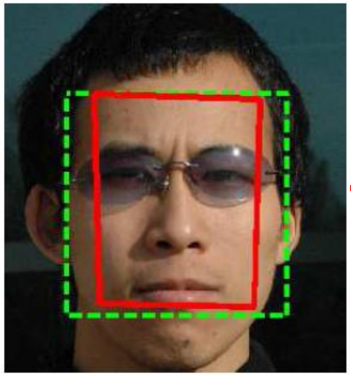
# Face and object recognition



*Viola and Jones' detector*

*Off-the-shelf alignment tools*

# Face and object recognition



*Viola and Jones' detector*

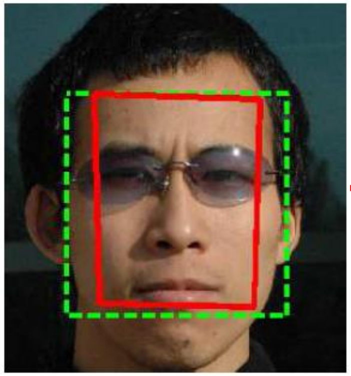
*Off-the-shelf alignment tools*


*Again, detection and alignment ?*

**VERY DIFFICULT!**

**ACTIVE AREAS!**

# Face and object recognition



*Viola and Jones' detector*

*Off-the-shelf alignment tools*

*Again, detection and alignment ?*

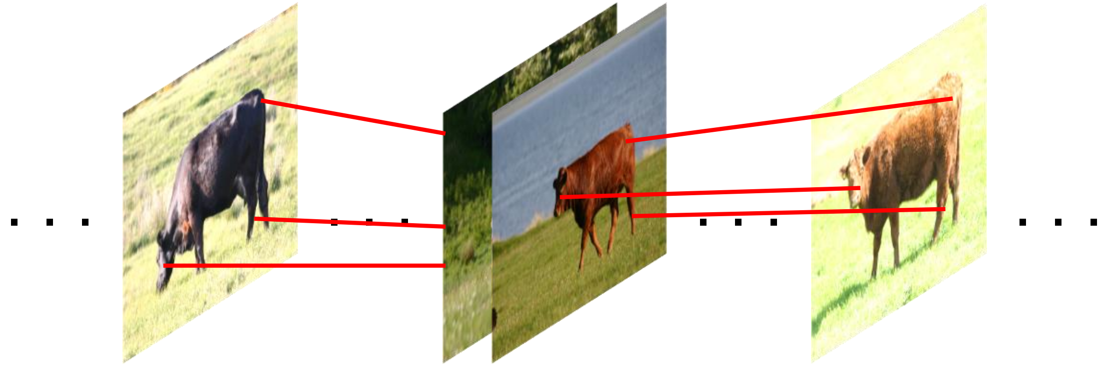
**VERY DIFFICULT!**

**ACTIVE AREAS!**

*Let's be back to the more traditional approach – MATCHING OF SALIENT INTEREST POINTS!*

# Matching of interest points in images

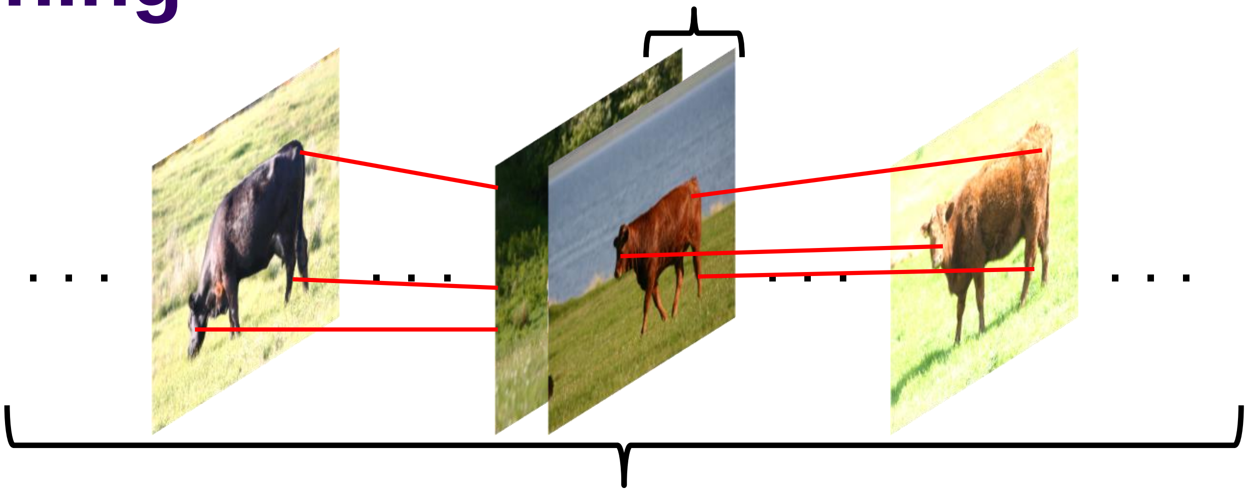
- Matching of interest points: a fundamental problem
- Applications: *object recognition, 3D reconstruction, tracking, motion segmentation …*
- Image coordinates based or feature based matching
- Challenges: *illumination change, viewpoint change, pose change, variability of same-category instances, occlusion …*
- *Global matching across a set of images*



# Pair-wise matching

- Point set based matching
  - *- e.g., shape context [Belongie et al. 02]*
- Matching using local appearance descriptor
  - *- e.g., SIFT, HOG, which are invariant and discriminative*
- Graph and hyper-graph matching
  - *- feature similarity and geometric compatibility*
  - *- formulated as NP-hard Quadratic Assignment Problem (QAP)*

# From pair-wise matching to global matching



More common and desirable to simultaneously match across a set of images

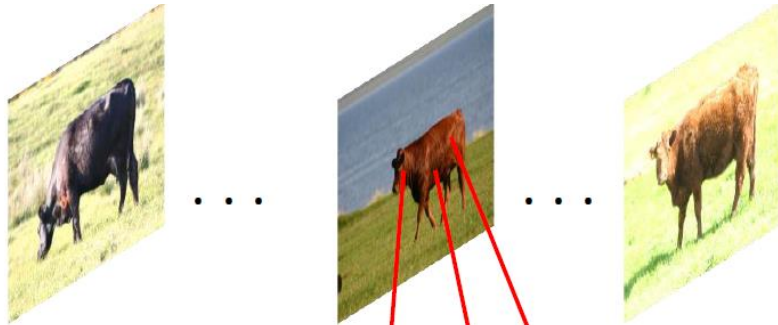
- *be able to establish a globally consistent matching*
- *more robust against outliers and occlusion of inlier features*

# Problem definition of ROML

*Given a set of images with both inlier and outlier features extracted from each image, **simultaneously** identify a given number of inlier features from each image and establish their **consistent** correspondences across the image set.*

*Jia, Chan, Zeng, Gao, Wang, Zhang, Ma, "ROML: A Robust Feature Correspondence Approach for Matching Objects in A Set of Images", arXiv:1403.7877, 2014.*

# The ROML formulation – motivation

$\cdots$  $\cdots$

$\overline{\mathbf{F}}^1 \quad \cdots \quad \overline{\mathbf{F}}^k = [\mathbf{f}_1^k, \ldots, \mathbf{f}_n^k] \in \mathbb{R}^{d \times n} \quad \cdots \quad \overline{\mathbf{F}}^K$

$n$ - the given number of inliers

$K$ - the number of images

$\mathbf{f}_i^k \in \mathbb{R}^d$ - feature vector

Inlier features in image k

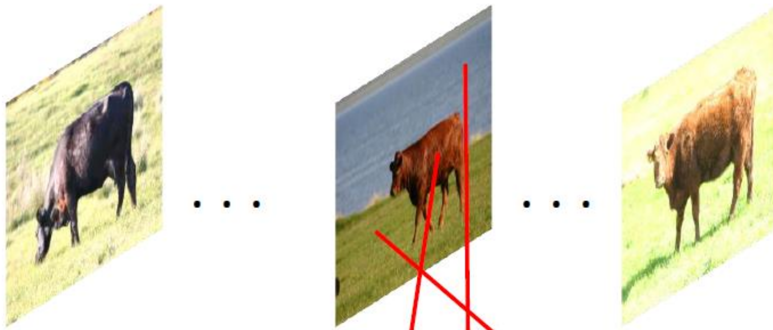
Inlier feature vectors in each of $\{\overline{\mathbf{F}}^1, \ldots, \overline{\mathbf{F}}^K\}$ are ordered.

$\overline{\mathbf{D}} = [\mathrm{vec}(\overline{\mathbf{F}}^1), \ldots, \mathrm{vec}(\overline{\mathbf{F}}^K)] \in \mathbb{R}^{dn \times K}$ is approximately low-rank, ideally rank-one.

*The underlying rationale*

- *Object pattern is determined by its associated inlier features and their geometric relations*
- *Inlier features repetitively appear in the image set, the corresponding ones in different images are correlated to each other*

# The ROML formulation – motivation

$\cdots$  $\cdots$

$n$ - the given number of inliers

$n_k$ - number of feature vectors in image k

$K$ - the number of images

$\mathbf{F}^1 \quad \cdots \quad \mathbf{F}^k = [\mathbf{f}_1^k, \ldots, \mathbf{f}_{n_k}^k] \in \mathbb{R}^{d \times n_k} \quad \cdots \quad \mathbf{F}^K$

$\mathbf{f}_i^k \in \mathbb{R}^d$ - feature vector

Both Inlier and outlier features in image k

Optimal partial permutation matrices (PPMs) $\{\mathbf{P}^k \in \mathcal{P}^k\}_{k=1}^K$

$\mathbf{D} = [\text{vec}(\mathbf{F}^1\mathbf{P}^1), \ldots, \text{vec}(\mathbf{F}^K\mathbf{P}^K)] \in \mathbb{R}^{dn \times K}$ is rank deficient.

$$\mathcal{P}^k = \{\mathbf{P}^k \in \mathbb{R}^{n_k \times n} | \mathbf{P}_{ij}^k \in \{0,1\}, \sum_i \mathbf{P}_{ij}^k = 1$$
$$\forall j = 1, \ldots, n, \sum_j \mathbf{P}_{ij}^k \le 1 \; \forall i = 1, \ldots, n_k\}$$

*The underlying rationale*

- *Object pattern is determined by its associated inlier features and their geometric relations*
- *Inlier features repetitively appear in the image set, the corresponding ones in different images are correlated to each other*
- *Outlier features appear in images in random, unstructured way*

# The ROML formulation

$$\min_{\{\mathbf{P}^k \in \mathcal{P}^k\}_{k=1}^K} \left\| [\mathrm{vec}(\mathbf{F}^1\mathbf{P}^1), \ldots, \mathrm{vec}(\mathbf{F}^K\mathbf{P}^K)] \right\|_*$$

- Jointly optimizing a set of PPMs
- An instance of multi-index assignment problem (MiAP) [Burkard, Dell'Amico, Martello 09]
- NP-hard, practically solved by approximate solution methods, e.g., classical greedy, GRASP methods …

# The ROML formulation

$$\min_{\{\mathbf{P}^k \in \mathcal{P}^k\}_{k=1}^K} \left\| [\mathrm{vec}(\mathbf{F}^1\mathbf{P}^1), \ldots, \mathrm{vec}(\mathbf{F}^K\mathbf{P}^K)] \right\|_*$$

- Jointly optimizing a set of PPMs
- An instance of multi-index assignment problem (MiAP) [Burkard, Dell'Amico, Martello 09]
- NP-hard, practically solved by approximate solution methods

$$\begin{aligned} \min_{\{\mathbf{P}^k \in \mathcal{P}^k\}_{k=1}^K, \mathbf{L}, \mathbf{E}} \quad & \|\mathbf{L}\|_* + \lambda \|\mathbf{E}\|_1 \\ \text{s.t.} \quad & [\mathrm{vec}(\mathbf{F}^1\mathbf{P}^1), \ldots, \mathrm{vec}(\mathbf{F}^K\mathbf{P}^K)] = \mathbf{L} + \mathbf{E}, \\ & \mathcal{P}^k = \left\{ \mathbf{P}^k \in \{0,1\}^{n_k \times n} \middle| \mathbf{1}_{n_k}^\top \mathbf{P}^k = \mathbf{1}_n^\top, \right. \\ & \left. \quad \mathbf{P}^k \mathbf{1}_n \le \mathbf{1}_{n_k} \right\}, \ \forall\ k = 1, \ldots, K, \end{aligned}$$

- Introducing auxiliary variables **L** and **E** (modelling sparse errors)
- Termed Robust Object Matching using Low-rank and sparse constraints (**ROML**)
- A formulation of regularized consensus problem in distributed optimization [Bertsekas & Tsitsiklis 89]
- Alternating Direction Method of Multipliers (ADMM) for such kind of distributed optimization

# Algorithm for approximate ROML solution

The augmented Lagrangian of ROML

$$\mathcal{L}_\rho(\mathbf{L}, \mathbf{E}, \{\mathbf{P}^k \in \mathcal{P}^k\}_{k=1}^K, \mathbf{Y}) = \|\mathbf{L}\|_* + \lambda\|\mathbf{E}\|_1 + \langle \mathbf{Y}, \mathbf{L} + \mathbf{E} - \mathbf{D}\rangle + \frac{\rho}{2}\|\mathbf{L} + \mathbf{E} - \mathbf{D}\|_F^2$$

where $\mathbf{D} = [\mathrm{vec}(\mathbf{F}^1\mathbf{P}^1), \ldots, \mathrm{vec}(\mathbf{F}^K\mathbf{P}^K)]$

ADMM procedure

$$\left.\begin{aligned}
\mathbf{L}_{t+1} &= \arg\min_{\mathbf{L}} \mathcal{L}_\rho(\mathbf{L}, \mathbf{E}_t, \{\mathbf{P}_t^k\}_{k=1}^K, \mathbf{Y}_t) \\
\mathbf{E}_{t+1} &= \arg\min_{\mathbf{E}} \mathcal{L}_\rho(\mathbf{L}_{t+1}, \mathbf{E}, \{\mathbf{P}_t^k\}_{k=1}^K, \mathbf{Y}_t)
\end{aligned}\right\} \text{Fusion steps}$$

$$\{\mathbf{P}_{t+1}^k\}_{k=1}^K = \arg\min_{\{\mathbf{P}^k \in \mathcal{P}^k\}_{k=1}^K} \mathcal{L}_\rho(\mathbf{L}_{t+1}, \mathbf{E}_{t+1}, \{\mathbf{P}^k\}_{k=1}^K, \mathbf{Y}_t) \rightarrow \text{Broadcast step, K independent subproblems}$$

$$\mathbf{Y}_{t+1} = \mathbf{Y}_t + \rho(\mathbf{L}_{t+1} + \mathbf{E}_{t+1} - \mathbf{D}_{t+1})$$

- *A "fusion-and-broadcast" strategy*
- *Broadcast step boils down as independent optimization of individual* $\mathbf{P}^k$, $k = 1, \ldots, K$

# Algorithm for approximate ROML solution

ADMM procedure

$$\begin{aligned} \mathbf{L}_{t+1} &= \arg\min_{\mathbf{L}} \mathcal{L}_\rho\big(\mathbf{L}, \mathbf{E}_t, \{\mathbf{P}_t^k\}_{k=1}^K, \mathbf{Y}_t\big) \\ \mathbf{E}_{t+1} &= \arg\min_{\mathbf{E}} \mathcal{L}_\rho\big(\mathbf{L}_{t+1}, \mathbf{E}, \{\mathbf{P}_t^k\}_{k=1}^K, \mathbf{Y}_t\big) \end{aligned} \Bigg\} \text{ Fusion steps}$$

$$\{\mathbf{P}_{t+1}^k\}_{k=1}^K = \arg\min_{\{\mathbf{P}^k \in \mathcal{P}^k\}_{k=1}^K} \mathcal{L}_\rho\big(\mathbf{L}_{t+1}, \mathbf{E}_{t+1}, \{\mathbf{P}^k\}_{k=1}^K, \mathbf{Y}_t\big) \rightarrow \text{Broadcast step, K independent subproblems}$$

$$\begin{aligned} &\min_{\theta^k} \frac{\rho}{2}\theta^{k\top}\mathbf{G}^{k\top}\mathbf{G}^k\theta^k - \mathbf{e}_k^\top\big[\mathbf{Y}_t^\top + \rho(\mathbf{L}_{t+1}+\mathbf{E}_{t+1})^\top\big]\mathbf{G}^k\theta^k \\ &\text{s.t. } \mathbf{J}^k\theta^k = \mathbf{1}_n,\ \ \mathbf{H}^k\theta^k \le \mathbf{1}_{n_k},\ \ \theta^k \in \{0,1\}^{nn_k} \end{aligned}$$

$$\begin{aligned} \theta^k &= \mathrm{vec}(\mathbf{P}^k) \in \mathbb{R}^{nn_k} \\ \mathbf{G}^k &= \mathbf{I}_n \otimes \mathbf{F}^k \in \mathbb{R}^{dn\times nn_k} \\ \mathbf{J}^k &= \mathbf{I}_n \otimes \mathbf{1}_{n_k}^\top \in \mathbb{R}^{n\times nn_k} \\ \mathbf{H}^k &= \mathbf{1}_n^\top \otimes \mathbf{I}_{n_k} \in \mathbb{R}^{n_k\times nn_k} \end{aligned}$$

*A difficult integer constrained quadratic program (IQP)*

$$\mathbf{Y}_{t+1} = \mathbf{Y}_t + \rho\big(\mathbf{L}_{t+1} + \mathbf{E}_{t+1} - \mathbf{D}_{t+1}\big)$$

# Algorithm for approximate ROML solution

IQP: $$\min_{\theta^k} \frac{\rho}{2}\theta^{k\top}\mathbf{G}^{k\top}\mathbf{G}^k\theta^k - \mathbf{e}_k^\top\left[\mathbf{Y}_t^\top + \rho\left(\mathbf{L}_{t+1} + \mathbf{E}_{t+1}\right)^\top\right]\mathbf{G}^k\theta^k$$
$$\text{s.t. } \mathbf{J}^k\theta^k = \mathbf{1}_n,\ \mathbf{H}^k\theta^k \leq \mathbf{1}_{n_k},\ \theta^k \in \{0,1\}^{nn_k}$$

**Theorem 1**

*For the proposed ROML problem, assume distinctive information of each column vector in any* $\mathbf{F}^k$ *of* $\{\mathbf{F}^k\}_{k=1}^K$ *is represented by the relative values of its elements.*

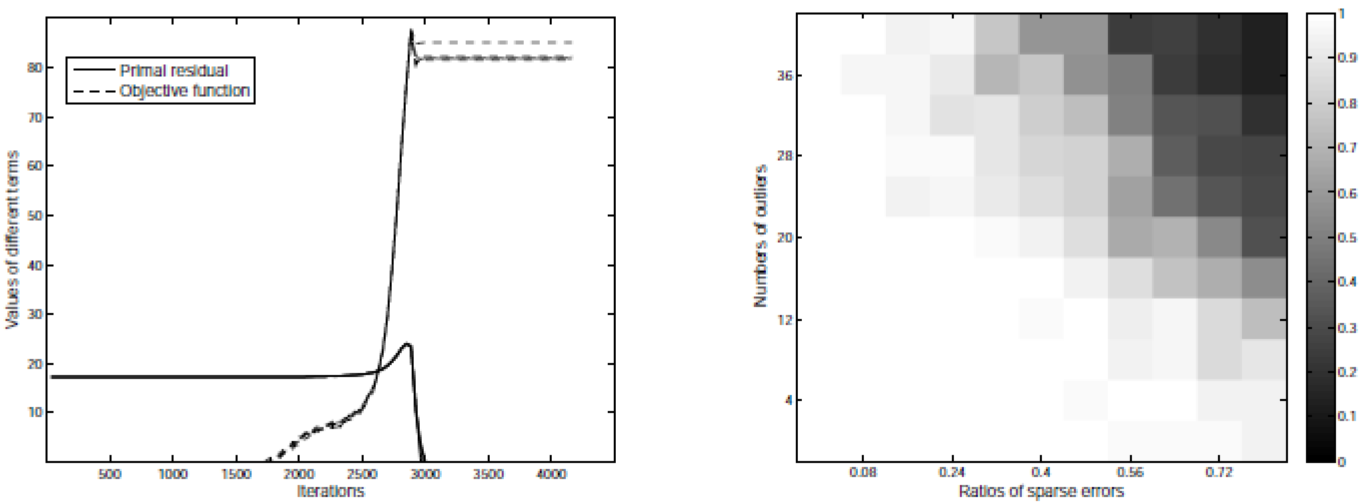
*The IQP subproblem is always equivalent to the following formulation of linear sum assignment problem (LSAP)*

$$\min_{\theta^k} -\mathbf{e}_k^\top\left[\mathbf{Y}_t^\top + \rho\left(\mathbf{L}_{t+1} + \mathbf{E}_{t+1}\right)^\top\right]\mathbf{G}^k\theta^k$$
$$\text{s.t. } \mathbf{J}^k\theta^k = \mathbf{1}_n,\ \mathbf{H}^k\theta^k \leq \mathbf{1}_{n_k},\ \theta^k \in \{0,1\}^{nn_k}$$

- *LSAP can be exactly and efficiently solved using a rectangular-matrix variant of the Hungarian algorithm*

# Convergence analysis

- Convergence property of ADMM for nonconvex problems such as ROML is still an open question
- Simulation




(a) (b)

(a) convergence plot in terms of the primal residual, objective function, and dual variable; (b) recovery precisions under varying numbers of outliers and ratios of sparse errors.

# Choices of feature types in ROML

- Image coordinates

  - *formation* $\mathbf{D}' = \left[(\mathbf{F}^1\mathbf{P}^1)^\top, \ldots, (\mathbf{F}^K\mathbf{P}^K)^\top\right]^\top \in \mathbb{R}^{2K\times n}$

  - *different from* $\mathbf{D} = [\mathrm{vec}(\mathbf{F}^1\mathbf{P}^1), \ldots, \mathrm{vec}(\mathbf{F}^K\mathbf{P}^K)] \in \mathbb{R}^{dn\times K}$

  - *Conditions of use: rigid object, no outliers*

- Local region descriptors

  - *SIFT, HOG, GIST …*

  - *Conditions of use: localizing object with a bounding box*

- Combination of image coordinates and region descriptors

  - *realized by low-dimensional embedding [Torki & Elgammal 10]*

  - *applying in most general settings: non-rigid object, instances of a same object category*

# Experiments – rigid object with 3D motion

Results of different methods on the "Hotel" sequence. Accuracies are measured by the match ratio criteria.

| Methods | DD [4] | SMAC [29] | LGM [3] | RankCon. [17] | One-Shot [20] | Prev [44] | ROML-Pair | ROML |
|---|---|---|---|---|---|---|---|---|
| Accuracies | 99.8% | 84% | 90% | 57% | **100%** | 72% | **100%** | **100%** |

- Matching 15 out of the total 101 frames (every 7th frame), 30 interest points
- DD, SMAC, LGM are pair-wise graph matching methods

  *- enumerating and matching all possible frame pairs for these methods*
- One-Shot *[Torki & Elgammal 10]* is able to match all frames simultaneously

  *- using advanced Shape Context features (computed from image coordinates)*
- ROML performs perfectly even in pair-wise setting

Feature type used in ROML: image coordinates

# Experiments – object instances of a common category

Match ratios of different methods on 6 image sets of different object categories

| Methods | RRWM [16] | SM [37] | TM [24] | RRWHM [35] | ProbHM [62] | DD [60] | OneShot [59] | Prev [63] | ROML |
|---|---|---|---|---|---|---|---|---|---|
| Airplanes | 28% | 54% | 17% | 54% | 32% | 70% | 65% | 87% | **95**% |
| Face | 40% | 57% | 26% | 54% | 14% | 64% | 61% | 53% | **89**% |
| Motorbike | 50% | 46% | 23% | 58% | 28% | 73% | 68% | 89% | **99**% |
| Car | 26% | 39% | 12% | 23% | 12% | 51% | 50% | 59% | **81**% |
| Bus | 13% | 25% | 24% | 43% | 18% | 52% | 44% | 64% | **79**% |
| BoA | 7% | 12% | 6% | 15% | 7% | 12% | 16% | 35% | **75**% |

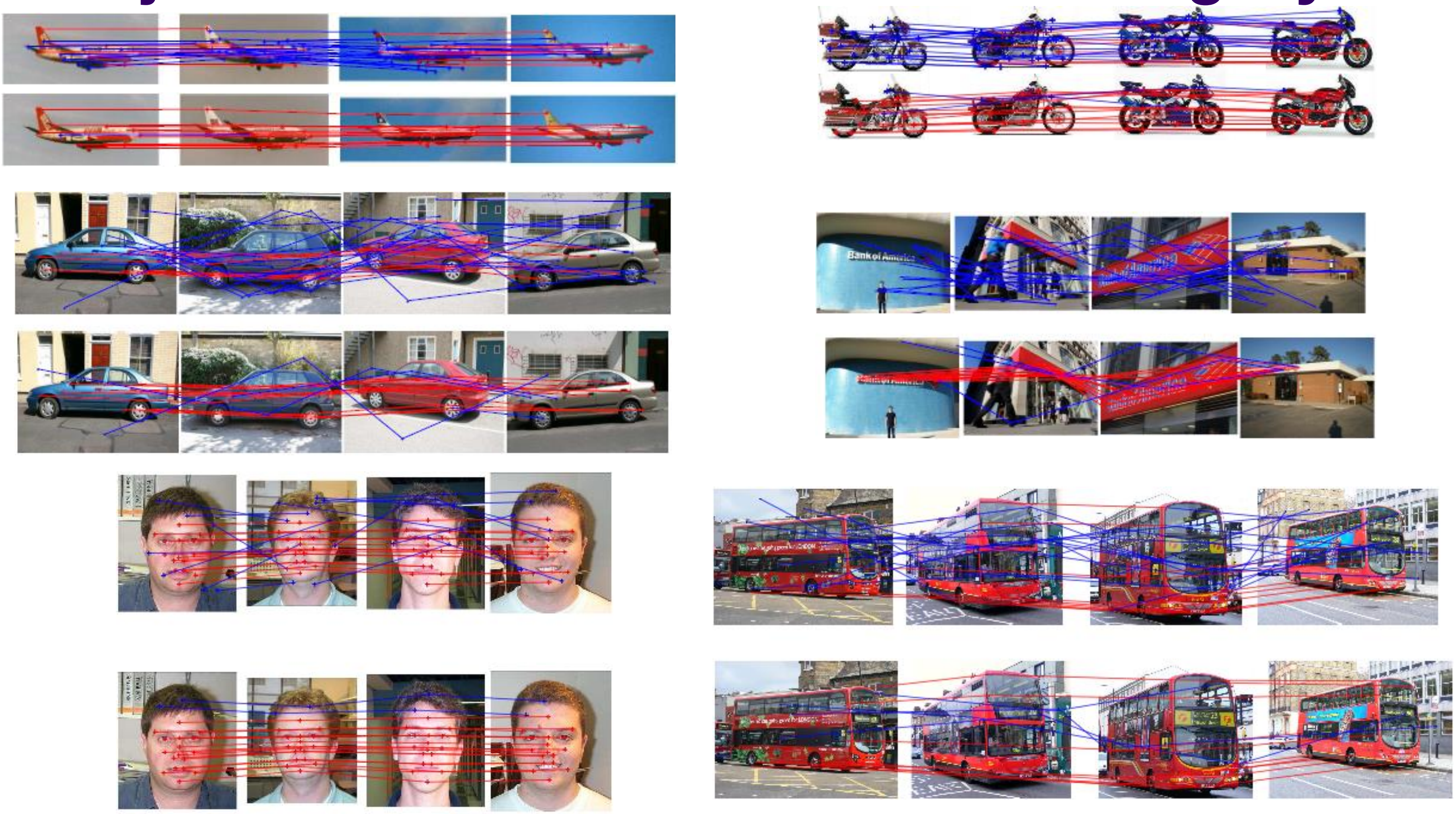
- Number of images per set: 16 ~ 25
- Number of interest points in each image: 26 ~ 174
- Pair-wise graph matching methods: DD, RRWM, SM
- Pair-wise hypergraph matching methods: TM, RRWHM, ProbHM

  *- for graph and hypergraph methods, enumerating and matching all possible image pairs*
- One-Shot is able to match all frames simultaneously

  *- ROML uses the exactly same feature to characterize each interest point as One-Shot does*
- ROML greatly outperforms exiting methods

Feature type used in ROML: learning low-dim. embedding feature by *[Torki & Elgammal 10]*, using Geometric Blur descriptor and image coordinates of each interest point

# Experiments –
# object instances of a common category



For every pair, top: DD [*Torresani et al. 08*], bottom: ROML, red lines: identified ground truth correspondences, blue lines: false correspondences

# Experiments – non-rigid object moving in a video sequence

Match ratios of different methods on the "Tennis" and "Marple" sequences

| Methods | KLT [32] | RRWM [6] | SM [2] | TM [5] | RRWHM [7] | ProbHM [8] | DD [4] | ROML |
|---|---|---|---|---|---|---|---|---|
| Tennis | 3% | 23% | 43% | 13% | 18% | 16% | 57% | **73**% |
| Marple | 4% | 3% | 25% | 8% | 13% | 14% | 23% | **51**% |

- Adapting ROML to object tracking scenario

  *- simply fixing* $\mathbf{P}^1$*, while optimizing the other PPMs* $\{\mathbf{P}^k\}_{k=2}^{K}$

  *- normalizing feature vectors of interest points in the first frame to a larger value of L2 norm*
- Detecting interest points using KLT tracker, labeling inlier points
- KLT tracker generally fails due to abrupt motion or occlusion
- Pair-wise graph matching methods: DD, RRWM, SM
- Pair-wise hypergraph matching methods: TM, RRWHM, ProbHM

  *- for graph and hypergraph methods, matching between the 1st frame and each of the other frames*

Feature type used in ROML: learning low-dim. embedding feature by *[Torki & Elgammal 10]*, using Geometric Blur descriptor and image coordinates of each interest point

# Experiments – non-rigid object moving in a video sequence

For every pair, top: DD [Torresani, Kolmogorov, Rother 08], bottom: ROML, red lines: identified ground truth correspondences, blue lines: false correspondences

# Experiments – non-rigid object moving in a video sequence

Failure cases without adapting ROML to the tracking scenario

with inlier labelling in the 1st frame

without inlier labelling in the 1st frame

with inlier labelling in the 1st frame

without inlier labelling in the 1st frame

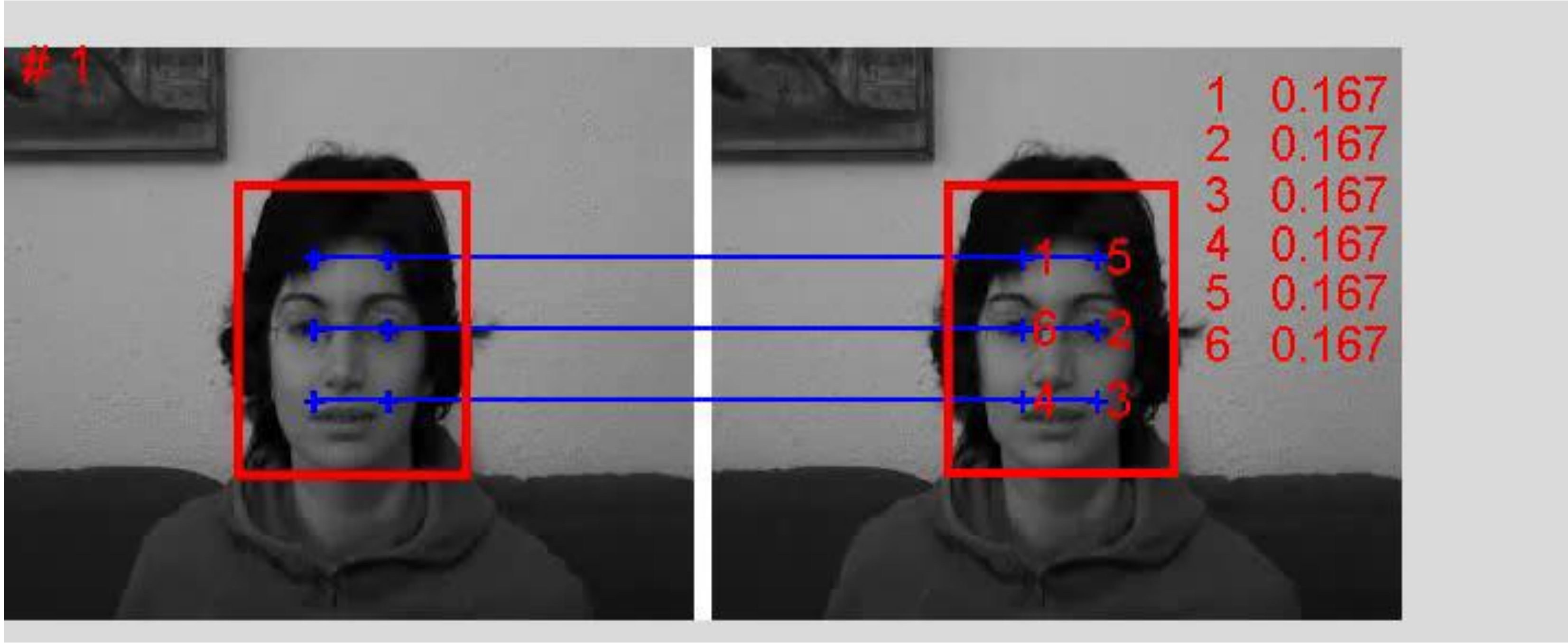
Red lines: identified ground truth correspondences

Blue lines: false correspondences

# Outline

- Background knowledge and motivation
- ROML: Robust Object Matching using Low-rank constraint
  - *Formulation*
  - *Solving algorithm*
  - *Results*
- Applications of ROML to other data problems
  - *Tracking*
  - *Ambiguous learning*

# Applications of ROML: Handling partial occlusion in visual tracking




*Zhang*, Jia*, Xu, Ma, Ahuja, "Partial Occlusion Handling for Visual Tracking via Robust Part Matching", CVPR, 2014*

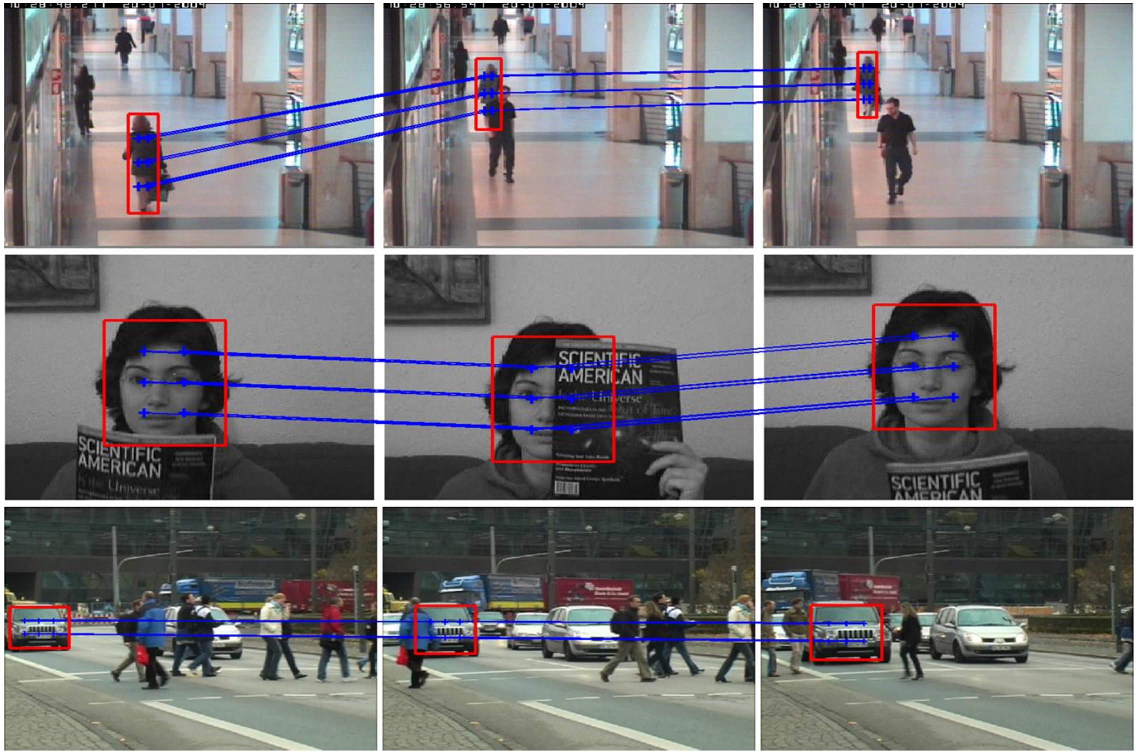
# Part Matching Tracker

- The challenge of occlusion


*Frames of two different video sequences with partial occlusion*

# Part Matching Tracker

- Partial occlusion can be addressed via robust *part matching across multiple frames overtime*




*+ denotes the positions of parts, and the blue lines show their correspondences*

# Part Matching Tracker

Incorporating spatial-temporal locality constraints

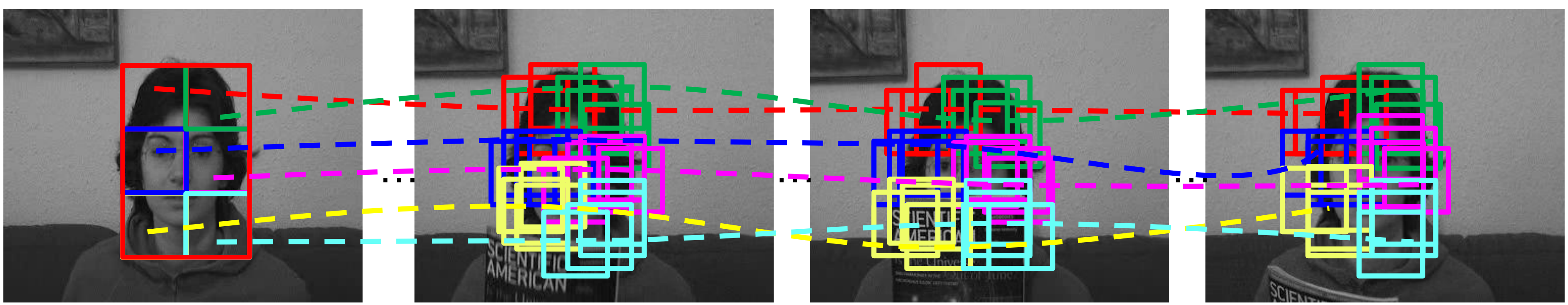
- The formulation

$$\min_{\{\mathbf{P}^k \in \mathcal{P}^k\}_{k=1}^K, \{\mathbf{L}_i, \mathbf{E}_i\}_{i=1}^n} \sum_i \|\mathbf{L}_i\|_* + \lambda \|\mathbf{E}_i\|_1$$

$$\text{s.t.} \quad \mathbf{D}_i = \mathbf{L}_i + \mathbf{E}_i,\ i = 1, \ldots, n.$$

$$\mathcal{P}^k = \{\mathbf{P}^k \big| \mathbf{P}^k \in \{0,1\}^{n_k \times n}, \mathbf{1}_{n_k}^\top \mathbf{P}^k = \mathbf{1}_n^\top, \mathbf{P}^k \mathbf{1}_n \leq \mathbf{1}_{n_k}, \mathbf{A}^k \mathbf{P}^k \mathbf{1}_n = \mathbf{1}_n\},$$

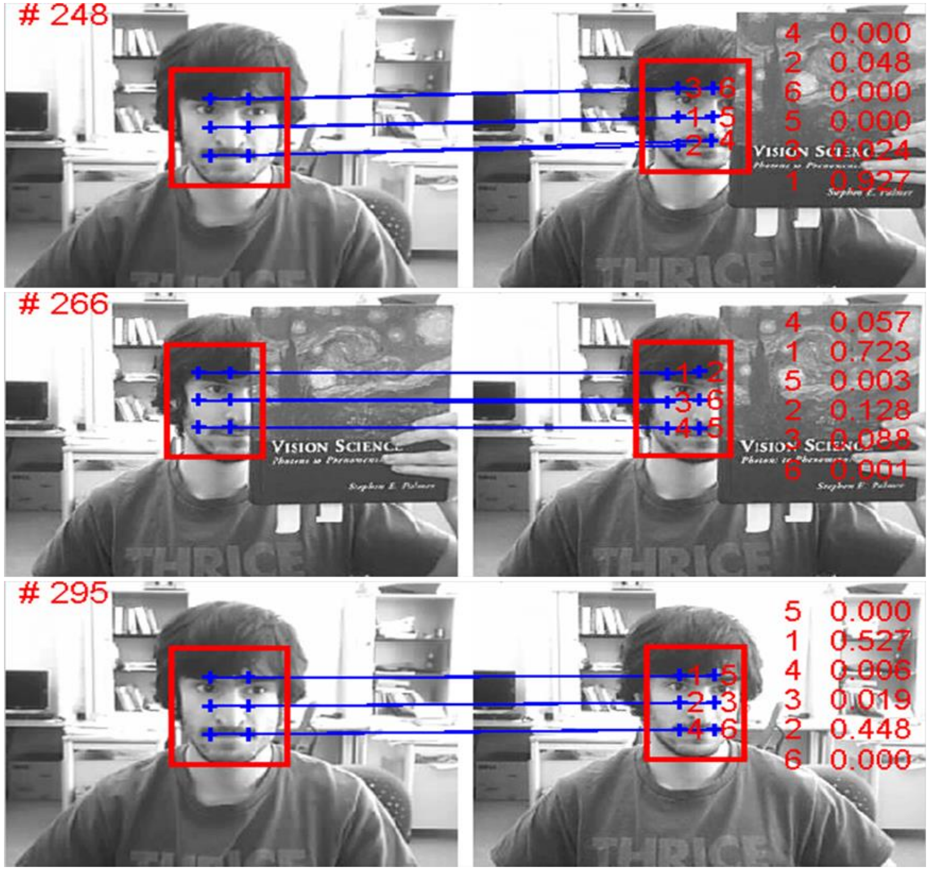
$$\mathbf{D_i} = [\mathbf{F}^1 \mathbf{p_i}^1, \ldots, \mathbf{F}^K \mathbf{p_i}^K] \qquad \mathbf{p_i}^k = \mathbf{P}^k \mathbf{e}_i$$



K frames

# Part Matching Tracker

- Illustration of PMT's robustness against partial occlusion



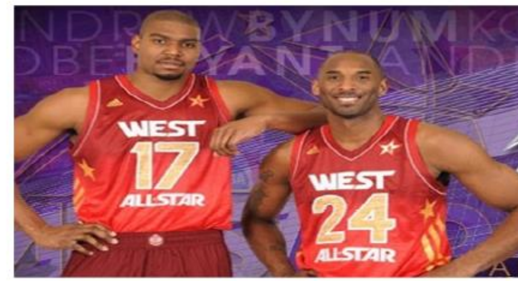

- *The numbers of "1" to "6" index different parts of the face.*
- *"1" ranks highest and "6" ranks lowest in terms of confidence score of part matching.*

# Applications of ROML: Learning from ambiguously labelled images

- A motivating example

A forceful President ***Barack Obama*** put Republican challenger ***Mitt Romney*** on the defensive on foreign policy issues on Monday night, scoring a solid victory in their third and final debate just 15 days before Election Day. [ News From CNN ]

President ***Barack Obama***, Italian Prime Minister ***Silvio Berlusconi***, center, and Russian President ***Dmitry Medvedev***, right, smile during a group photo at the G20 Summit in London. [ News From Washington Post ]




***Bryant*** and ***Andrew Bynum*** have been named Western Conference All-Star starters at guard and center respectively. This is ***Bryant***'s 14th time starting the league's annual showcase game. All-Star nod. [ News from NBA ]

- Each image contains some samples of interest (e.g., human faces).
- Each caption has labels with the true ones included.
- Task: *to learn classifiers from these ambiguously labelled images*

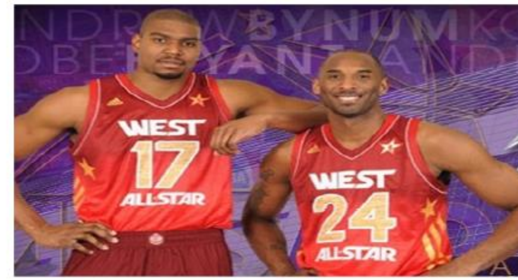
*Zeng, Xiao, Jia, Chan, Gao, Xu, Ma, "Learning by Associating Ambiguously Labeled Images", CVPR, 2013*

# Applications of ROML: Learning from ambiguously labelled images

- A motivating example

A forceful President ***Barack Obama*** put Republican challenger ***Mitt Romney*** on the defensive on foreign policy issues on Monday night, scoring a solid victory in their third and final debate just 15 days before Election Day. [ News From CNN ]

President ***Barack Obama***, Italian Prime Minister ***Silvio Berlusconi***, center, and Russian President ***Dmitry Medvedev***, right, smile during a group photo at the G20 Summit in London. [ News From Washington Post ]




***Bryant*** and ***Andrew Bynum*** have been named Western Conference All-Star starters at guard and center respectively. This is ***Bryant***'s 14th time starting the league's annual showcase game. All-Star nod. [ News from NBA ]

- Each image contains some samples of interest (e.g., human faces).
- Each caption has labels with the true ones included.
- Task: *to learn classifiers from these ambiguously labelled images*

*Make use of the information embedded in the relations between samples and labels, both within each image and across the image set.*

# Ambiguous learning by class-wise low-rank assumption

- **Observation**

  Samples of the same class, assuming they are similar, repetitively appear in the ambiguously labelled image set.

  → *Class-wise low-rank assumption*

- **Leveraging class-wise low-rank assumption**

  To identify samples of the same class from each image

  To associate them across the image set

  → *Ambiguous learning problem solved*

# Ambiguous learning by class-wise low-rank assumption

- **Observation**

  Samples of the same class, assuming they are similar, repetitively appear in the ambiguously labelled image set.

  → *Class-wise low-rank assumption*

- **Leveraging class-wise low-rank assumption**

  To identify samples of the same class from each image

  To associate them across the image set

  → *Ambiguous learning problem solved*

*Again, PPM optimization for sample-label correspondences!*

# Ambiguous learning by class-wise low-rank assumption

- **Formal definition of PPM** (for the image $n$ of the total $N$ images)

$$\mathbf{P}_n \in \{0,1\}^{K_n \times \bar{K}}$$

s.t. $$\mathbf{1}_{K_n}^T \mathbf{P}_n (\mathbf{1}_{\bar{K}} - \mathbf{t}_n) = 0$$

*Enforcing samples in the image n can only be associated with classes that have labels appearing in the caption.*

$$\mathbf{P}_n \mathbf{1}_{\bar{K}} = \mathbf{1}_{K_n}$$

*Enforcing the constraint/assumption that every sample in the image belongs to a class.*

$$\mathbf{1}_{K_n}^T \mathbf{P}_n \leq \mathbf{1}_{\bar{K}}^T$$

*Enforcing the constraint/assumption that samples of the same class cannot appear in the same image.*

$K_n$ no. of samples in image *n*

$\bar{K}$ the assumed no. of classes

$\mathbf{t}_n \in \{0,1\}^{\bar{K}}$ Binary vector indicating label appearance in the caption of image *n*

# Ambiguous learning by class-wise low-rank assumption

- **Formulation** – *an extension of ROML*

*Class-wise low-rank matrices*

$$\min_{\substack{\{\mathbf{L}_i,\mathbf{E}_i\}_{i=1}^{\bar{K}} \\ \{\mathbf{P}_n \in \bar{\mathcal{P}}_n\}_{n=1}^{N}}} \sum_{i=1}^{\bar{K}} \|\mathbf{L}_i\|_* + \lambda \|\mathbf{E}_i\|_1,$$

$$\text{s.t. } \mathcal{F}(\{\mathbf{P}_n\}_{n=1}^{N}) = [(\mathbf{L}_1 + \mathbf{E}_1)^T, \ldots, (\mathbf{L}_{\bar{K}} + \mathbf{E}_{\bar{K}})^T]^T$$

*Again, ADMM based solving algorithm, with feature normalized!*

# *Thank you! And questions?*

# References


- Kui Jia, Tsung-Han Chan, Zinan Zeng, Shenghua Gao, Gang Wang, Tianzhu Zhang, and Yi Ma, *"**ROML: A Robust Feature Correspondence Approach for Matching Objects in A Set of Images**"*, arXiv:1403.7877, 2014.

  *Data and code available at [https://sites.google.com/site/kuijia/research/roml](https://sites.google.com/site/kuijia/research/roml)*

- Tianzhu Zhang*, Kui Jia*, Changsheng Xu, Yi Ma, and N. Ahuja, *"**Partial Occlusion Handling for Visual Tracking via Robust Part Matching**"*, IEEE Conference on Computer Vision and Pattern Recognition, 2014. (* indicates equal contributions)

- Zinan Zeng, Shijie Xiao, Kui Jia, Tsung-Han Chan, Shenghua Gao, Dong Xu, and Yi Ma, "***Learning by Associating Ambiguously Labeled Images***", IEEE Conference on Computer Vision and Pattern Recognition, 2013.

- Zinan Zeng, Tsung-Han Chan, Kui Jia, and Dong Xu, "***Finding Correspondence from Multiple Images via Sparse and Low-rank Decomposition***", European Conference on Computer Vision, 2012.

# A job post

- *A few Research Assistant positions are available in my group at University of Macau, Macau SAR, China*
- *Payment is similar/identical to RA jobs in universities in Hong Kong (e.g., 1,5000 HKD per month)*
- *Interested students may send your CV to [kuijia@umac.mo](mailto:kuijia@umac.mo) or contact me via QQ (124401525) for a casual discussion of the potential research topics*