Low-Rank Analytics for Explosive Detection

Yun Raymond Fu, Assistant Professor
ECE, Northeastern University

Conclusions

- Explosives-related sensory data are often under uncertainties: noises, cross-modalities, lack of training samples, large-scale data, over-fitting of models, etc.
- Low-rank analytics crates a promising algorithmic tool set to mitigate these uncertainties.
- Low-rank analytics based transfer learning, manifold learning, and subspace learning are demonstrated to be effective feature extraction methods of ATR.

Explosive Detection, Many Ways

- Explosive detection --- A non-destructive inspection process to determine whether a container contains explosive materials
- Many possible ways to approach


Dogs


Explosive detection


Mechanical scent detection


Honey bees


X-ray machines


Spectrometry

Multi-Sensor Cross-Modality Problem

- Data source:

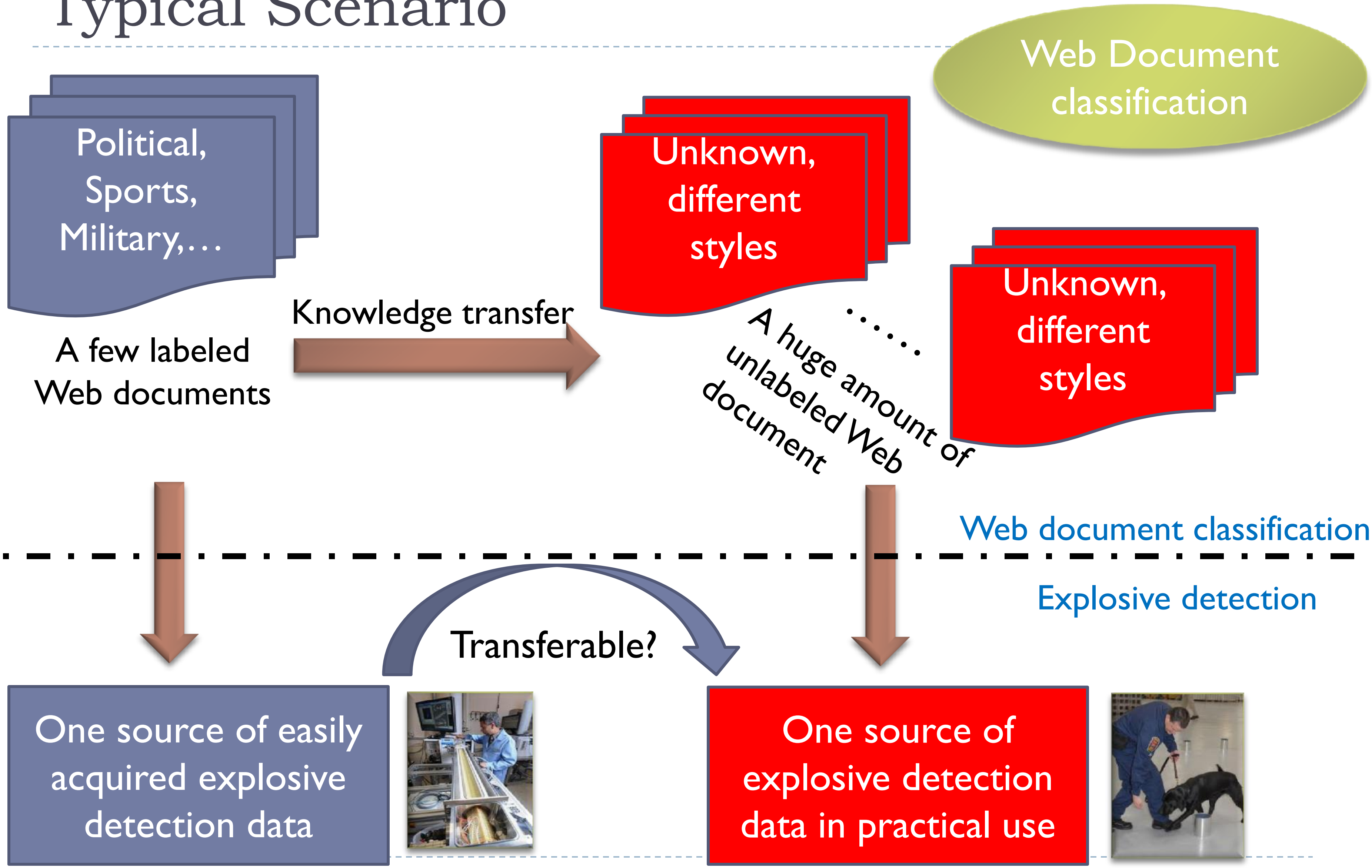
Sense of smell of dog	Bee's reaction	Machine olfaction	Images from different spectrals	X-ray images

- Sensor: CT, XBS, MMW, Trace, QR, XRD, Fused system

- How to better use multi-sensor cross-modality data?
  - Noise/outlier(anomaly) detection
  - Feature selection
  - Knowledge transfer

One source of data is easily acquired for training, but not applicable in test, while another source of data is opposite. Can we transfer the knowledge from the former to the latter?

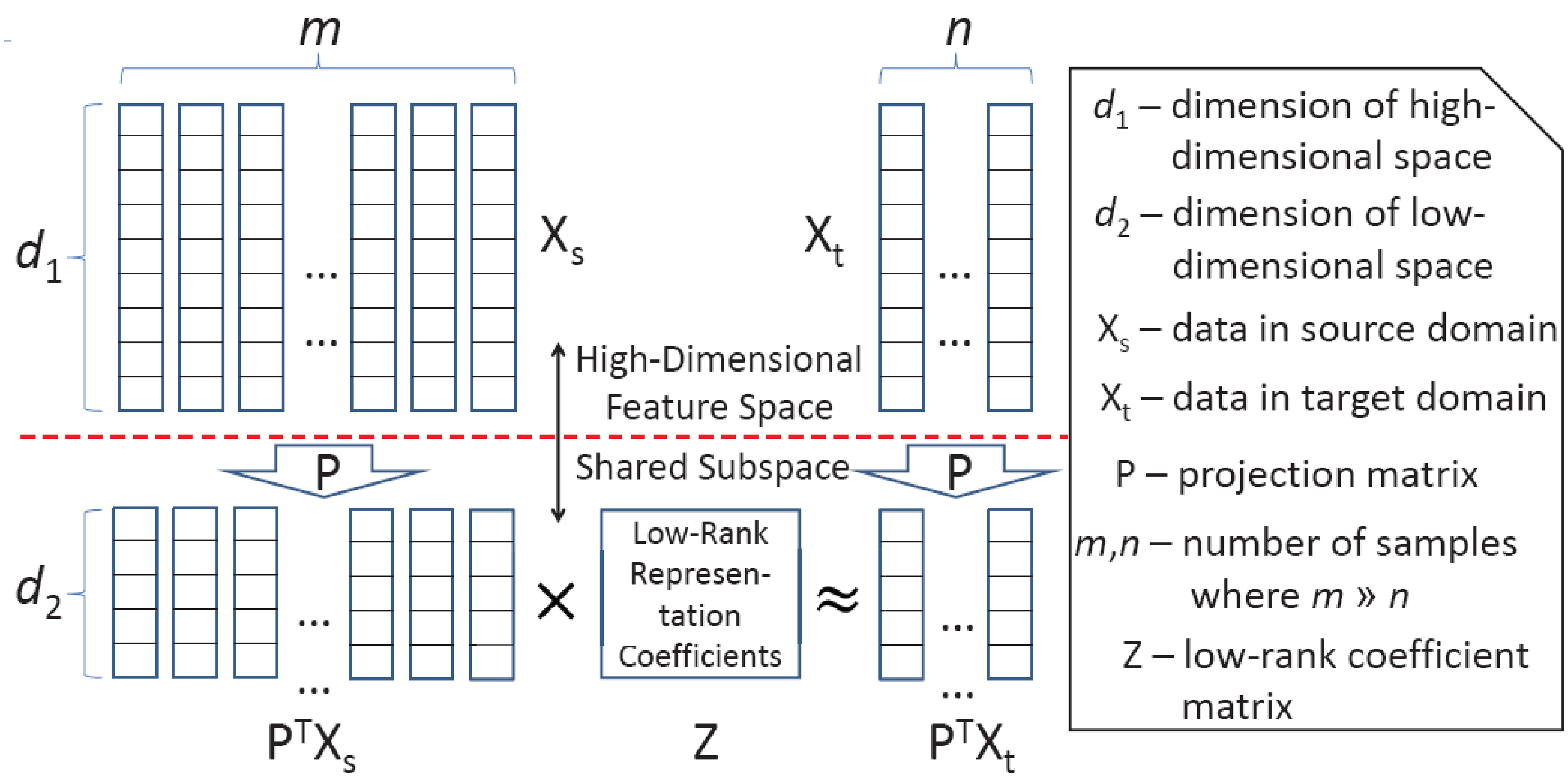
Knowledge Transfer in Machine Learning, A Typical Scenario



Why Transfer?

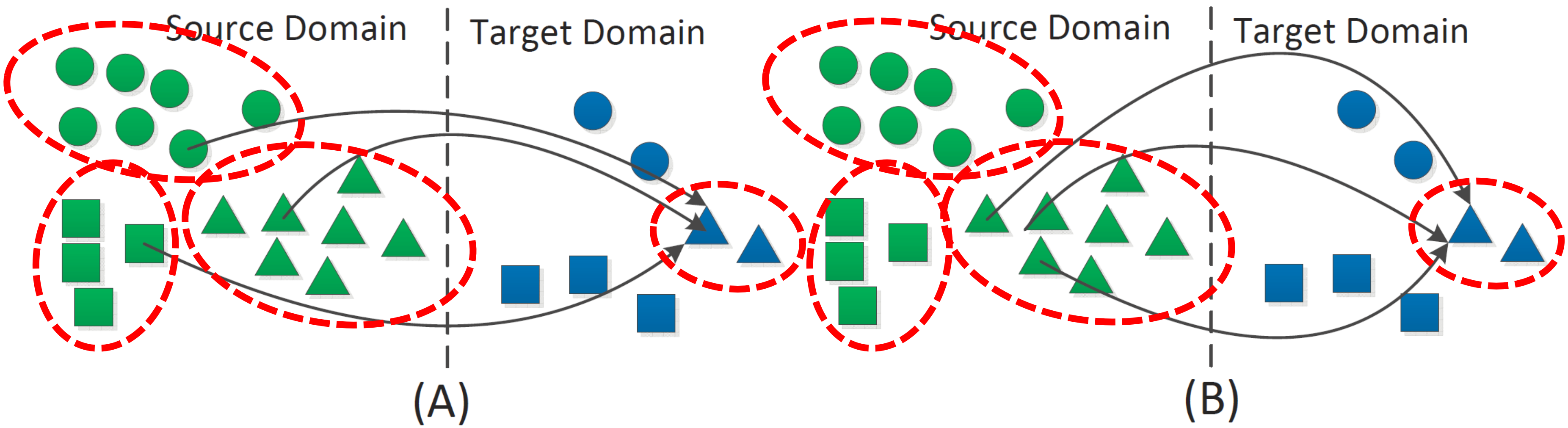
- One common assumption in classification problems is the training/testing consistency of the data.
- This cannot be always satisfied, especially in complex applications common in many areas:
  - web document classification,
  - sentiment analysis,
  - image annotation,
  - face recognition.
- How to apply previous well-labeled data to a huge amount of unseen data with possibly different distributions?
- The correct way might be using only a few data in the source domain within an appropriate subspace to reconstruct a specific target data, as shown in the above figure.

Our Contribution



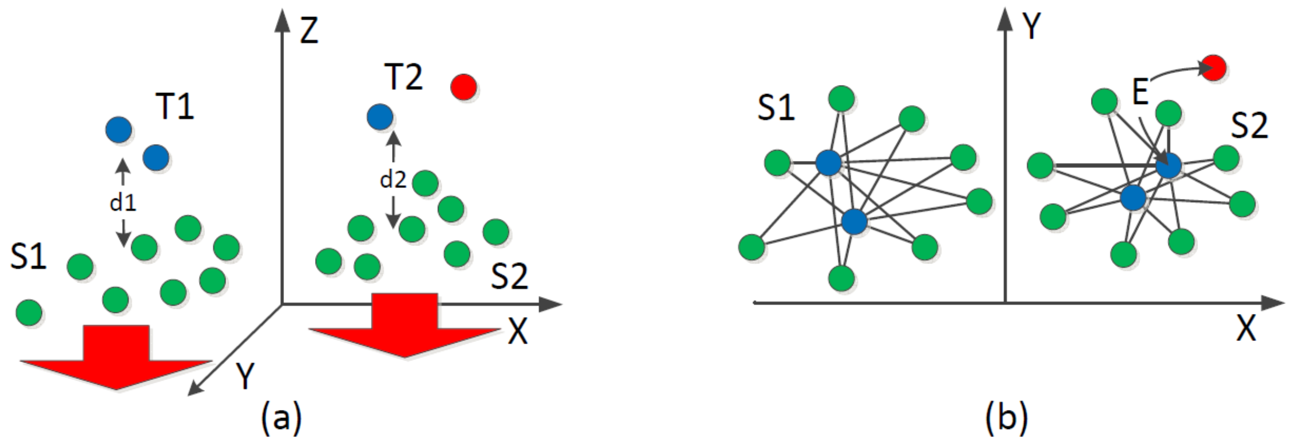
- Contributions
  - A novel method for transfer learning via low-rank representation, which we call low-rank transfer subspace learning (LTSL).

Problem Formulation



- A given data set is seldom well described by a single subspace, rather, data are more likely lying in several subspaces.
- Suppose we adopt source data to linearly represent target data to achieve the purpose of knowledge transfer.
- For over-complete source data that span the entire feature space, however, we could always obtain trivial solutions.
- The correct way might be using only a few data in the source domain within an appropriate subspace to reconstruct a specific target data, as shown in the above figure.

Problem Formulation

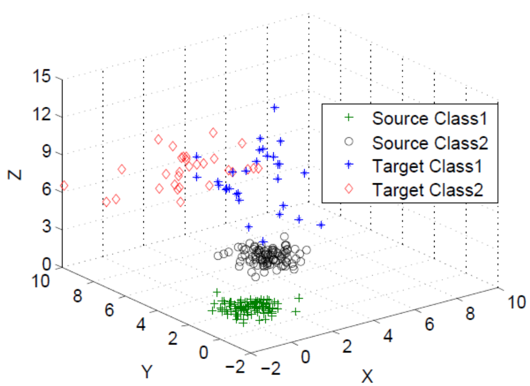


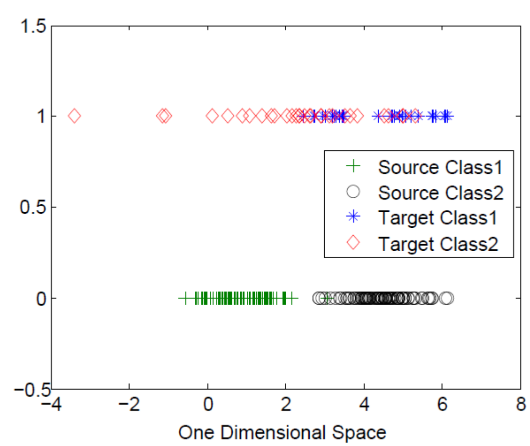
- In the original data space, the mapping between source and target domain are not necessarily the best!
- Extreme case in above figure is blue points in (a) are hardly represented by green ones
- We consider the knowledge transfer in some subspace spanned by P, plus an error term E, where mapping are clearly shown in (b)

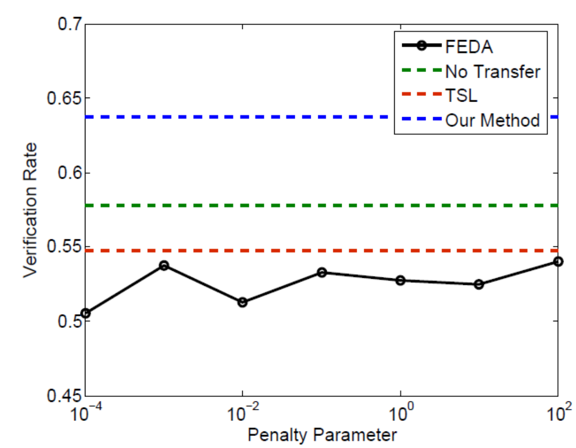
$$
\begin{aligned}
P, Z, E = \underset{P,Z,E}{\arg\min}\, F(P, X_s) + \operatorname{rank}(Z) + \lambda \|E\|_{2,1}, \\
\text{s.t., } P^T X_t = P^T X_s Z + E.
\end{aligned}
$$

Solution and Results

- The former problem can be solved by augmented Lagrangian multipliers (ALU).
- Experiment 1, synthetic data
  - Two classes in the source domain, each class has 100 samples!
  - Two classes in the target domain, each class has 30 samples!
  - Mess target data in figure (left) are now separable in figure (middle) by mapping them to corresponding source data.







Experimental Results

- Experiment 2, : Kinship verification, UB KinFace database

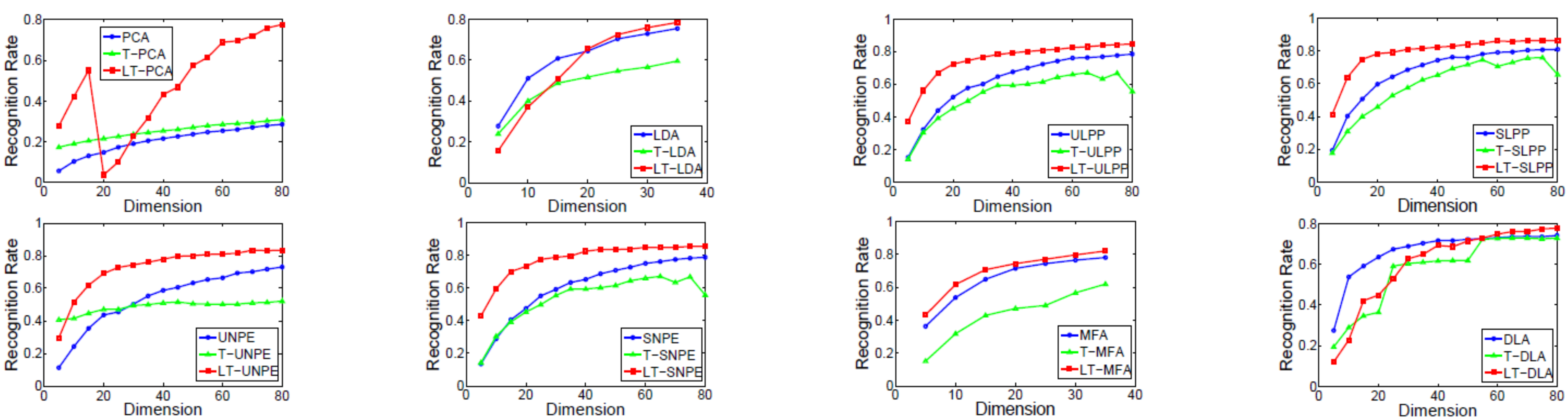
BEST RESULTS AND DIMENSIONS OF KINSHIP VERIFICATION.

Method	PCA	SLPP	ULPP	SNPE	UNPE	MFA	DLA
No Transfer	53.98%(11)	55.00%(9)	57.74%(11)	53.26%(9)	54.26%(21)	52.74%(17)	54.74%(35)
TSL	54.78%(25)	54.02%(3)	54.02%(11)	50.74%(9)	53.26%(9)	52.24%(3)	53.98%(39)
Our Method	**56.57**%(19)	**57.17**%(17)	**63.72**%(11)	**54.60**%(11)	**58.80**%(3)	**54.50**%(35)	**55.00**%(33)

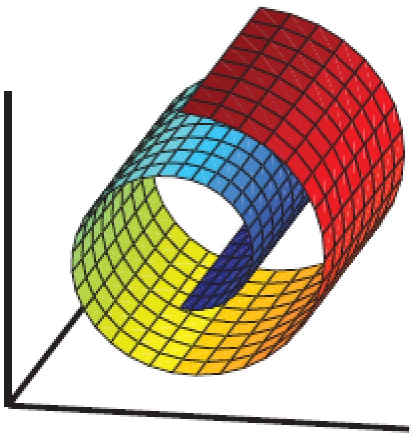
- Experiment 3: Face recognition, from Yale B to CMU PIE

BEST RESULTS AND DIMENSIONS OF PROBLEM Y2P.

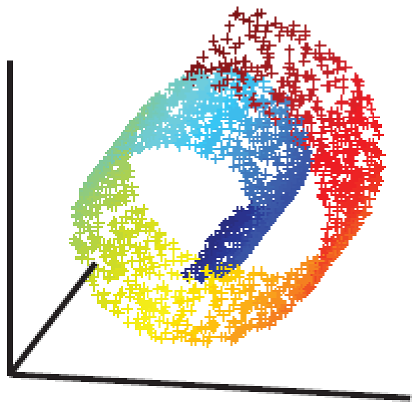
Method	PCA	SLPP	ULPP	SNPE	UNPE	LDA	MFA	DLA
No Transfer	28.6%(80)	80.7%(80)	78.3%(80)	78.6%(80)	73.1%(80)	75.4%(35)	78.1%(35)	74.1%(80)
TSL	30.9%(80)	75.7%(75)	67.0%(65)	67.0%(65)	52.1%(80)	59.6%(35)	62.0%(35)	72.8%(80)
Our Method	**77.6**%(80)	**86.1**%(75)	**84.6**%(80)	**85.2**%(75)	**83.5**%(80)	**78.4**%(35)	**82.2**%(35)	**77.8**%(80)



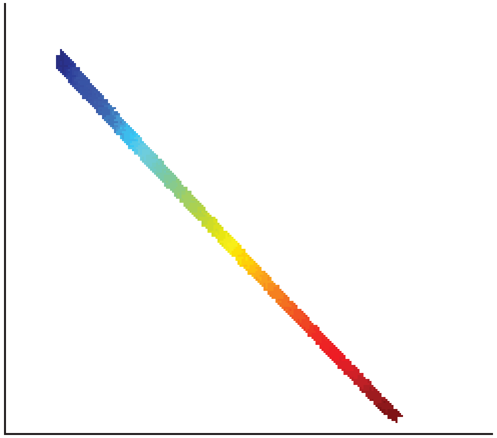
Manifold with Noise Effect



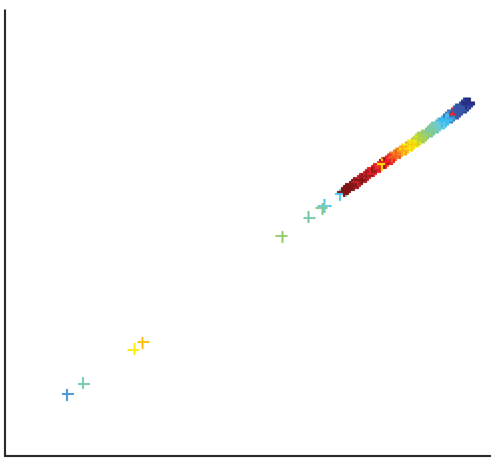
Original Manifold



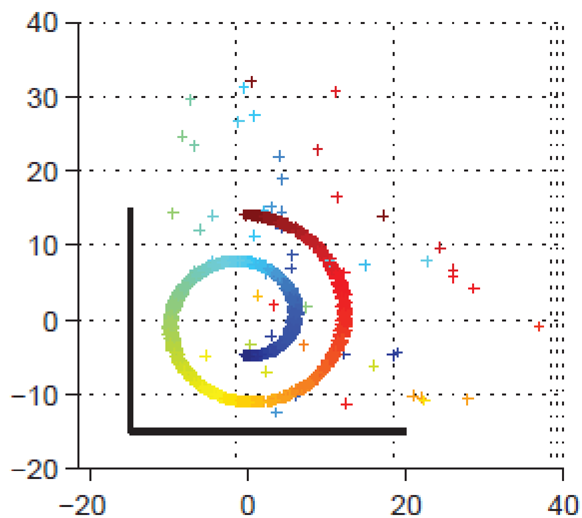
Original Manifold Sampling



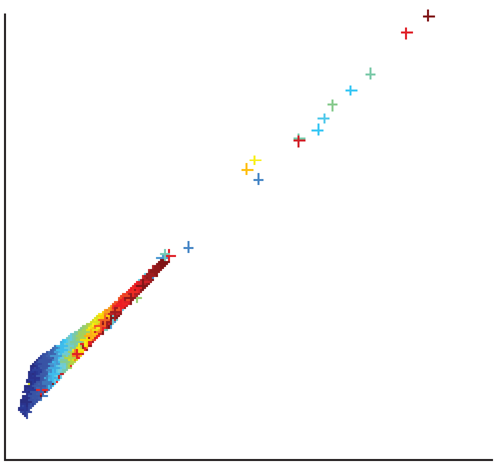
LLE Result (Original Manifold)



LLE with Low Rank Recovery



Perturbed Manifold



LLE without Recovery

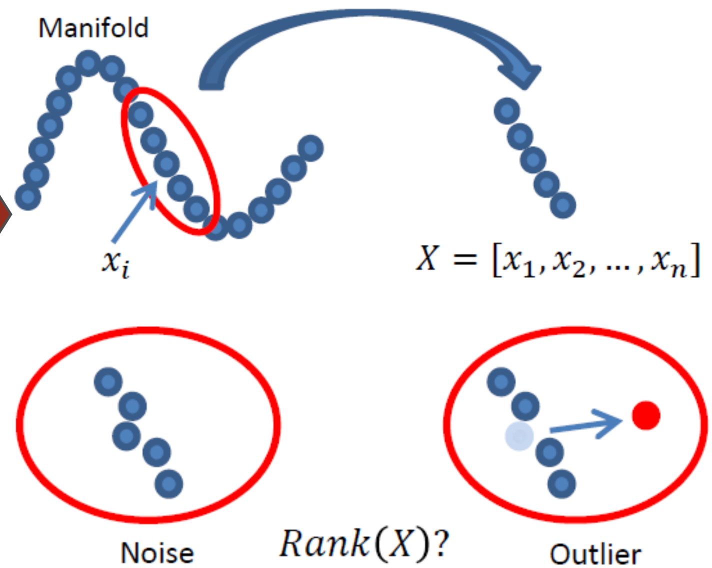
Northeastern University

Robust Manifold by Low-Rank Recovery

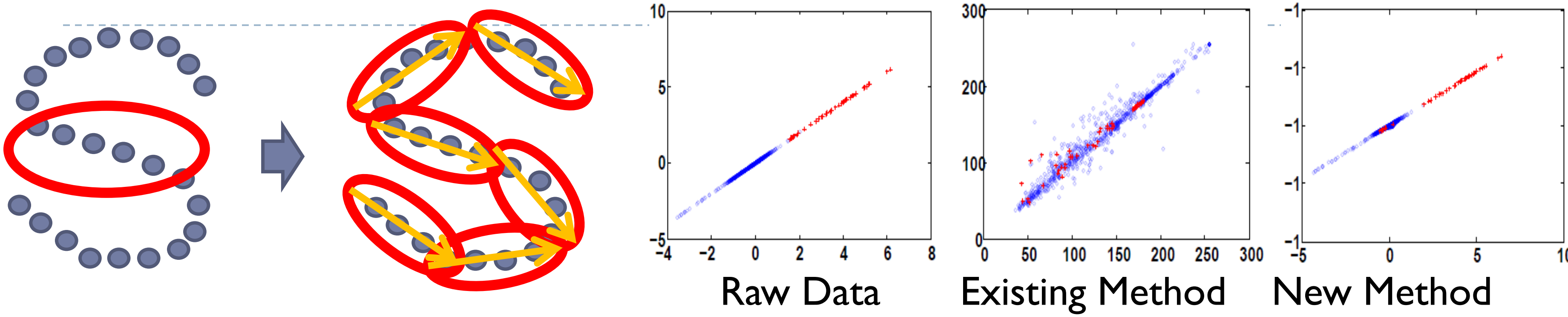
Real-world ATR data are large scale, unbalanced in dynamic sampling, and easily affected by noises and outliers, which are difficult to represent.

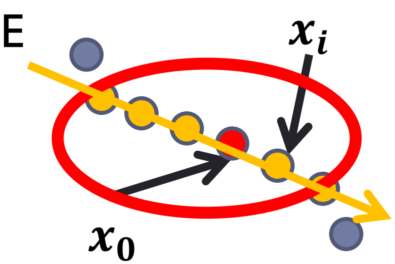
Low-rank matrix recovery can deal with noises and outliers for data reconstruction.

Automated, real-time, and robust description of ATR data space under uncertainty.

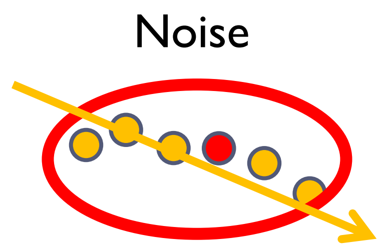


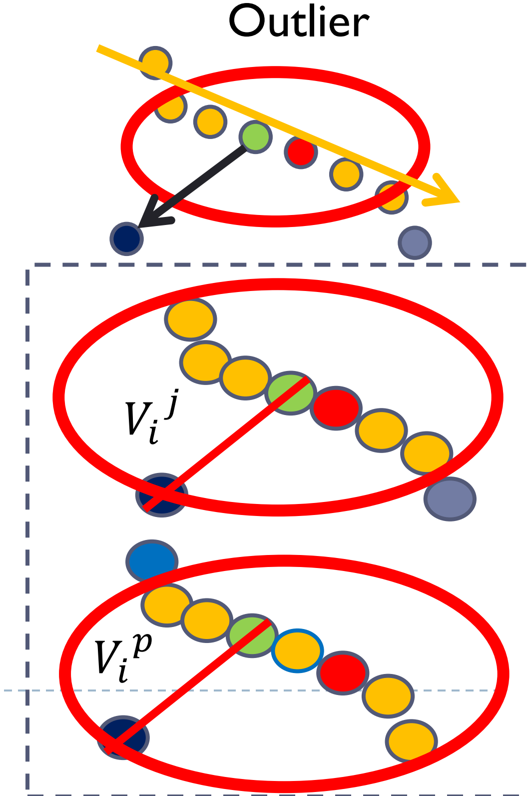
Stabilized Manifold Learning





$$x_0 = \sum_{i=1}^{k} \omega_i x_i$$



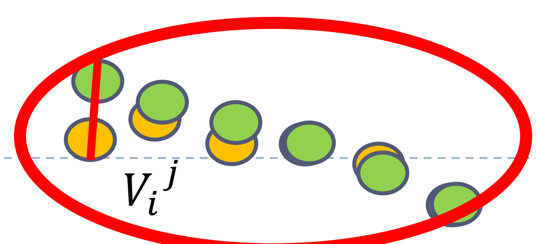


Voting for Outlier Detection

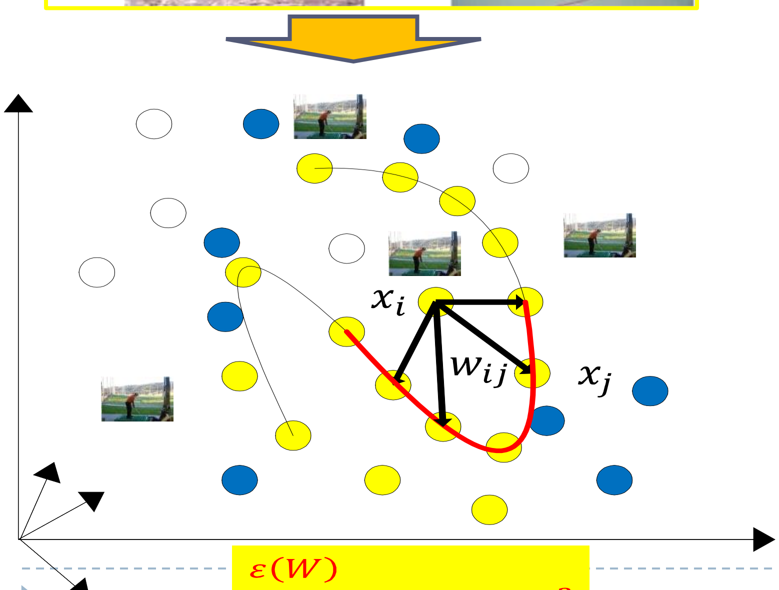
$$R_i^{\,j} = 1 - {V_i^{\,j}} \Big/ {\sum_{j=1}^{k} V_i^{\,j}}$$

$$R_i = \sum_{i=1}^{n} R_i^{\,j}$$

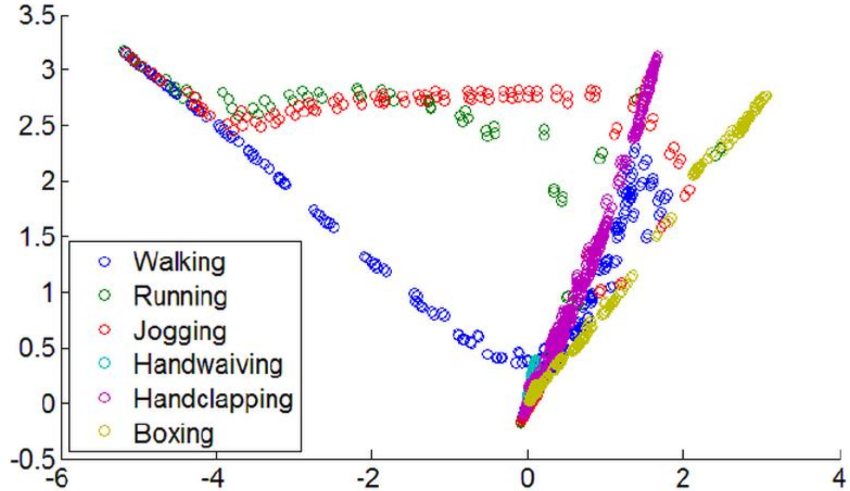
$\min rank(M_T) + \eta E$
s.t. $M = M_T + E$

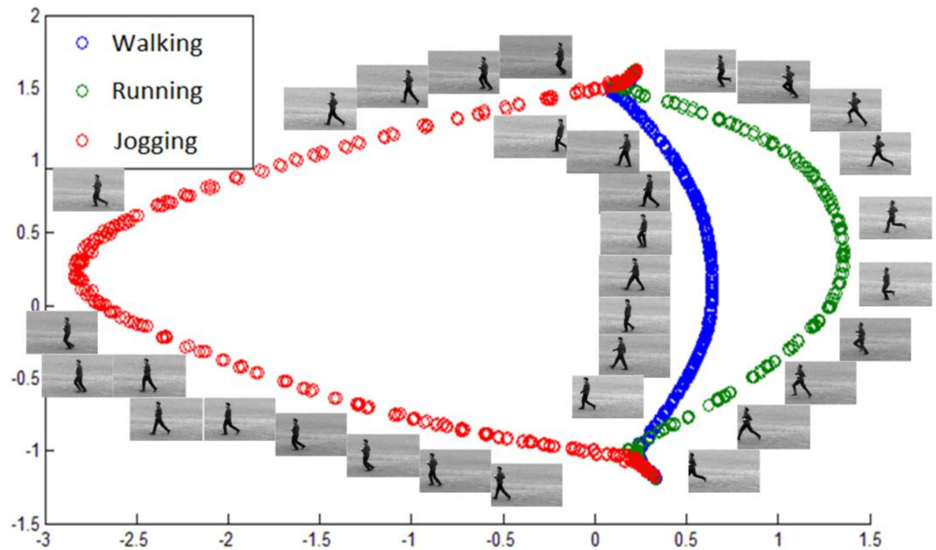


Stabilized Manifold Learning



$\varepsilon(W)$
$= \Sigma_i |x_i - \Sigma_j W_{ij} x_j|^2$





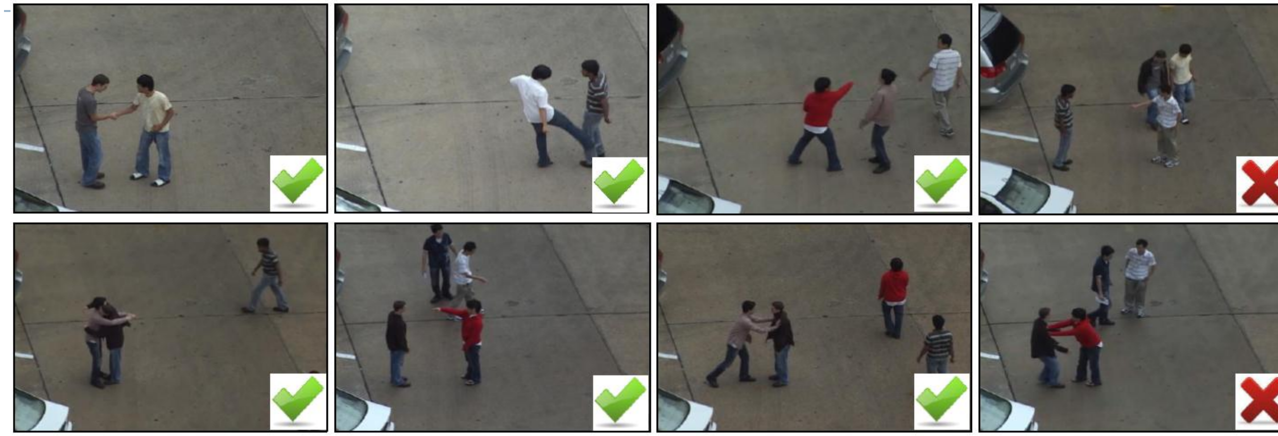
Northeastern University

Results on UT-Interaction dataset

- 6 interaction classes, 60 videos, 23 interactive phrases, 16 motion attributes

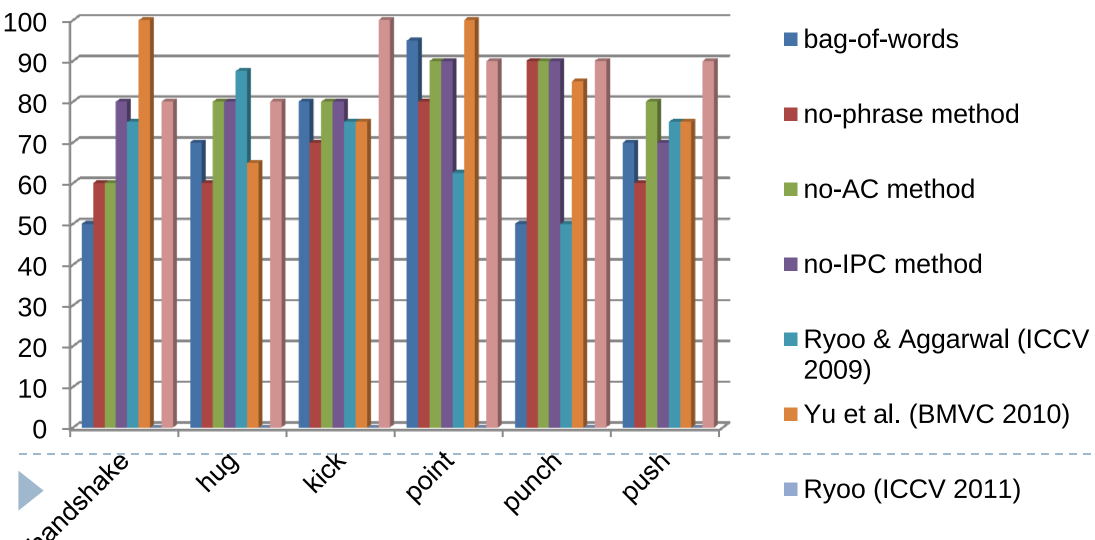
	handshake	hug	kick	point	punch	push
handshake	0.80	0.00	0.10	0.00	0.00	0.10
hug	0.00	0.80	0.20	0.00	0.00	0.00
kick	0.00	0.00	1.00	0.00	0.00	0.00
point	0.00	0.00	0.00	0.90	0.10	0.00
punch	0.00	0.00	0.00	0.00	0.90	0.10
push	0.00	0.00	0.10	0.00	0.00	0.90

Confusion matrix of our method
Accuracy = 88.33%



Classification examples of our method

Recognition accuracy (%) of methods



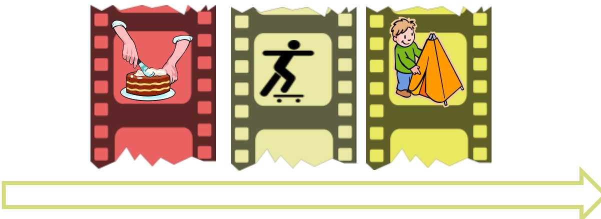
Recognition accuracy (%) of methods

Methods	Overall
bag-of-words	68.33
no-phrase method	70
no-AC method	80
no-IPC method	81.67
Ryoo & Aggarwal (ICCV 2009)	70.8
Yu et al.(BMVC 2010)	83.33
Ryoo (ICCV 2011)	85
Our method	**88.33**

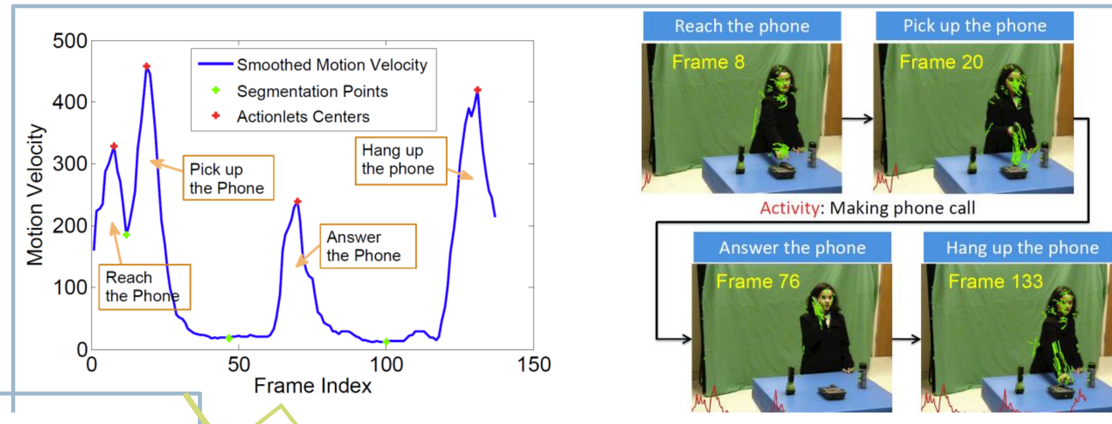
Northeastern University

Activity Prediction

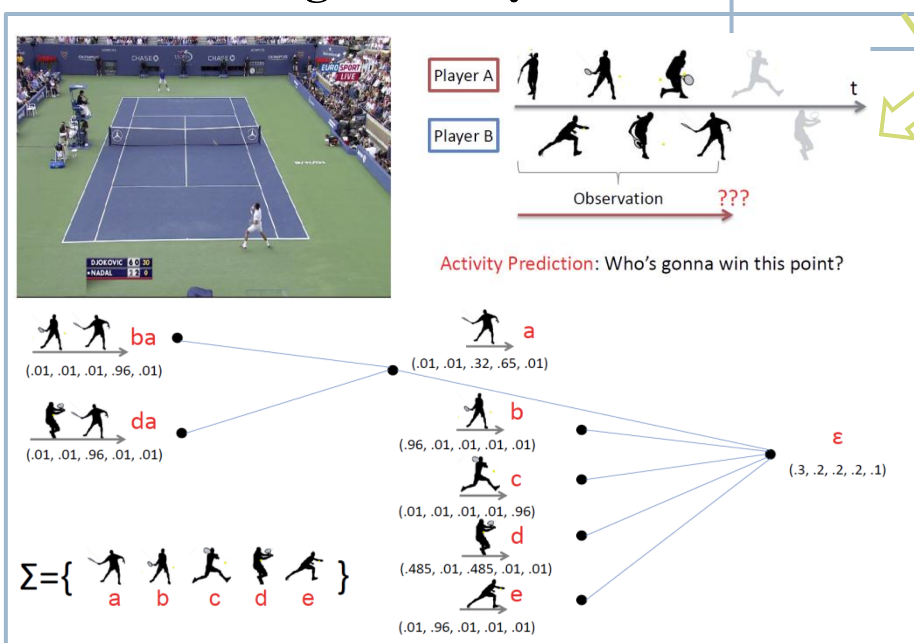
Activity videos



Actionlets: Activity Decomposition

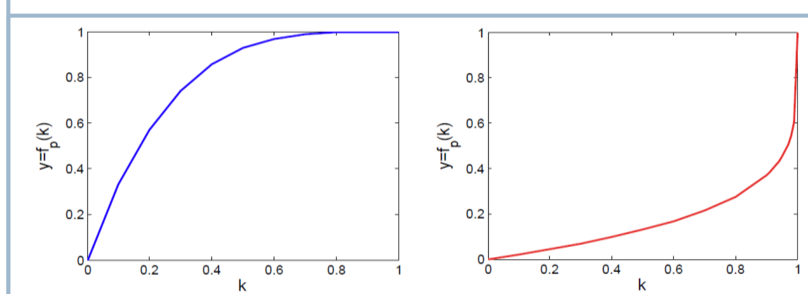


PST: Modeling Causality



Prediction Model

PAF: Modeling Predictability



(a) Early predictable problem (b) Late predictable problem

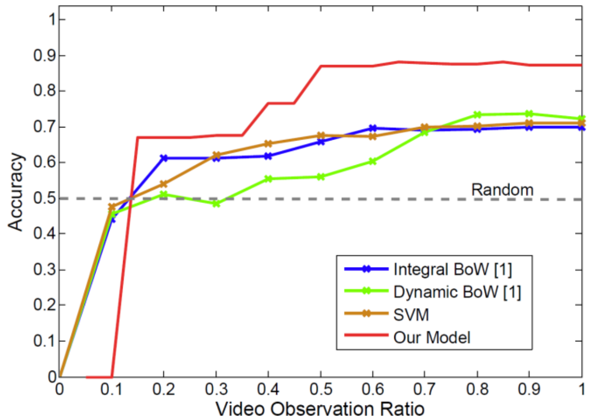
Activity Prediction- ECCV2012

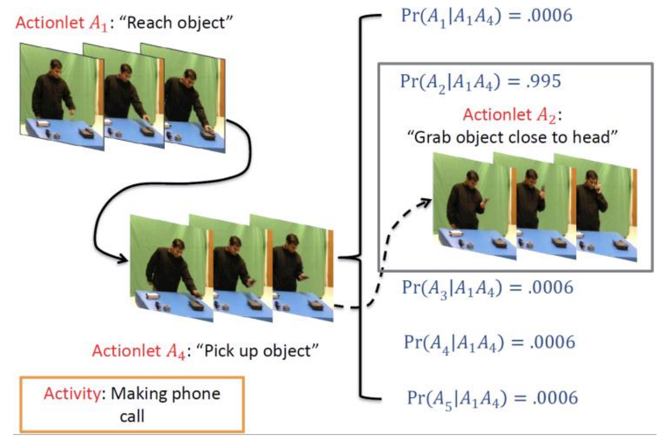
Kang Li, Jie Hu, and Yun Fu

Northeastern University

Results on Activity Prediction

On Daily Activity Dataset (Mid-level complex)



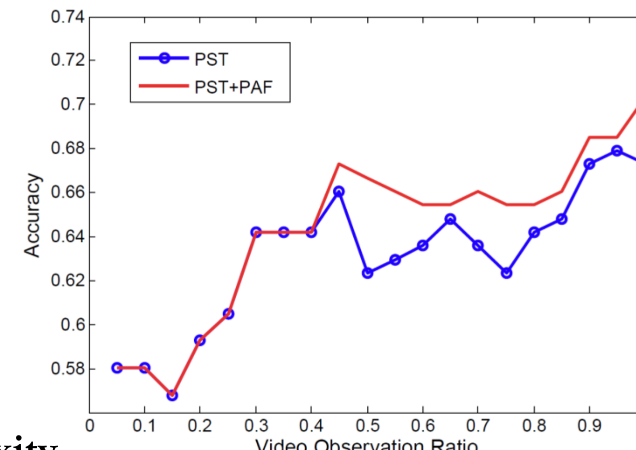


Activity class prediction:
Outperform state of the art with a large margin

Our method can also predict NEXT move

On Tennis Game Dataset (High-level complex)

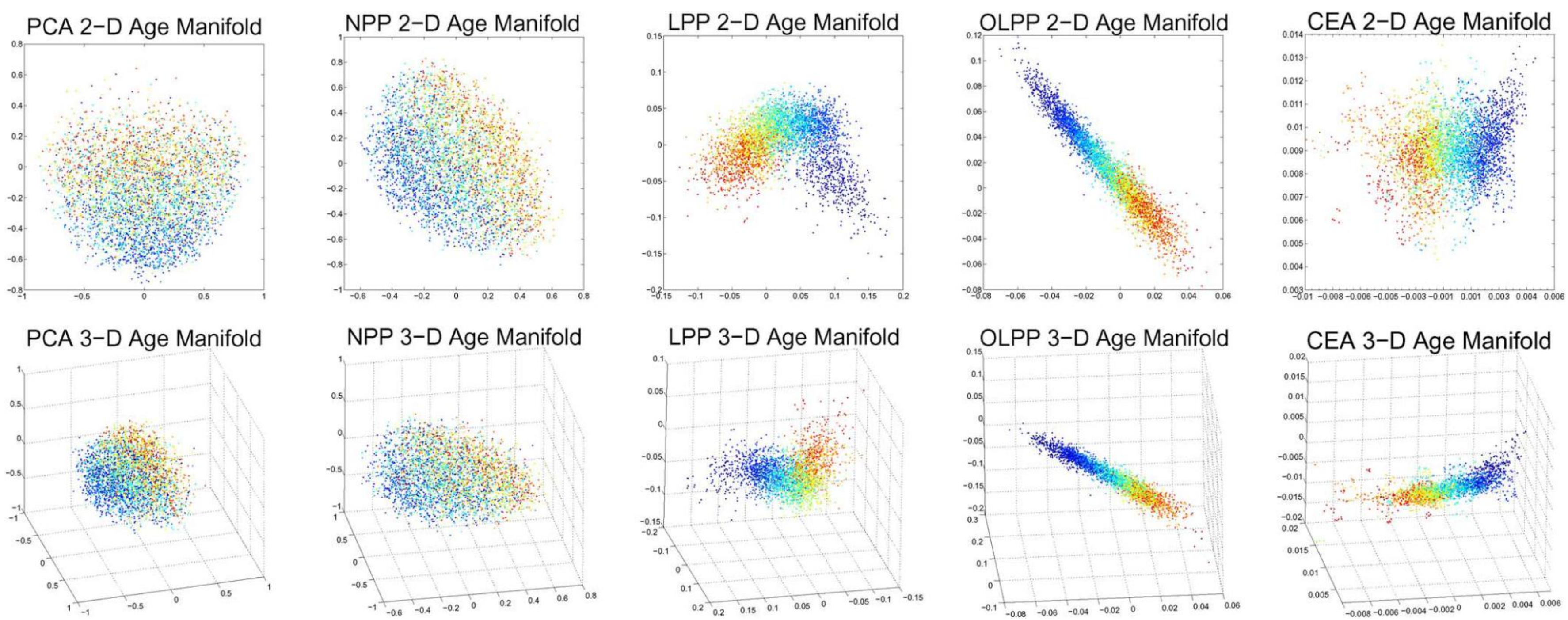
Methods	Tennis Game Dataset 20% observed	40% observed	60% observed	80% observed	100% observed
Integral BoW [1]	0.47	0.44	0.53	0.47	0.51
Dynamic BoW [1]	0.53	0.55	0.49	0.44	0.48
SVM	0.56	0.52	0.51	0.48	0.49
Our Model	**0.59**	**0.64**	**0.65**	**0.65**	**0.70**



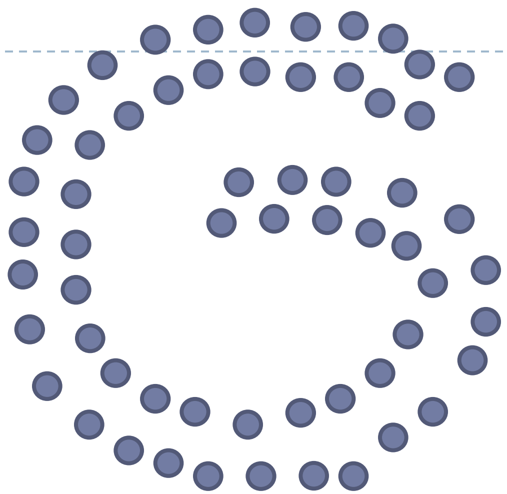
Only our method can predict activities with this kind of complexity.

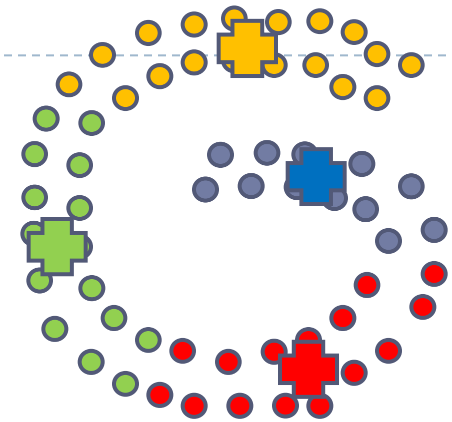
Large Scale Manifold Learning

- Graph based methods require spectral decomposition of matrices of n x n, where n denotes the number of samples.
- The storage cost and computational cost of building neighborhood maps are $O(n^2)$ and $O(n^3)$, it is almost intractable to apply these methods to large-scale scenarios.
- Neighborhood search is also a large scale aspect.

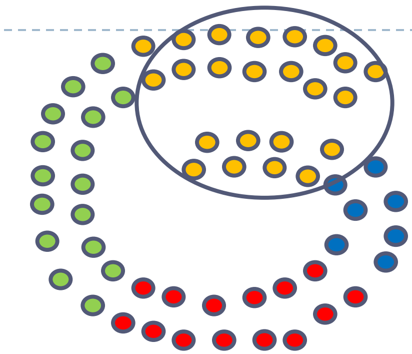


Large Scale Manifold Learning

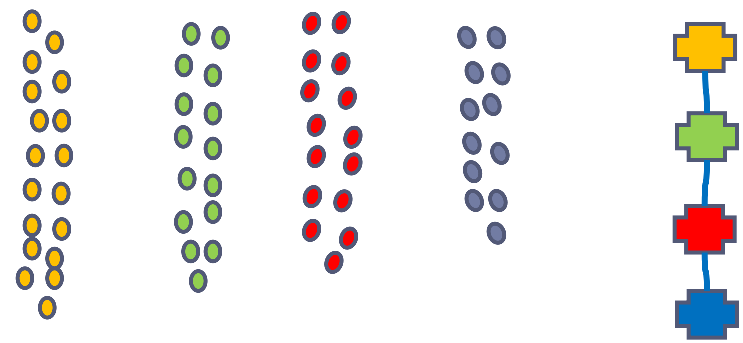




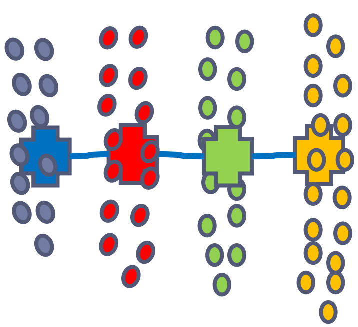
Graph oriented clustering



K-means clustering



$$G = (V, W)$$



$$[\widehat{W}_{Inter} \backslash \widehat{W}_{Intra}]$$
$$= LowRankApproximation(W_{Inter} \backslash W_{Intra})$$
$$s.t. \left\| [\widehat{W}_{Inter} \backslash \widehat{W}_{Intra} \right\|_F \leq \varepsilon$$

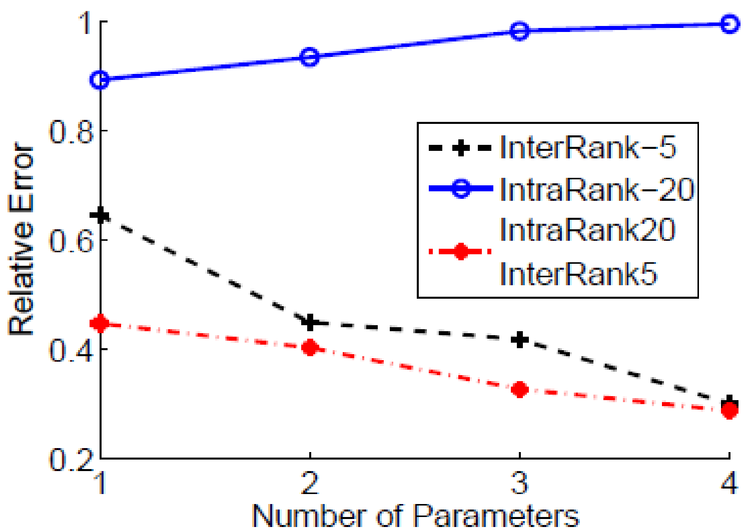
$$W_{Intra} = diag(W_1, \cdots, W_k)$$

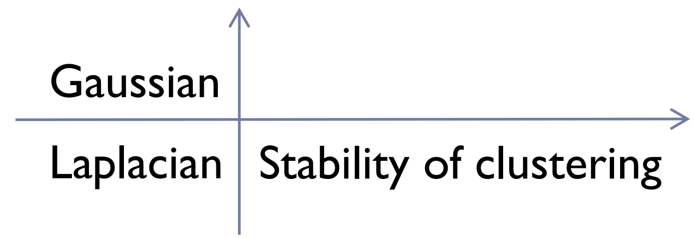
$$W_{Inter} = \begin{bmatrix} 0 & W_{12} & \cdots & W_{1k} \\ W_{21} & 0 & \cdots & \vdots \\ \vdots & \vdots & \ddots & W_{1k-1} \\ W_{k1} & \cdots & W_{k-11} & 0 \end{bmatrix}$$

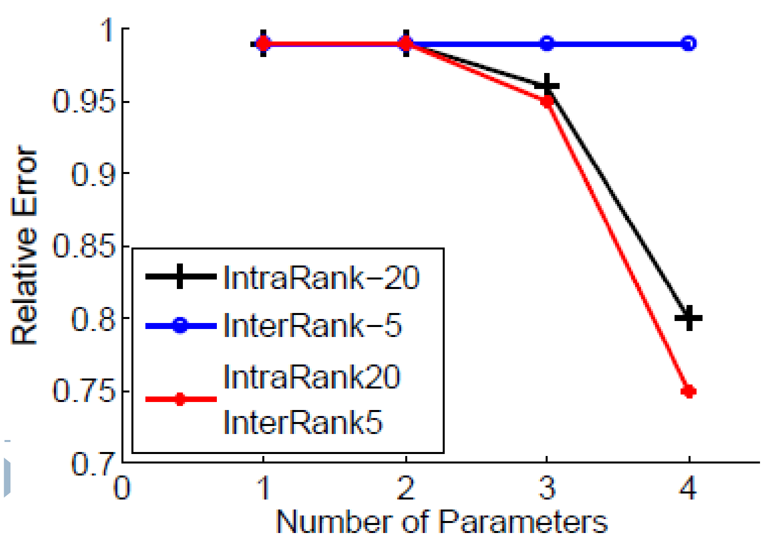
$$G = (V, \widehat{W} = \widehat{W}_{Inter} + \widehat{W}_{Intra})$$

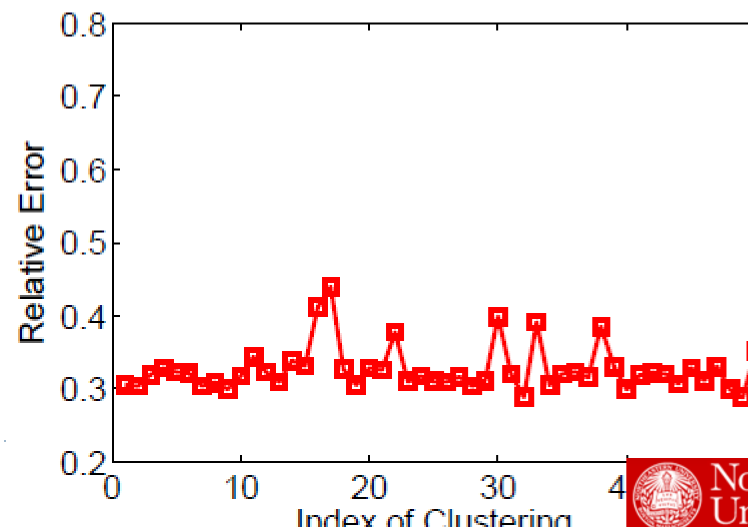
Experiments

- Poker-Hand data set



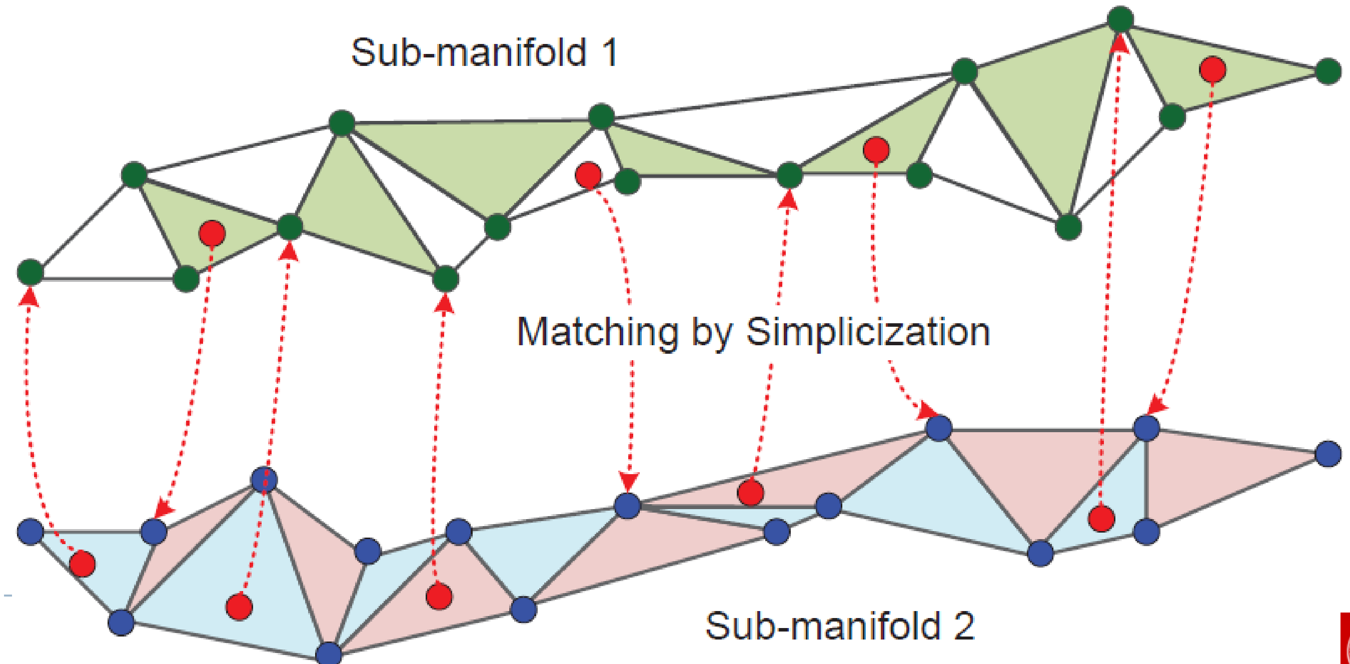






Robust Matching of Sub-Manifolds

- A robust visual representation must be insensitive to durations in the case of dynamics or time series, such as action/activity videos.
- A generalized manifold can be considered as a union of sub-manifolds with different durations which characterize different instances with similar structures, such as different individuals performing the same action, instead of a single continuous manifold as conventionally regarded.
- Robust matching of these sub-manifolds can be achieved through both low-rank matrix recovery and simplex synchronization.



Conclusion

- It is all about data!
- Low-rank analytics based algorithmic tool set is general and promising for explosives-related data representation.
- Transfer learning, manifold learning, and subspace learning are feasible extensions for uncertainty analysis.
- This ATR framework is certainly beyond the visual surveillance scenarios.

- Thank you!