1 KAUST

2

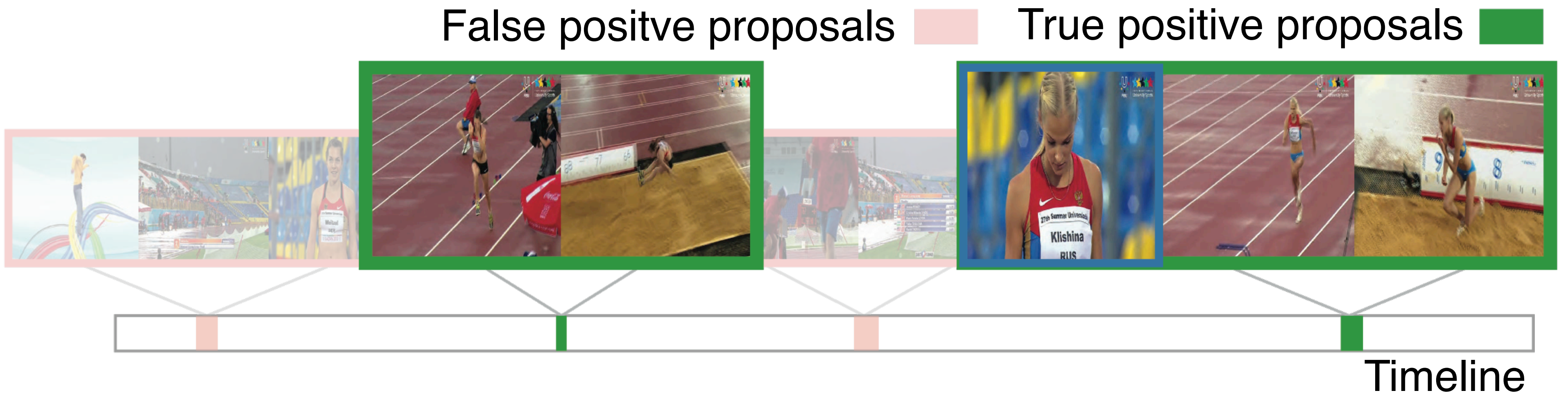
Fast Temporal Activity Proposals for Efficient Detection of Human Actions in Untrimmed Videos

Fabian Caba Heilbron[1], Juan Carlos Niebles[2], Bernard Ghanem[1]

CVPR2016

1. Introduction

Goal. Efficiently retrieve temporal segments from untrimmed videos, which are likely to contain human actions



Motivation:

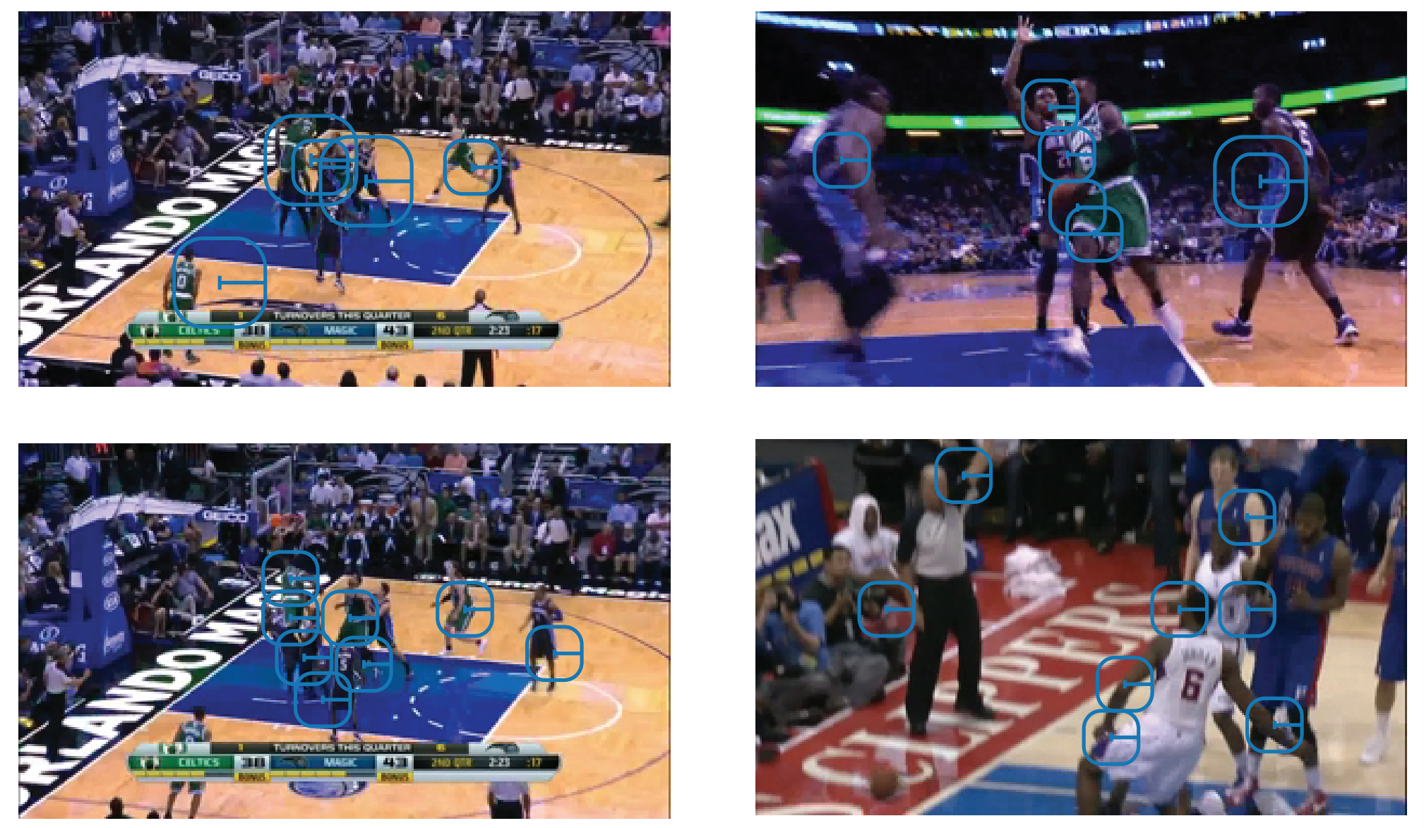
- Video data is inherently untrimmed
- Spatio-Temporal proposals are computationally expensive and ineffective on untrimmed scenarios

Contributions:

- Sparse learning framework to represent human actions
- Efficient and high recall action proposal generation method

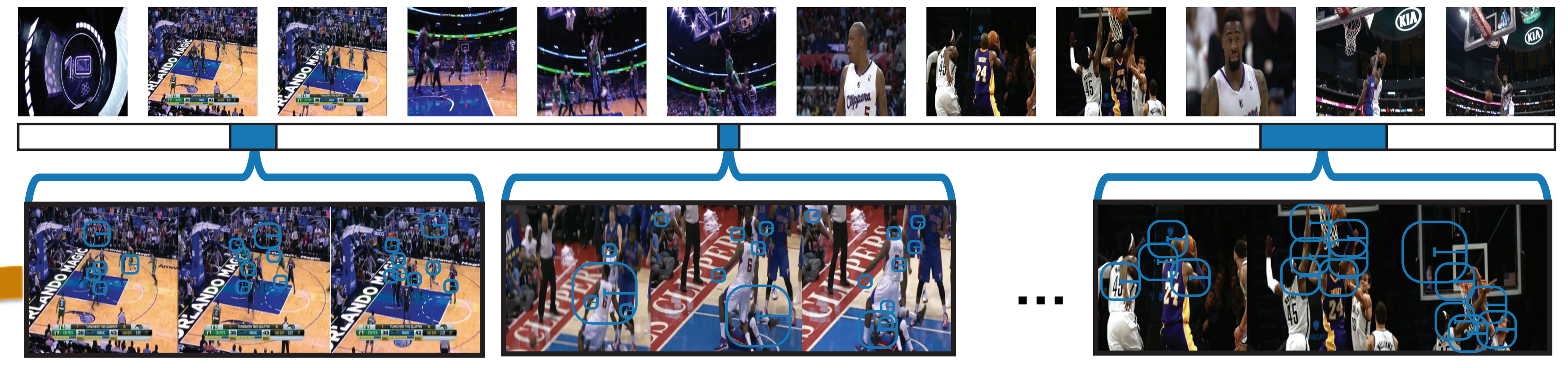
2. Candidate Proposals

- Video segments are represented using STIPs
- A set of candidate proposals is generated using Sliding Windows
- A set of typical action lengths are used.



3. Retrieving Proposals

- Our aim is to **retrieve segments that likely contain actions**
- Proposals are ranked using the **reconstruction error**

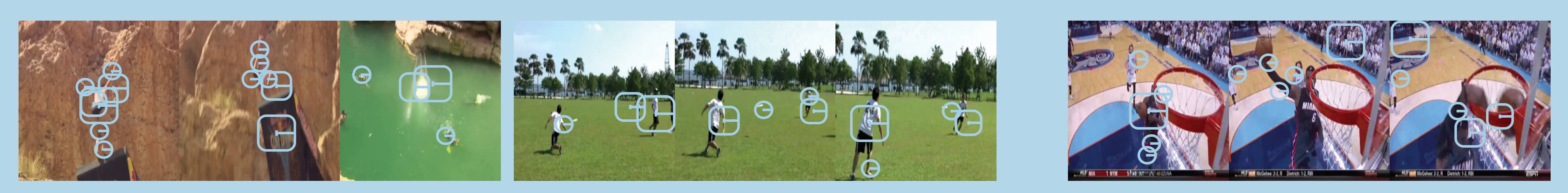


$$\mathbf{A}_k^* = \underset{\mathbf{A}_k}{\arg\min}\ \underbrace{\frac{1}{n_k}\|\mathbf{X}_k - \mathbf{D}\mathbf{A}_k\|_F^2}_{\text{Reconstruction Error}} + \lambda\|\mathbf{A}_k\|_{2,1}$$

$\mathbf{A}_k^*$: Sparse Codes **Reconstruction Error** **D: Dictionary***

4. Learning to Represent Proposals

- An over-complete dictionary jointly represents STIP features
- The dictionary is empowered with discriminative capabilities



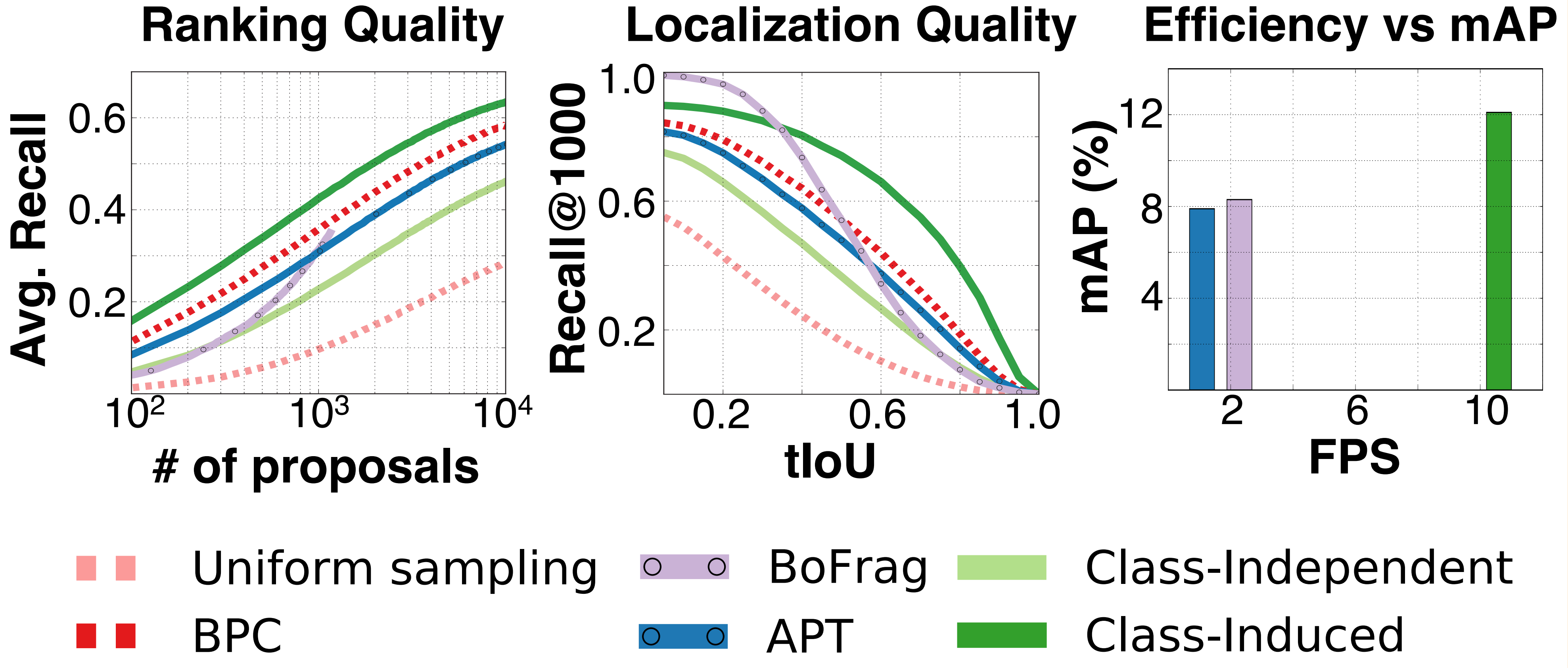
CliffDiving FrisbeeCatch BasketballDunk

$$\underset{\mathbf{D},\mathbf{A},\mathbf{W}}{\arg\min}\ \underbrace{\underbrace{\frac{1}{n}\|\mathbf{X}-\mathbf{D}\mathbf{A}\|_F^2+\lambda_1\|\mathbf{A}\|_{2,1}}_{\text{Class-Independent}}+\lambda_2\|\mathbf{W}^T\mathbf{A}-\mathbf{Y}\|_F^2+\lambda_3\|\mathbf{W}\|_F^2}_{\text{Class-Induced}}$$

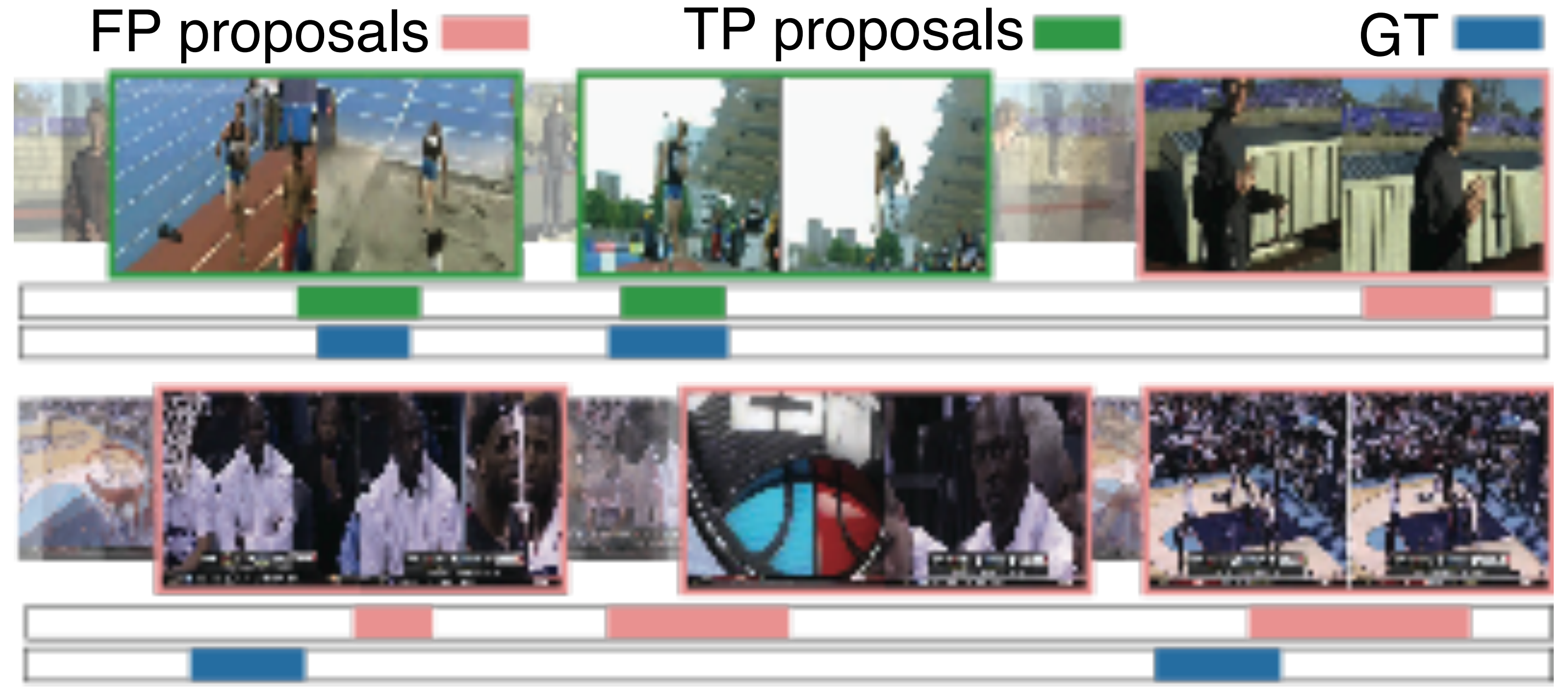
*Fixed point optimization **D is initialized with K-Means

5. Recall and Efficiency Analysis

- Experiments conducted on THUMOS14



6. Temporal Proposal Visualization



cabaf.net/temporalproposals/demo

Acknowledgments: This work was supported by KAUST Office of Sponsored Research, the Stanford AI Lab Toyota Center for Artificial Intelligence Research, and a Google Faculty Research Award (2015).