

Problem Statement

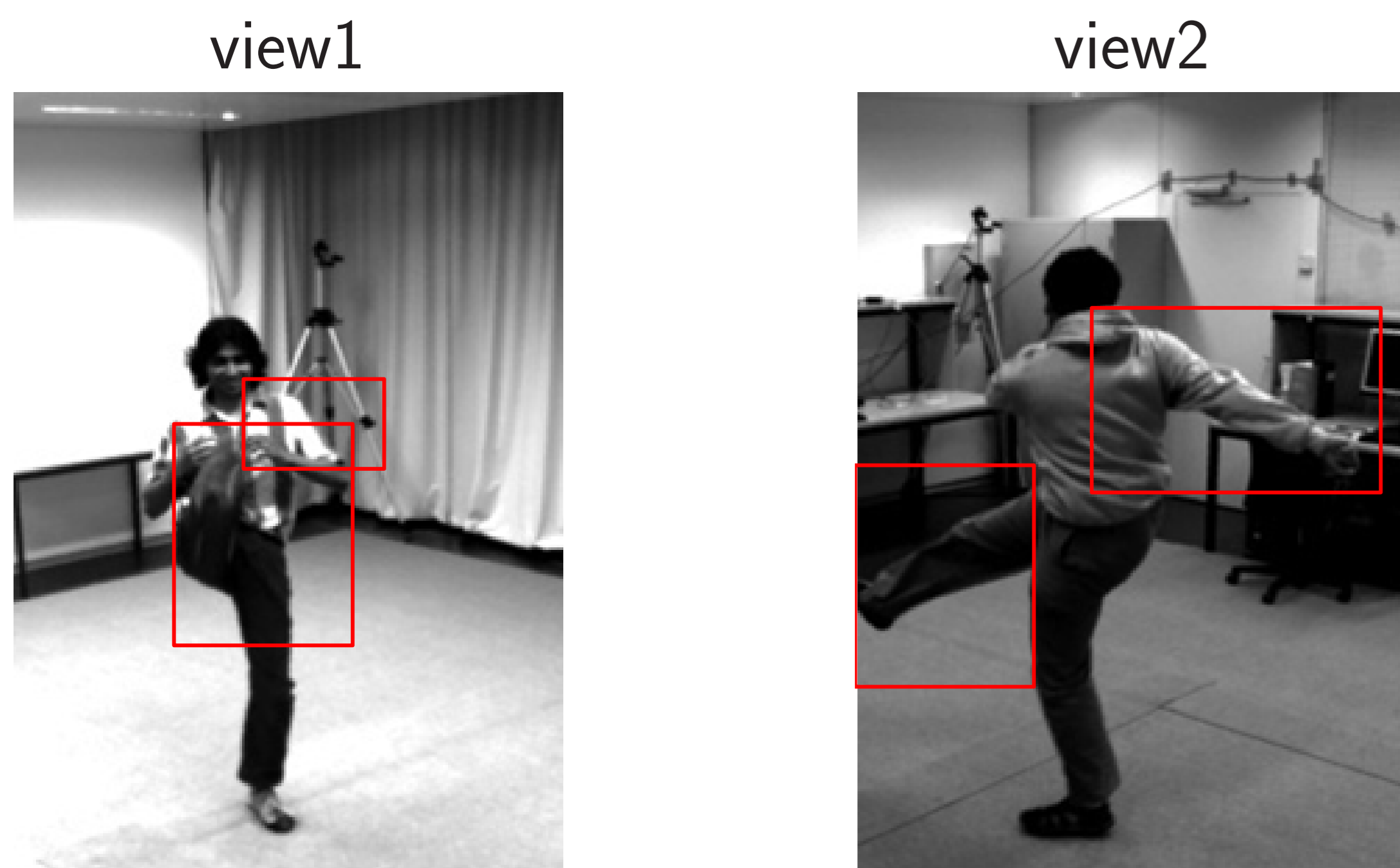
Action recognition across multiple views



Same activity from different viewpoints

- Actions
 - Temporally structured activities
 - Interaction of two actors
- Challenges
 - Occlusion and self-occlusion
 - Different viewpoints

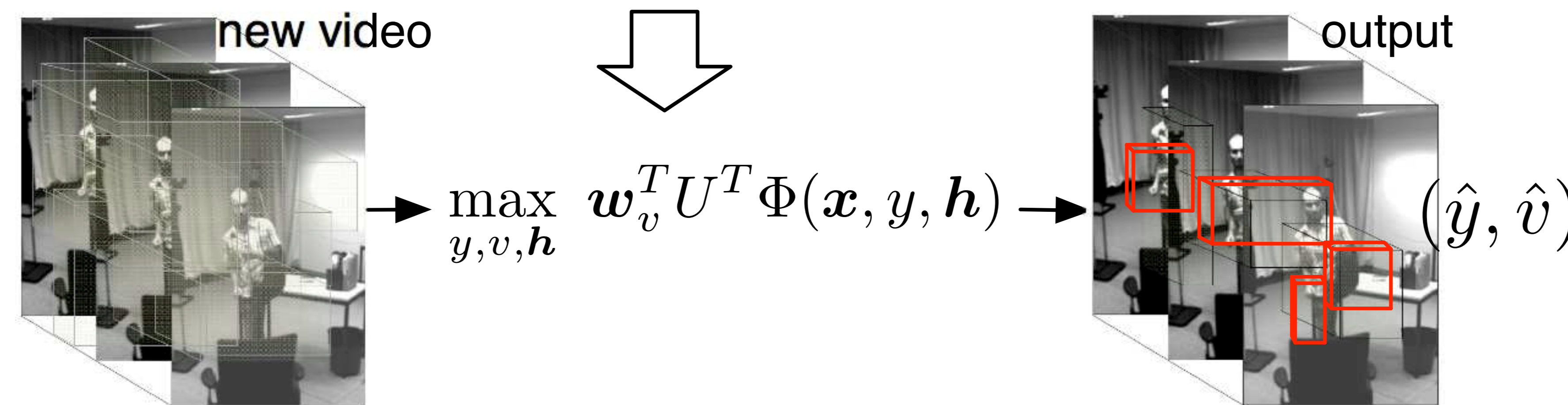
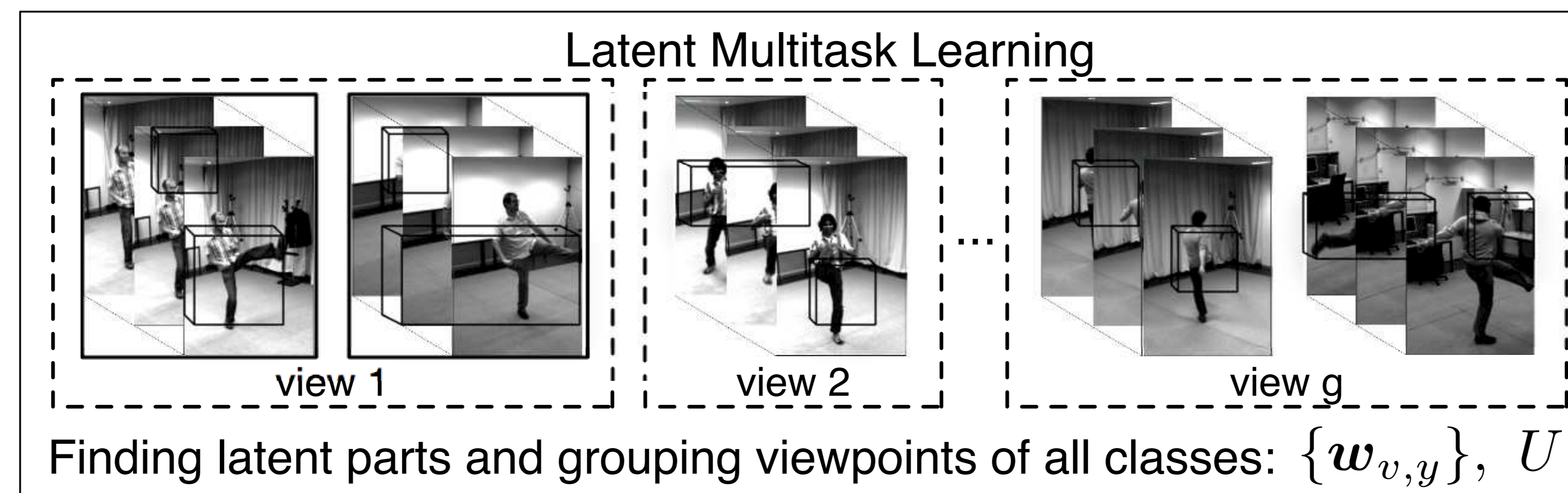
Key Ideas



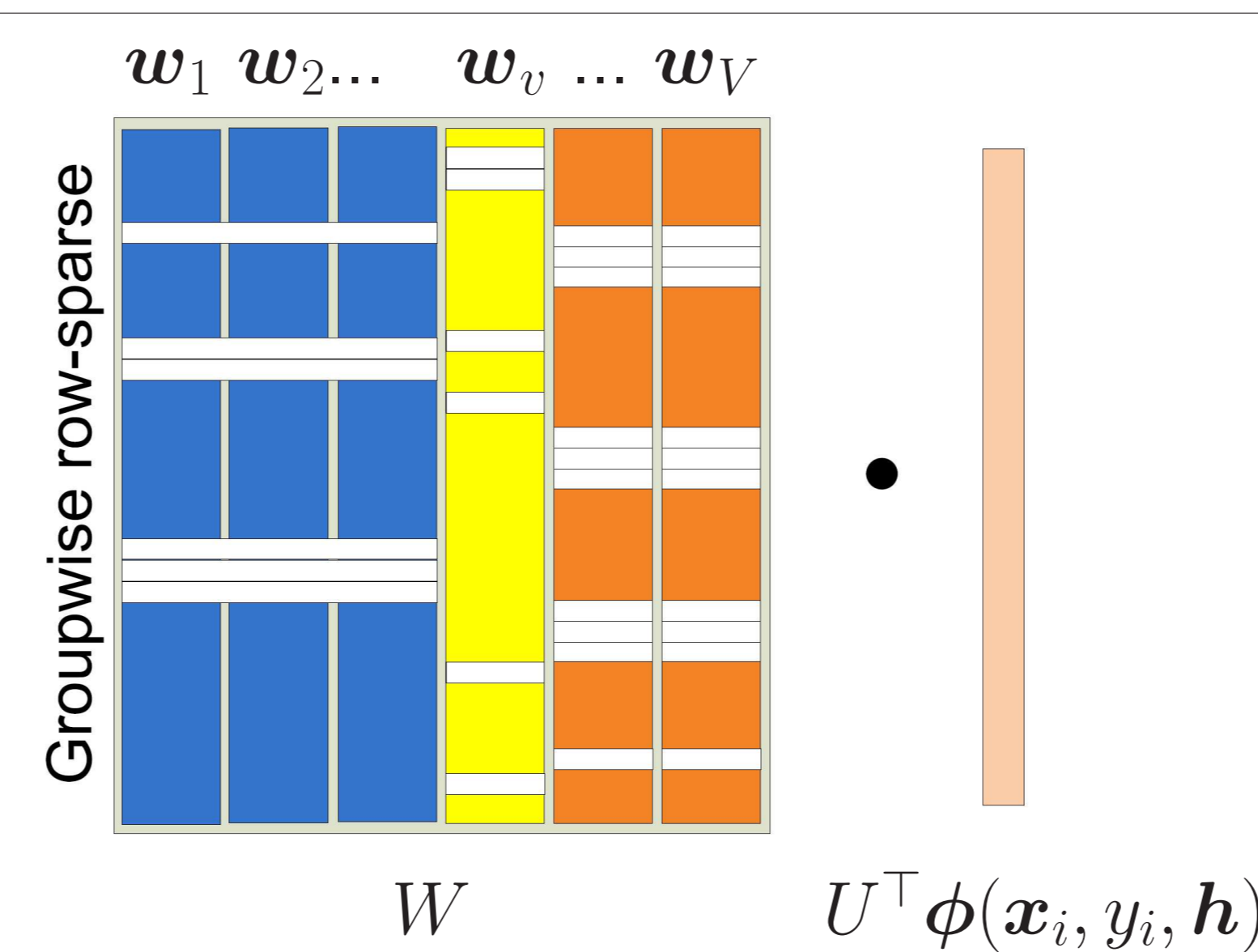
- Latent-parts model of spatiotemporal structure
- View = Task \Rightarrow Multitask Learning (MTL)
- But not all views share the same action parts

Goal: Find correlated groupings of views.

LMTL



Learning



Substitution of the variables : $\Theta = UW$

Q_g = assignment matrix of tasks in group g .

▸ **Step 1:** Find $\Theta = \{\theta_v\}$:

$$\min_{\theta_v} \frac{1}{2} \|\theta_v\|^2 + c \sum_i \max_{\hat{y}, \hat{h}} [\theta_v \Phi(x_i, \hat{y}, \hat{h}) + \Delta(y_i, \hat{y})] - c \sum_i \max_h [\theta_v \Phi(x_i, y_i, h)].$$

▸ **Step 2:** Find $\{Q_g\}$:

$$\min_{\{Q_g\}} \sum_g \|\Theta Q_g\|_*^2 \quad \text{s.t.} \quad \sum_g Q_g = I.$$

Results

Target View	Cam0	Cam1	Cam2	Cam3	Cam4	Avg
[11]	62.0	65.5	64.5	69.5	57.9	63.9
[26]	86.1	93.1	73.6	80.6	-	83.3
LMTL	89.2	87.7	86.8	90.5	79.6	86.8

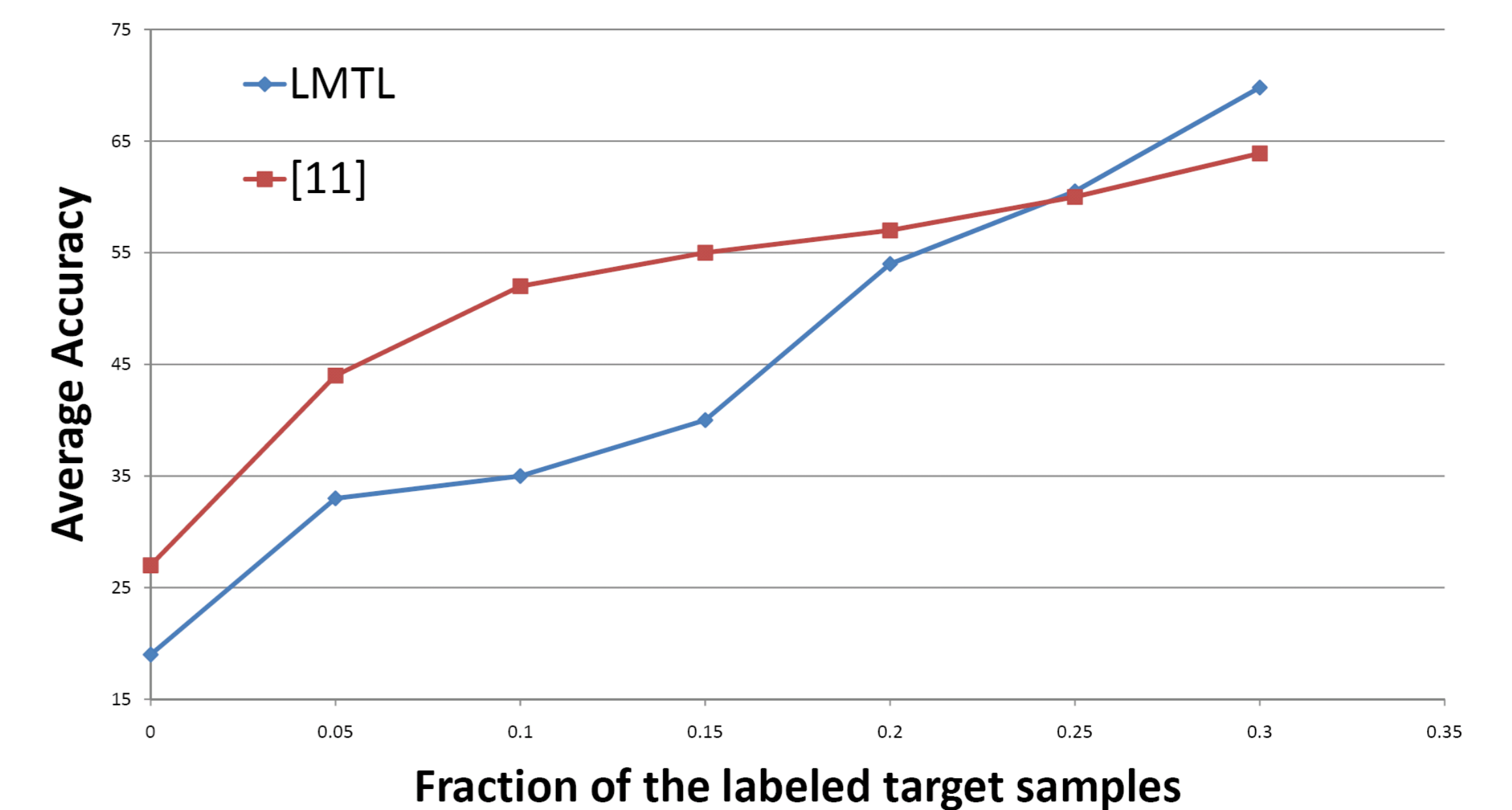
Avg. accuracy in [%] on IXMAS; 30% of videos in the target view used in learning.

Target View	Cam0	Cam1	Cam2	Cam3	Cam4	Avg
[9]	74.8	74.5	74.8	70.6	61.2	71.2
[12]	86.6	81.1	80.1	83.6	82.8	82.8
[26]	95.1	89.6	91.7	90.3	-	91.7
LMTL	95.9	96.8	94.5	96.9	89.9	94.8

Avg. accuracy in [%] on IXMAS; 100% of videos in the target view used in learning.

Target View	Cam0	Cam1	Cam2	Cam3	Cam4	Avg
[24]	87.0	88.3	85.6	87.0	69.7	83.5
LMTL	90.2	91.4	88.7	88.1	84.4	88.6

Avg. accuracy in [%] on IXMAS (new).



Avg. accuracy in [%] on IXMAS for different fractions of training examples on IXMAS.

References

- [9] I. N. Junejo, E. Dexter, I. Laptev, and P. Pérez. PAMI, 2011.
- [11] R. Li and T. Zickler. CVPR, 2012.
- [12] J. Liu, M. Shah, B. Kuipers, and S. Savarese. CVPR, 2011.
- [24] D. Weinland, M. Ozuysal, and P. Fua. ECCV, 2010.
- [26] X. Wu and Y. Jia. ECCV, 2012.

Acknowledgment