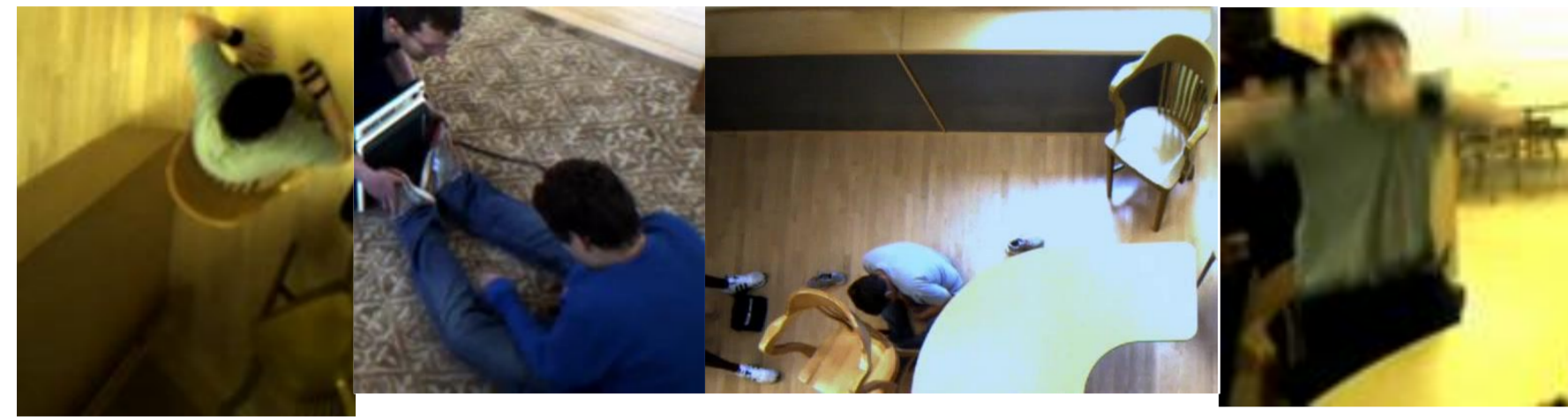


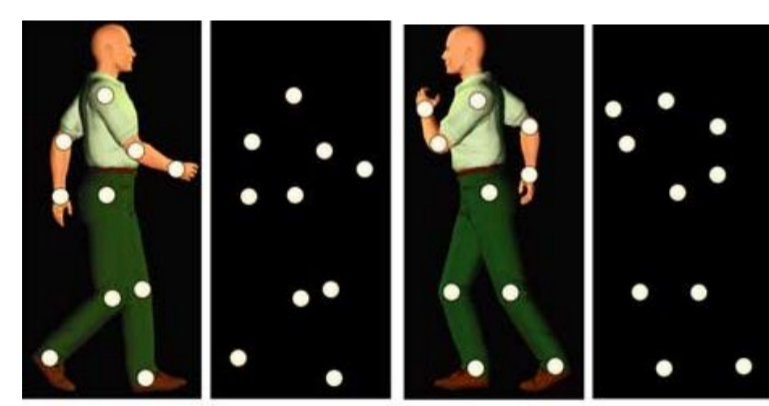
Motivation

Many real-world applications can benefit from action representation that is robust to variations in viewpoint, scaling, and appearance (clothing, lighting, etc.).

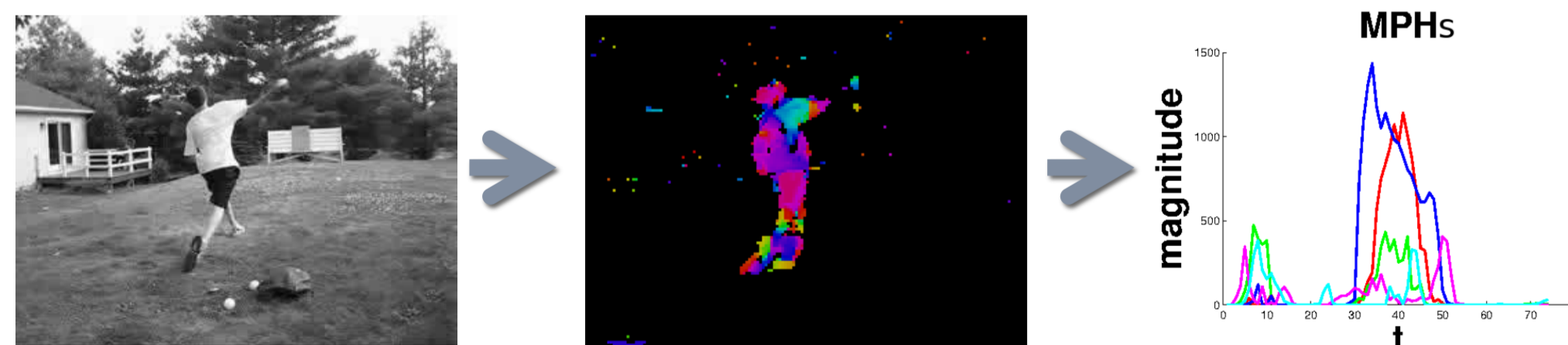


Intuition

Johansson's moving light display
→ Dynamics of an action provide a powerful cue for discrimination.



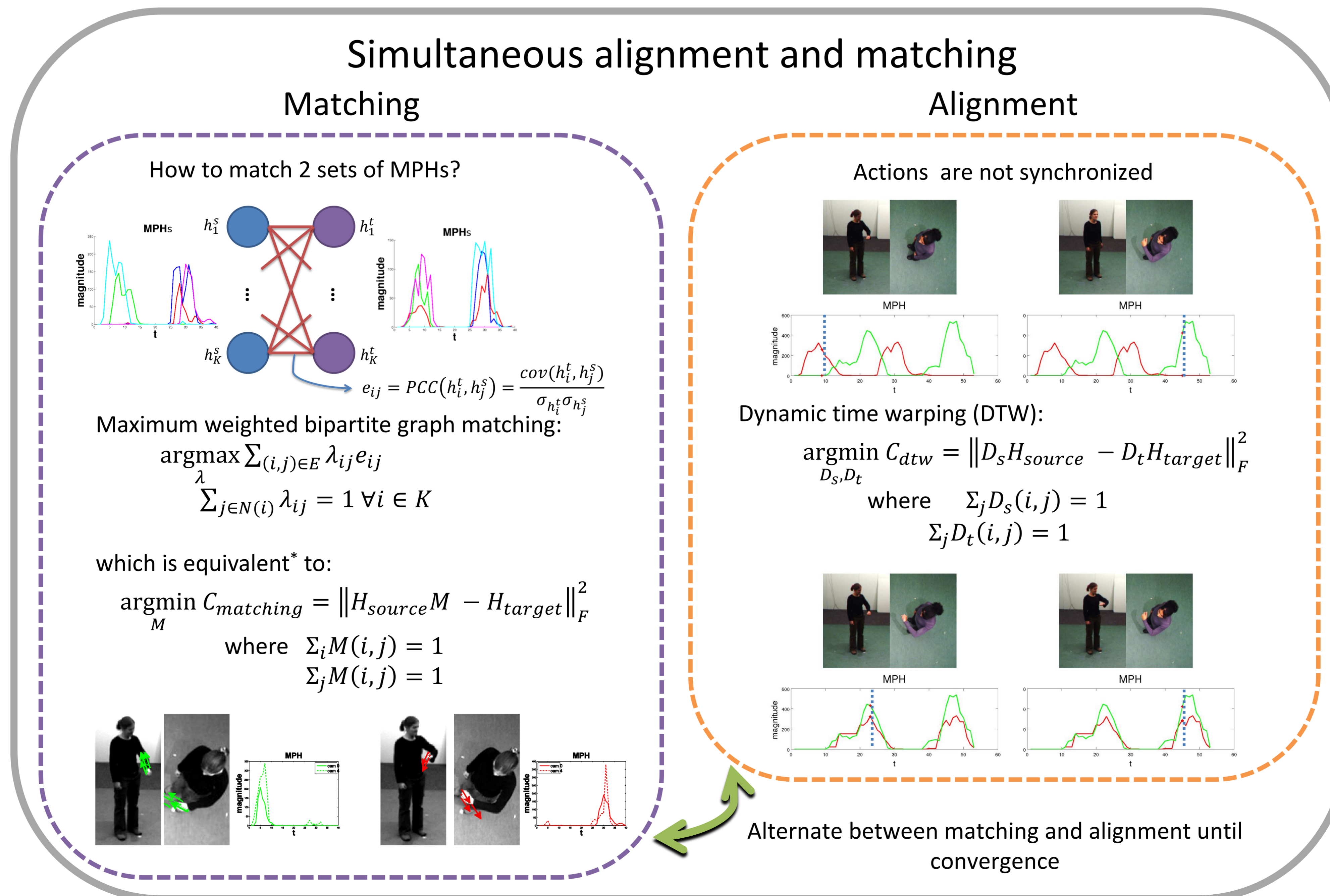
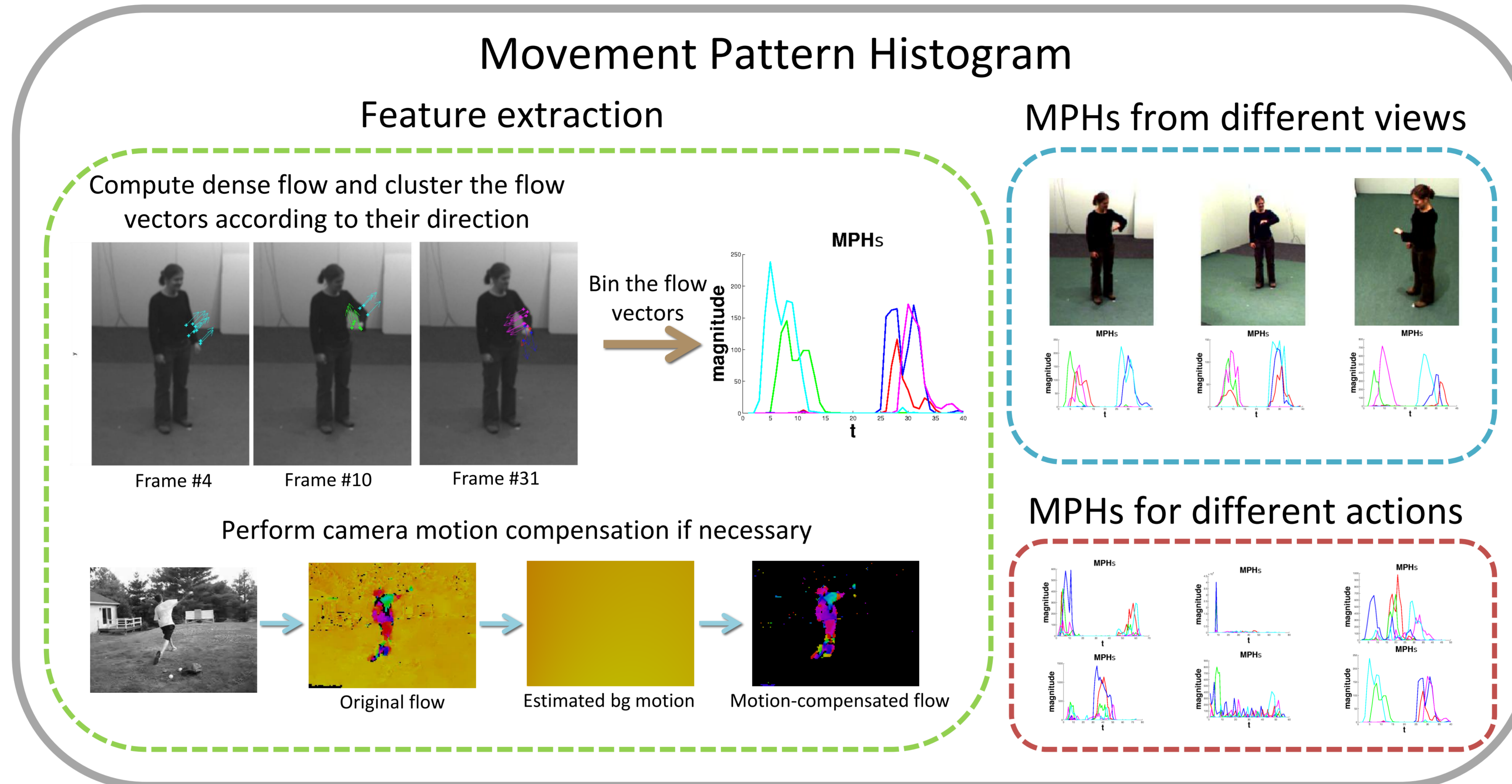
Decompose an action into movement primitives (roughly corresponds to body parts) and encode the temporal dynamics of each primitives.



Contributions

- The movement pattern histogram (MPH), a multi-channel temporal distribution of movement primitives.
- Novel optimization approach to simultaneously match and align MPH across videos.
- Stereotypy dataset: annotated video corpus of children with autism in a classroom setting.

Overall Approach

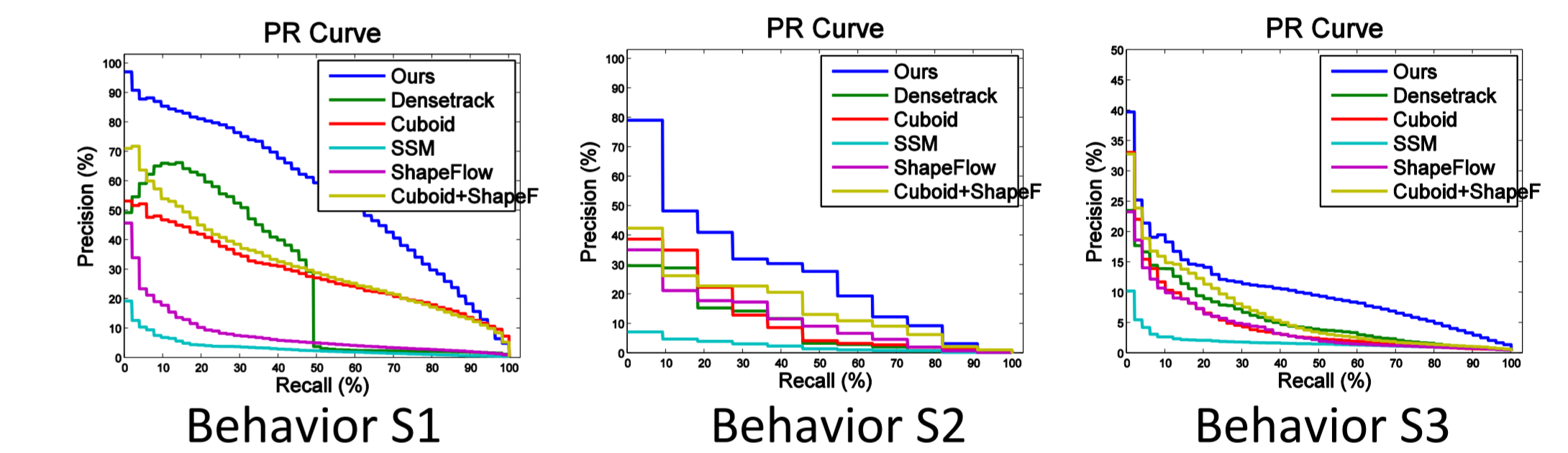
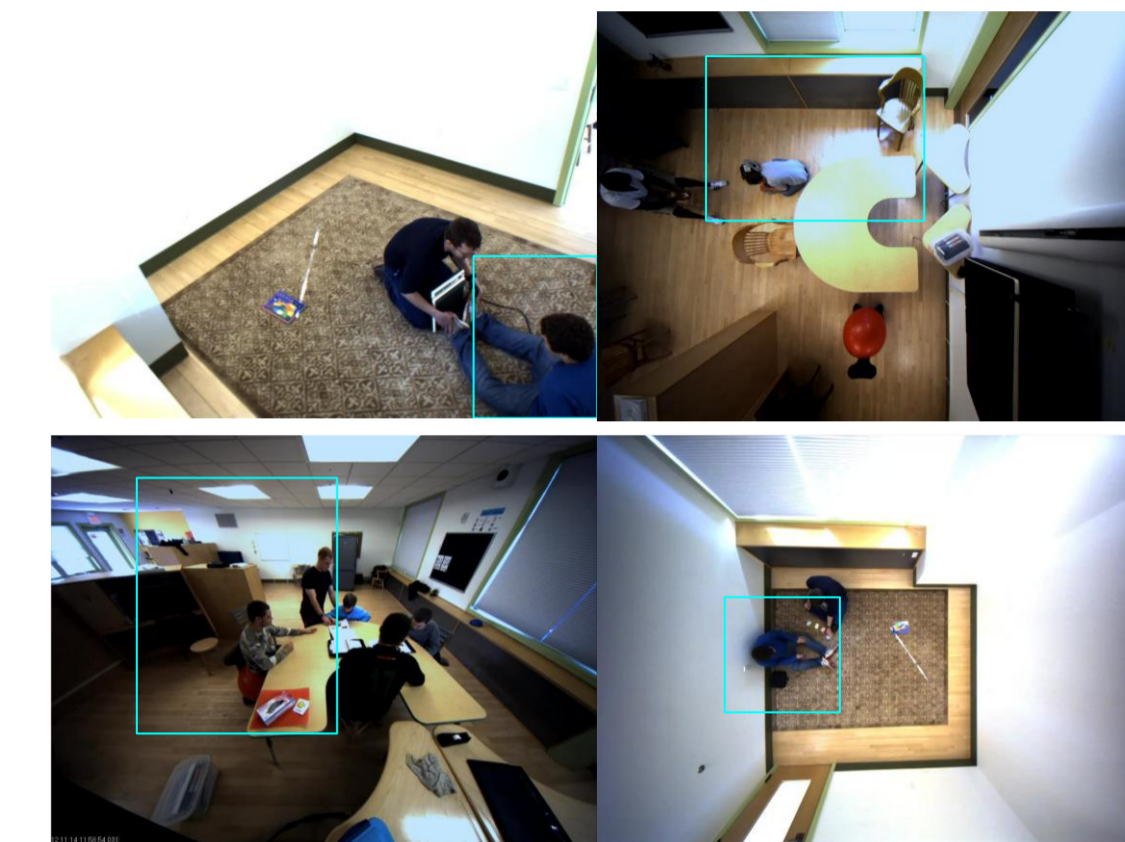


Results Action Retrieval

Given a single example video clip containing an action of interest, retrieve all matching instances of that action from an unstructured video collection.

Stereotypy Dataset

- Real-world recording of children with autism in a classroom setting.
- A collection of 10-20 minute classroom sessions with total length of 2 hours.
- 3 behaviors of interest annotated by psychologist with autism expertise.



Action Recognition

IXMAS (11 actions, 5 viewpoints, 1650 total clips)

Method	Test View					Avg.
	c0	c1	c2	c3	c4	
MPH	83.9	81.8	87.6	83.0	73.6	82.0
[1]	66.4	73.5	71.0	75.4	66.4	70.5
[2]	62.0	65.5	64.5	69.5	57.9	63.9
[3]	77.0	78.8	80.0	73.9	63.6	74.6

UCF50 (50 actions from Youtube, 6618 total clips)

Method	Accuracy (LoGo)
MPH + FV of DenseTrajectories [4]	90.5
DenseTrajectories [4] w/ FV encoding	88.9
MBH + scene context [5]	76.9
GIST3D + STIP [6]	73.7
MIP [7]	72.7

References:
 [1] Zhang, Z., Wang, C., Xiao, B., Zhou, W., Liu, S., Shi, C.: Cross-view action recognition via a continuous virtual path. CVPR (2013).
 [2] Li, R., Zickler, T.: Discriminative Virtual Views for Cross-View Action Recognition. CVPR (2012).
 [3] Junejo, I., Dexter, E., Laptev, I., P. Perez, P.: View-Independent Action Recognition from Temporal Self-Similarities. PAMI (2010).
 [4] Wang, H., Klaser, A., Schmid, C., Liu, C.L.: Dense trajectories and motion boundary descriptors for action recognition. IJCV (2013).
 [5] Reddy, K.K., Shah, M.: Recognizing 50 human action categories of web videos. Machine Vision and Applications (2012).
 [6] Solmaz, B., Assari, S.M., Shah, M.: Classifying web videos using a global video descriptor. Machine Vision and Applications (2012).
 [7] Kliper-Gross, O., Gurovich, Y., Hassner, T., Wolf, L.: Motion interchange patterns for action recognition in unconstrained videos. ECCV (2012)