





Goal

Given a collection of first person videos, predict the future location and gaze direction of the basketball players.



Input: First person videos



Output

Physical Connection Cue

Other players' distance and orientation to the camera wearer.







Highly selective retrieval results





Trajectory retrieval via (x, y, θ, v)

Trajectory retrieval via physical connection cue

Predicting Basketball Behaviors from First Person Videos

Shan Su¹, Jung Pyo Hong², Jianbo Shi¹, and Hyun Soo Park³

Mental Connection Cue

Joint attention: the location where all players pay attention to.



Geometric information



Statistic information

CNN-LSTM to learn joint attention from team formation statistically.



Group Selection via Social Compatibility





Social compatibility:

Alignment of the gaze direction with respect to estimated joint attention

$$\eta = \frac{1}{n} \sum_{\mathbf{x}} \frac{(\hat{\mathbf{s}} - \mathbf{x}_i)^{\mathrm{T}} \mathbf{g}_i}{\|\hat{\mathbf{s}} - \mathbf{x}_i\|}$$



Qualitative Result

