Look-ahead before you leap:
End-to-end active recognition by forecasting the effect of motion
Dinesh Jayaraman and Kristen Grauman
UT Austin

Active recognition setting
passive 1-view setting
active setting
mug? bowl? pan?
Difficulty: unconstrained visual input
Opportunity:
• Not restricted to a single snapshot.
• Strategically acquiring new views.

Our idea
Prior art: independent, often heuristic components

Our idea: Multi-task joint training of components for active recognition + auxiliary internally supervised “look-ahead” task.

Towards real-world active recognition
SCENE CATEGORIES
SUN 360 panoramas
HAND-HELD OBJECTS
GERMS toy manipulation
3-D CAD MODELS
ModelNet-10 3-D models
Complex real-world categories + easily benchmarkable setups.

Quantitative evaluation results
SUN 360
GERMS
ModelNet-10

Ablation studies
SUN 360
GERMS
Training all 3 components jointly is most critical to performance.

Component module architectures