

# Supplementary Material For rCRF: Recursive Belief Estimation over CRFs in RGB-D Activity Videos

Ozan Sener

School of Electrical & Computer Eng.  
Cornell University

Ashutosh Saxena

Department of Computer Science  
Cornell University

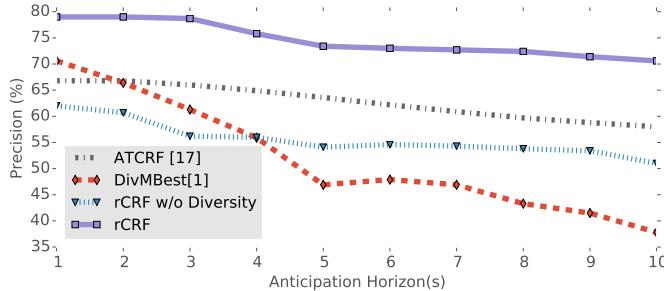


Fig. 2: Precision vs. anticipation horizon for object affordance. We run all possible experiments for  $\tau$  seconds into the future experiments as showing  $t$  seconds of the video and anticipating  $t + \tau$  for all  $t < T - \tau$

## I. QUALITATIVE RESULTS

## II. QUANTITATIVE RESULTS

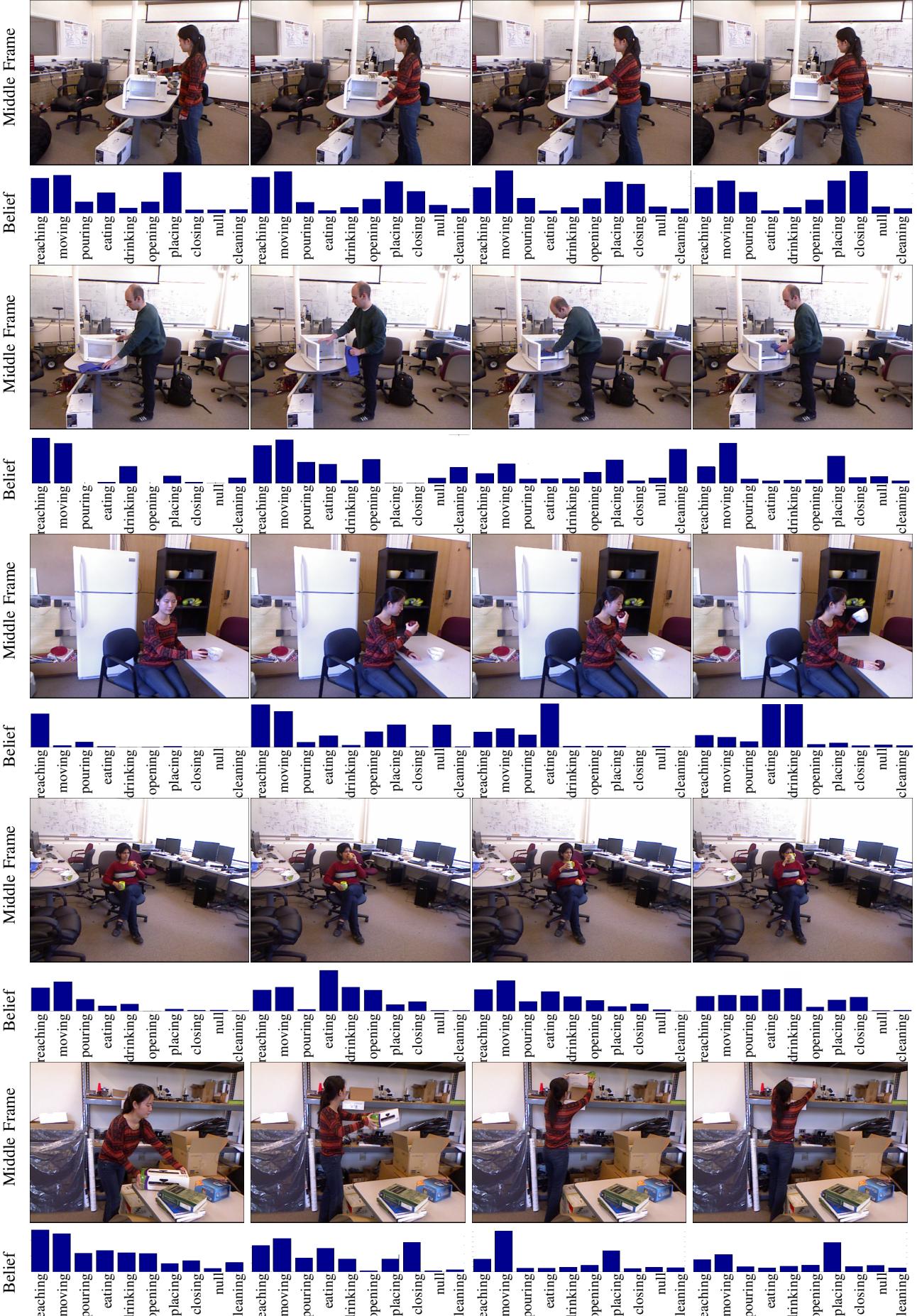
**TABLE I: Anticipation performance for the anticipating 3 seconds in the future.** We compare rCRF with state-of-the-art anticipation algorithm and baselines for anticipation accuracy. rCRF outperforms the state-of-the-art heuristic method [?] and the GP-LCRF method [?] significantly as well as all other baselines. We believe this result is due to the accurate joint-modeling of the temporal relations and the CRF model.

Method	Sub-activity			Object Affordance		
	micro prec(%)	macro f1-scr(%)	robot ant. metric(%)	micro prec(%)	macro f1-scr(%)	robot ant. metric(%)
Chance	10.0±0.1	10.0±0.1	30.0±0.1	8.3±0.1	8.3±0.1	24.9±0.1
GP-LCRF [?]	52.1±1.2	43.2±1.5	76.1±1.5	68.1±1.0	44.2±1.2	74.9±1.1
ATCRF [?]	47.7±1.6	37.9±2.6	69.2±2.1	66.1±1.9	36.7±2.3	71.3±1.7
DivMBest[?]	47.9±1.4	43.2±3.6	71.5±2.7	61.3±1.4	56.3±2.1	73.3±0.5
DCRF[?]	48.3±2.6	35.4±1.8	66.6±1.1	55.2±3.1	48.5±3.1	71.24±2.2
rCRF w/o div	49.6±2.1	39.7±2.6	65.1±1.1	56.2±1.9	47.4±3.1	70.8±2.5
rCRF	<b>54.3±3.9</b>	<b>45.8±2.7</b>	<b>76.5±2.6</b>	<b>78.7±3.4</b>	<b>74.9±3.8</b>	<b>82.1±2.9</b>

**Computationally-efficient inference:** We evaluated the computational efficiency by computing the average computation time for anticipating 3 second in the future via rCRF and the fastest available anticipation algorithm (the ATCRF[?]). Within our experiments, we did not include any pre-processing or feature extraction computation (they are same for all algorithms).

**TABLE II: Computation time for anticipating 3 seconds in the future excluding pre-processing (see supplementary material for details).**

ATCRF [?]	34.1s	rCRF	1.41s



**Fig. 1: Anticipated belief over activity.** In the odd numbered rows, we show a middle frame of the temporal segment. In the even numbered rows, we show the anticipated belief. Note that frames are not visible to the algorithm and only included for evaluation.