An Hybrid Approach for Robots Learning Folding Tasks

Benjamin Balaguer and Stefano Carpin

Contribution

We use cooperative manipulators to perform towel folding tasks. Our method executes a *momentum fold*: a swinging motion that exploits the object's dynamics. We propose a new learning algorithm that combines imitation and reinforcement learning.
The strengths of the algorithm come from its efficient processing, fast learning capabilities, absence of a deformable object model, and applicability to other problems exhibiting temporally incoherent parameter spaces.

Reward Function

Computing $R(O^c)$**:**

$minAvgError \leftarrow 1000$
for all $O_i \in O^t$ **do**
  $Training \leftarrow LastFrame(O_i)$
  $Current \leftarrow LastFrame(O^c)$
  $AvgError \leftarrow$ ICP$(Trainning, Current)$
  if $AvgError \leq minAvgError$ **then**
    $minAvgError \leftarrow AvgError$
  end if
end for
return $\exp(-minAvgError/100)$

Imitation Layer: Exploratory

- Explore trained action space with robot
- Looking for most diverse set of examples
- Lloyd's k-means clustering algorithm

$$\arg\min_{C} \sum_{j=1}^{M} \sum_{\theta_i \in C_j} ||\theta_i - \mu_j||$$

$$\theta^{Epl} = [\theta_1^{Epl} \dots \theta_M^{Epl}] \;\; s.t. \;\; \theta_i^{Epl} \in C_i \in \theta^t$$

Imitation Layer: Expansion

- Expand best robot action based on θ_{Best}^{Epl}
- Run PCA on O^t to get $\hat{O}^t \in \mathbb{R}^{80\times 29}$
- Train Radial Basis Functions (RBF) on $\hat{O}^t$
- **Generate** l **new actions,** θ_s^{Epa}**, using:**

$n =$ NumColumns$(\hat{O}^t)$ // n=29
$\hat{O}^t \sim [\hat{O}_1^t \;\; \hat{O}_2^t \dots \hat{O}_n^t]$
$\mu = [E[\hat{O}_1^t] \;\; E[\hat{O}_2^t] \dots E[\hat{O}_n^t]]$
$\Sigma = [\mathrm{Cov}(\hat{O}_i^t, \hat{O}_j^t)]_{i=1,2,\dots,n; j=1,2,\dots,n}$
repeat
  Sample $\hat{O}^s$ from $\mathcal{N}(\mu, \Sigma)$
until $|$Time$(\hat{O}^s)$-Time$(\theta_{Best}^{Epl})| \leq \varepsilon$
θ_s^{Epa} =RBF$(\hat{O}^s)$
return θ_s^{Epa}

$$\theta^{Epa} = [\theta_1^{Epa} \theta_2^{Epa} \dots \theta_l^{Epa}]$$

Reinforcement Learning Layer

- θ_n^{RL} is updated to produce a new action θ_{n+1}^{RL}

$$\theta_{n+1}^{RL} = \theta_n^{RL} + \left(\theta_{Top} - \theta_n^{RL}\right)\left[R(O_{Top}) - R(O_n)\right]$$

- θ_{Top} corresponds to the best rewards of:

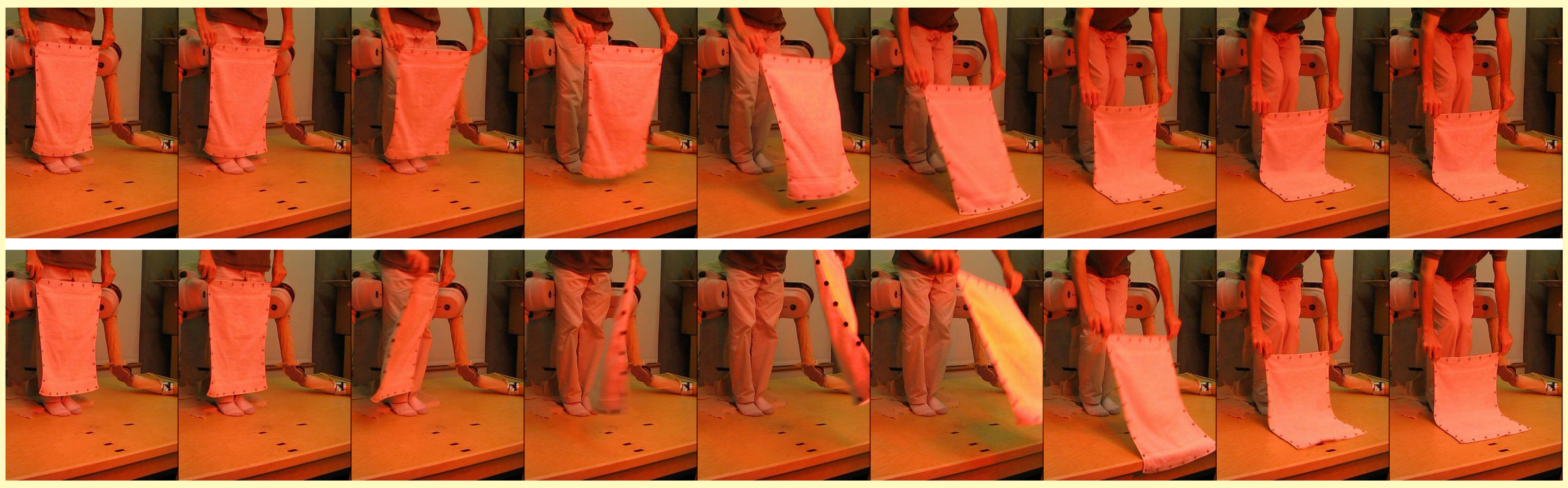
$$[\theta_{Best}^{Epl} \theta_1^{Epa} \dots \theta_l^{Epa} \theta_1^{RL} \dots \theta_{n-1}^{RL}]$$

Acknowledgement

National Science Foundation grant BCS-0821766.

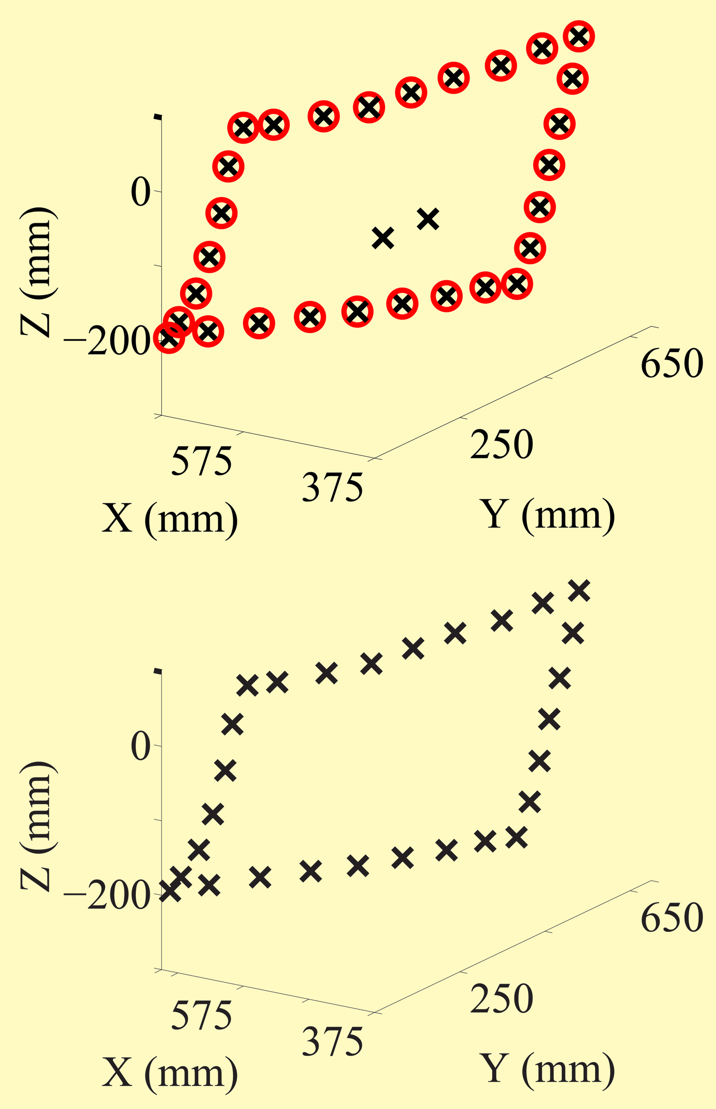
Imitation Learning

- **Motion capture system:** Cartesian coordinates and time stamp of 28 markers recorded at 30 Hz
- **Action/Observation pairs:** $\theta_i \in \mathbb{R}^{1050}$ and $O_i \in \mathbb{R}^{12750}$
- **Complete training data set:** $\theta^t \in \mathbb{R}^{80\times 1050}$ and $O^t \in \mathbb{R}^{80\times 12750}$



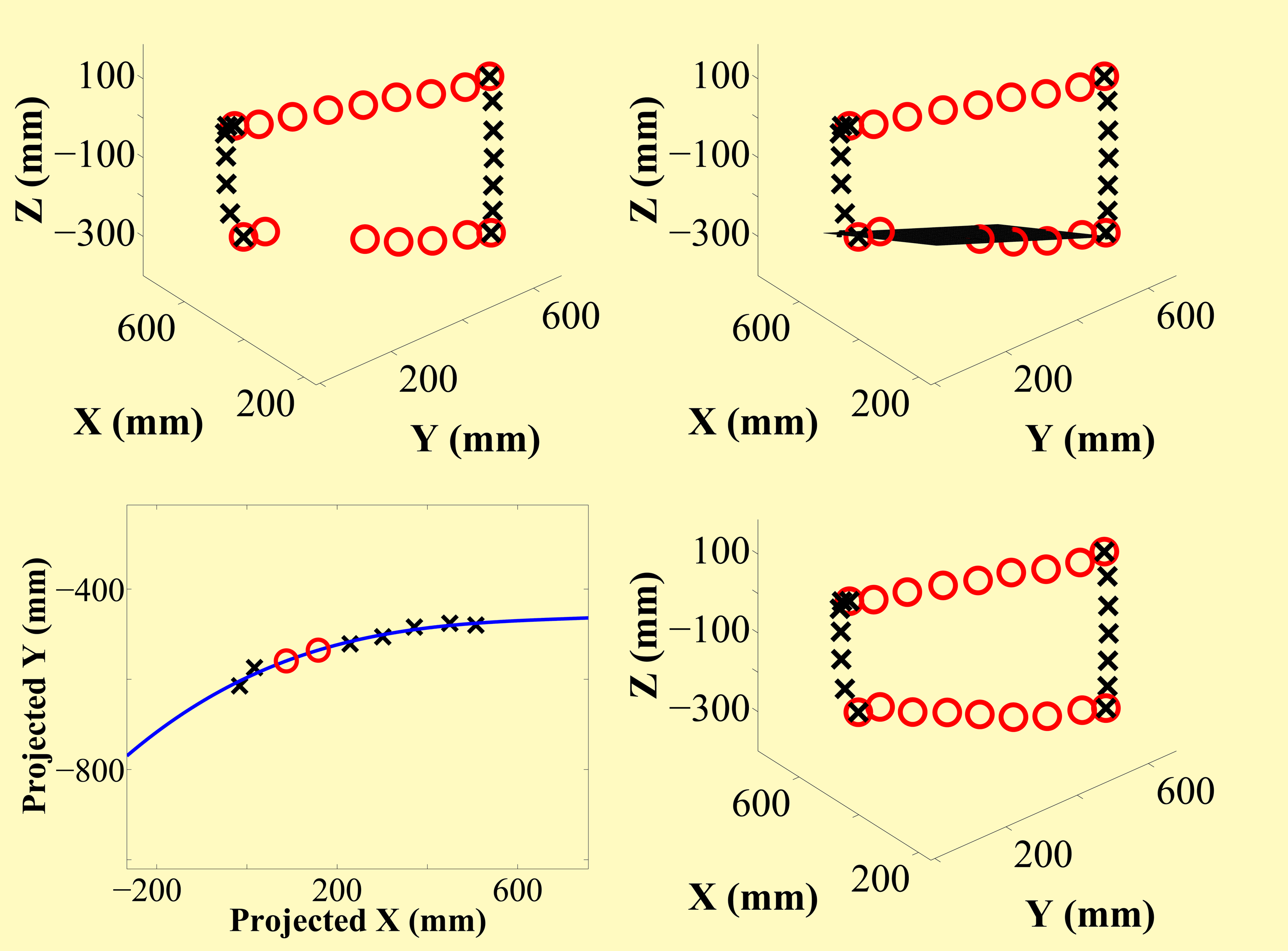
Correcting False Positives and False Negatives

False Positives



100% Correction Rate

False Negatives



88.89% Correction Rate

Results and Experiments

- Robot configurations acquired through 7-DoF IK
- Joint minimization to reduce mechanical loads

- **Training Data:** 80 examples
- **Exploratory Layer:** 10 runs
- **Expansion Layer:** 5 runs
- **Reinforcement Learning:** until convergence
  - Difference in last 3 rewards $\leq$ 0.1%
  - Converges in only 4 trials
  - Good seed due to exploratory and expansion layers