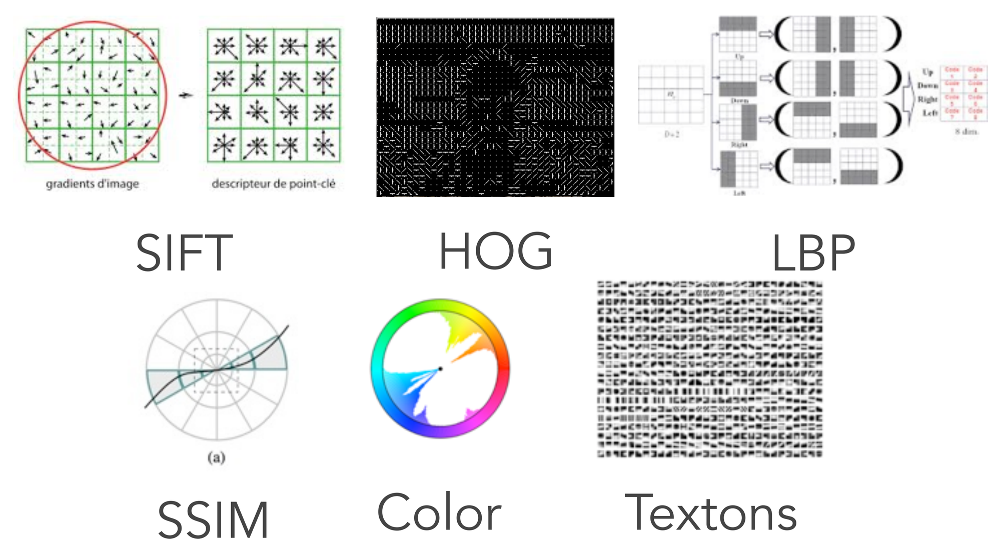
Dynamic Recognition on a Budget

Sergey Karayev · Mario Fritz
Trevor Darrell

Motivation

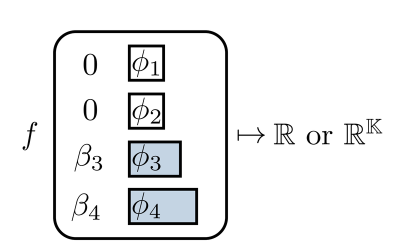
With a test time budget, cannot compute all features.



With different instances benefiting from different features, selection needs to be dynamic.

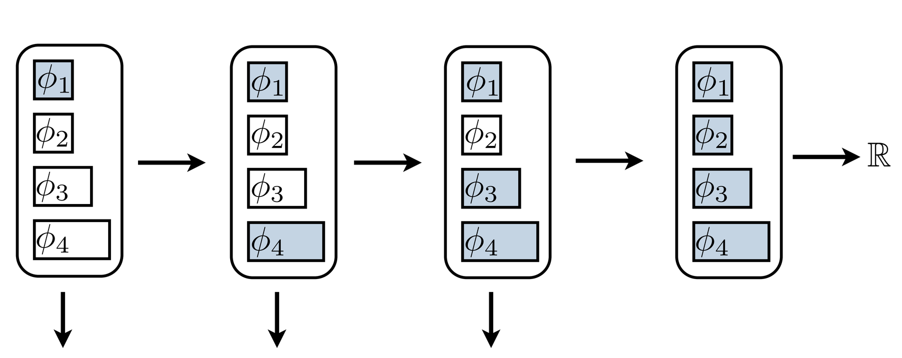
Related Work

Feature selection



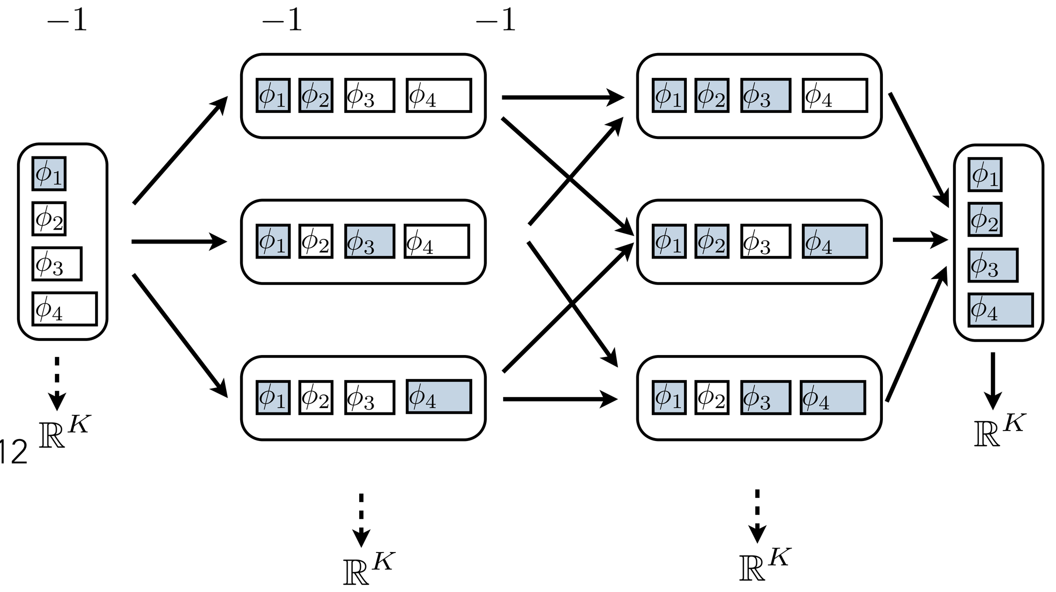
Cascade

Viola & Jones CVPR 2001
Chen et al. ICML 2012



Dynamic feature selection

Gao & Koller NIPS 2011
Benbouzid et al. NIPS 2012
Karayev et al. NIPS 2012

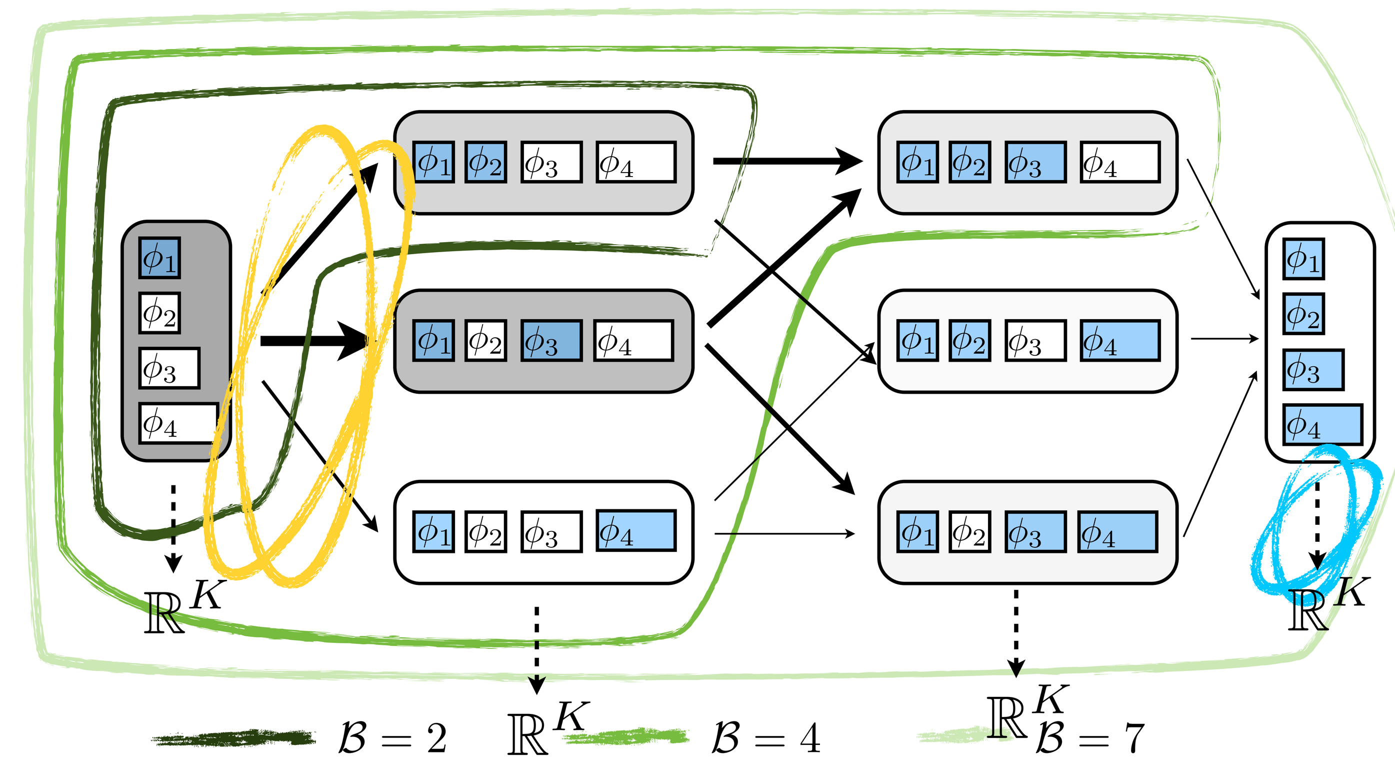


Model

State view

Action selection: non-myopic policy learned by MDP.

Feature combination: linear for efficiency.



Action selection

policy:

$$\pi(s) = \arg\max_{a_i \in \mathcal{A}\setminus\mathcal{O}} Q(s, a_i)$$

action-value function:

$$Q^\pi(s, a_i) = \mathbb{E}_{s'}\left[\underbrace{R(s', a_i)}_{\text{reward definition}} + \gamma Q^\pi(s', \pi(s'))\right]$$

assume linearity:

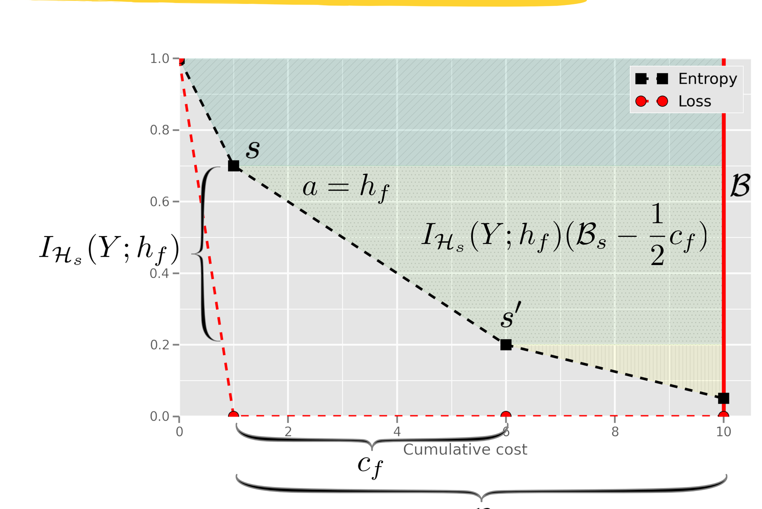
$$Q^\pi(s, a_i) = \underbrace{\theta_\pi^\top}_{\text{learning the policy}} \phi(s, a_i)$$

Learning

Input: $\mathcal{D} = \{x_n, y_n\}_{n=1}^N$; $\mathcal{L}_\mathcal{B}$
Result: Trained π, g

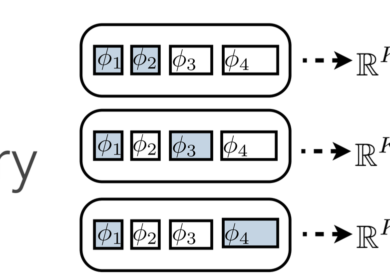
$\pi_0 \leftarrow$ random;
for $i \leftarrow 1$ **to** *max_iterations* **do**
  States, Actions, Costs, Labels $\leftarrow$ `GatherSamples` ($\mathcal{D}$, π_{i-1}) ;
  $g_i \leftarrow$ `UpdateClassifier` (*States, Labels*) ;
  Rewards $\leftarrow$ `ComputeRewards` (*States, Costs, Labels*, g_i, $\mathcal{L}_\mathcal{B}$, γ) ;
  $\pi_i \leftarrow$ `UpdatePolicy` (*States, Actions, Rewards*) ;
end

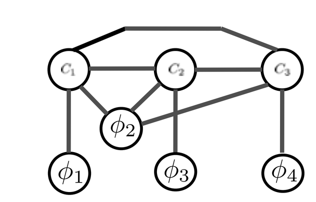
Reward definition



Feature combination

- How to combine arbitrary subsets of features?





Karayev et al. NIPS 2012

	Accuracy	Subset?	Cost
MRF model	✓	✓	✗
Naive Bayes	✗	✓	✓
Linear model	✓	?	✓

- Value imputation
- Learning a mixture of linear models

Baselines

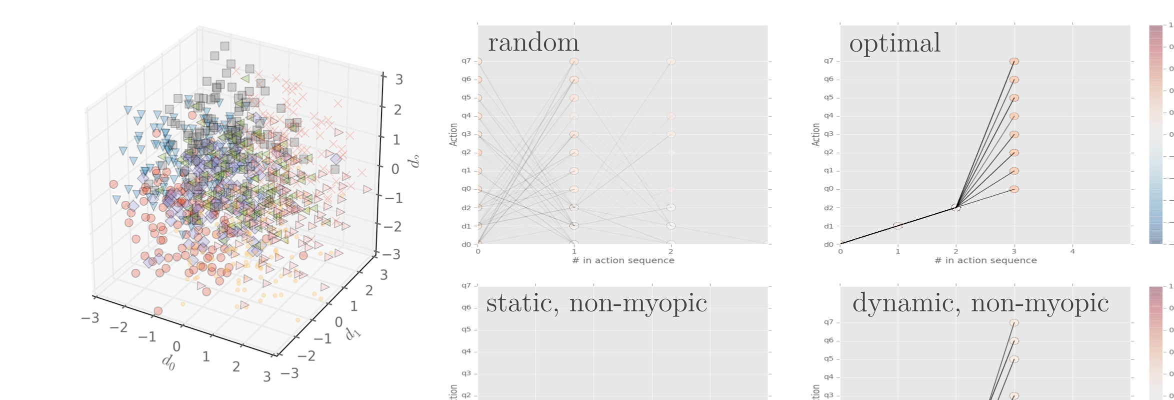
We evaluate the following baselines:

- **Static, greedy**: corresponds to best performance of a policy that does not observe feature values and selects actions greedily ($\gamma = 0$).
- **Static, non-myopic**: policy that does not observe values but considers future action rewards ($\gamma = 1$).
- **Dynamic, greedy**: policy that observes feature values, but selects actions greedily.

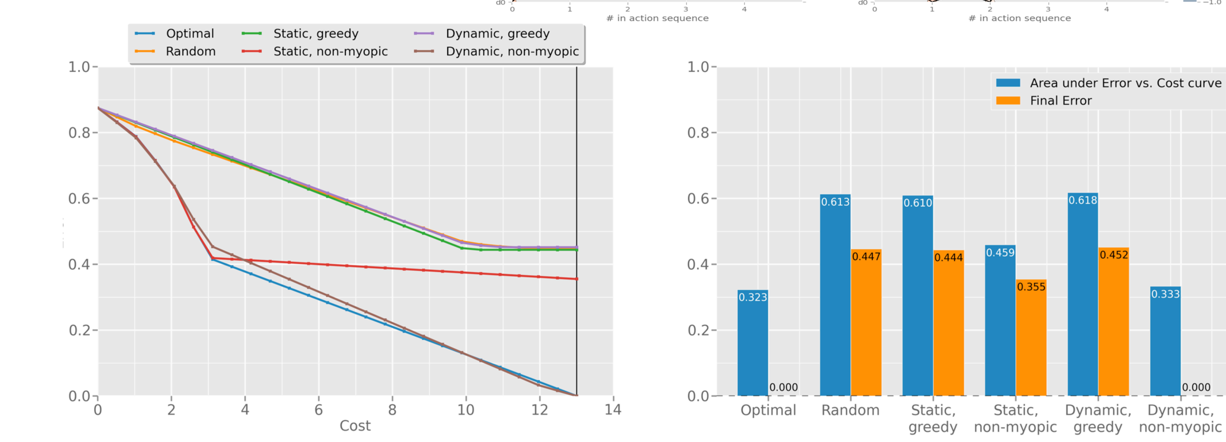
Our method is the **Dynamic, non-myopic** policy: feature values are observed, with full lookahead.

Evaluation

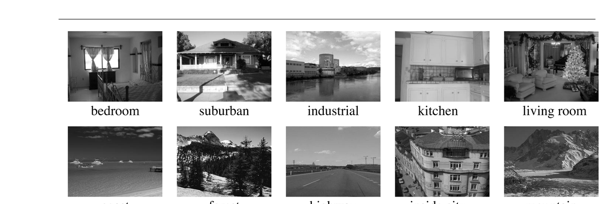
Synthetic Example

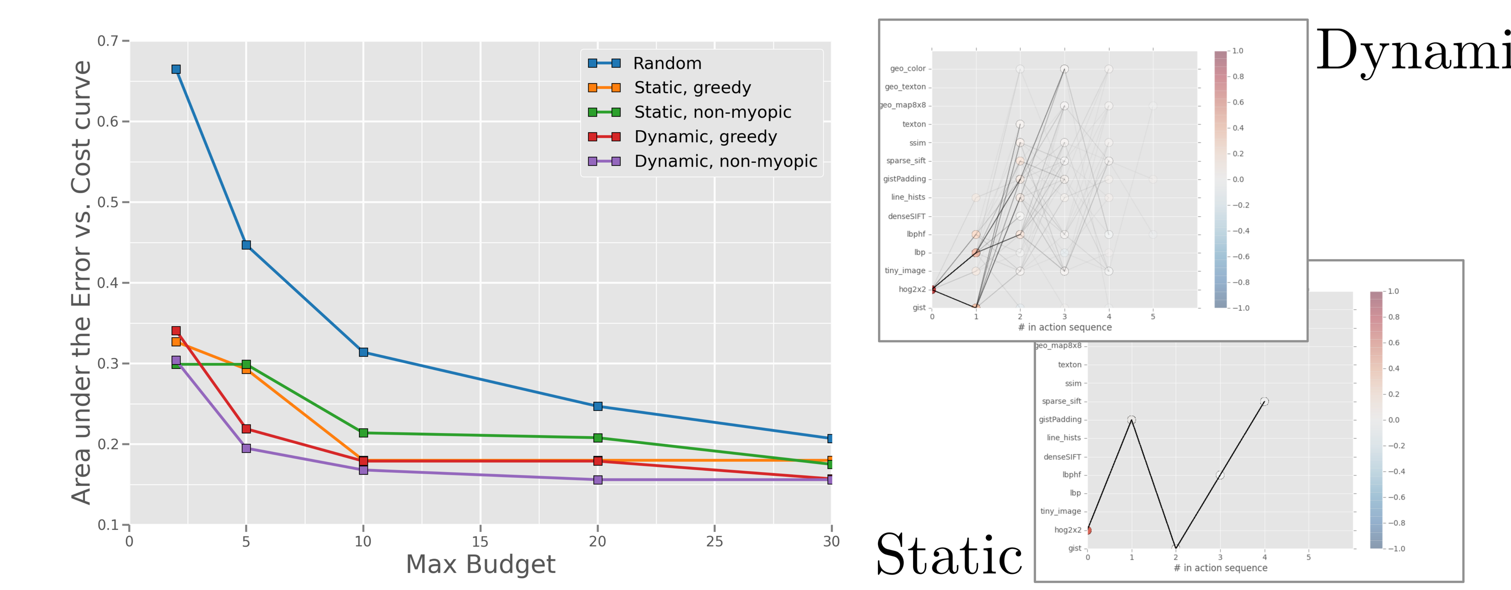


Feature	Number	Cost
d_i: sign of dimension i	D	1
q_o: label of datapoint, if in quadrant o	2^D	10

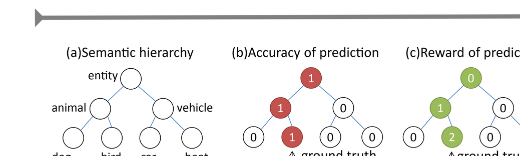


Scenes





Imagenet



- Progressive enhancement of specificity with increasing budget.