Visual meta-learning for planning and control

Seminar on Current Works in Computer Vision
@ Chair of Pattern Recognition and Image Processing.

Samuel Roth

Winter Semester 2018/19

Albert-Ludwigs-Universität Freiburg

Content

- Motivation, Goal.
- Meta learning.
- Concept acquisition through meta learning (CAML).
- Implementation of CAML.
- Use observation classifier as reward functions.
- Experimental results.

Introduction - Motivation

Motivation

- Reinforcement Learning needs rewards for learning.

Introduction - Motivation

Motivation

- Reinforcement Learning needs rewards for learning.

Example Tasks

Introduction - Motivation

Motivation

- Reinforcement Learning needs rewards for learning.
- Specifying rewards are often hand crafted and domain specific.

Example Tasks

Introduction - Motivation

Motivation

- Reinforcement Learning needs rewards for learning.
- Specifying rewards are often hand crafted and domain specific.
- Learning rewards for a specific tasks requires large labeled dataset for training.

Example Tasks

Introduction - Goal and Approach

Goal

- Derive reward function.
- From camera input.
- Only few positive examples of success.
- Utilize information given by other tasks.

Introduction - Goal and Approach

Goal

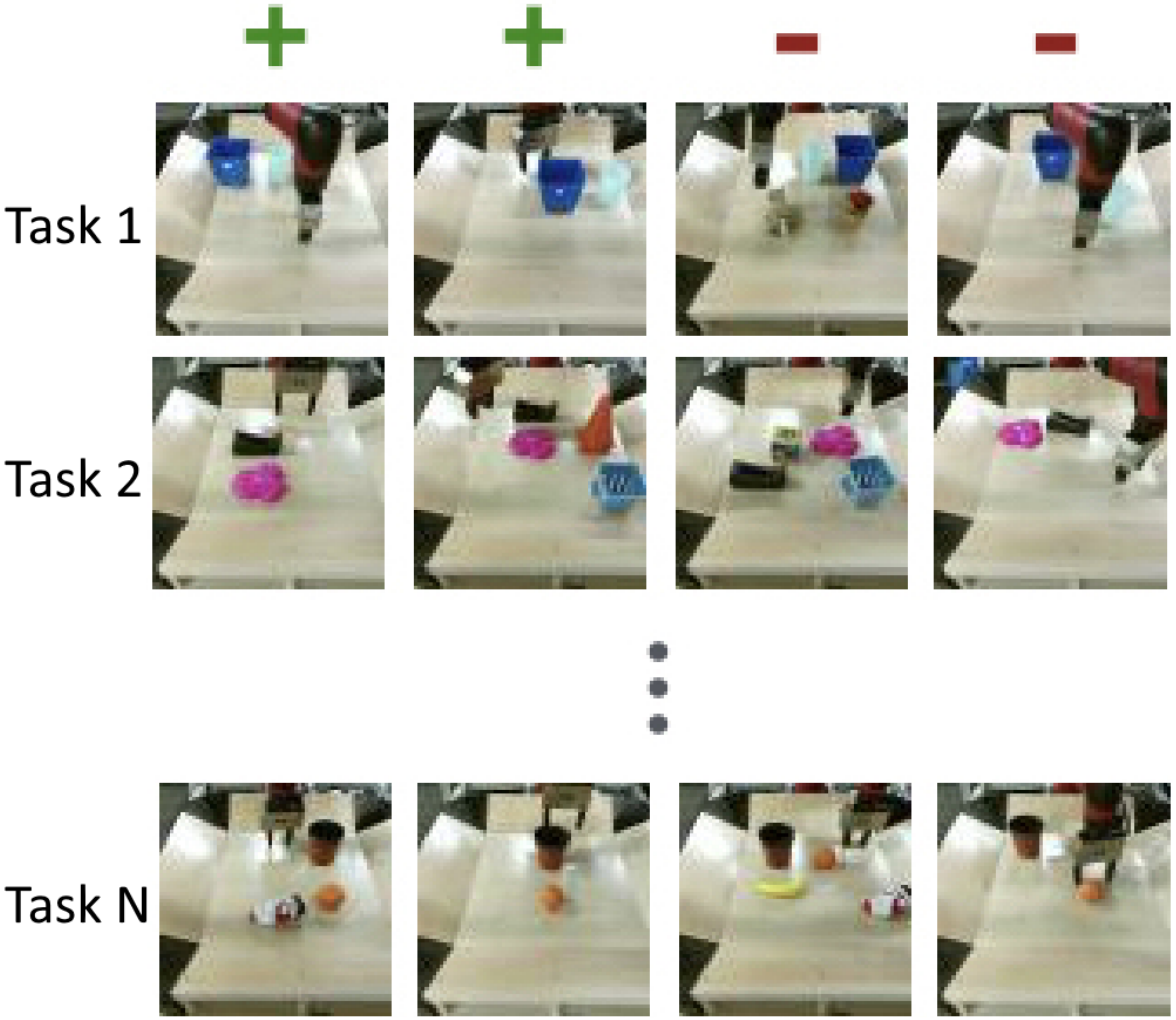
- Derive reward function.
- From camera input.
- Only few positive examples of success.
- Utilize information given by other tasks.

Meta learning approach

- "Learn to learn"
- Train a model to classify tasks for multiple different tasks.
- Expect the model to learn internal representations.
- For a new task is given $\Rightarrow$ Model has a good initialization.

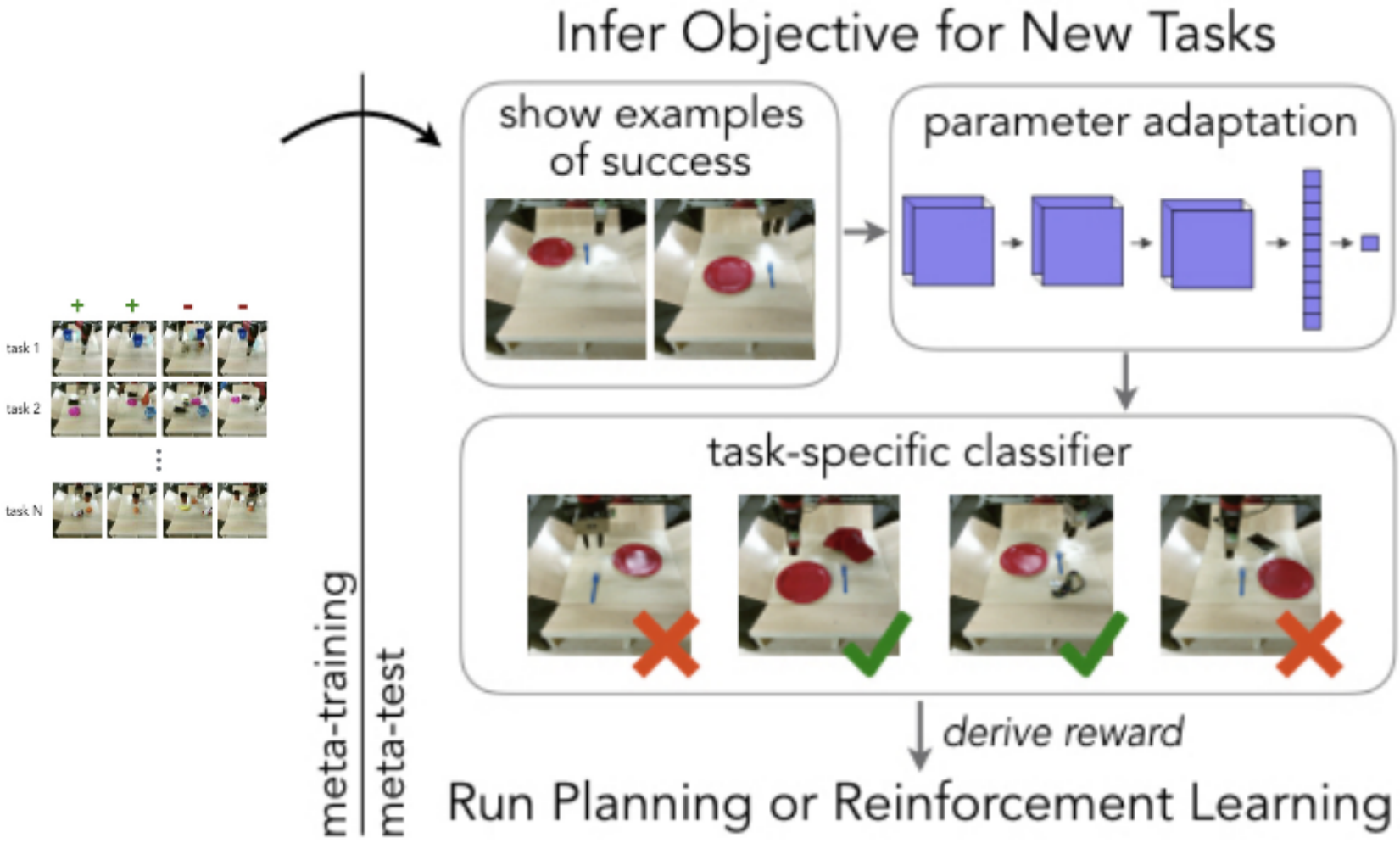
Introduction - Framework

Learn Structure among relevant tasks



Meta-training

Introduction - Framework



Meta Learning for few shot objective learning

Concept

- Given a dataset of labeled observations over multiple tasks

$$D_i := \{(o_j, y_j) | j = 1, \cdots, k; y_j \in \{0, 1\}\}$$

- A meta-learner f_L and a classifier f_C

$$f_C = f_L(D^+_{new}; \theta) : Observation \rightarrow [0, 1]$$

- With the objective to

$$\min_\theta \sum_i^n \sum_{(o_j, y_j) \in D_i^{train}} \mathcal{L}\big(y_j, f_L(D_i^+; \theta)(o_j)\big)$$

CAML - Concept Acquisition through meta learning.

CAML - Concept Acquisition through meta learning.

- Incorporate gradient descent for parameter adaption.

CAML - Concept Acquisition through meta learning.

- Incorporate gradient descent for parameter adaption.
- Adapt θ to new task τ:

$$g(\theta, D_\tau^+) := \theta - \alpha \nabla_\theta \sum_{(o_i, y_i) \in D_\tau^+} \mathcal{L}\big(y_i, f_{CAML}(\theta)(o_i)\big) = \theta'_\tau$$

CAML - Concept Acquisition through meta learning.

- Incorporate gradient descent for parameter adaption.
- Adapt θ to new task τ:

$$g(\theta, D_\tau^+) := \theta - \alpha \nabla_\theta \sum_{(o_i, y_i) \in D_\tau^+} \mathcal{L}\big(y_i, f_{CAML}(\theta)(o_i)\big) = \theta'_\tau$$

- Initial θ learned in meta training:

$$\min_\theta \sum_i^n \sum_{(o_j, y_j) \in D_{\tau_i}^{train}} \mathcal{L}\big(y_j, f_{CAML}(g(\theta, D_{\tau_i}^+))(o_j)\big)$$

CAML - Algorithm.

Algorithm 1 CAML training phase

Require: α, β learning rate Hyper-parameters. k examples size (k-shot).

Randomly initialize θ
while not done **do**
    Sample batch of tasks of size k.
    for all τ_i in batch **do**
        Sample D_i^{train} and D_i^{+} s.t. $D_i^{train} \cap D_i^{+} = \emptyset$
        Calculate task parameters with gradient decent:
        $\theta_i = \theta - \alpha \nabla_\theta \sum_{(o_j, y_j) \in D_i^{+}} \mathcal{L}\big(y_j, f_{CAML}(\theta)(o_j)\big)$
    end for
    Update $\theta \leftarrow \theta - \beta \nabla_\theta \sum_{\tau_i \in \mathrm{batch}} \sum_{(o_j, y_j) \in D_i^{train}} \mathcal{L}\big(y_j, f_{CAML}(\theta_i)(o_j)\big)$
end while

CAML - Algorithm.

Algorithm 2 CAML test phase

Require: D^{+}_{new} positive examples of success for a new task.

Require: θ Pre-trained meta-learner parameters.

Require: α learning rate Hyper-parameters.

1: Adapt parameters with gradient decent:

2: $\theta' = \theta - \alpha \nabla_{\theta} \sum_{(o_j, y_j) \in D^{+}_{new}} \mathcal{L}\big(y_j, f_{CAML}(\theta)(o_j)\big)$

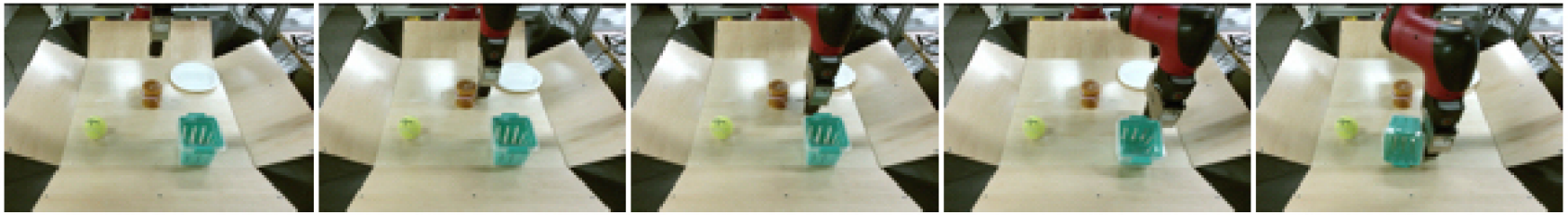
3: **return** $f_{CAML}(\theta')$

From observation classifier to rewards.

- $f_C(o) \in [0, 1]$ interpreted as probability of o to ba a successful execution of the task.
- Directly use the f_C output as reward.
- To reduce the effect of false positives.
- Use a threshold $t \in (0, 1)$ and set $f_C(o) \leftarrow 0$ if $f_C(o) < t$.

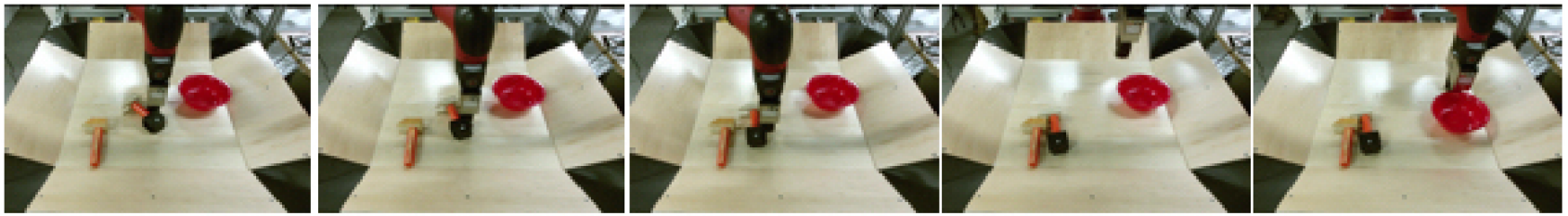
Experimental results - Single task relative position

example 1 example 2



Experimental results - Multi task relative position

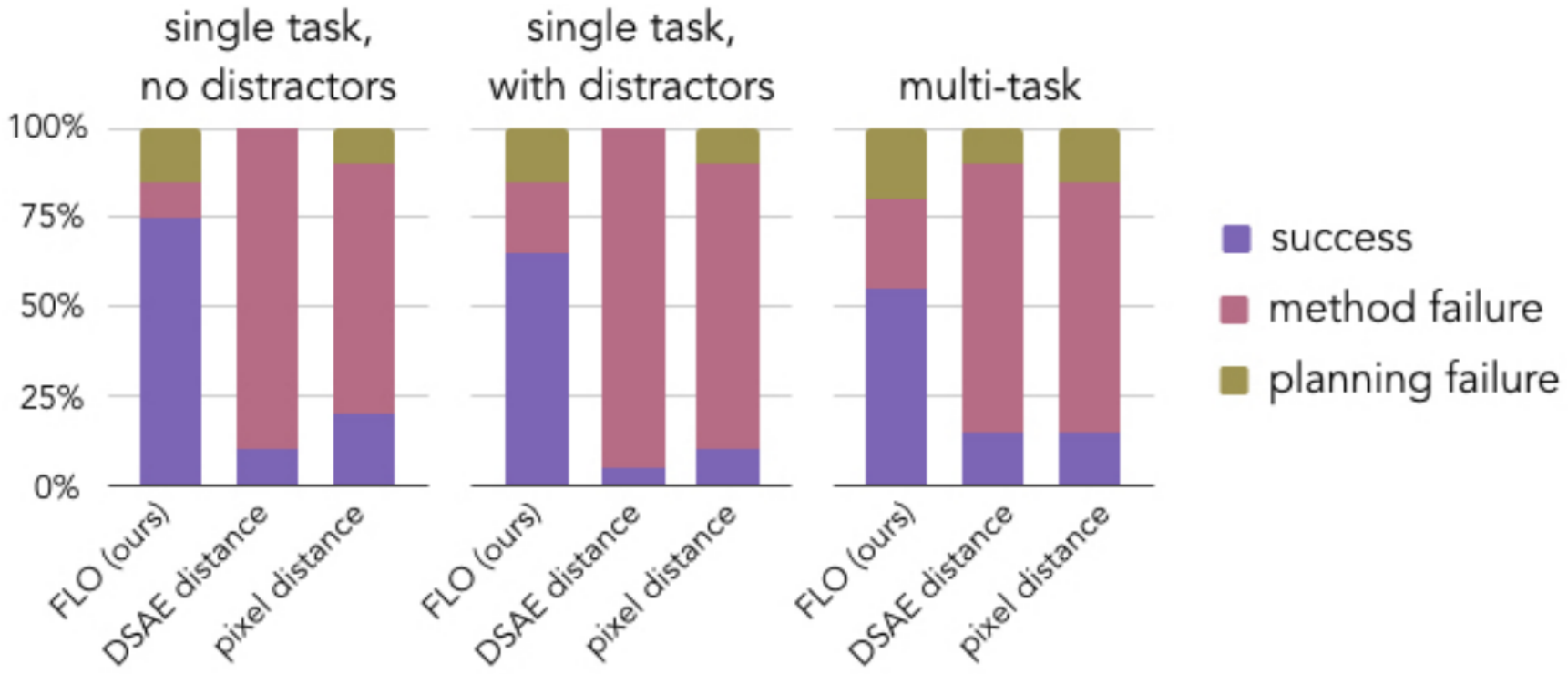
task 1 task 2



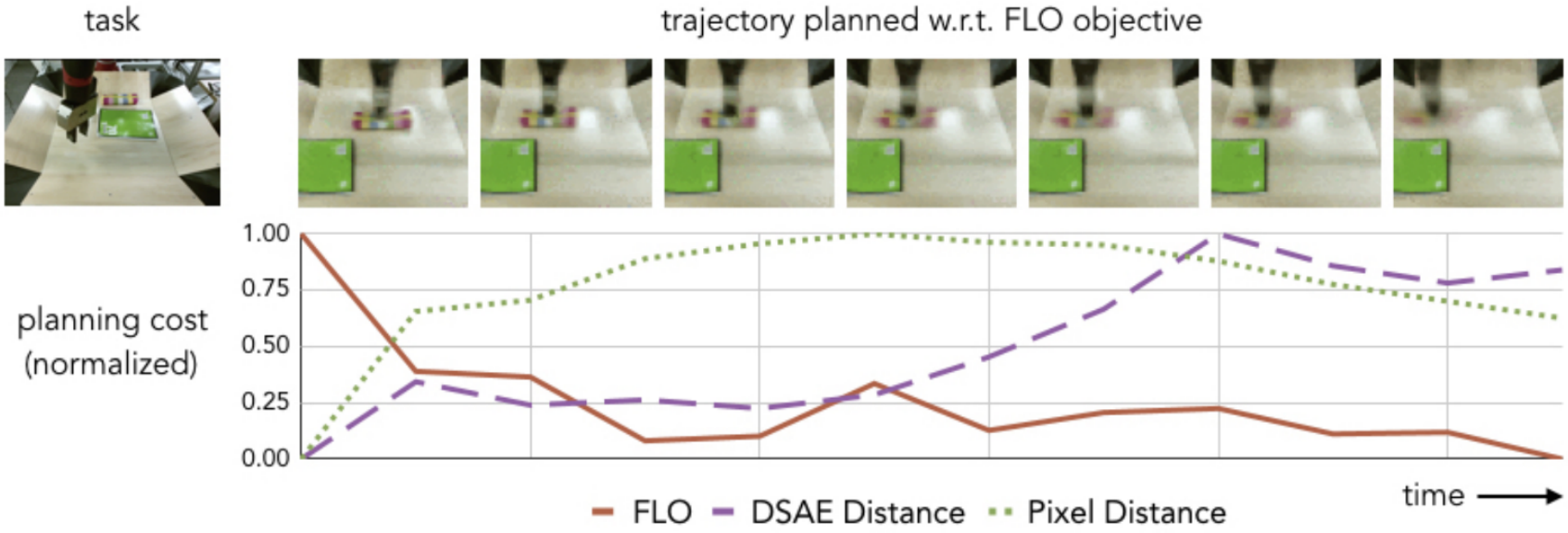
Experimental evaluation competitors

- **Pixel distance:** Measure the ℓ_2 distance between observation and successful observation.
- **Latent space distance:** Measure distance in a learned latent space using an auto encoder. Here "Deep spatial autoencoders for visuomotor learning" by Finn et al. 2016.
- **Oracle:** Use the ground truth if available to get an upper bound of performance (only applied in the rope manipulation domain).

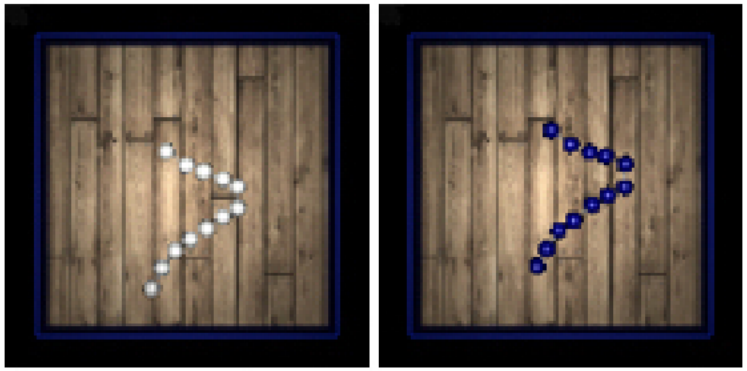
Experimental results - Relative position



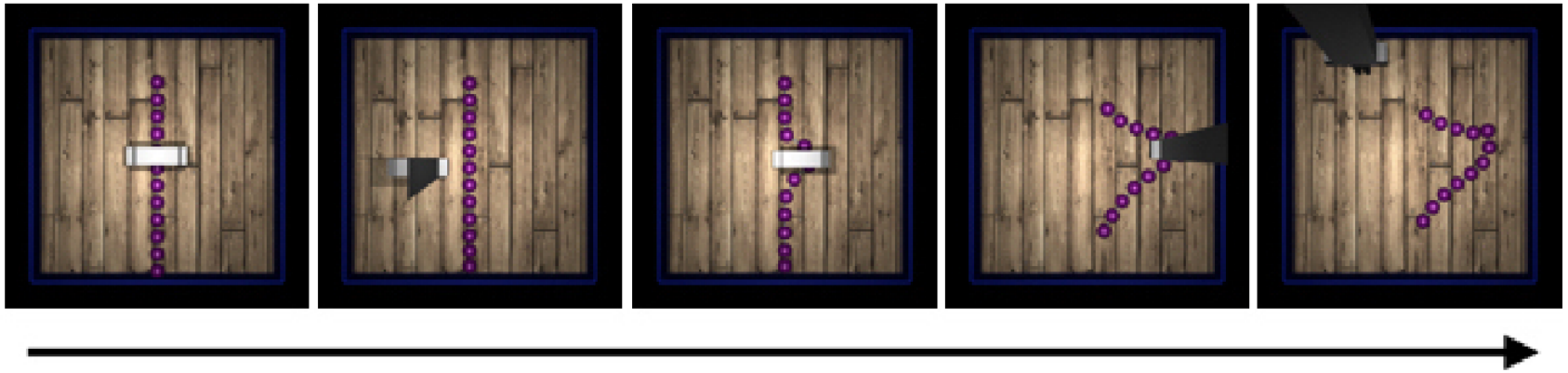
Experimental results - Relative position planning costs



Experimental results - Rope manipulation

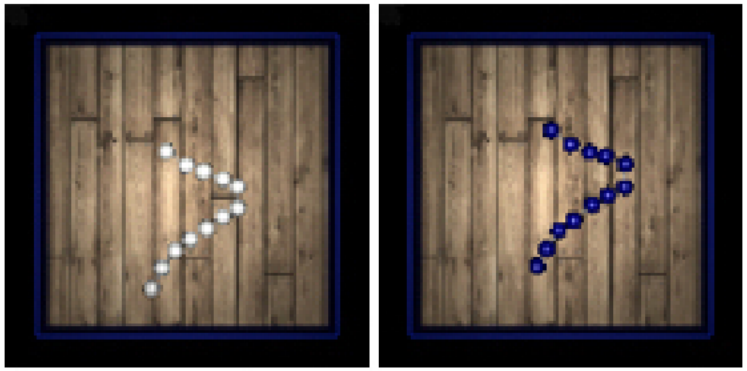


example 1 example 2



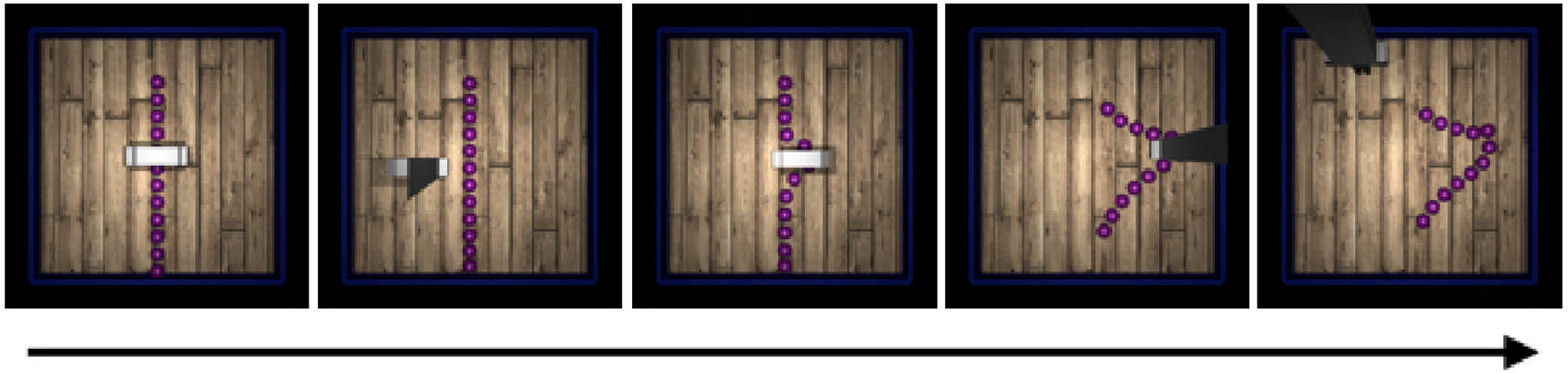
time

Experimental results - Rope manipulation

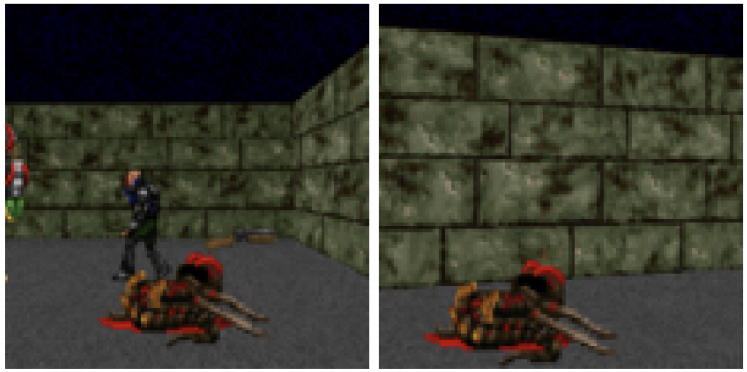


example 1 example 2

method	median distance
pixel distance	0.54 (0.34, 0.70)
AE distance	0.59 (0.50, 0.69)
FLO (ours)	0.27 (0.18, 0.34)
oracle	0.16 (0.13, 0.23)



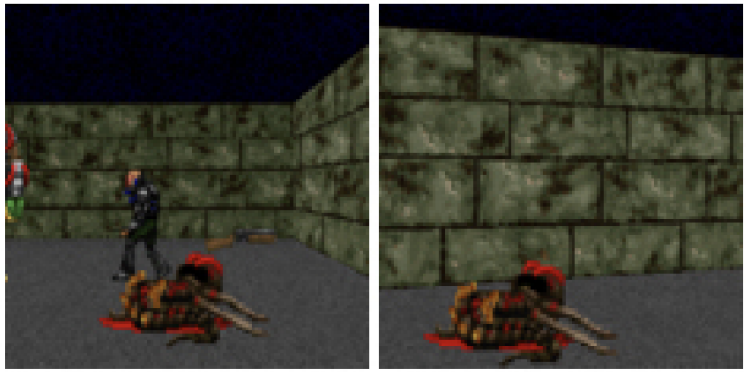
Experimental results - 3D navigation



example 1 example 2

time

Experimental results - 3D navigation



example 1 example 2

method	success rate
pixel distance	33.0%
AE distance	44.0%
FLO (ours)	77.8%

time

Conclusion

Conclusion

- Model that can achieve rapid adaptation.

Conclusion

- Model that can achieve rapid adaptation.
- Internal representation that is broadly suitable to many tasks.

Conclusion

- Model that can achieve rapid adaptation.
- Internal representation that is broadly suitable to many tasks.
- If tasks have distinct visual concept unlikely that meta-learning will acquire useful knowledge.

Conclusion

- Model that can achieve rapid adaptation.
- Internal representation that is broadly suitable to many tasks.
- If tasks have distinct visual concept unlikely that meta-learning will acquire useful knowledge.
- Experiments showed good performance on vision-based manipulation skills.

Thanks!

Bibliography

Finn, C., Abbeel, P., and Levine, S. (2017).
Model-agnostic meta-learning for fast adaptation of deep networks.
arXiv preprint arXiv:1703.03400.

Xie, A., Singh, A., Levine, S., and Finn, C. (2018).
Few-shot goal inference for visuomotor learning and planning.
CoRR, abs/1810.00482.