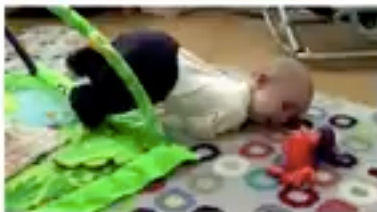


Learning to Poke by Poking: Experiential Learning of Intuitive Physics

2017.06.29

Introduction

Babies learn by experimentation



0:27 / 21:01

The Scientist in a Crib, Gopnik, Meltzoff & Kuhl, 1999



Six Lessons from Babies, Smith & Gasser, 2005

겉보기에 아기가 무작위로 행동한다고 생각할 수도 있지만 행동을 통해 그 결과를 관찰하고 배운다.
즉, 물체의 물성을 추론하고 역학을 예측할 수 있다.
(ex. 해머를 사용할 수없는 경우 못을 치기 위해 돌이나 드라이버 뒤를 사용할 수 있다.)

Like babies, the robot performs experiments ...



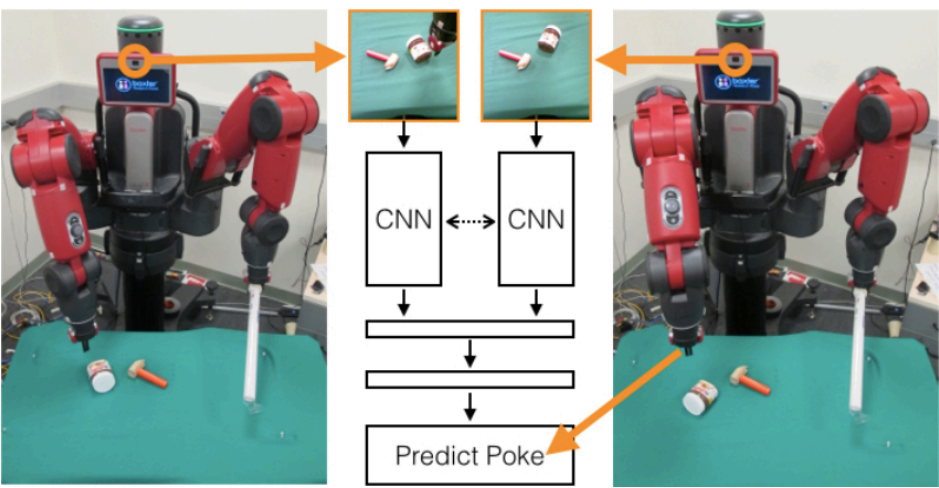
무작위 실험을 통해 로봇이 자신의 경험을 사용하여 직관적인 물리학 모델을 학습할 수 있는지 연구한다.

학습할 모델은 현재의 visual 상태 및 적용된 힘 (즉, forward dynamics 모델)으로부터 다음 visual 상태를 예측하는 모델이다.

실제로는 다음 이미지의 모든 픽셀 값을 예측하는 것이 쉽지 않고, 대부분의 경우 정확한 픽셀 값 보다 추상적인 이벤트에 관심이 있다. (ex. 유리병이 땅에 떨어질 때 부서질 것이라고 예측하는 것이 부서진 유리 조각의 픽셀을 정확하게 예측하는 것보다 더 큰 관심사)

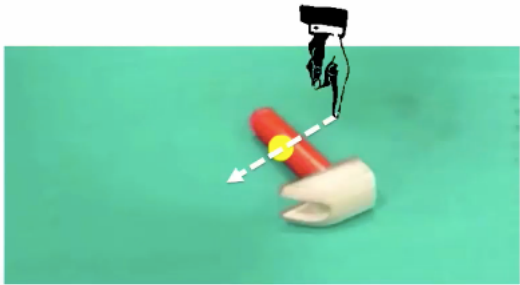
문제는 추상 개념이나 사건에 대한 supervision이 unsupervised 환경에서는 쉽게 이용할 수 없는 것인데, 이 연구에서는 forward dynamics 모델과 inverse dynamics 모델을 공동으로 training하여 문제에 대한 솔루션을 제안한다.

Data

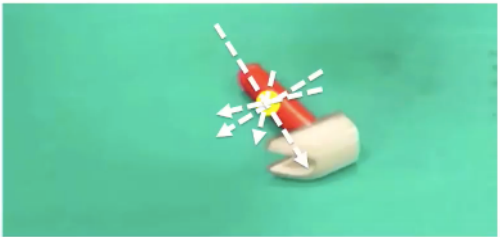


로봇에는 Kinect 카메라와 앞에 있는 탁자 위에 물건을 poking 할 수있는 gripper 가 장착되어 있다.
주어진 시간에 테이블에 존재하는 16 개의 별개의 물체들의 집합으로부터 1-3개의 물체를 선택한다. 로봇의 좌표계 X 축과 Y 축은 수평 및 수직 축을 나타내고 Z 축은 로봇 반대방향을 가리킨다.
로봇은 테이블에서 고정 된 높이로 xz 평면을 따라 finger를 움직여 물체를 찌른다.

Robot makes random pokes ...



Robot makes random pokes ...



빠른 데이터 수집을 위해 Kinect depth 카메라의 point cloud(각 픽셀마다 rgb데이터와 3차원 좌표)를 사용하여 모든 물체에있는 점만 선택했다. 포인트 클라우드 정보는 데이터 수집중에만 사용되었으며 테스트시 시스템은 RGB 이미지 데이터만 필요로 한다. 임의의 점을 선택하여 물체에 poke(p)하기위해, 로봇은 poke방향(θ)과 길이 (l)를 random하게 샘플링한다.

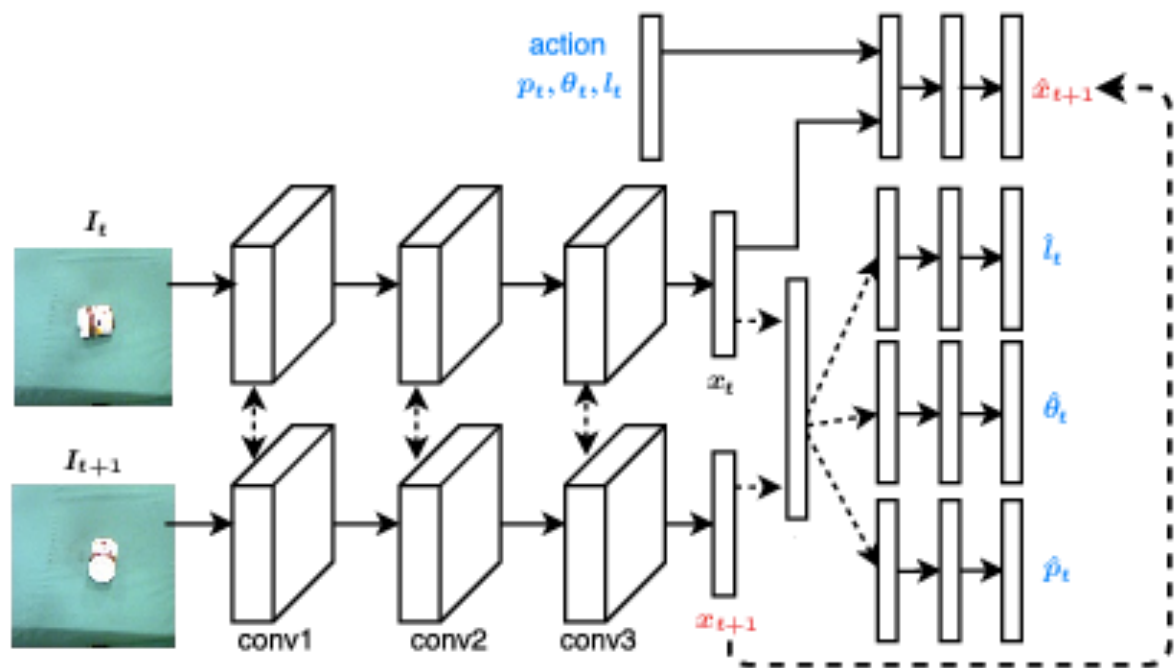
Method

$$\hat{x}_{t+1} = F(x_t, u_t; W_{fwd}) \tag{1}$$

$$\hat{u}_t = G(x_t, x_{t+1}; W_{inv}) \tag{2}$$

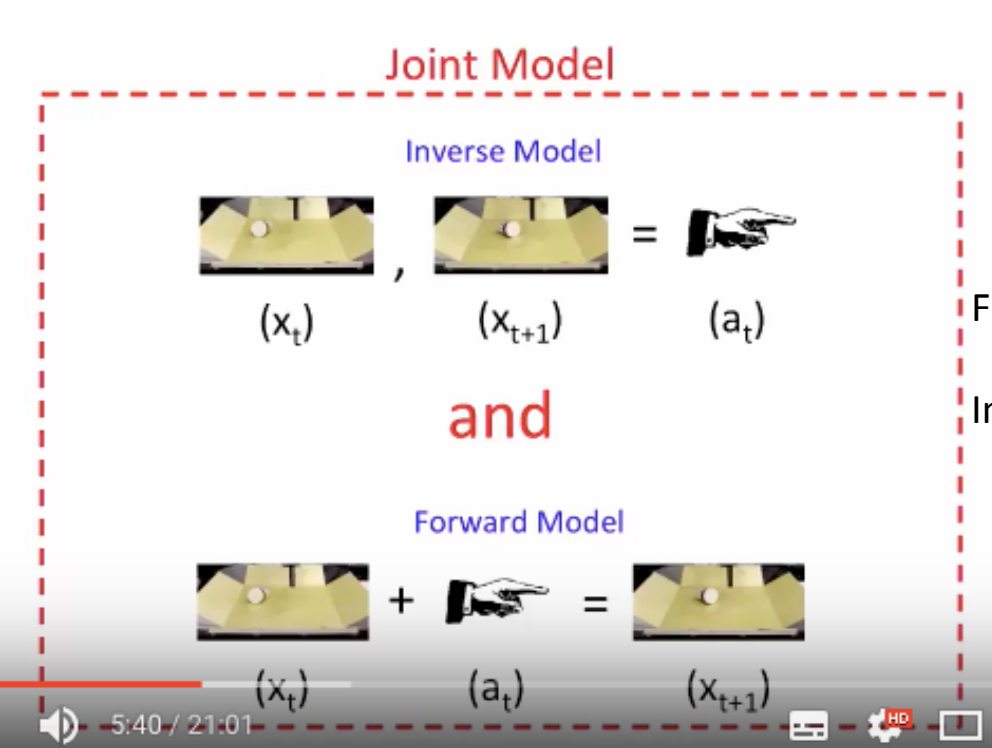
forward 및 inverse 모델은 각각 식 1 및 2로 나타낼 수 있다.

- xt : 상태 (time step t)
- ut : action (time step t)
- xt + 1 : 예측된 상태
- ut + 1 : 예측된 action
- Wfwd : 함수 F의 parameter
- Winv : 함수 G의 parameter



Inverse model :
It에서 이미지를 It + 1 상태로 변환하
는데 필요한
위치 (pt), 각도 (θt) , 길이 (lt)를 예
측하도록 훈련.

Model



Forward 모델 : 현재 상태 및 동작에서 다음 상태 예측

Inverse 모델 : 초기 상태 및 목표 상태에서의 동작 예측

$$L_{joint} = L_{inv}(u_t, \hat{u}_t, W) + \lambda L_{fwd}(x_{t+1}, \hat{x}_{t+1}, W) \quad (3)$$

L_{inv} : 실제 및 예상 된 poke의 위치, 각도, 길이 사이 3 가지 cross 엔트로피 손실의 합.

L_{fwd} : $t + 1$ 의 예측된 상태(\hat{x}_{t+1})와 실제 상태(x_{t+1}) feature representation 사이의 L1 loss.

W : parameter of the neural network

식 3에서 람다를 0으로 설정하여 inverse 모델만 학습한 것과 성능을 비교한다.

Inverse Model

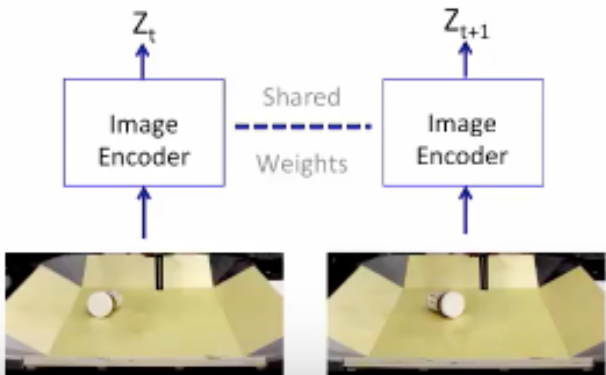
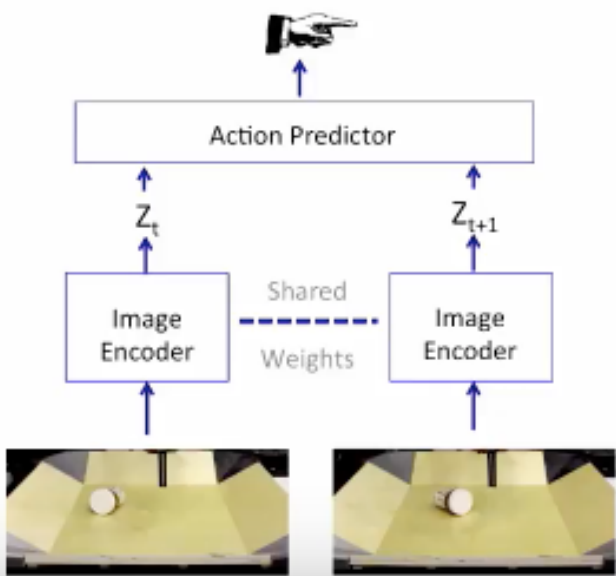


Image Encoder is
a ConvNet with
AlexNet architecture

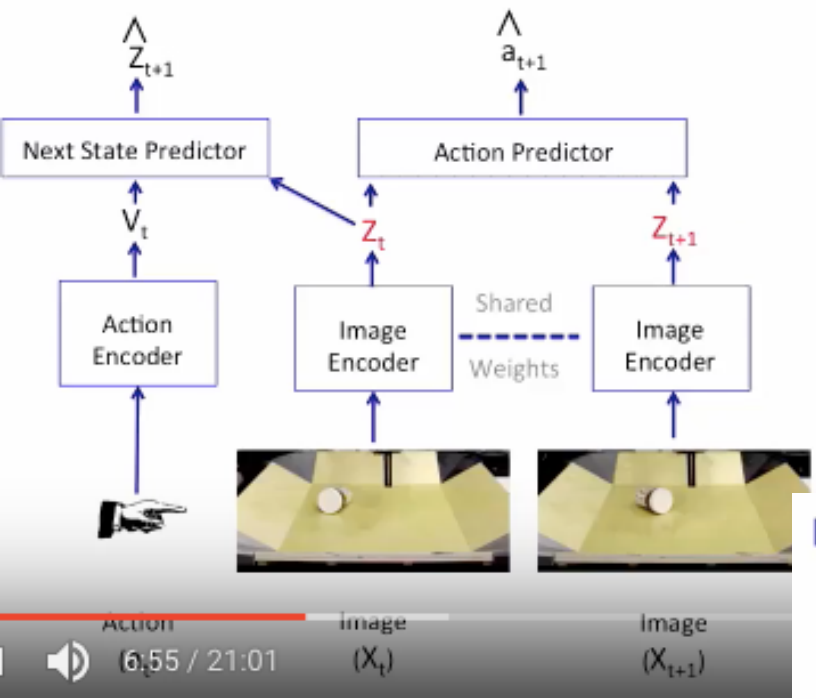
(random initialization)

Inverse Model



Action Predictor
is a two layer
neural network

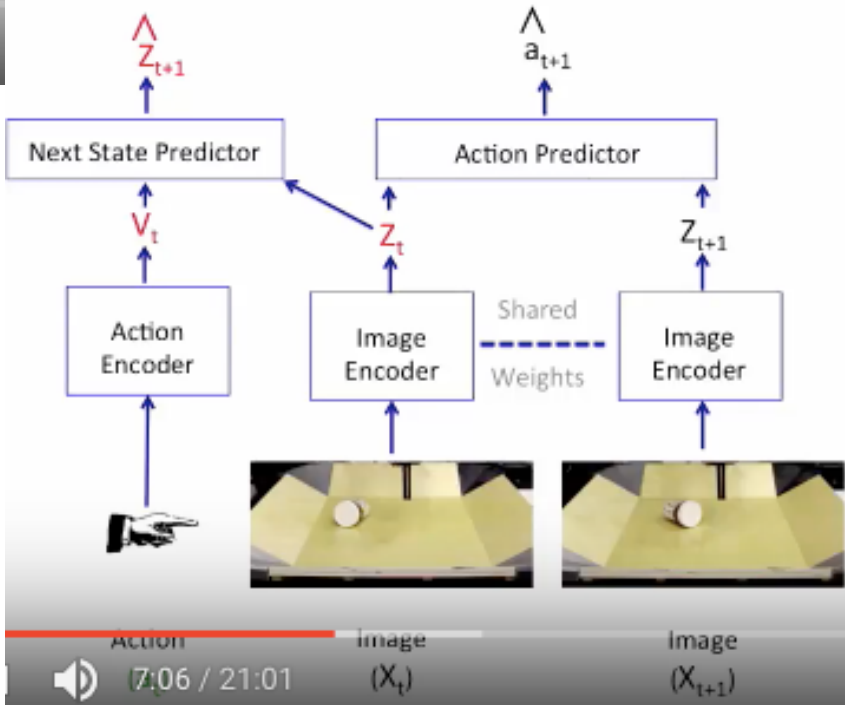
Inverse model learns a feature space (Z)



Inverse model에서 학습된 X_t 의 feature space를 forward model에서 다음 state를 예측하는데 사용.

Pixel space가 아닌 abstract space에서 dynamics model을 만듦.

Inverse model learns a feature space

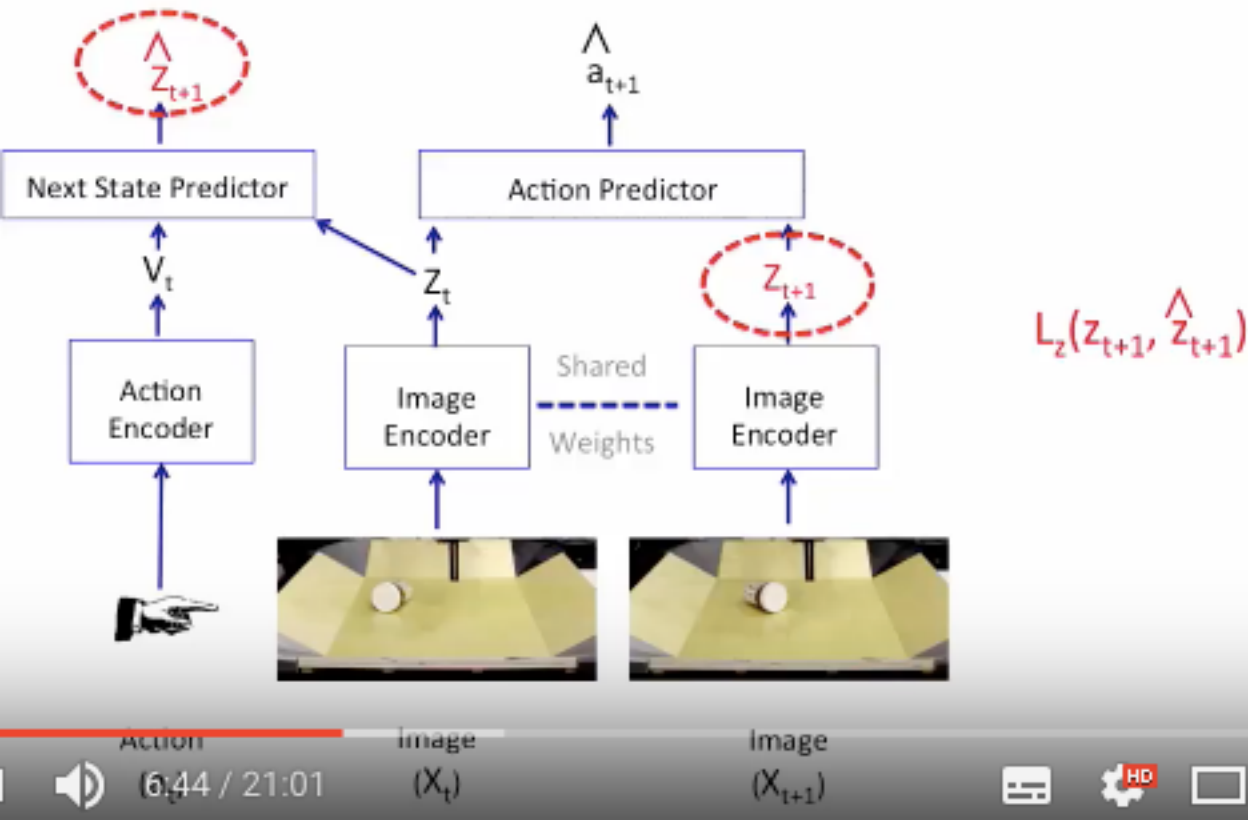


$\hat{Z}_{t+1} = f(Z_t, V_t)$

Forward model makes predictions in the feature space learnt by the inverse model!

An alternative to making predictions in the pixel space !!!

Learning joint Forward + Inverse Model



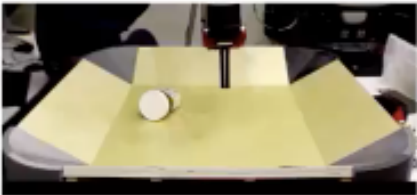
Joint training에서, inverse 모델은 이미지 픽셀을 추상 피쳐 공간으로 변환하기위한 supervision을 제공하고, forward 모델은 다음상태를 예측할 수 있다.

Inverse 모델은 픽셀 공간에서 예측을 위한forward 모델의 필요성을 완화시키고, forward 모델은 차례로 inverse 모델의 특징 공간을 정규화한다.


Evaluation Procedure

학습 된 모델을 테스트 하기 위해 로봇에게 초기 및 목표 이미지를 제공하고 목표 이미지에 표시된 구성으로 object를 이동시키는지 확인한다.

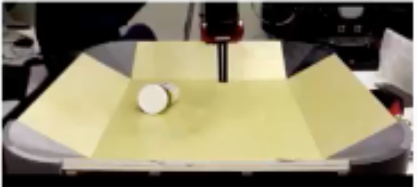
X_t




X_{t+1}



Initial Image

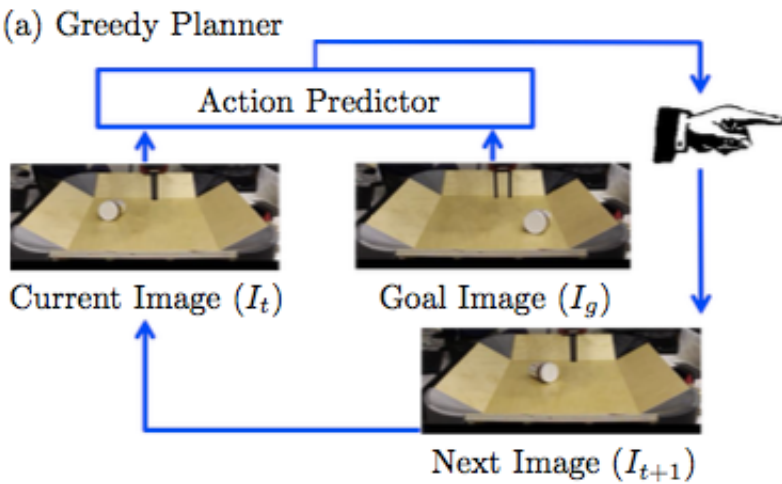


Goal Image



Training

Testing



초기이미지와 목표이미지에서 물체가 single poke로 밀수있는 최대 거리보다 멀리 떨어져 있는 경우 greedy planning방법을 사용한다.
이 과정은 로봇이 no-poke를 예측하지 않거나, 최대 10번 pokes에 도달하지 않는 한 반복된다.

(c) Pose Error Evaluation



Error metrics:

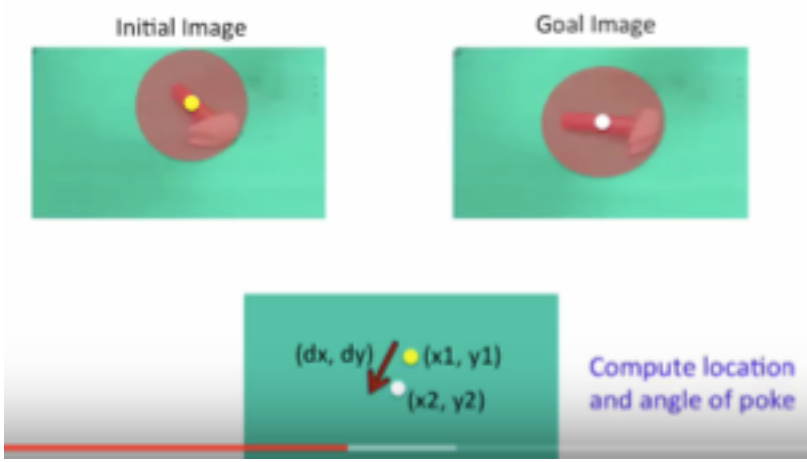
모든 실험에서 초기이미지와 목표 이미지는 물체의 위치만 다르다.

위치 오차는 물체위치 사이의 유클리드 거리이다.

상대적인 위치 error(어느 지점에서 얼마나 이동했는지)를 사용한다.

포즈에러(c)는 최종 이미지와 목표이미지에서 물체의 major axis 간의 angle로 정의된다.

3.3 Blob Model



Blob 모델은 template based object detector를 사용하여 현재 및 목표 이미지의 object 위치를 평가한다. 그런다음 벡터 차이를 사용하여 로봇에 의해 실행된 poke의 길이, 각도, 위치를 계산한다. 오브젝트가 목표이미지에서 미리 정의된 threshold나 최대 pokes수에 도달함으로 원하는 위치에 가까워질 때까지 프로세스가 반복된다.

Results

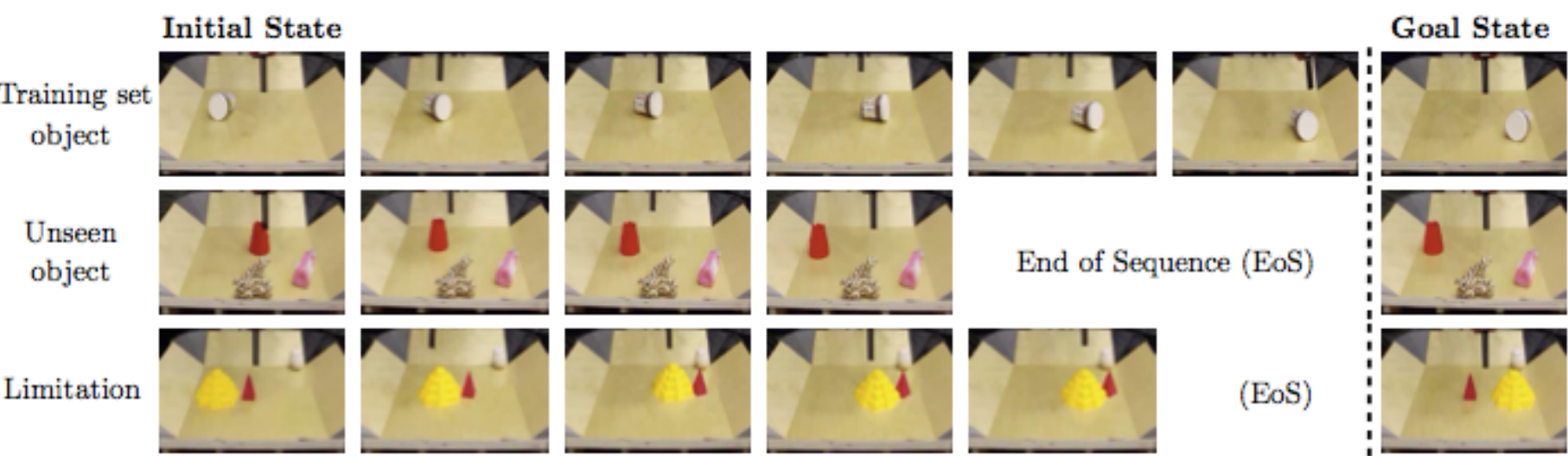


Figure 5: The robot is able to successfully displace objects in the training set (row 1; Nutella bottle) and objects with previously unseen geometry (row 2; red cup) into goal locations that are significantly farther than pair of before and after images used in the training set. The robot is unable to push objects around obstacles (row 3; limitation of greedy planning).

Figure 5에서 로봇은 초기 상태에서 목표 상태로 물체를 이동하는 작업을 했다.

로봇은 트레이닝셋에 있는 물체(누텔라병),트레이닝 셋에 없는 물체(빨간컵)를 성공적으로 이동시킬 수 있다. Greedy planning의 한계로 장애물 주변으로 물체를 움직일 수 없다.

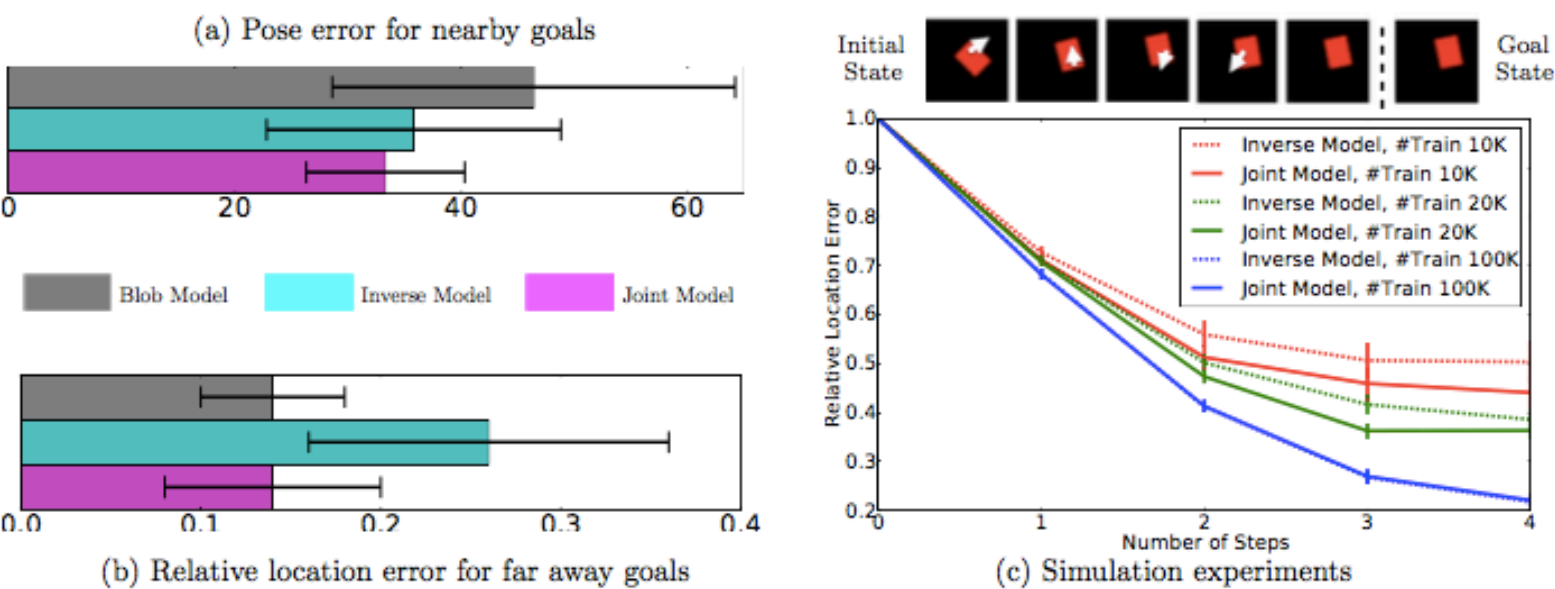


Figure 6: (a) Inverse and Joint model are more accurate than the blob model at pushing objects towards the desired pose. (b) The joint model outperforms the inverse-only model when the robot is tasked to push objects by distances that are significantly larger than object distance in before and after images used in the training set (i.e. a test of generalization). (c) Simulation studies reveal that when less number of training examples (10K, 20K) are available the joint model outperforms the inverse model and the performance is comparable with larger amount of data (100K). This result indicates that the forward model regularizes the inverse model.

- (a) Inverse모델과 joint모델은 물체를 원하는 포즈로 바꿀때 blob모델보다 더 정확하다.
- (b) Joint모델은 로봇이 training set에서 사용된 before와 after이미지에서 이동거리보다 먼 거리로 물체를 이동시켜야 할 때 inverse-only모델보다 우수하다.
- (c) 시뮬레이션 연구에 따르면 적은 수의 훈련데이터(10K,20K)를 사용할 경우 joint모델이 inverse 모델보다 뛰어나고 더 적은 action으로 목표에 가까이 도달함을 보여준다.
이 결과는 forward모델이 inverse모델을 정규화 하는 것을 나타낸다.
그러나 training 수가 100k로 증가하면 두 모델의 성능은 동등한 수준이다.

Summary

- Jointly model forward + inverse dynamics
 - Learn an abstract feature space for forecasting
 - Forward model regularizes inverse model
- Multistep planning
- Generalization
 - To unseen objects and locations
 - Easily extends to deformable objects



14:34 / 21:00



참고링크

1. <http://ashvin.me/pokebot-website/>
2. <https://www.youtube.com/watch?v=9Gu6AkgttwI>