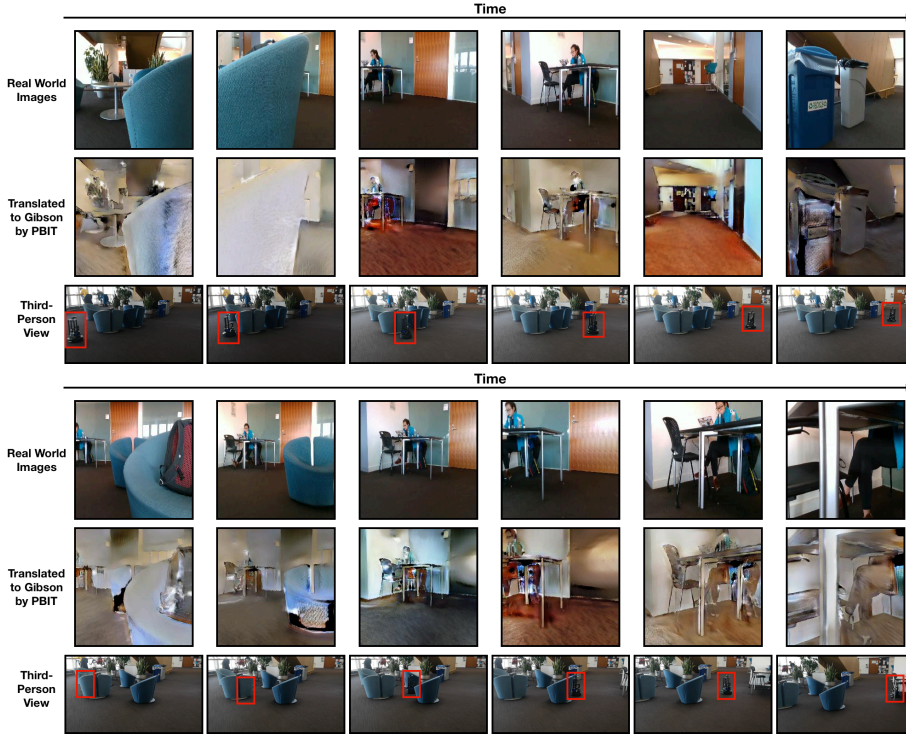
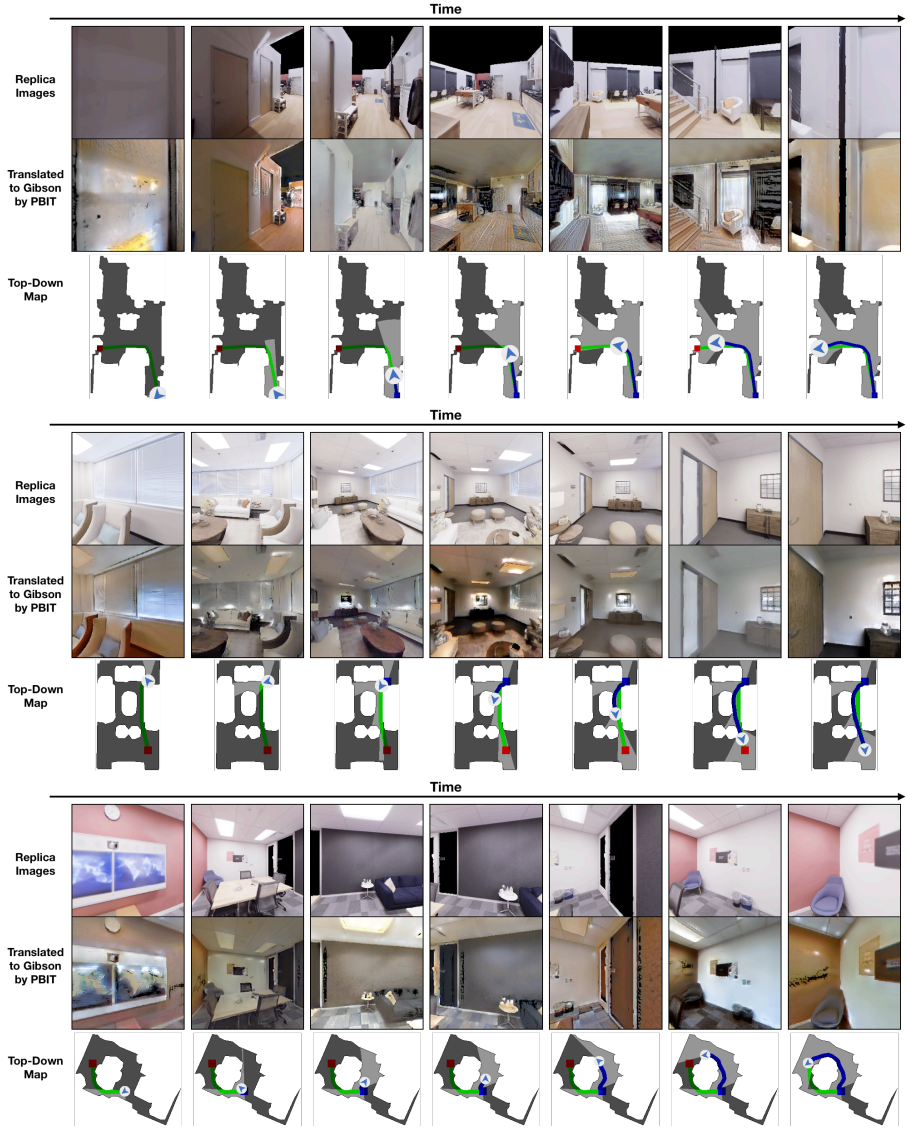
Unsupervised Domain Adaptation for Visual Navigation: Supplementary Material

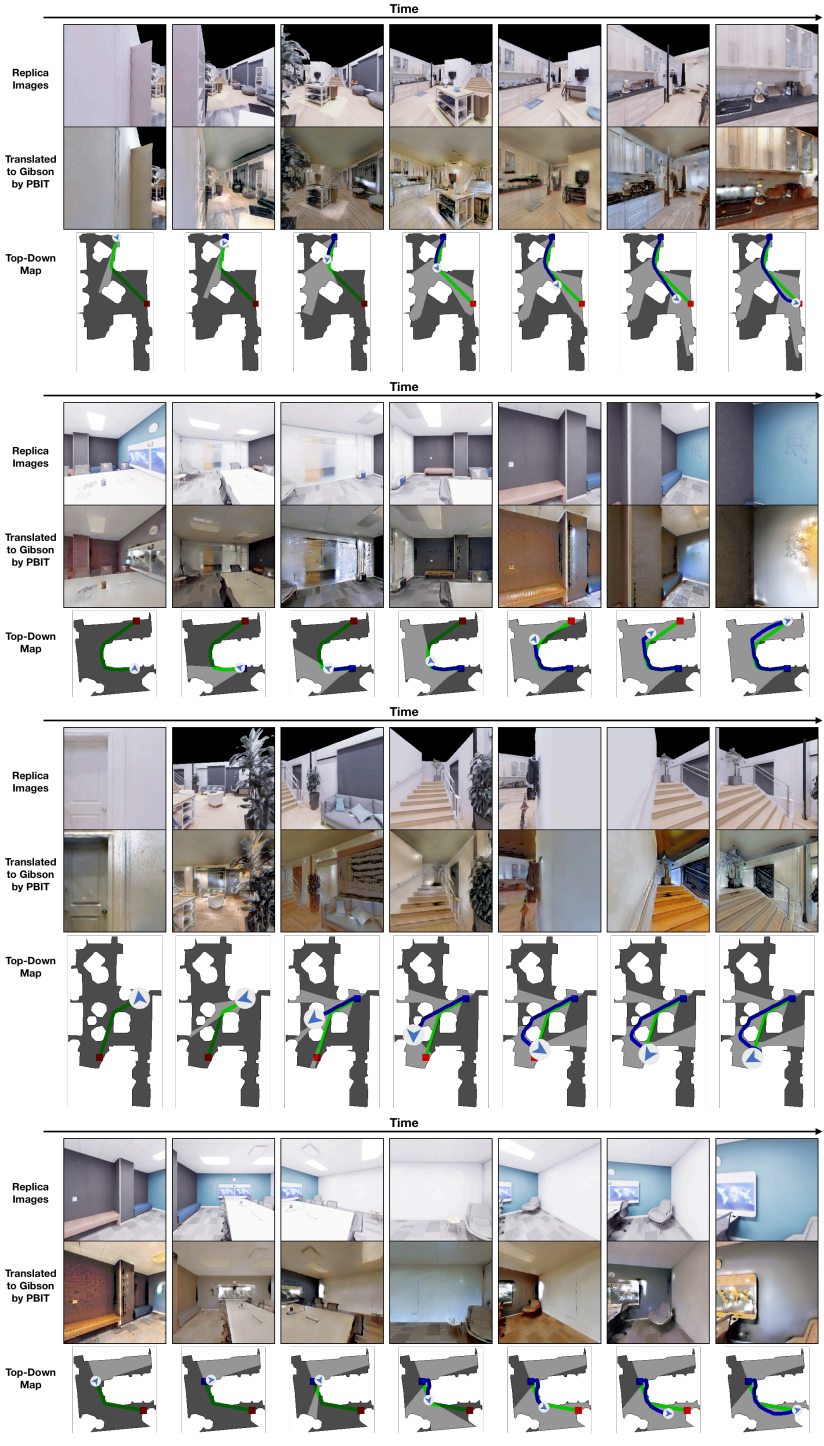
Anonymous ECCV submission

Paper ID 1995

A Additional Trajectory Visualizations




Time
Replica Images
Translated to Gibson by PBIT
Top-Down Map
Time
Replica Images
Translated to Gibson by PBIT
Top-Down Map
Time
Replica Images
Translated to Gibson by PBIT
Top-Down Map


Time
Replica Images
Translated to Gibson by PBIT
Top-Down Map
Time
Replica Images
Translated to Gibson by PBIT
Top-Down Map
Time
Replica Images
Translated to Gibson by PBIT
Top-Down Map
Time
Replica Images
Translated to Gibson by PBIT
Top-Down Map

B Visualization of Policy Representations

We analyze the policy representation before and after translation by reducing the dimensionality of the policy representations using Principle Component Analysis (PCA) [1]. In Figures 1 and 2, we visualize the policy representations reduced to 2 dimensions using PCA in Replica and Real-World respectively. Both figures show that PBIT brings the representations of target domain Replica/Real-World images closer to the distribution of representations of Gibson images.

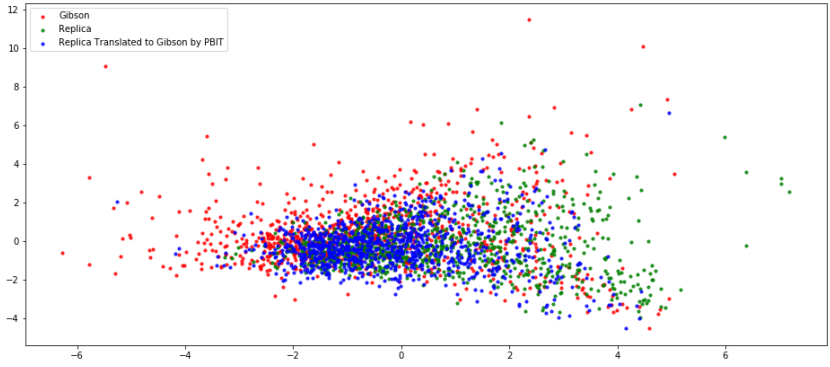


Fig. 1: We use PCA to visualize and compare the 128-dimensional task-specific feature vectors produced by PointGoal agent trained in Gibson, when given Gibson images, Replica images, or Replica images translated to Gibson by PBIT as input. The figure shows the translated images by PBIT bridge the policy domain gap between Gibson and Replica.

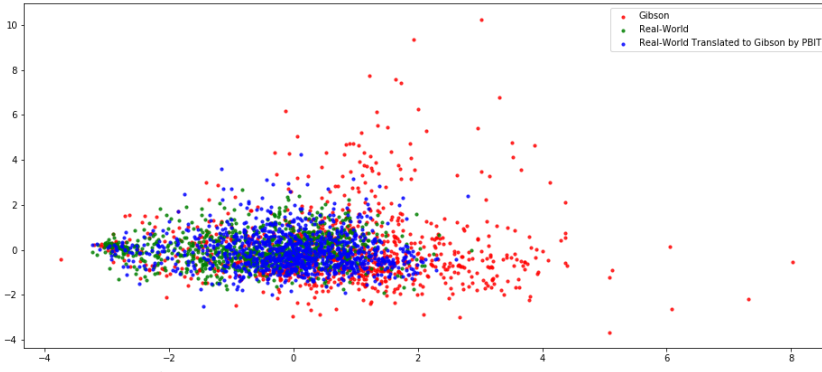


Fig. 2: We use PCA to visualize and compare the 128-dimensional task-specific feature vectors produced by PointGoal agent trained in Gibson, when given Gibson images, Real-World images, or Real-World images translated to Gibson by PBIT as input. The figure shows the translated images by PBIT bridge the policy domain gap between Gibson and Real-World.

C Model Architecture Details

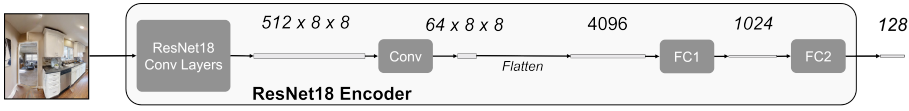


Fig. 3: An illustration of the network architecture of the Visual Policy Encoder.

Table 1: Architecture specification of different parts our PBIT model

Style Encoder

(256×256) RGB
Conv(32,(7x7),1)
Conv(64,(4x4),2)
Conv(128,(4x4),2)
Conv(128,(4x4),2)
Conv(128,(4x4),2)
global_avg_pool
dense(8)
$\rightarrow s \in \mathbb{R}^8$

Content Encoder

(256×256) RGB
Conv(32,(7x7),1)
Conv(64,(4x4),2)
Conv(128,(4x4),2)
Residual(128)
Residual(128)
Residual(128)
Residual(128)
$\rightarrow c \in \mathbb{R}^{128 \times 64 \times 64}$

Decoder

c and AdaIN(s)
Residual(128)
Residual(128)
Residual(128)
Residual(128)
UpSampling(2x2)
Conv(64,(5x5),1)
UpSampling(2x2)
Conv(32,(5x5),1)
UpSampling(2x2)
Conv(3,(7x7),1)
(256×256) RGB

Discriminator

Multi-scale Input
Conv(64,(4x4),2)
Conv(128,(4x4),2)
Conv(256,(4x4),2)
Conv(512,(4x4),2)

References

1. Wold, S., Esbensen, K., Geladi, P.: Principal component analysis. Chemometrics and intelligent laboratory systems **2**(1-3), 37–52 (1987)