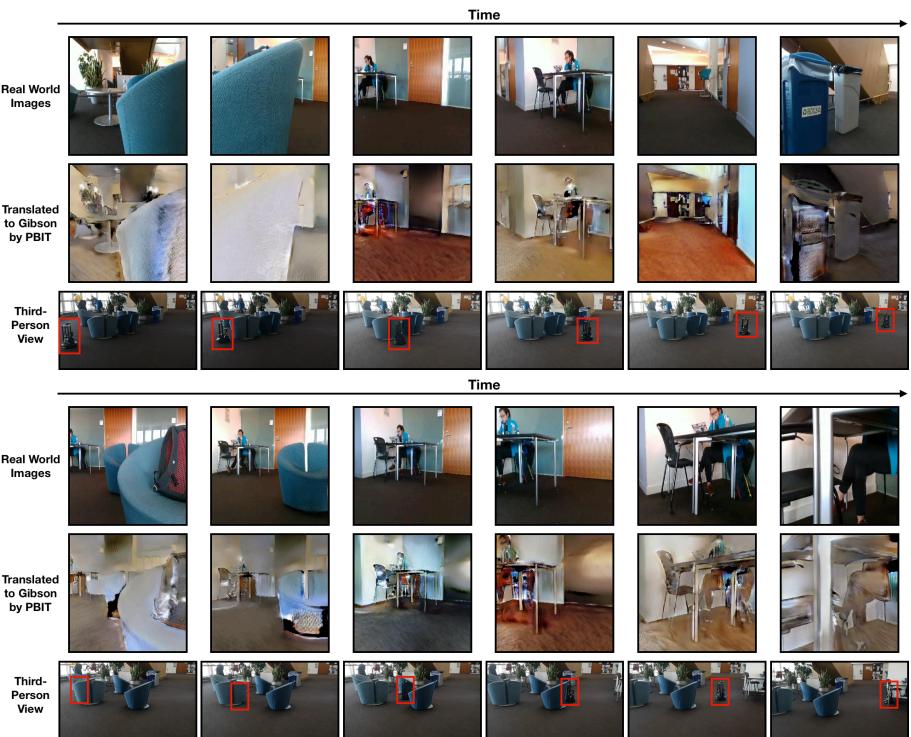


000
001 **Unsupervised Domain Adaptation**
002 **for Visual Navigation: Supplementary Material**
003

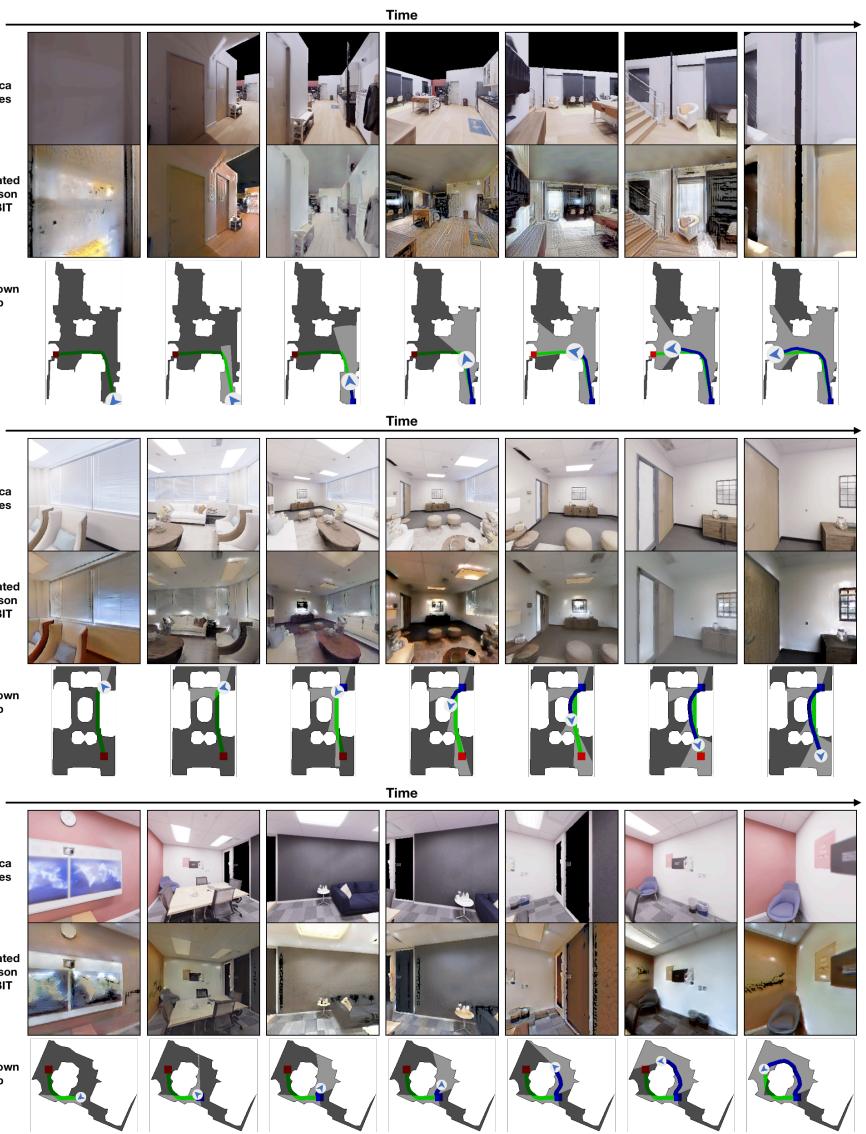
004 Anonymous ECCV submission
005

006 Paper ID 1995
007

009 **A Additional Trajectory Visualizations**
010



045
046
047
048
049
050
051
052
053
054
055
056
057
058
059
060
061
062
063
064
065
066
067
068
069
070
071
072
073
074
075
076
077
078
079
080
081
082
083
084
085
086
087
088
089



045
046
047
048
049
050
051
052
053
054
055
056
057
058
059
060
061
062
063
064
065
066
067
068
069
070
071
072
073
074
075
076
077
078
079
080
081
082
083
084
085
086
087
088
089



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.

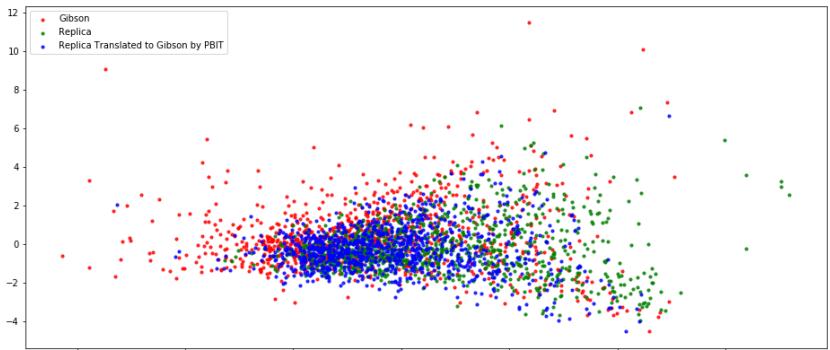


Fig. 1: We use PCA to visualize and compare the 128-dimensional task-specific feature vectors produced by PointGoal agent trained in Gibson, when given Gibson images, Replica images, or Replica images translated to Gibson by PBIT as input. The figure shows the translated images by PBIT bridge the policy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

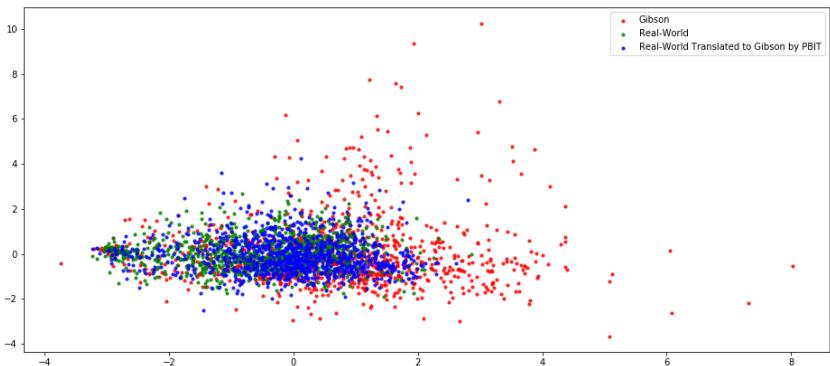


Fig. 2: We use PCA to visualize and compare the 128-dimensional task-specific feature vectors produced by PointGoal agent trained in Gibson, when given Gibson images, Real-World images, or Real-World images translated to Gibson by PBIT as input. The figure shows the translated images by PBIT bridge the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

C Model Architecture Details

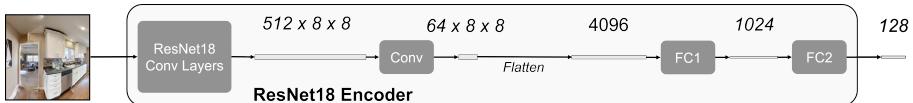


Fig. 3: An illustration of the network architecture of the Visual Policy Encoder.

Table 1: Architecture specification of different parts our PBIT model

		Decoder	Discriminator
Style Encoder	Content Encoder	c and AdaIN(s)	
(256 x 256) RGB	(256 x 256) RGB	Residual(128)	
Conv(32,(7x7),1)	Conv(32,(7x7),1)	Residual(128)	
Conv(64,(4x4),2)	Conv(64,(4x4),2)	Residual(128)	
Conv(128,(4x4),2)	Conv(128,(4x4),2)	Residual(128)	
Conv(128,(4x4),2)	Residual(128)	UpSampling(2x2)	Multi-scale Input
Conv(128,(4x4),2)	Residual(128)	Conv(64,(5x5),1)	Conv(64,(4x4),2)
global_avg_pool	Residual(128)	UpSampling(2x2)	Conv(128,(4x4),2)
dense(8)	Residual(128)	Conv(32,(5x5),1)	Conv(256,(4x4),2)
$\rightarrow s \in \mathbb{R}^8$	$\rightarrow c \in \mathbb{R}^{128 \times 64 \times 64}$	UpSampling(2x2)	Conv(512,(4x4),2)
		Conv(3,(7x7),1)	
		(256 x 256) RGB	

225 References

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

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269