

1. 研究背景・目標・内容

・ロボットの動作ソフトウェアを開発者が一つ一つ作成するのは困難

人間が物体を操作する様子をロボットが観察することにより、その動作を自動で獲得できるようにすることを目指す

研究内容:

物体の形やシーンの状態を見ることにより、ロボットがその物体を操作したりシーンに働きかけたりして目標状態にするための手順を想起する

予備実験:

- ・現時刻の人物スケルトンと室内シーンからシーンに応じた人物動作を想起
- ・現時刻の室内シーンから目標となるシーンに至るまでのシーンの変遷を想起

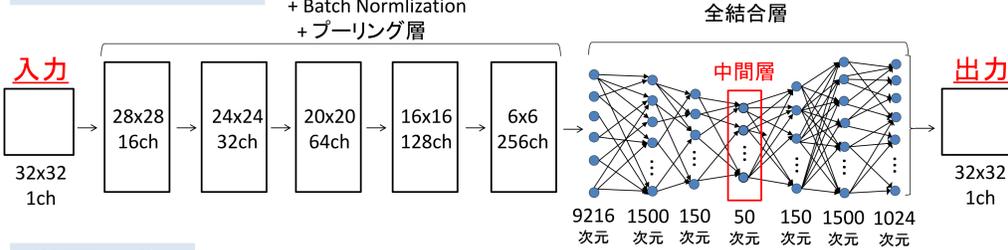
2. Auto-Encoderを用いた空間状態の記述

LSTMモデルを学習するために、Sparse Auto-Encoder[1]を用いて室内のシーン環境(机、椅子、人物の位置など)を50次元のベクトルで記述

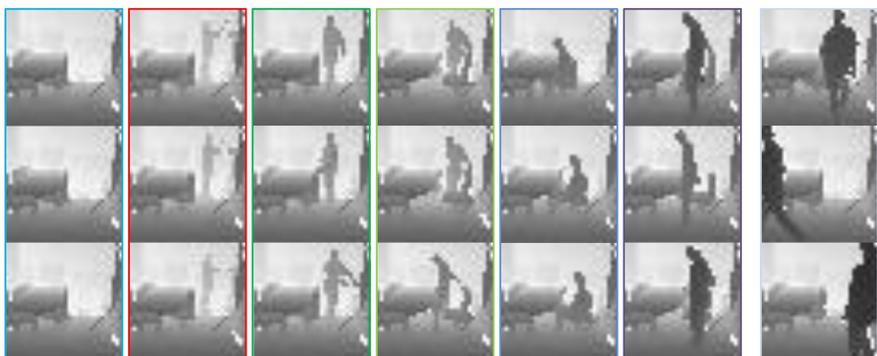
[1] Tadashi Matsuo, Nobutaka Shimada "Construction of Latent Descriptor Space of Hand-Object Interaction", The 22nd Joint Workshop on Frontiers of Computer Vision (FCV2016): pp. 117-122

モデルの構造:

畳み込み
+ Batch Normalization
+ プーリング層

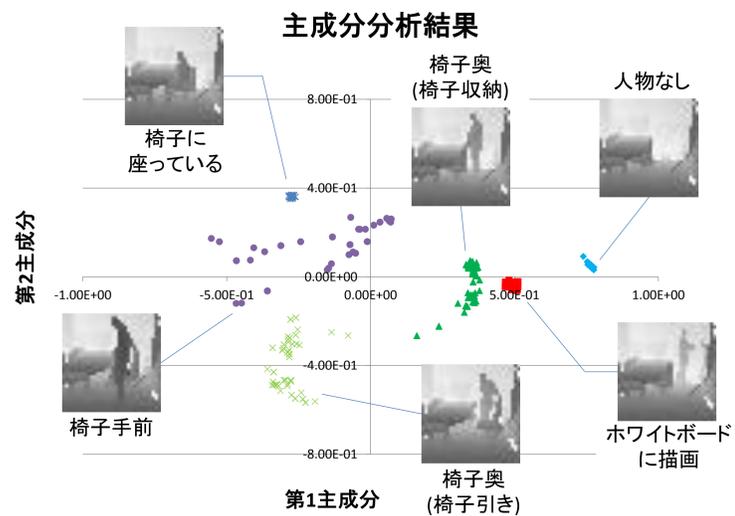


学習画像:



学習結果:

入力画像 復元画像



似た特徴を持つ画像同士が近くにプロットされ、異なる特徴を持つ画像同士が離れてプロットされている。

似た特徴を持つシーン画像同士が似たベクトルで表現されているといえる。

復元画像: 入力画像をエンコード + デコードした結果

5. まとめ

- ・室内における物体使用動作をLSTM(Long Short-Term Memory)を用いて記述し、シーンに応じた動作を想起した
- ・現在の状態から目標状態に至るまでのシーン変化を想起した

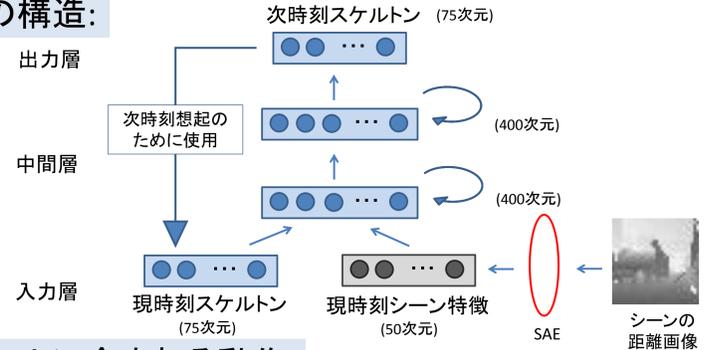
謝辞 本研究はJSPS 科研費24500224, 15H02764 の助成を受けたものです

3. スケルトンとシーン特徴を用いた人物動作の記述と想起

スケルトンとシーン特徴を用いて人物動作モデルを記述

- 椅子が引かれていれば座る、椅子が引かれてなければ椅子を引いてから座るなど、現在のシーンに応じた人物動作が想起される

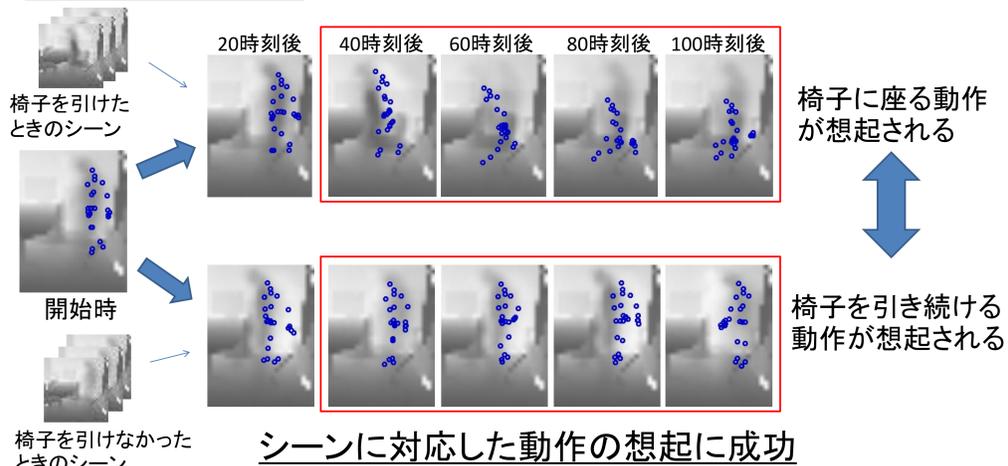
モデルの構造:



データセットに含まれる動作:



動作の想起結果:

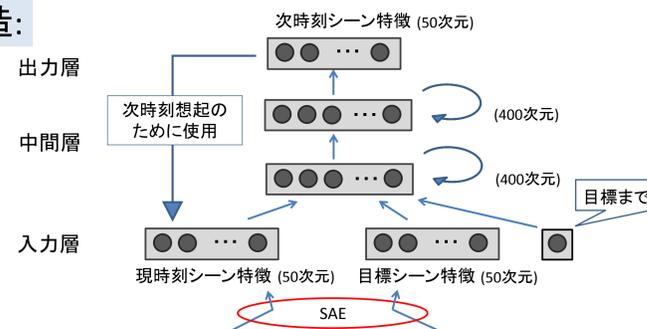


シーンに対応した動作の想起に成功

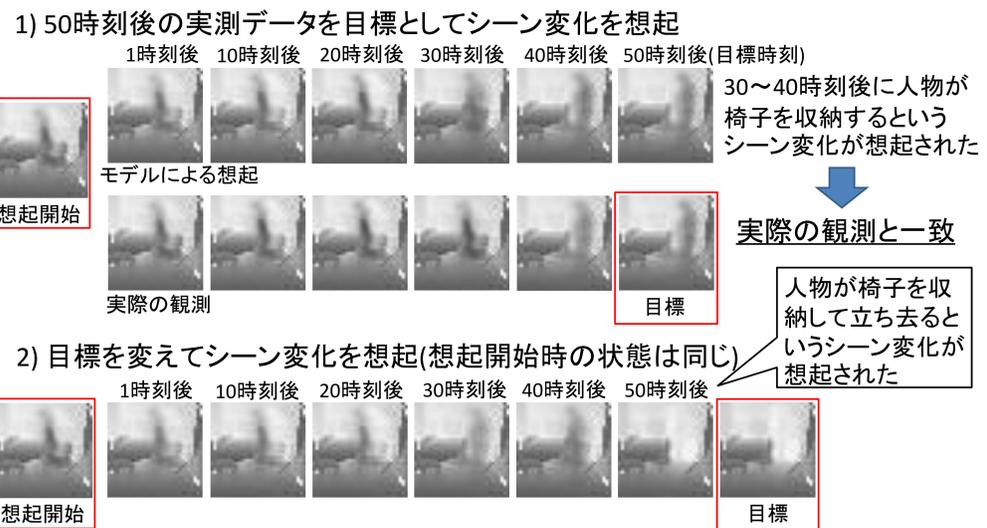
4. 現在から目標に至るまでのシーン変化の想起

現在と目標のシーンを入力すると、目標に至るまでのシーンの変遷を想起

モデルの構造:



シーン変化の想起結果:



6. 今後の課題

シーンを目標まで変化させるための人物動作と動作に伴うシーン変化の想起を目指す

