Learning and Predicting How Humans Will Move

Taku Komura

School of Informatics
University of Edinburgh
Motivation

- Predicting how humans will move in the future given their movements in the past and the surrounding environment
- Learn from the data
- Let robots assist humans for some tasks
Predicting Human Motion

- Easier to predict the motion if the motion subspace is known.
- We build the subspace using neural networks
Motion Data

- Motion data is represented as joint positions.
- Valid motion data only exists in a small subspace of this representation.
- This subspace is called the Motion Manifold and is what we wish to construct.
Projection

- Projection onto the manifold fixes invalid motion.
Interpolation

- Interpolation along the manifold avoids invalid motion.
Distance

- Distance along the manifold is more natural.
Prior Belief

- Manifold represents a *Prior Belief* about motion data.
- Data closer to the manifold is likely to be correct.
Data Capture

- Motion capture data of locomotion
- Interaction with the environment
Unsupervised Learning of the Motion Manifold

• Using the large human motion database, we learn the motion manifold

• Constructing a representation using the Autoencoder
Unsupervised Learning

Learning the space of data

Face Model

Can design any faces

training
Unsupervised Learning

- Synthesizing new faces
Autoencoder

- *Neural Network* used for learning efficient encodings.
- Unsupervised method.
- Trained to reconstruct input data.
- Network learns “natural” features of the input data.
Autoencoders + Dropout

- Optimizing the network such that the following loss is minimized

\[
\text{Loss} = ||X - w^{T}(w(X_c))||^2
\]
Convolutional Neural Networks

- *Convolutional Neural Networks* (CNNs) convolve over dimensions with local coherence.

- Great successes in Image Classification, Video Classification, Speech Recognition.
Automatic feature extraction

- Previously features had to be provided by humans
- Deep nets can automatically extract features
Convolution

- *Filters* convolve over the temporal dimension.
Convolution

- *Filters* convolve over the temporal dimension.
Convolution

- Filters convolve over the temporal dimension.
Convolution

- *Filters* convolve over the temporal dimension.
Convolution

- Filters convolve over the temporal dimension.
First Layer Filters

- Many strong correlations can be seen in the filters.
- High values across multiple frames and joints.
Example Layer 2 Feature Maps
Naive Distance

Reference  5 Most Similar Motions
Motion Synthesis in the Motion Subspace

- Regressing high level control signals to the motion subspace
Network Architecture

- Combination of a feedforward network and an autoencoder
- Mapping the low dimensional control signal to the hidden units of the autoencoder
Results
Walking over uneven terrains
Evaluation

- Can learn from a large amount of data
  - 24 hours of motion capture data
- The network can compress the data a lot
  - 1.5 GB motion capture data → 8MB Network
- Reproduction can be done very fast
  - 1 frame ~nano seconds on the CPU
  - Even faster on the GPU
Video for crowd animation
Make Use of the Motion Subspace for Assistance Technology

- Use the motion manifold as a prior to predict the human motion from videos
- Recognizing the action that the person is doing
- Predicting what the person will do next
Construction of the Motion Manifold in Special Tasks

• For some special tasks such as manipulation, we may wish to produce a “specialized manifold”

• Capturing some special tasks such as manipulation in constrained areas may be difficult
Construction of the Motion Manifold in Special Tasks

- Make use of micro UAS and ground robots for motion capture and learning human motions
- Make use of the data for simultaneously predicting the next human action
Summary

- Application of neural networks for human motion capture data
- Introducing novel architectures for satisfying the requirements: real-time, compact, etc.
- Can import many novel ideas from the machine learning and computer vision community
- Computer graphics can provide novel challenging problems to the machine learning community
Bibliography


- Daniel Holden, Jun Saito, Taku Komura, Tom Joyce "Learning Motion Manifolds with Convolutional Autoencoders", ACM SIGGRAPH Asia 2015, Technical Briefs