Deep Learning in Video Games

Frank Xing
Robert Chang

Why we need deep learning

Movement of NPCs By Hand

- Animations before were mostly done by hand
- Animation by hand works well for 2D, cartoonish games such as Mario, Cuphead, etc.
- Not realistic and lacking details regarding surface and dynamic of hair and cloth
- Ultimately, animation by hand consumes memory, processor time, and is hard to create on a large scale

Example of Animation by Hand
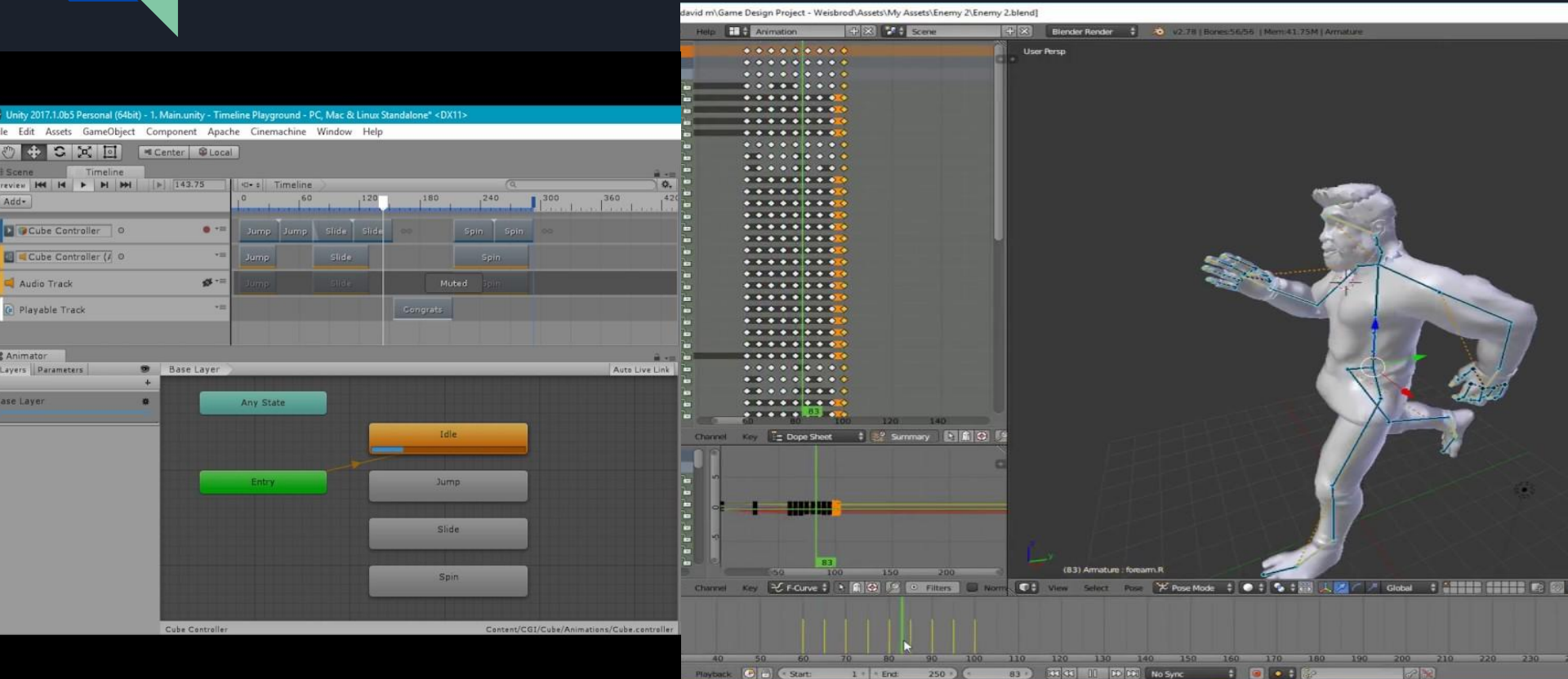
Movement of NPCs by Motion Capture

- Once motion capture technology was refined, it became new option for digital animation/model creation
- Works great for small scenes such as cinematic cutscenes
- Still can't plan for all interactions characters will have with environment in complex games
- Some movements may be difficult or unfeasible to create with live actors

Example of Animation by Human Performance

14:20:59:10

Image courtesy of Naughty Dog.

And so computer scientists at University of Edinburgh along with Adobe Research devised a solution...

Neural Networks

- Artificial Neural Networks (ANN) - computational nonlinear model that learns how to perform tasks through an input and output layer that involves artificial neurons which send signals with different weights to one another
- Nonlinear because you can have multiple inputs with different weights thus creating a complex relationship
- Deep Neural Network (DNN) - is a type of ANN with many layers between the input and output layer thus allowing the generation of more complex relations between data
- Train this deep neural network by inputting database of motions and actions

Neural Network Examples
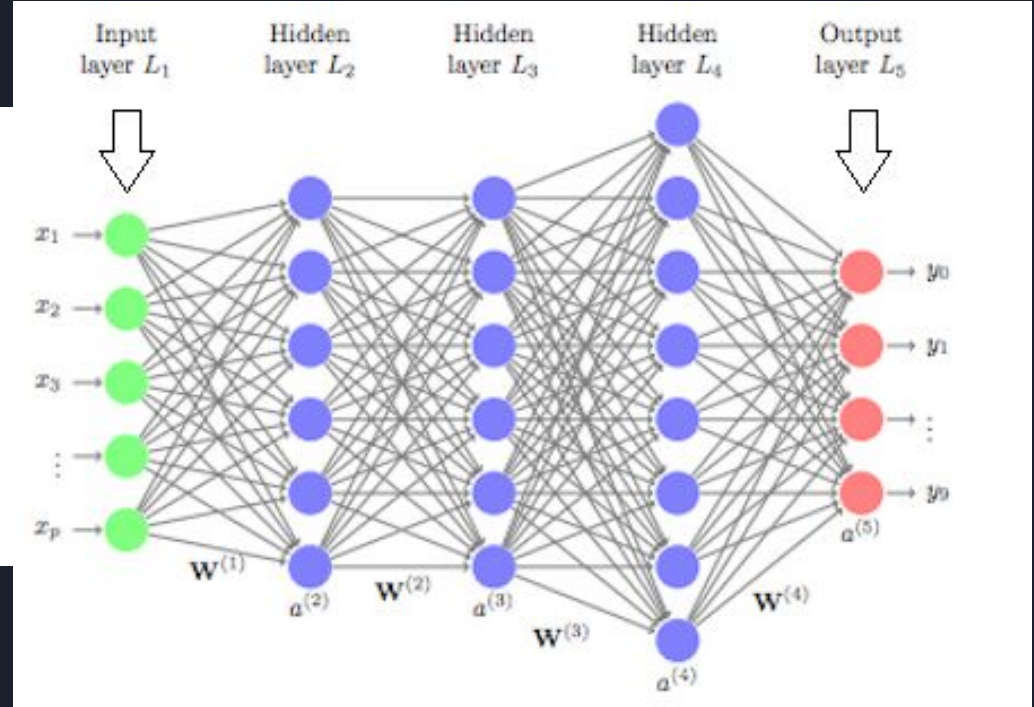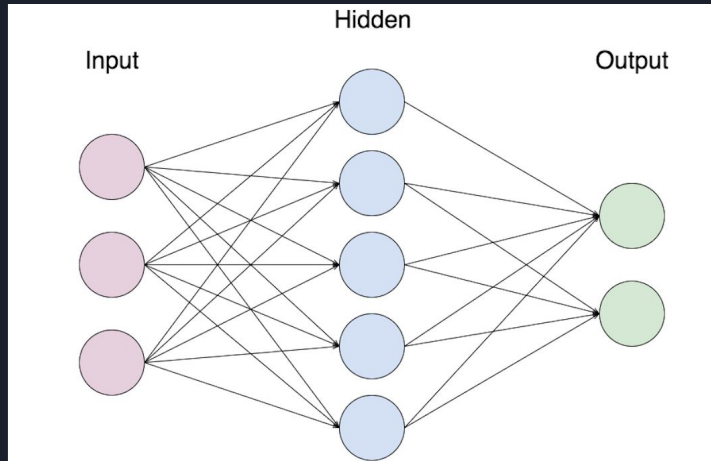
Movement of NPCs by Deep Learning

- Uses deep-neural and data-driven techniques to precisely guide characters through a variety of motions
- Achieves this using simple control commands
  - Examples of movements include sitting in chairs, picking up objects, side-stepping, running etc
- Planning and adaptation are two important aspects to consider for character motions
- Does require a large database with various movements from the performers, but can infer other movements

The Technique

- Planning: plan what kinds of different movements need to be put together to create a certain action
- Adapting: have ability to adapt to changes in environment such as obstacles or geometry
- The Neural State Machine that they developed learns animation by studying geometry of environment and the features of the desired action
- Learns how to create transitions by inferring next pose based on current state of scene and current pose
- One network is able to create multiple animations of different actions besides specified one

Head Motion Synthesis

- Predict head movement of a speaker based on their speech
  - Learn from a DNN with audio-visual broadcast news data
- A trained network outperforms a randomly initialized one
- Extra training data from other speakers in a stage can improve the prediction accuracy of a specific target speaker
- Results can be used in talking avatar animation in cutscenes
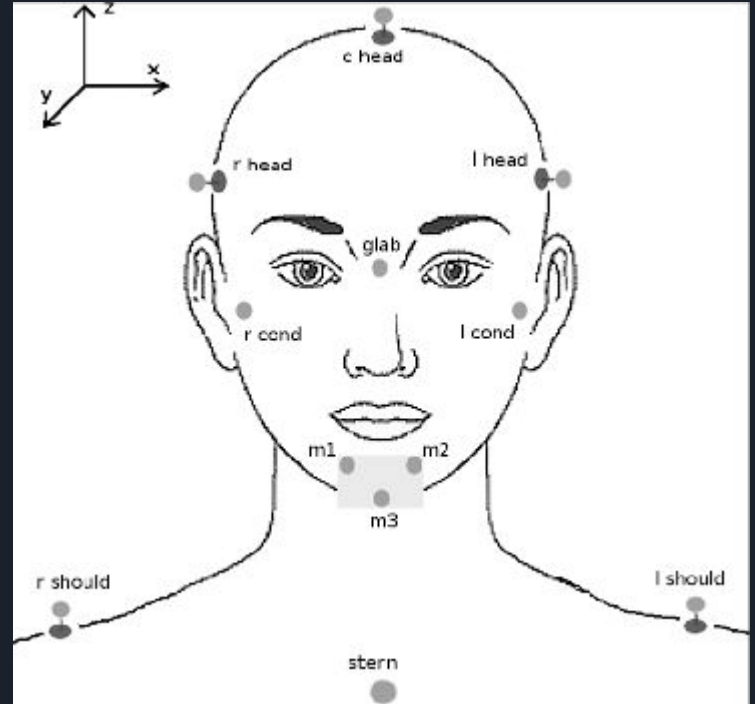
Generalized Speech Animation

- Focuses on the motion of mouths in accordance to different sounds
- Advantages include the following:
  - Runs in real-time, minimal tuning, generalizes well to speech sequences, can be edited for emotional speech and compatible with existing animation approaches


Generalized Speech Animation
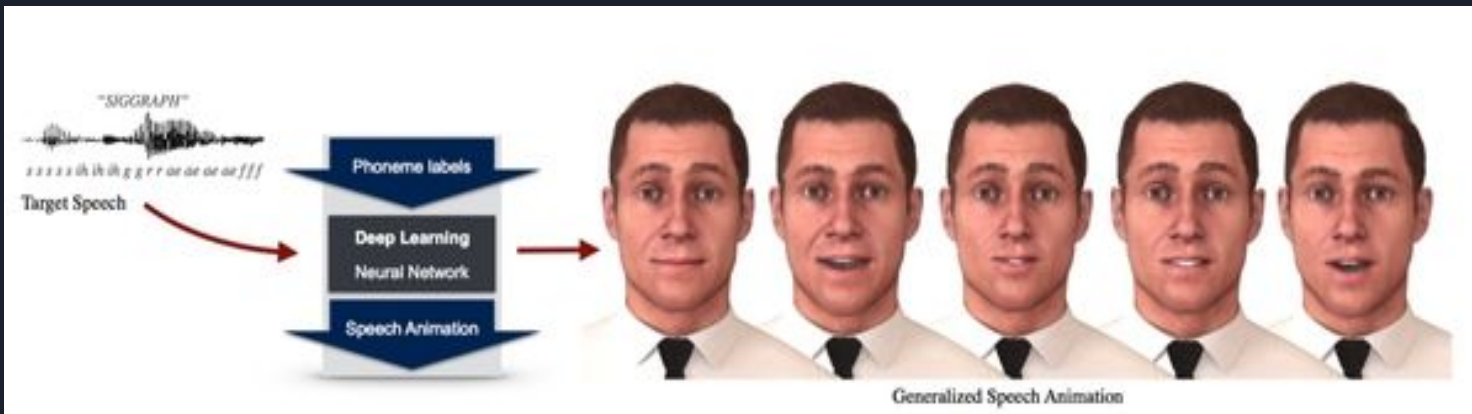
Facial Performance Capture

- User recorded in a natural environment with non-intrusive 3-D sensors
- Novel face tracking algorithm that combines pre recordings, textures, and geometry for a single optimization
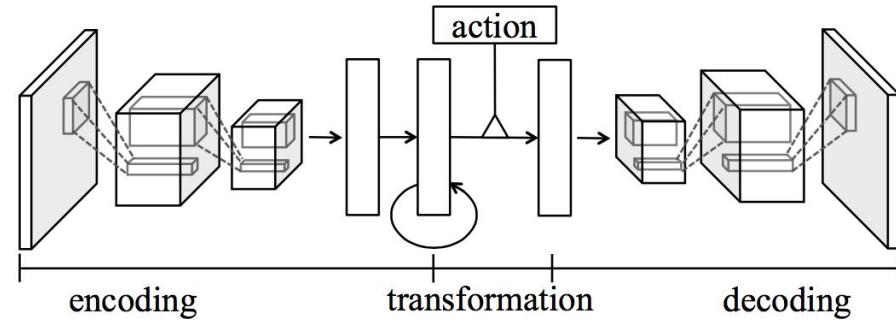- Easy to deploy system and widely applicable in digital gaming

Action-Conditional Video Prediction

- Future images dependent on past image frames and control variables/actions
- Process includes encoding, action-conditional transformation, and decoding based on layers of neural networks
- Able to generate visually-realistic frames that are also useful for control in the next 100 or so future frames
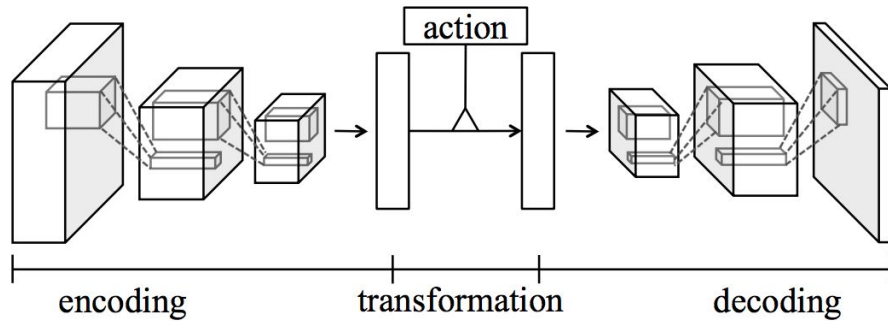
(a) Feedforward encoding   (b) Recurrent encoding

Our Review

Immediate Reactions

- Very interesting, gave us new knowledge about game animation and rough idea about future direction of game design
- Not that much technical information about how the Neural State Machine will work
- More about why and how it will be a significant development for multiples applications/industries
- New insight about new possibilities and technological capabilities of entertainment industry
- Chose this topic because we were unaware of deep learning application in video game design

Future Implications
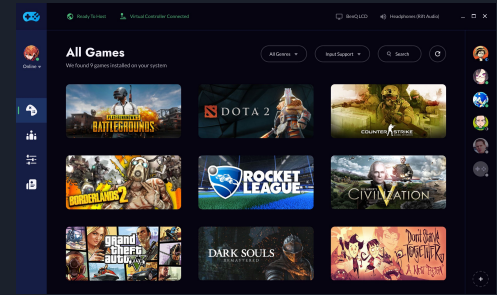
- Teaching games how characters should be animated and letting game do it instead of us making it all
- Saves time, money, resources in the game development process
- Can be expanded to other fields such as movie, tv shows, entertainment
- Reduce file sizes of games
- Reduce amount of data processed
- Maybe make streaming games more viable by reducing amount of data being transferred
- Create more complex video game ais that can do more on their own
- Create more realistic world with actual independently acting and animated figures
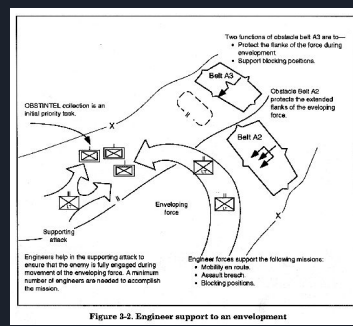
Possible Applications

Figure 3-2. Engineer support to an envelopment

30+ Weapons, Customizable Loadouts and Real-Life Gear.

- Other forms of entertainment that involves visuals/graphics
- CGI for movies, animated shorts, tv shows
- Improved, more complete, intelligent AI
- Self learning NPCs in games
- Self-driving car navigational/camera system intelligence
- Anything that relies on the construction of visuals from a set of patterns
- Virtual reality, augmented reality applications
- Environment analysis
- Military technology
- Development of tactical plans, strategy for military ops (take in examples of movement patterns, routes and based on current position plan next move)
- Simulation softwares

Citations

Liszewski, Andrew. "Deep Learning Is Making Video Game Characters Move Like Real People." *Gizmodo Australia*, Gizmodo Australia, 31 Oct. 2019, https://www.gizmodo.com.au/2019/10/deep-learning-is-making-video-game-characters-move-like-real -people/.

Theofficialacm. "Deep Neural Network Generates Realistic Character-Scene Interactions." *EurekAlert!*, 29 Oct. 2019, https://www.eurekalert.org/pub_releases/2019-10/afcm-dnn102919.php.

Codrina. "Skeletal versus Sprite Sheet Animations: Advantages and Disadvantages ." *Marionette Studio*, 20 Oct. 2016, https://marionettestudio.com/advantages-and-disadvantages-of-skeletal-versus-sprite-sheet-animatio ns/.

Starck, Jonathan, and Adrian Hilton. "Surface Capture for Performance-Based Animation." *Surface Capture for Performance-Based Animation - IEEE Journals & Magazine*, 7 May 2007, https://ieeexplore.ieee.org/abstract/document/4178157, https://doi.org/10.1109/mcg.2007.68.

Citations Part 2

Ding, Chuang, et al. "Head Motion Synthesis from Speech Using Deep Neural Networks." *SpringerLink*, Springer US, 24 July 2014, https://link.springer.com/article/10.1007/s11042-014-2156-2, https://doi.org/10.1007/s11042-014-2156-2.

"A Deep Learning Approach for Generalized Speech Animation." *ACM Transactions on Graphics (TOG)*, ACM, July 2017, https://dl.acm.org/citation.cfm?id=3073699, https://doi.org/10.1145/3072959.3073699.

"Realtime Performance-Based Facial Animation." *ACM Transactions on Graphics (TOG)*, ACM, July 2011, https://dl.acm.org/citation.cfm?id=1964972, https://doi.org/10.1145/1964921.1964972.

Guo, Xiaoxiao, et al. "Action-Conditional Video Prediction Using Deep Networks in Atari Games." *Action-Conditional Video Prediction Using Deep Networks in Atari Games*, 1 Jan. 2015, https://papers.nips.cc/paper/5859-action-conditional-video-prediction-using-deep-networks-in-atari-games.pdf.