Data-driven Approaches to Simulation (Motion Capture)

Ting-Chun Sun

tingchun.sun@usc.edu

Preface

- The lecture slides [1] are made by Jessica Hodgins [2], who is a professor in Computer Science Department at Carnegie Mellon University. She is also a world expert in motion capture.
- As implied by the name, motion capture is a technology to capture motion. In order to do that, usually special purpose cameras which are specific for motion capture are utilized to record some motion performed by a human actor. The final output is joints' trajectory to time.
- Some body joints are one-dimensional, e.g. elbow; some other joints are three-dimensional, e.g. shoulder joint.

Example

- CMU Motion Capture Database [3] is public available on the web and free for all uses. There are 2605 motion capture data, which were all recorded in the motion capture studio at Carnegie Mellon University.
- A typical frame rate is 120 for motion capture. High frame rate is needed since human motion is highly detailed; however, noisy images tend to be produced because of less amount of light captured during a smaller frame time.

- The input [4] is a person wearing a black suit attached by white markers and performing motion in front of 12 special custom made cameras, which are monotone with high frame rates. The output [5] is a motion data based on a three-dimensional model of a skeleton.

- .AMC file [6] is a text format file with motion data of each frame; .ASF file [7] contains hierarchy of the body skeleton. The combination of both contributes to a set of motion capture data. Please check [8] for detailed format of both files.
- Root joint, the center of the body, has six degrees of freedom, whereas other joints only have three.
- The resolution of the motion capture data depends on each motion capture system. Meter is the most natural choice since it belongs to International System of Units.
- Sometimes not all cameras can see all markers because of occlusion; therefore, some skilled people are needed to clean up the raw motion data first. Next, the body skeleton needed to be produced. The most common way to achieve that is to take one frame from the captured video to manually create new link with the segment. Once skeleton proportion is acquired, three-dimensional trajectory of the markers can be transformed to motion capture data via a semi-automated process. However, there is no best solution to pick joint angles which are able to

completely explaining the marker position. Some least squares optimization methods are utilized to find out the joint angles for every frame.

- Currently one drawback of motion capture is missing the contact information. For example, motion capture system would not know whether the foot of a walking character is in the air or not in a particular frame. This is also one research topic which consists of computer vision and computer animation to capture character motion and world object simultaneously.

Usage

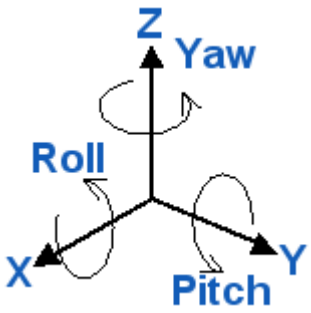
- Animation, video games, and robot control.
- Motion capture produces natural human motion which is hard to be designed by animator. One reason is that dynamics of mass in different parts of body affect other parts of body, and therefore there is so much subtlety in human motion. Another reason is that people are sensitive to human motion since they do and watch it all the time.

What is Motion Capture

- Track motion of reference points. (usually white dots)
- Convert to joint angles. (usually using least squares optimization method)
- Use angles to drive an articulated model.

- Motion paths can then be combined to give greater control. (kinematics blending)
- Angle is easier to be interpolated than position since the skeleton will not be changed. However, motion interpolation is still a difficult problem unless two motions are similar.

What is Captured



- X, Y, Z positions of the root joint.
- Roll, pitch, yaw angles. Yaw is the most important one in computer animation since it denotes which direction the character faces.
- Sometimes joints tend to come apart or links tend to grow or shrink since the marker positions are not consistent with the skeleton model. Even though the markers are carefully handled, it is still unable to explain the position well since human muscle is not rigid.

How to Use the Data

- Off-line: after capture and clean-up process, motion can be modified or blended via computer animation algorithm.
- On-line: the motion is analyzed in real-time.
- Performance animation refers to animation that somebody performs in front of motion capture system in real-time. All the clean-up process has to be done automatically, and therefore some limitations exist. The result is usually worse than the off-line technique.

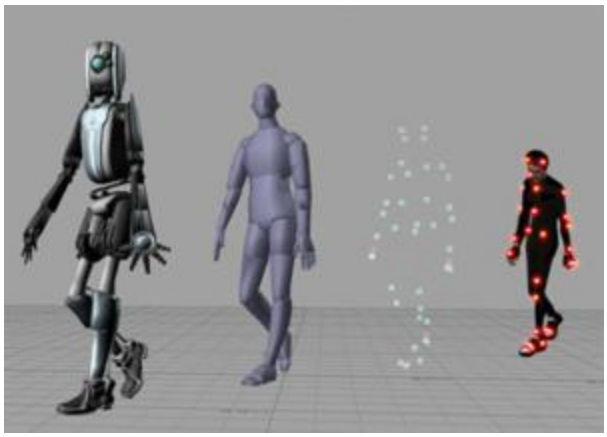
- Recently, depth camera helps the capture process since it gives not only color but also depth information, just like z-buffer in computer graphics.
- In practice, gray-scale camera is widely utilized to capture white markers mounted on a black suit without any depth information. For acquiring the three-dimensional position, multiple gray-scale cameras with different perspectives are calibrated based on vision algorithm.
- How to know which marker is which and where it is going if occlusion happens are both problems of motion capture. One suggestion is using RGB cameras and different color markers, but the three times of data may cause other problems in system organization and network architecture.

Production Pipeline

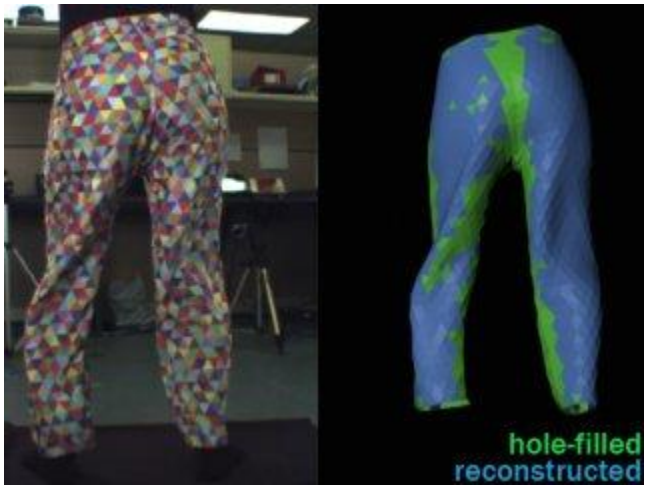
- Calibration: knowing positions of cameras.
- Capture: knowing positions of markers.
- Skeleton estimation: knowing skeleton of character.
- Inverse kinematics processing: knowing angles of joints.

What is Captured

- Dynamic motions: very fast, even 120 frame rates may not be enough.
- Various scales: large as elephants, small as fingers or face muscles.



- Non-rigid objects: cloth.
- One SIGGRAPH 2007 paper [9] implemented a technique to capture cloth motion by using different color patches instead of traditional markers. Another SIGGRAPH 2009 paper [10] designed a system to capture motion of a hand which wears a special glove with color patches in front of single camera in real-time.



- One problem is that some objects may interfere capture process, e.g. ping-pong ball which is like a marker. Another problem is that motion of passive object is hard to be captured since it is usually complicated, e.g. cloth and fluid.

Technologies: Optical Passive

- The most common technique with passive markers.
- Not outdoors because of the sunlight.
- Barely real-time now.

Technologies: Optical Active

- The markers actively emits different kinds of light.
- No correspondence problem.
- Not popular in practice because of the higher expense.

Technologies: Magnetic

- No cameras. No occlusion. Clean motion data.
- Good for complicated motion which may have heavy occlusion.
- Cheaper.
- Character must wear a bunch of wires and magnetic sensors, and therefore the character mobility is even more limited.
- Not just uncomfortable for actor but also changing the entire motion.

Technologies: Exoskeleton

- Extremely high precision.
- Too heavy.

Technologies: Monkey

- Animator grabs, bends, or moves the robot to pose it as arbitrary configuration.
- Not real motion capture. Kind of like stop motion [11].
- It is more intuitive to directly utilize a physical model rather than a virtual one.

Technology Issues

- Resolution/range of motion.
- Calibration.

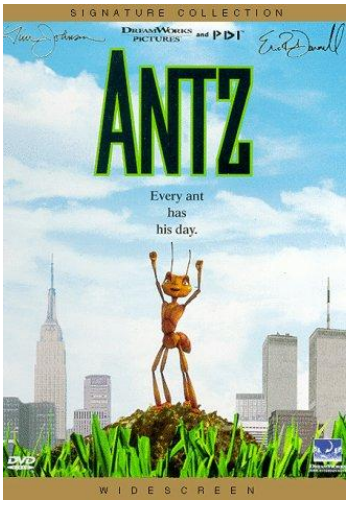
- Accuracy.
- Occlusion/correspondence.

Marker Set

- It really matters where to position the markers to get better motion result.
- It also relates to how to do the clean-up process.

Eric Darnell, Co-director of Antz

- The main problem with motion capture associated with characters has to do with mass distribution, weight, and exaggeration. It is impossible for a performer to produce the kind of motion exaggeration that a cartoon character needs, and the mass and weight of the performer almost never looks good when applied to a character of different proportions.



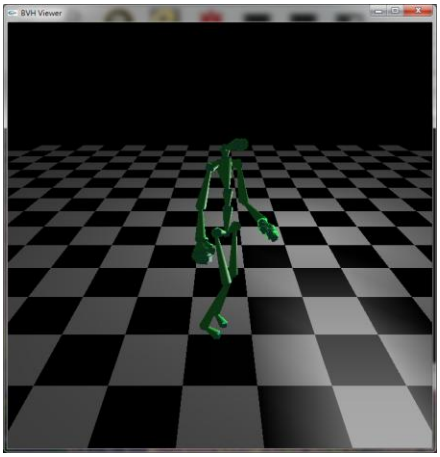
- That is why motion capture is not used more widely than keyframe technique in computer animation.

Richard Chuang, VP at PDI

- The mapping of human motion to a character with non-human proportions doesn't work, because the most important things you get out motion capture are the weight shifts and the subtleties and that balancing act of the human body. If the

proportions change, you throw all that out the door, so you might as well animate it.

Research Topics



- Marker placement.
- Capture of deforming skin rather than pretending it is a rigid chain.
- Retargeting, which means transforming motion from one character to another.
- Constraint satisfaction. Foot slide or penetration may sometimes happen in motion capture data since the inverse kinematics solver is not perfect.
- Generalization of data. How much data is needed to create any motion by combining, blending, or merging.

References

[1] http://www.cs.cmu.edu/~jkh/video_games/mocap.pdf

[2] http://www.cs.cmu.edu/~jkh/

[3] http://mocap.cs.cmu.edu/

[4] http://mocap.cs.cmu.edu:8080/subjects/01/01_01.mpg

[5] http://mocap.cs.cmu.edu:8080/subjects/01/01_01.avi

[6] http://mocap.cs.cmu.edu:8080/subjects/01/01_01.amc

[7] http://mocap.cs.cmu.edu:8080/subjects/01/01.asf

[8] http://www.darwin3d.com/gamedev/articles/col0198.pdf

[9] http://www.ryanmwhite.com/research/white2007siggraph.pdf

[10] http://www.cs.washington.edu/homes/jovan/papers/wang-2009-rth.pdf

[11] http://www.youtube.com/watch?v=V2fpgpanZAw