A Brief History on Motion Capture

Jinxiang Chai

TAMU
Computer
Science

Outline of Mocap (Motion Capture)

Mocap history

Mocap technologies

Mocap pipeline

Mocap data format

Motion Capture

“ …recording of motion for immediate or delayed analysis or playback…”

- David J. Sturman

“The creation of a 3d representation of a live performance”

- Alberto Menache

“…is a technique of digitally recording movements for entertainment, sports, and medical applications.”

- Wikipedia

History of Motion Capture

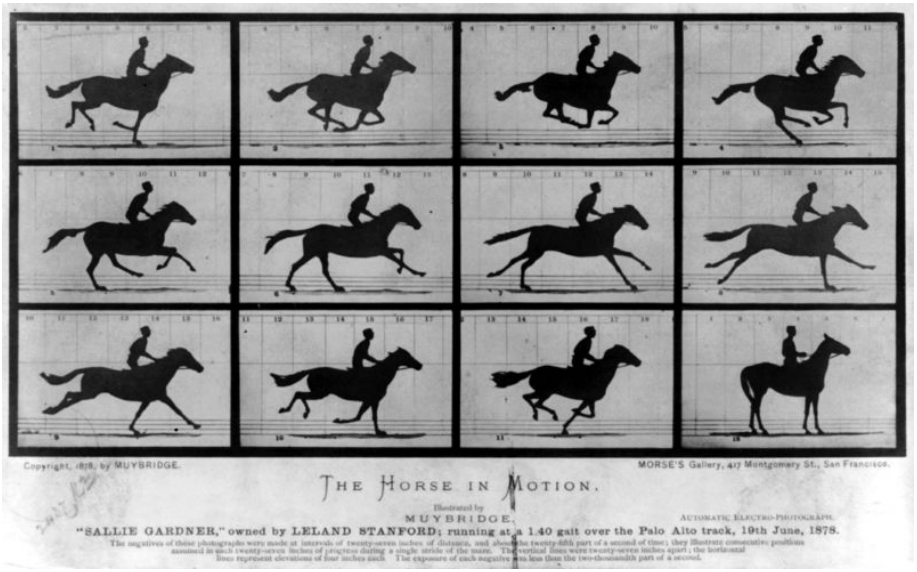
Eadweard Muybridge (1830-1904)

- first person to photograph movement sequences

History of Motion Capture

Eadweard Muybridge (1830-1904)

- first person to photograph movement sequences
- whether during a horse's trot, all four hooves were ever off the ground at the same time.
- the flying horse

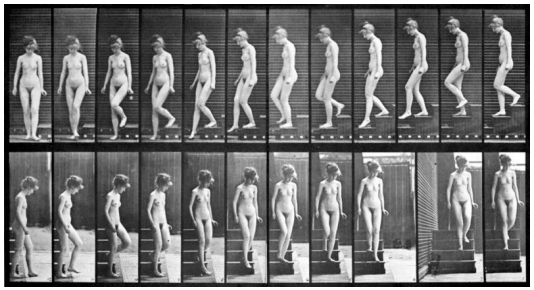


Sequence of a horse jumping
(courtesy of E. Muybridge)

History of Motion Capture

Eadweard Muybridge (1830-1904)

- first person to photograph movement sequences
- the flying horse
- animal locomotion (20k pictures about men, women, children, animals, and birds).

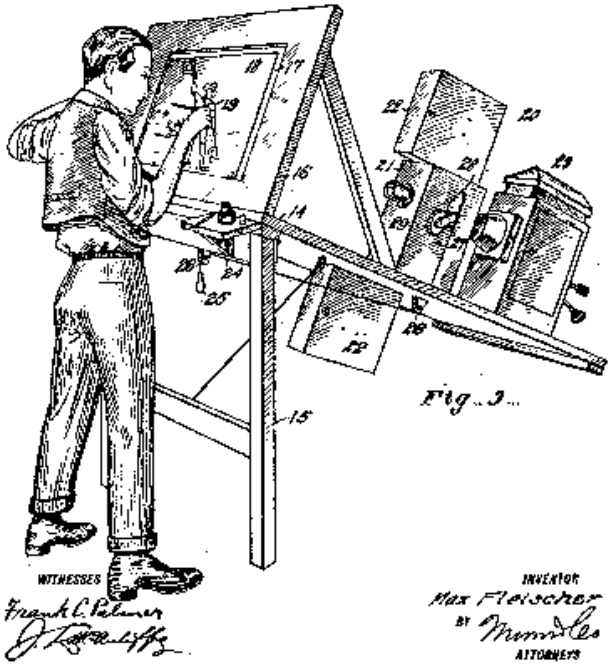


Woman walking downstairs
(courtesy of E. Muybridge)

Rotoscope

Allow animators to trace cartoon character over photographed frames of live performances

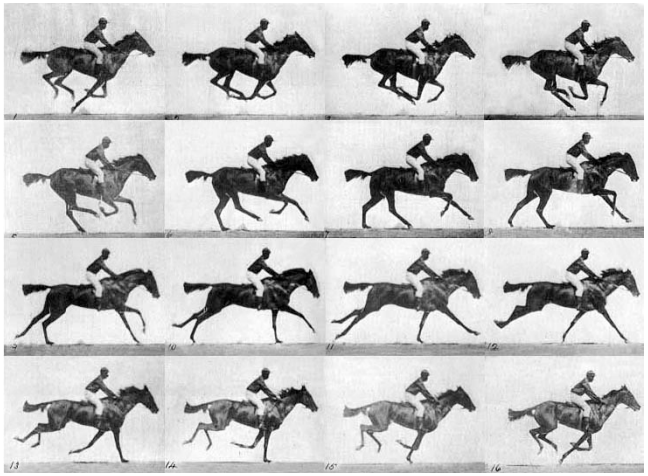
- invented by Max Fleischer in 1915



Rotoscope

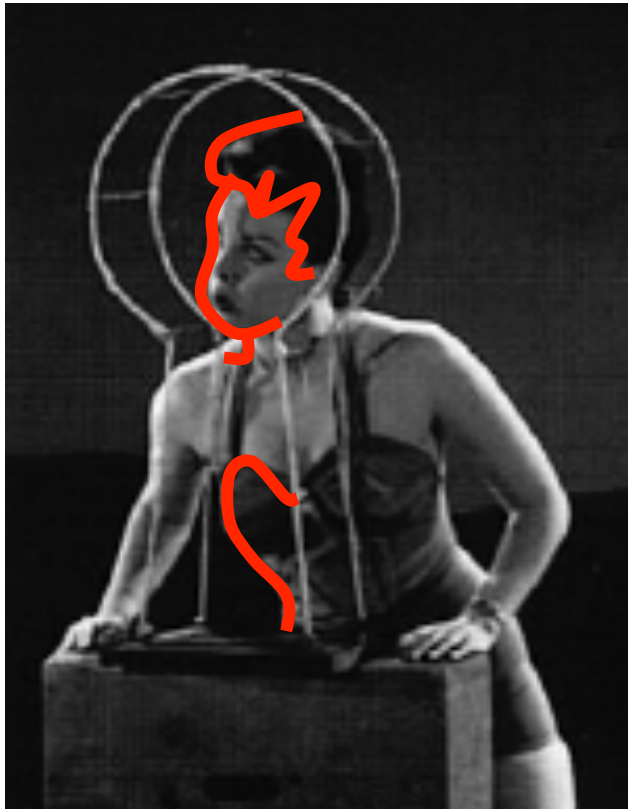
Allow animators to trace cartoon character over photographed frames of live performances

- invented by Max Fleischer in 1915
- 2D manual motion capture



A horse animated by rotoscoping from Muybridge's photos

Rotoscoping



“rotoscoping can be thought of as a primitive form or precursor to motion capture, where the motion is ‘captured’ painstakingly by hand.” - Sturman

Another example

Rotoscope

Allow animators to trace cartoon character over photographed frames of live performances

- invented by Max Fleischer in 1915
- 2D manual capture
- the first cartoon character to be rotoscoped -- “Koko the clown”
- the human character animation -- snow white and her prince (Walt Disney, 1937)

Current Motion Capture Technologies

"3D Rotoscoping": measuring 3D positions, orientations, velocities or accelerations automatically

Current motion capture systems

- Electromagnetic
- Electromechanical
- Fiber optic
- Optical

Electromagnetic Mocap

Each sensor record 3D position and orientation

Each sensor placed on joints of moving object

Full-body motion capture needs at least 15 sensors

Popular system:

http://www.ascension-tech.com/

Electromagnetic Mocap

See video demo [1, 2]!

Electromagnetic Mocap

Pros

- measure 3D positions and orientations
- no occlusion problems
- can capture multiple subjects simultaneously

Cons

- magnetic perturbations (metal)
- small capture volume
- cannot capture deformation (facial expression)
- hard to capture small bone movement (finger movement)
- not as accurate as optical mocap systems

Electromechanical Mocap

Each sensor measures 3D orientations

- including 3D accelerometers, 3D gyros, and 3D magnetometers

Electromechanical Mocap

Each sensor measures 3D orientations

Each sensor placed on joints of moving object

Full-body motion capture needs at least 15 sensors

Popular systems:

http://www.xsens.com/

Electromechanical Mocap

See video demo [1,2]!

Electromechanical Mocap

Pros

- measure 3D orientations
- no occlusion problems
- can capture multiple subjects simultaneously
- large capture volume
- portable and outdoors capture (e.g. skiing)

Cons

- getting 3D position info is not easy
- the root positions is often measured with ultrasonic position sensors
- cannot capture deformation (facial expression)
- hard to capture small bone movement (finger motion)
- not as accurate as optical mocap system

Fiber Optic Mocap

Measures 3D position and orientation of entire tape

Binding the tape to the body

Popular systems: http://www.measurand.com/

Fiber Optic Mocap

See video demo [1,2]!

Fiber Optic Mocap

Pros

- measure 3D orientations and positions
- no occlusion problems
- can capture multiple subjects simultaneously
- go anywhere mocap system
- can capture hand/finger motion

Cons

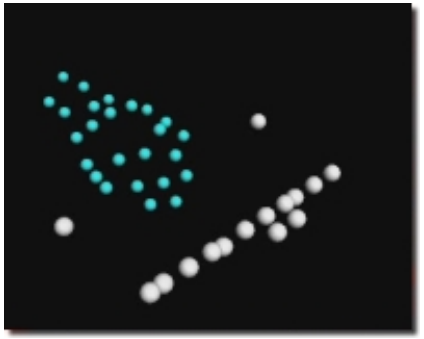
- intrusive capture
- cannot capture deformation (facial expression)
- not as accurate as optical mocap system

Optical Mocap

Multiple calibrated cameras (>=8) digitize different views of performance

Wears retro-reflective markers

Accurately measures 3D positions of markers



Optical Mocap

Vicon mocap system: http://www.vicon.com

See video demo [1,2,3,4]!

Optical Mocap

Pros

- measure 3D positions and orientations
- the most accurate capture method
- very high frame rate
- can capture very detailed motion (body, finger, facial deformation, etc.)

Cons

- has occlusion problems
- hard to capture interactions among multiple actors
- limited capture volume
- expensive

Video-based Mocap

- Capturing 3D performance from single-camera video streams

Video-based Mocap

- Capturing 3D performance from single-camera video streams

- Click video here

Video-based Mocap

Pros:

- capturing human motion anytime, anywhere
- very cheap
- zillions of films, sports footage, and internet videos.

Cons:

- not a mature technology
- quality is not as good as other capturing technologies.

Motion Capture Using Depth Sensors

- Kinect system [video]

- Our own system [see video here]

Mocap Pipeline

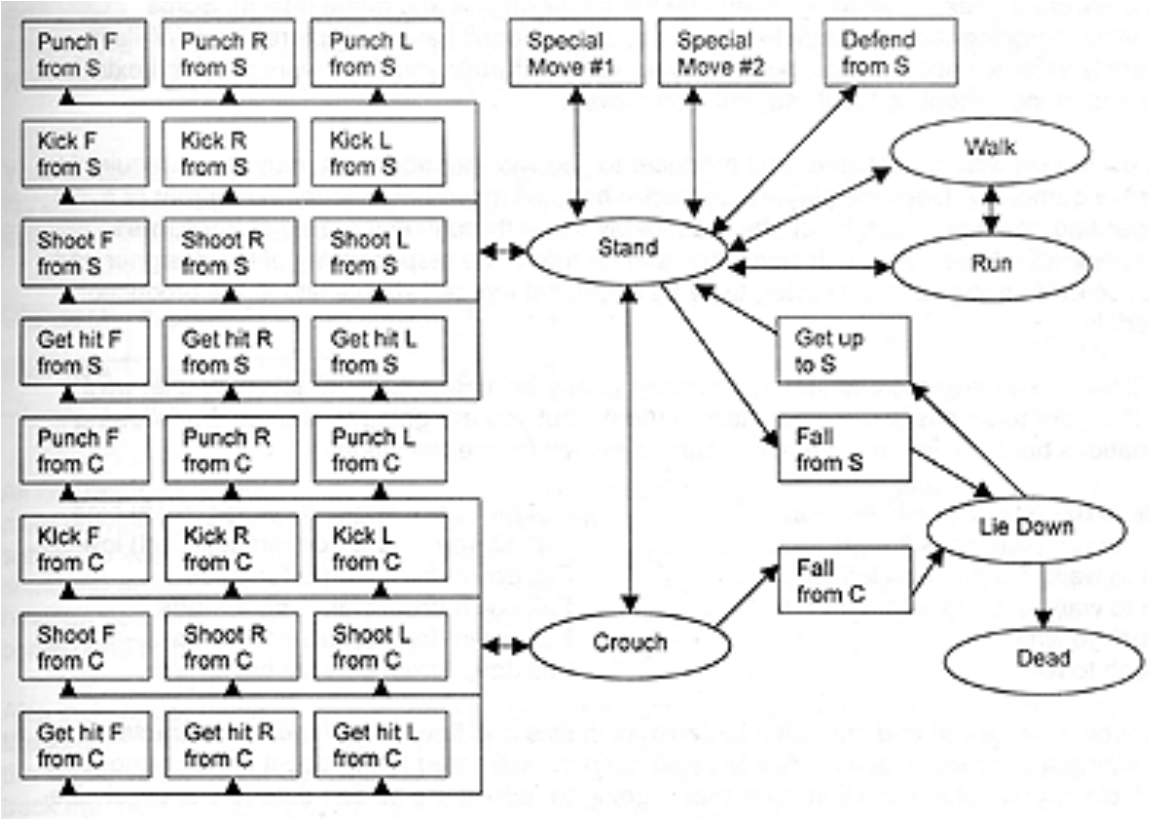
Optical Mocap pipeline

- Planning
- Calibration
- Data processing

Planning

- Motion capture requires serious planning
- Anticipate how the data will be used
- Garbage in garbage out
- Shot list
- Games
  - motions need to be able to blend into one an another
  - capture base motions and transitions
  - which motions transition into which other transitions
  - cycles/loops

Movement Flowchart for Games



•**Planning and Directing Motion Capture For Games**
By Melianthe Kines
Gamasutra
January 19, 2000
URL: http://www.gamasutra.com/features/20000119/kines_01.htm

Planning

Character/prop set up

- character skeleton topology (bones/joints number, Dofs for each bone)
- location and size of props

Marker Setup

- the number of markers
- where to place markers

Calibration

Camera Calibration:

determine the location and orientation of each camera

determine camera parameters (e.g. focal length)

Subject calibration

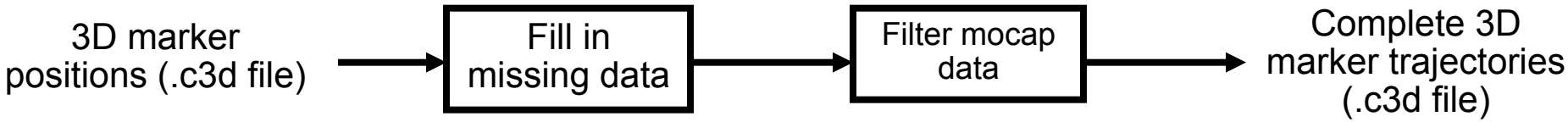
- determine the skeleton size of actors/actresses (.asf file)
- relative marker positions in terms of bones
- determine the size and location of props

Processing Markers

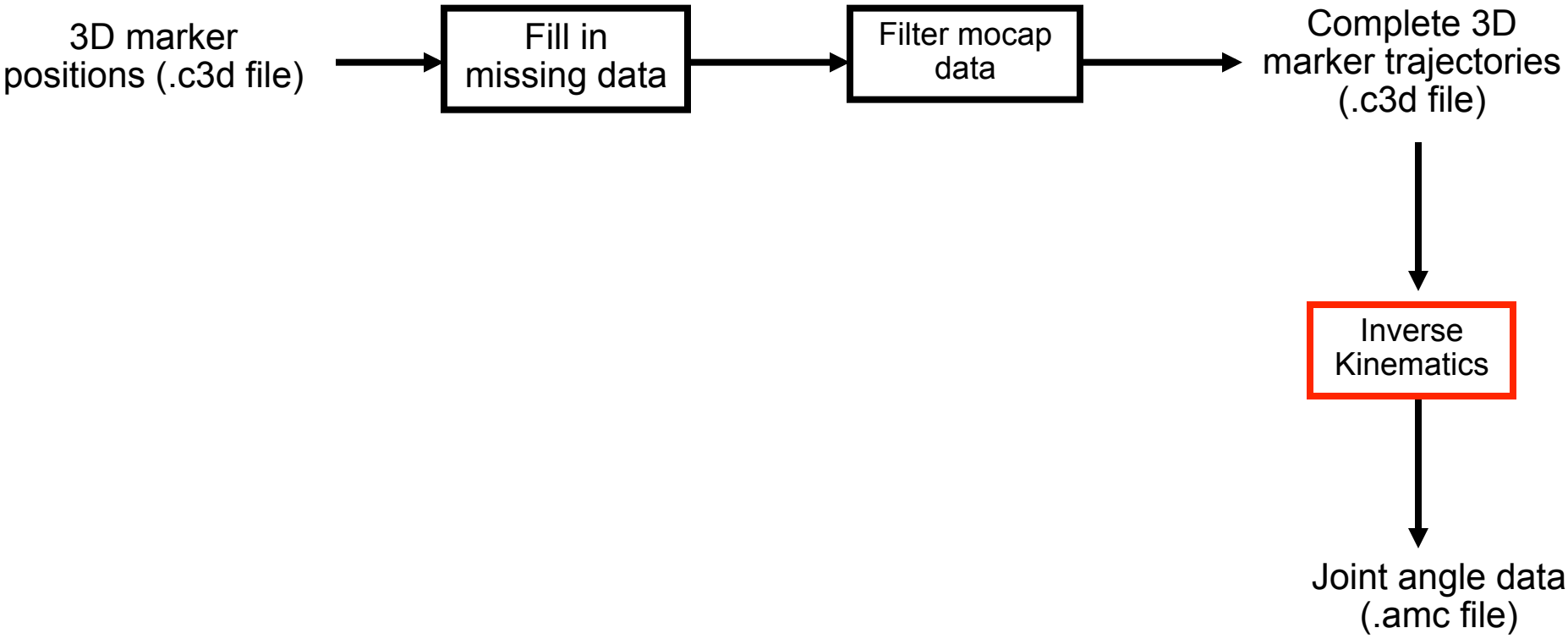
- Each camera records capture session
- Extraction: markers need to be identified in the image
  - determines 2d position
  - problem: **occlusion**, marker is not seen
    - use more cameras
- Markers need to be labeled
  - which marker is which?
  - problem: crossover, markers exchange labels
    - may require user intervention
- Compute 3d position: if a marker is seen by at least 2 cameras then its position in 3d space can be determined

http://www.xbox.com/NR/rdonlyres/3164D1BE-C1C4-46A1-90F0-26507CF2C9BD/0/ilmnflfever2003lightscam001.jpg

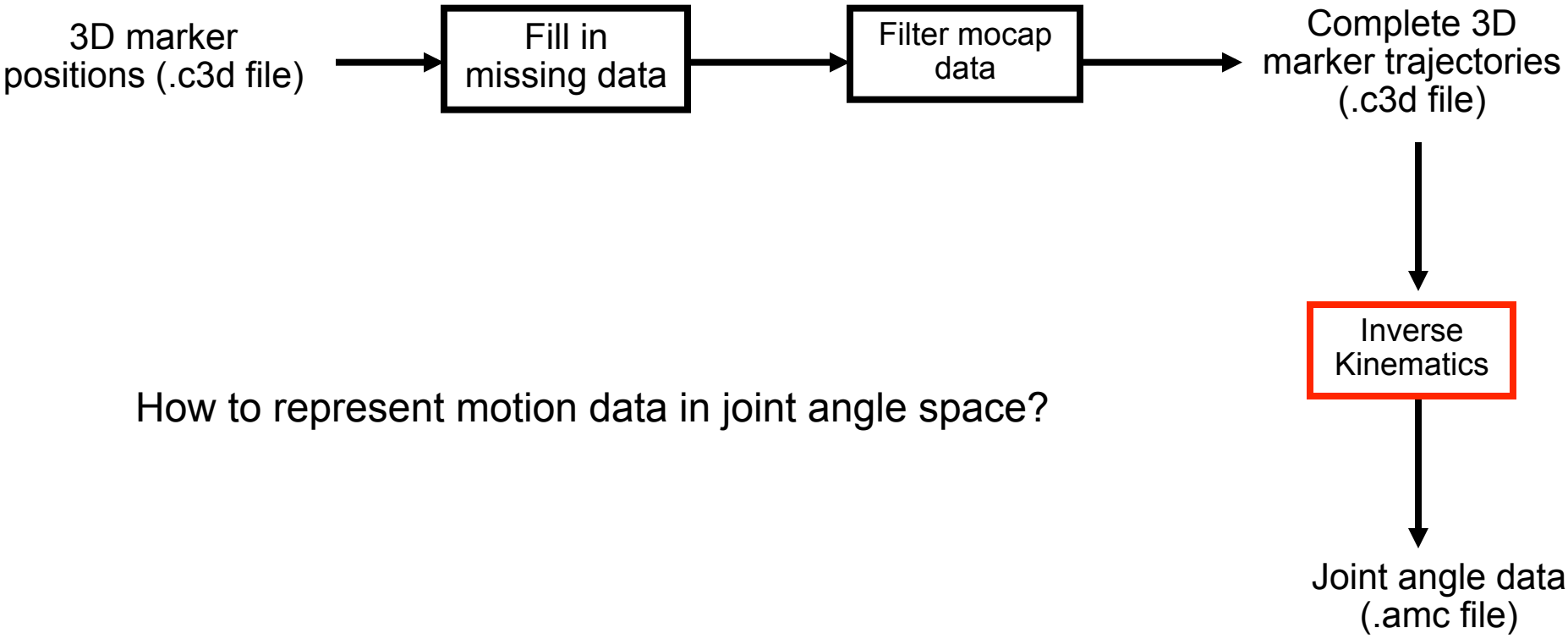
Data Process



Data Process



Data Process

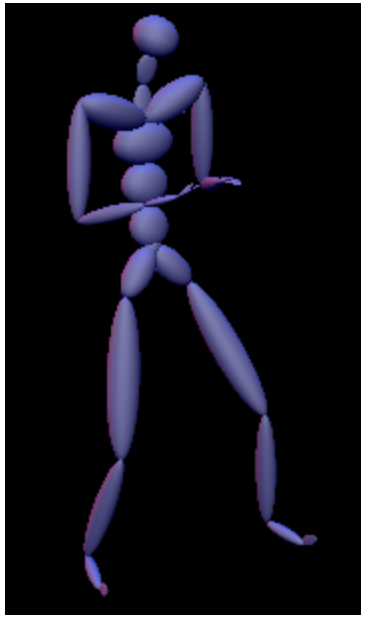


How to represent motion data in joint angle space?

Motion Capture Data Files

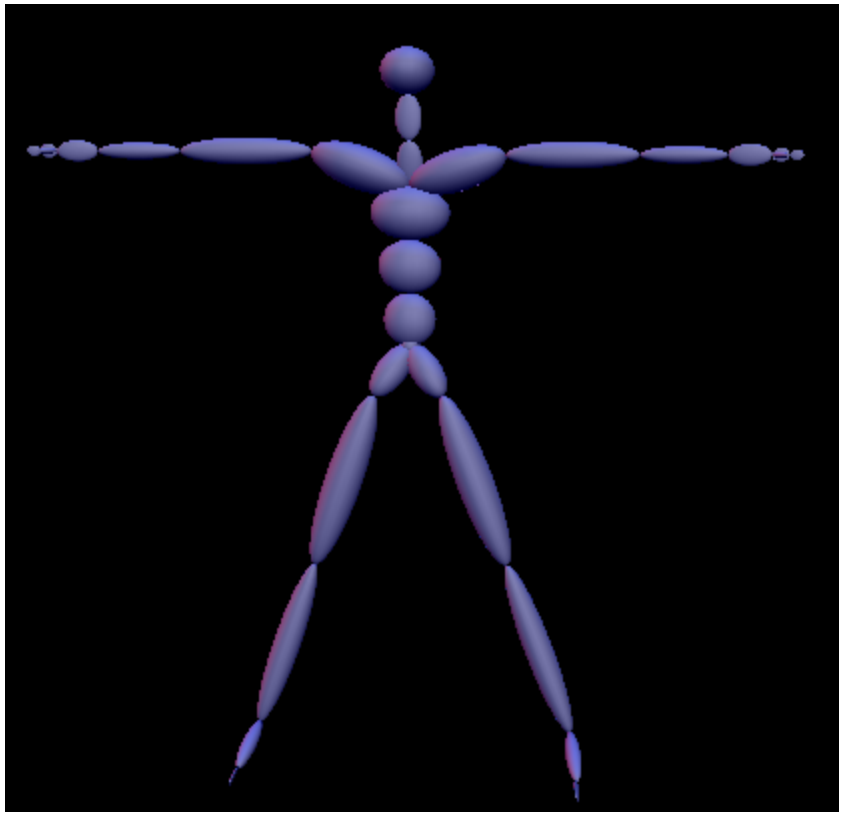
Each sequence of human motion data contains two files:

- Skeleton file (.asf): Specify the skeleton model of character
- Motion data file (.amc): Specify the joint angle values over the frame/time
- Both files are generated by Vicon software

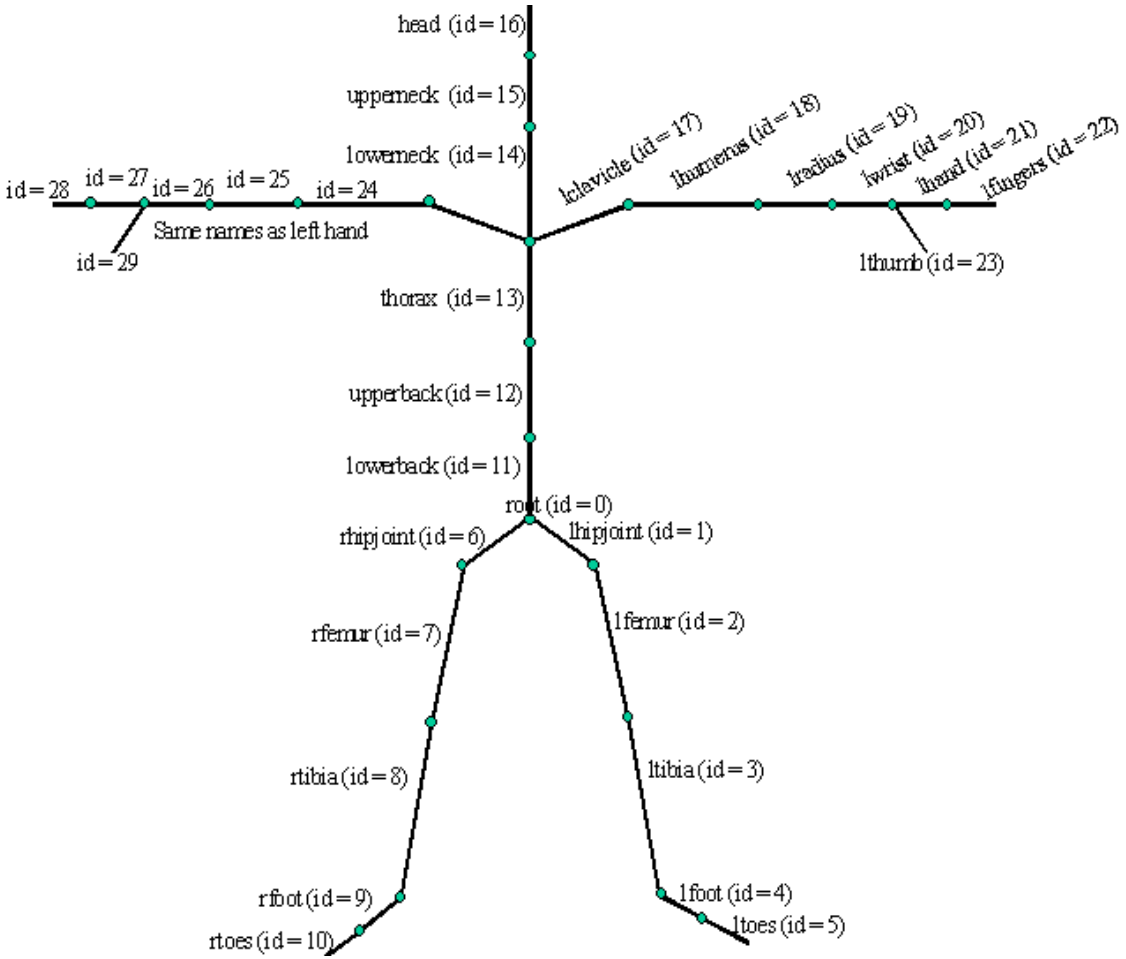


Human Skeletal File

Described in a default pose

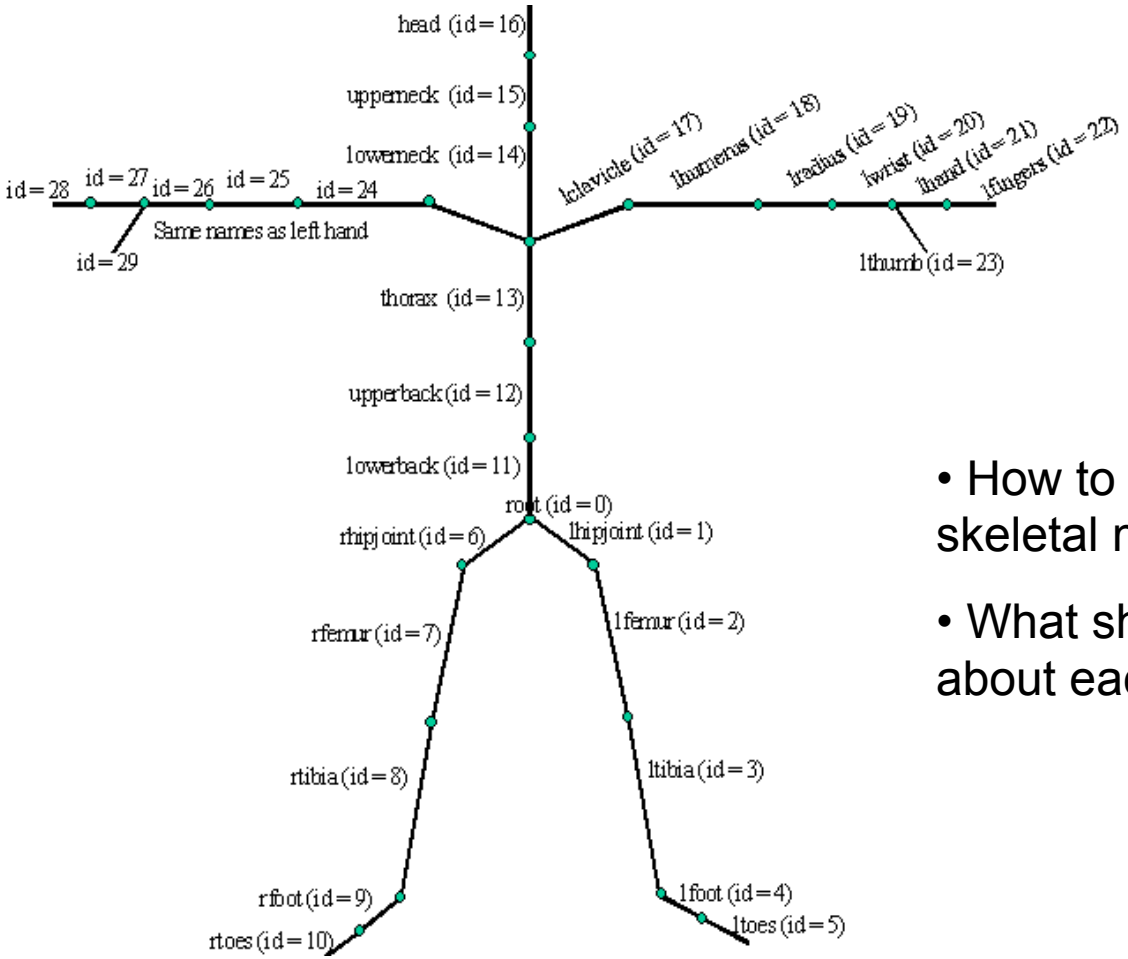


Human Skeletal Model



This is still a tree!

Human Skeletal Model

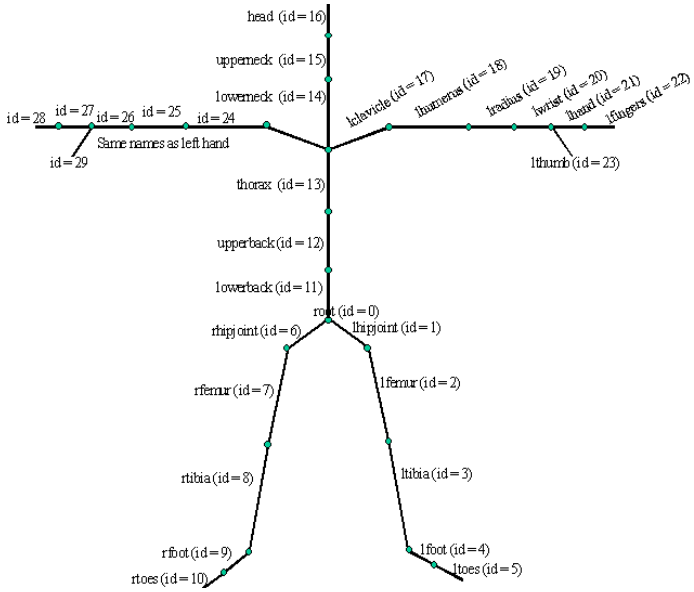


- How to describe the skeletal model?
- What should you know about each bone?

This is still a tree!

Human Skeletal File (.asf)

- individual bone information
  - length of the bone
  - direction of the bone
  - local coordinate frame
  - number of Dofs
  - joint limits



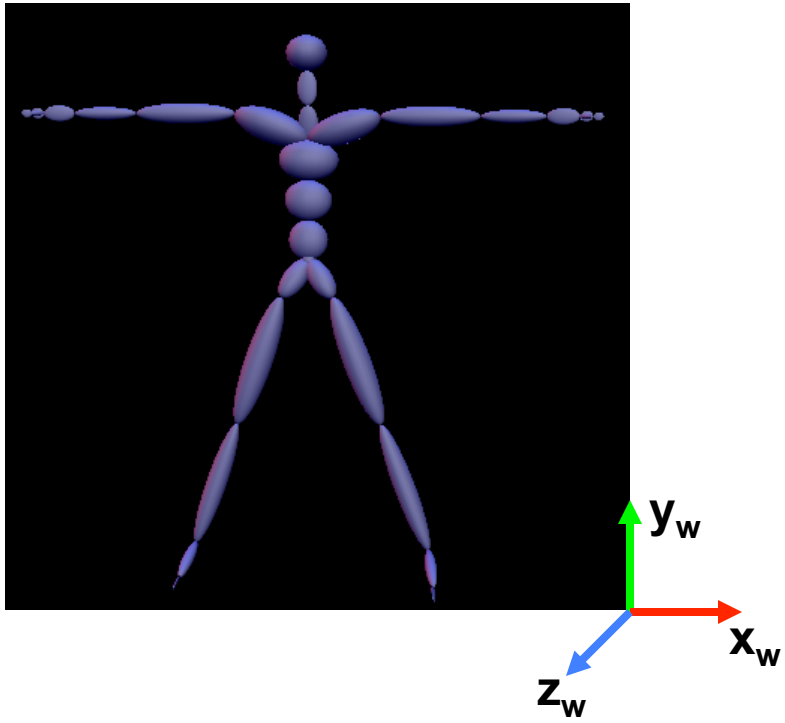
- bone hierarchy/connections

Individual Bone Information

```
begin
    id bone_id              /* Unique id for each bone */
    name bone_name          /* Unique name for each bone */
    direction dX dY dZ      /* Vector describing direction of the bone in world */
    coor. system
    length 7.01722          /* Length of the bone*/
    axis 0 0 20 XYZ         /* Rotation of local coordinate system for
                               this bone relative to the world coordinate
                               system. In .AMC file the rotation angles
                               for this bone for each time frame will be
                               defined relative to this local coordinate
                               system**/
    dof rx ry rz            /* Degrees of freedom for this bone.
    limits (-160.0 20.0)    /* joint limits*/
           (-70.0 70.0)
           (-60.0 70.0)
    end
```

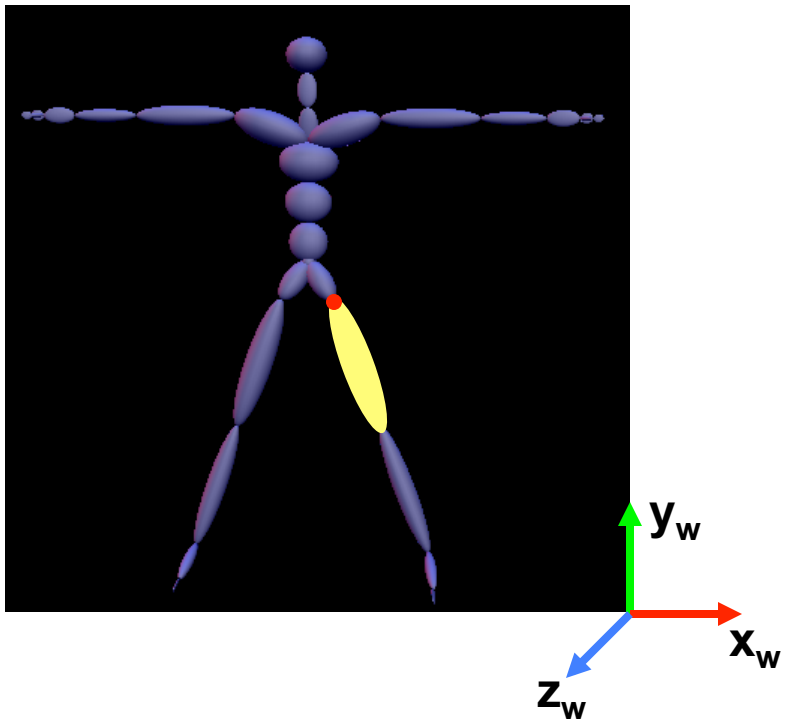
Individual Bone Information

```
begin
id 2
name lfemur
direction 0.34 -0.93 0
length 7.01722
axis 0 0 20 XYZ
dof rx ry rz
limits (-160.0 20.0)
          (-70.0 70.0)
          (-60.0 70.0)
end
```



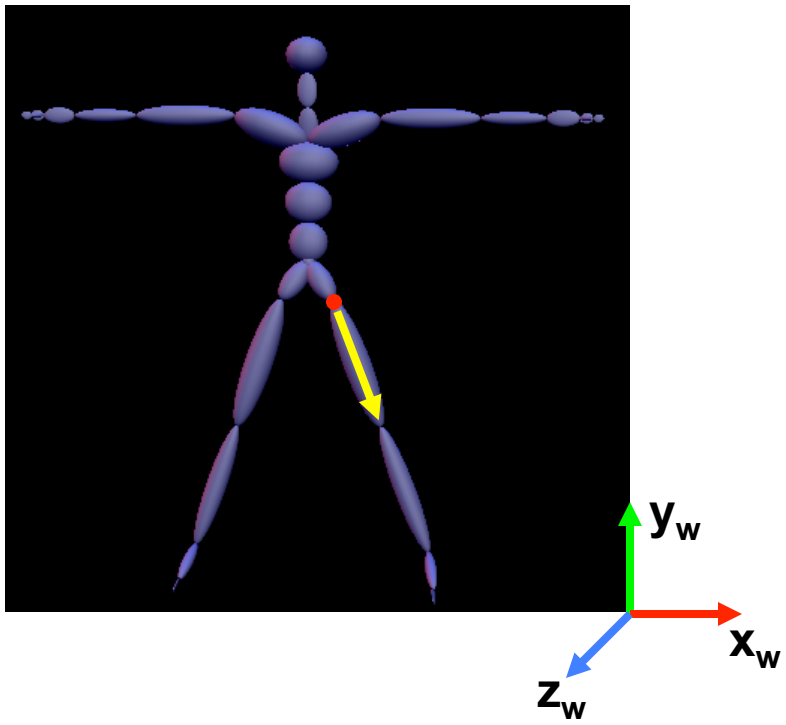
Individual Bone Information

begin
id 2
name lfemur
direction 0.34 -0.93 0
length 7.01722
axis 0 0 20 XYZ
dof rx ry rz
limits (-160.0 20.0)
(-70.0 70.0)
(-60.0 70.0)
end



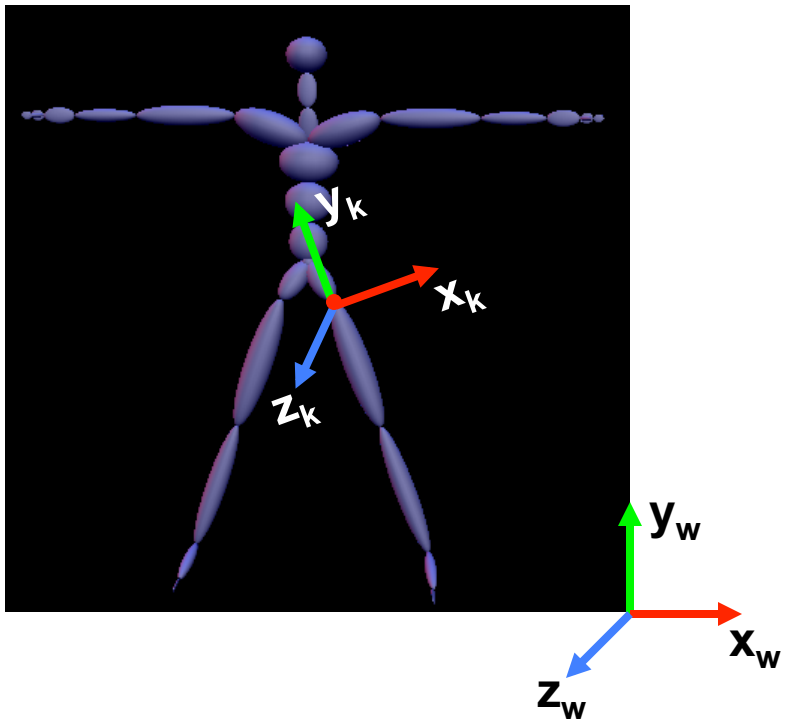
Individual Bone Information

begin
id 2
name lfemur
direction 0.34 -0.93 0
length 7.01722
axis 0 0 20 XYZ
dof rx ry rz
limits (-160.0 20.0)
(-70.0 70.0)
(-60.0 70.0)
end



Individual Bone Information

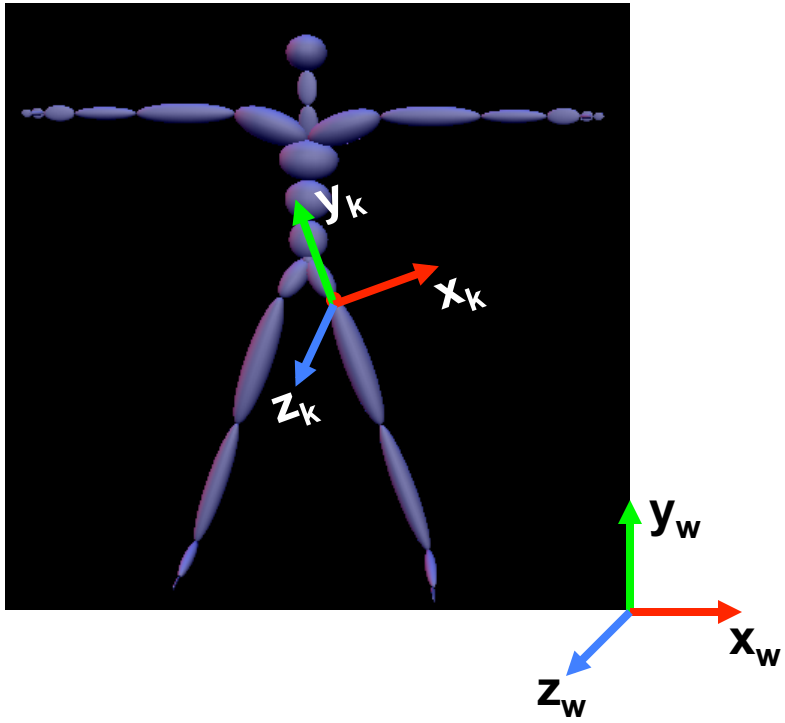
```
begin
id 2
name lfemur
direction 0.34 -0.93 0
length 7.01722
axis 0 0 20 XYZ
dof rx ry rz
limits (-160.0 20.0)
          (-70.0 70.0)
          (-60.0 70.0)
end
```



Euler angle representation: $R_k=R_z(\gamma)R_y(\beta)R_x(\alpha)$

Individual Bone Information

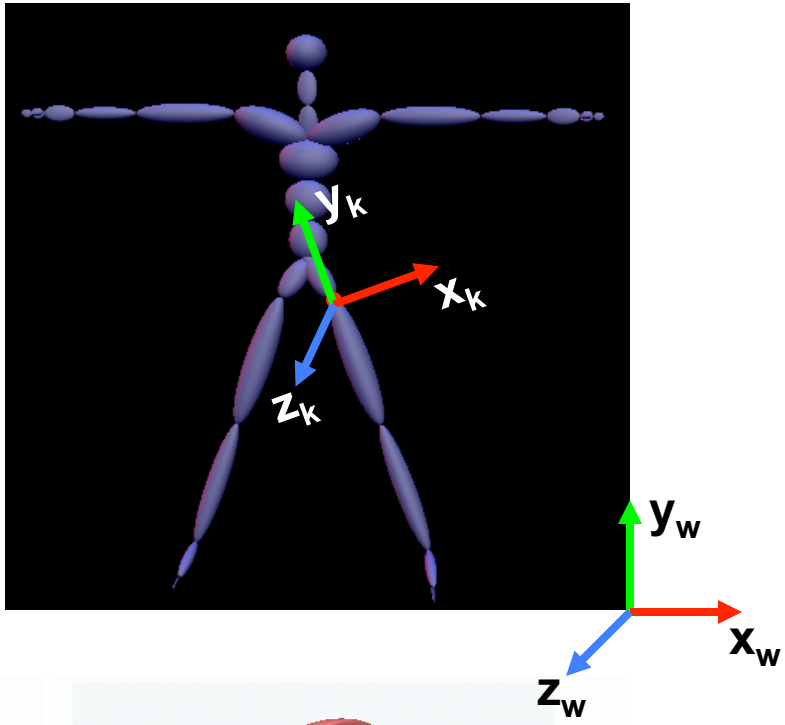
begin
id 2
name lfemur
direction 0.34 -0.93 0
length 7.01722
axis 0 0 20 XYZ
dof rx ry rz
limits (-160.0 20.0)
(-70.0 70.0)
(-60.0 70.0)
end

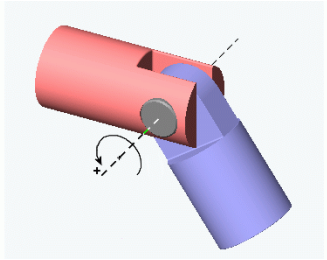


- The number of dof for this joint
- The minimal and maximum joint angle for each dof

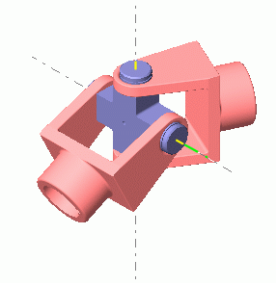
Individual Bone Information

begin
id 2
name lfemur
direction 0.34 -0.93 0
length 7.01722
axis 0 0 20 XYZ
dof rx ry rz
limits (-160.0 20.0)
(-70.0 70.0)
(-60.0 70.0)
end

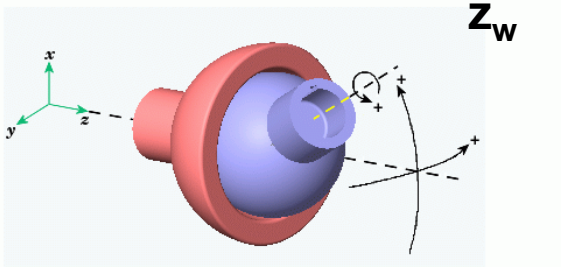




1-dof joint



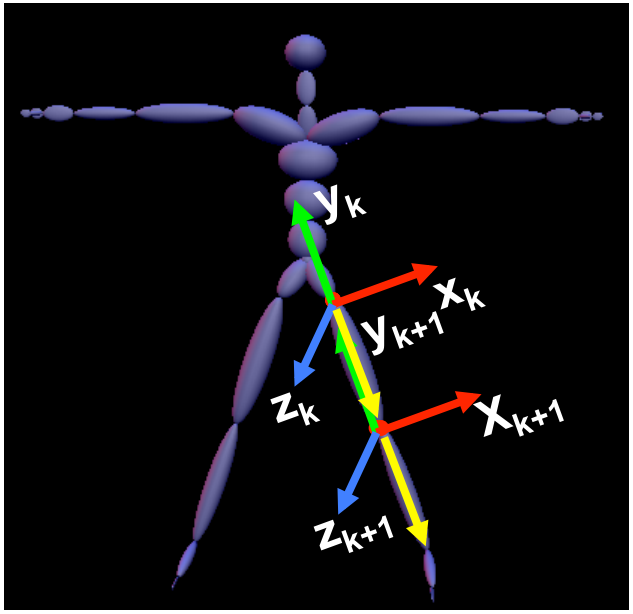
2-dof joint



3-dof joint

Individual Bone Information

```
begin
id 2
name lfemur
direction 0.34 -0.93 0
length 7.01722
axis 0 0 20 XYZ
dof rx ry rz
limits (-160.0 20.0)
          (-70.0 70.0)
          (-60.0 70.0)
end
begin
id 3
name ltibia
direction 0.34 -0.93 0
length 7.2138
axis 0 0 20 XYZ
dof rx
limits (-10.0 170.0)
end
```



Root Representation

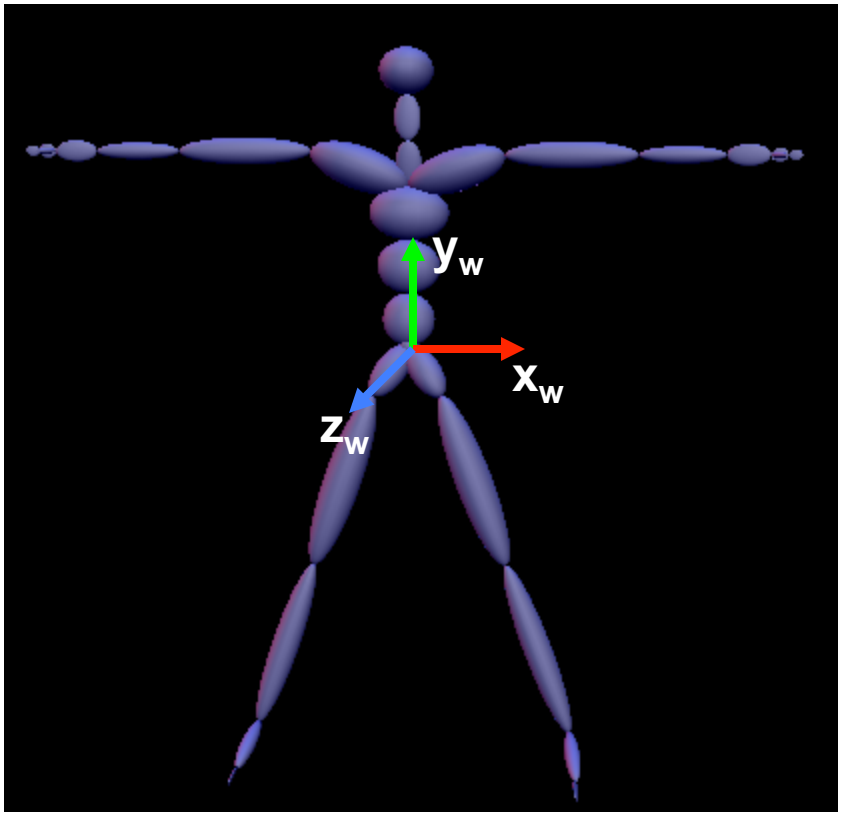
:root

order TX TY TZ RX RY RZ

axis XYZ

position 0 0 0

orientation 0 0 0



Root Representation

:root

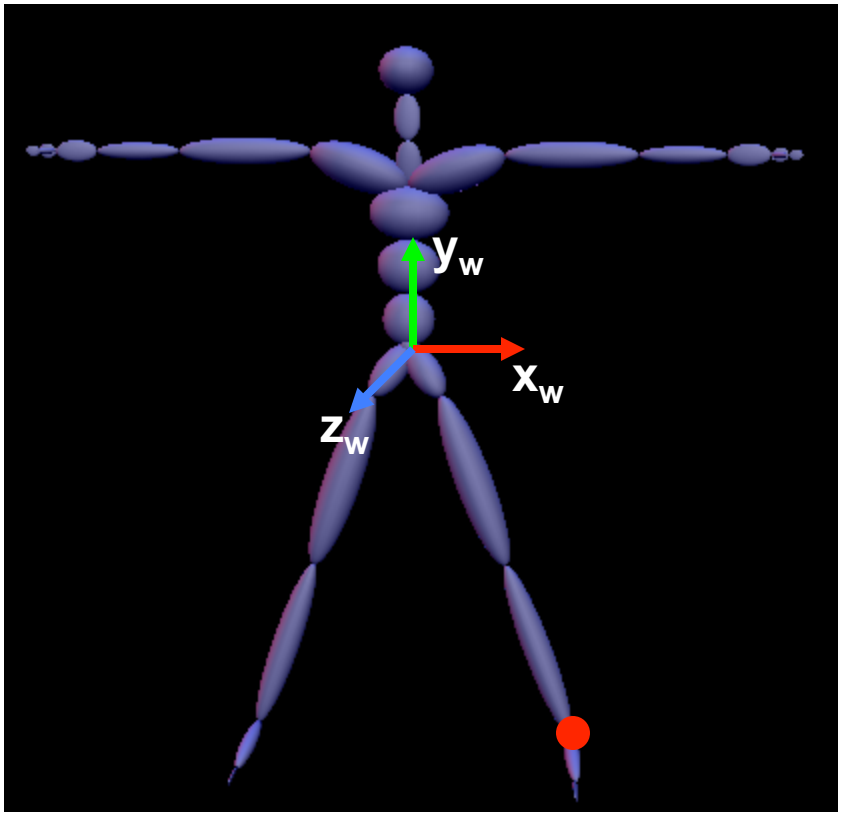
order TX TY TZ RX RY RZ

axis XYZ

position 0 0 0

orientation 0 0 0

How to compute the coordinate of a joint in the world coordinate frame?



Root Representation

:root

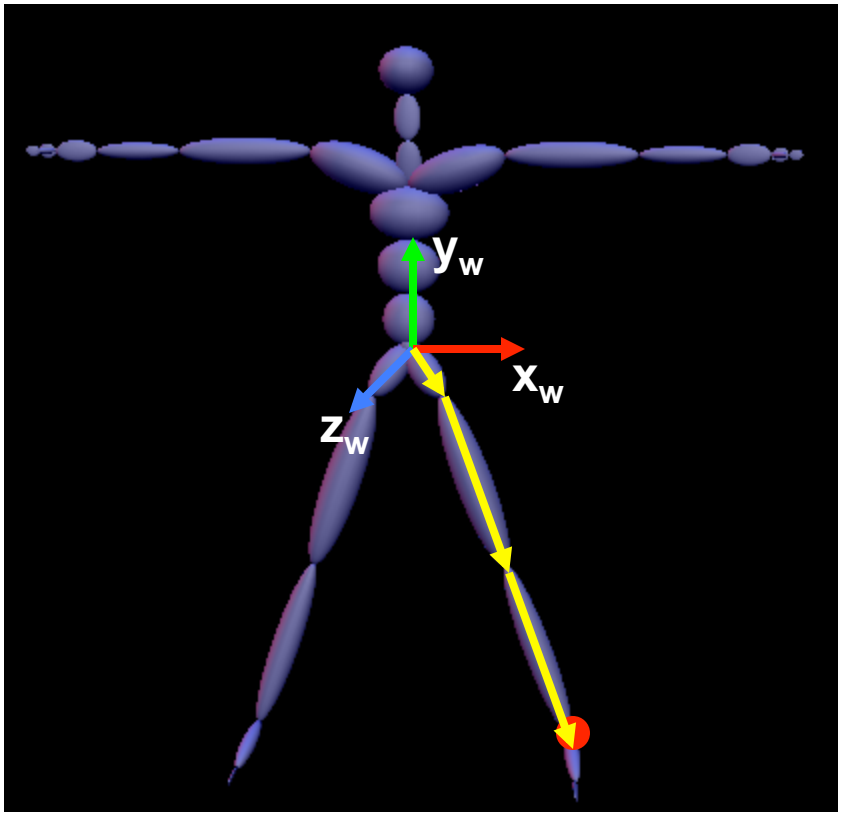
order TX TY TZ RX RY RZ

axis XYZ

position 0 0 0

orientation 0 0 0

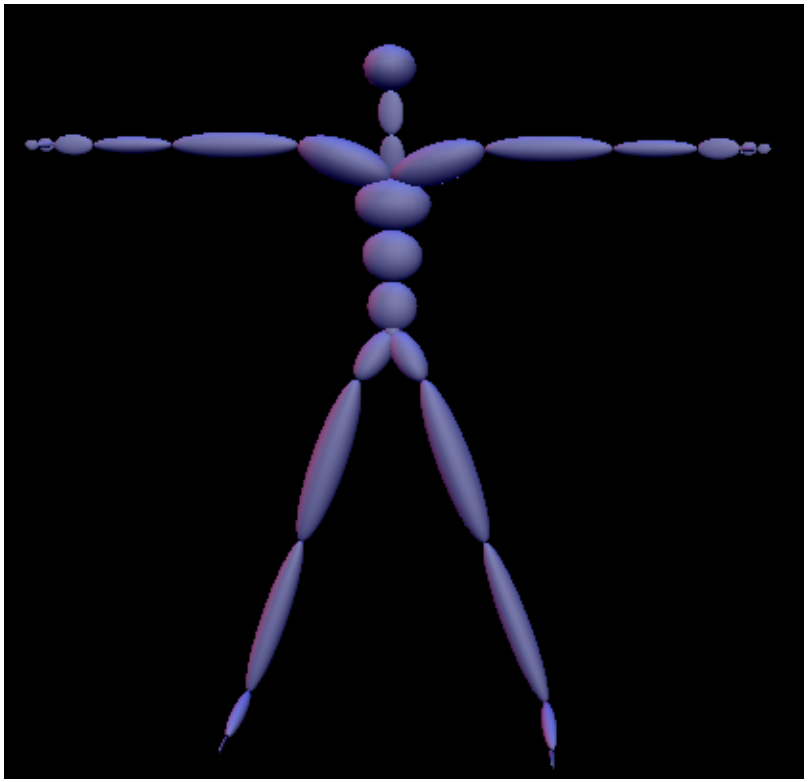
How to compute the coordinate of a joint in the world coordinate frame?



Hierarchy/Bone Connections

:hierarchy

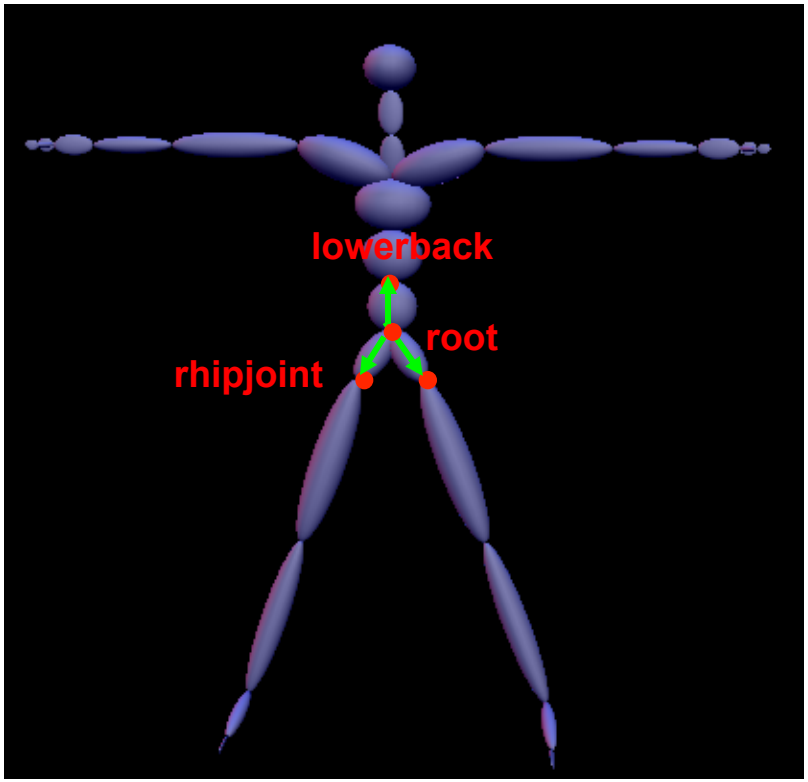
begin
root lhipjoint rhipjoint lowerback
lhipjoint lfemur
lfemur ltibia
ltibia lfoot
lfoot ltoes
rhipjoint rfemur
rfemur rtibia
rtibia rfoot
rfoot rtoes
lowerback upperback
upperback thorax
thorax lowerneck lclavicle rclavicle
…
end



Hierarchy/Bone Connections

:hierarchy

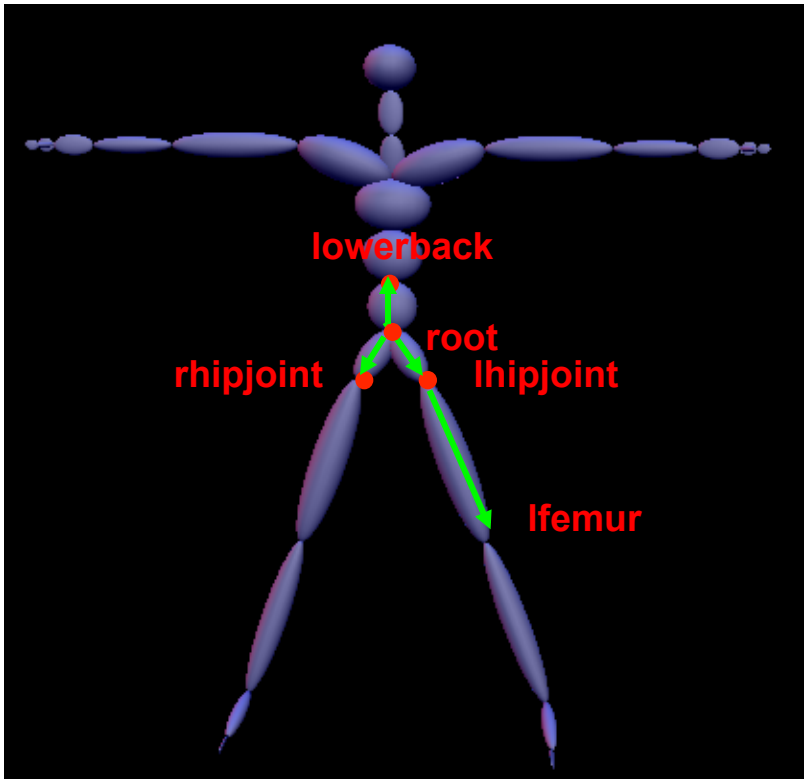
begin
root lhipjoint rhipjoint lowerback
lhipjoint lfemur
lfemur ltibia
ltibia lfoot
lfoot ltoes
rhipjoint rfemur
rfemur rtibia
rtibia rfoot
rfoot rtoes
lowerback upperback
upperback thorax
thorax lowerneck lclavicle rclavicle
…
end



Hierarchy/Bone Connections

:hierarchy

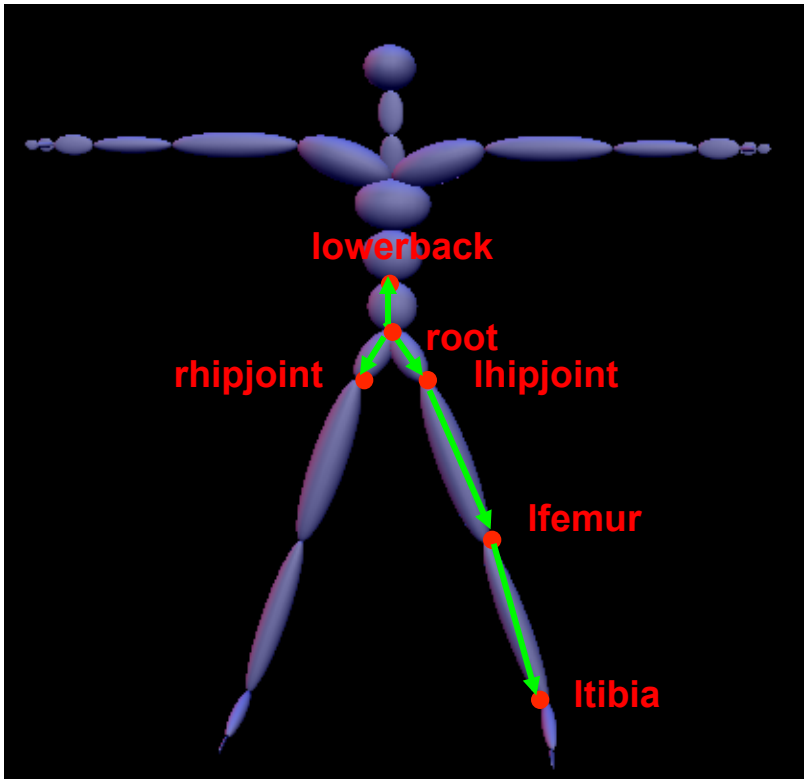
begin
root lhipjoint rhipjoint lowerback
lhipjoint lfemur
lfemur ltibia
ltibia lfoot
lfoot ltoes
rhipjoint rfemur
rfemur rtibia
rtibia rfoot
rfoot rtoes
lowerback upperback
upperback thorax
thorax lowerneck lclavicle rclavicle
…
end



Hierarchy/Bone Connections

:hierarchy

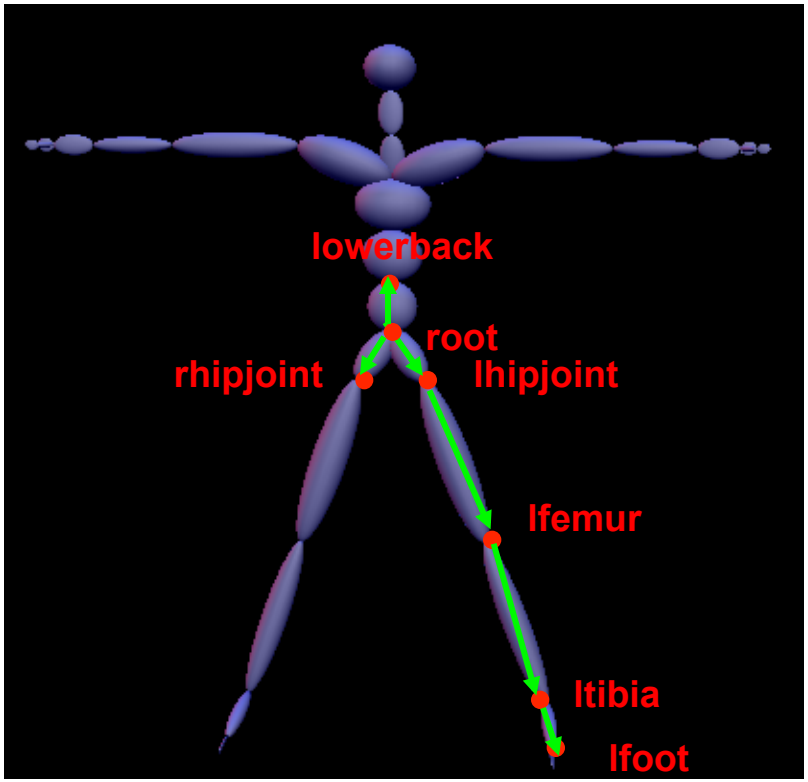
begin
root lhipjoint rhipjoint lowerback
lhipjoint lfemur
lfemur ltibia
ltibia lfoot
lfoot ltoes
rhipjoint rfemur
rfemur rtibia
rtibia rfoot
rfoot rtoes
lowerback upperback
upperback thorax
thorax lowerneck lclavicle rclavicle
…
end



Hierarchy/Bone Connections

:hierarchy

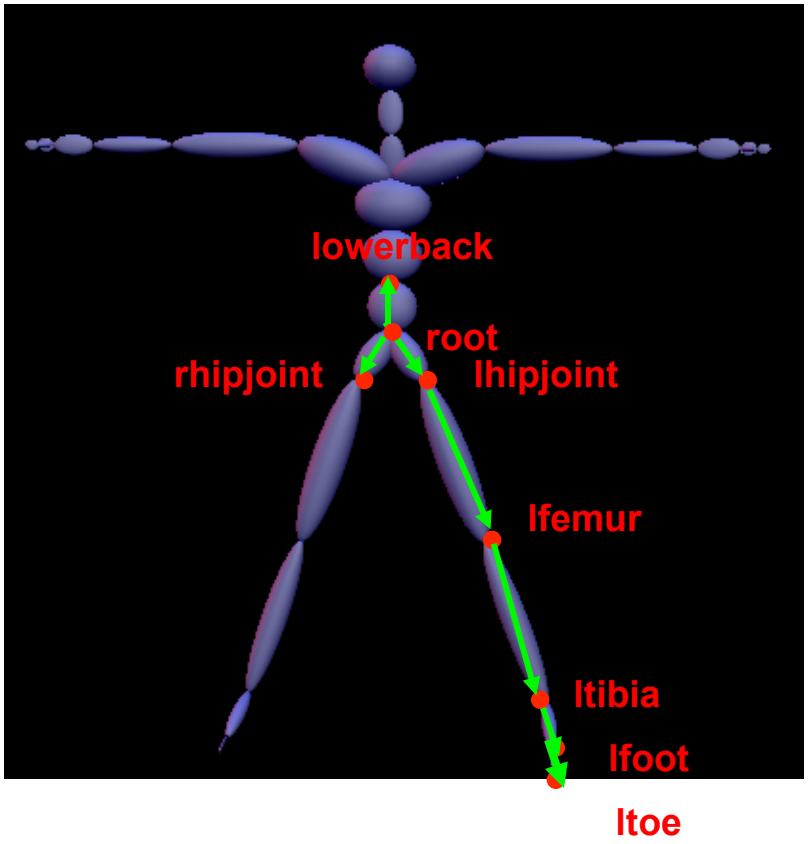
begin
root lhipjoint rhipjoint lowerback
lhipjoint lfemur
lfemur ltibia
ltibia lfoot
lfoot ltoes
rhipjoint rfemur
rfemur rtibia
rtibia rfoot
rfoot rtoes
lowerback upperback
upperback thorax
thorax lowerneck lclavicle rclavicle
…
end



Hierarchy/Bone Connections

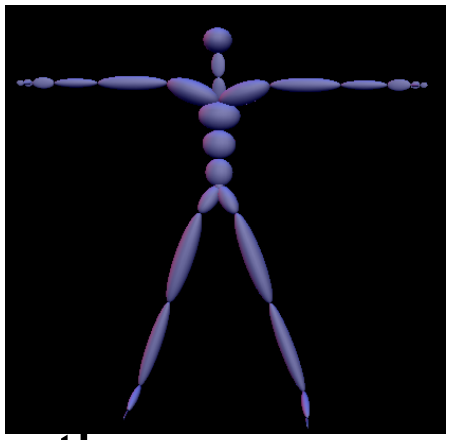
:hierarchy

begin
root lhipjoint rhipjoint lowerback
lhipjoint lfemur
lfemur ltibia
ltibia lfoot
lfoot ltoes
rhipjoint rfemur
rfemur rtibia
rtibia rfoot
rfoot rtoes
lowerback upperback
upperback thorax
thorax lowerneck lclavicle rclavicle
…
end



What Can We Do With .asf File?

We can visualize the default pose

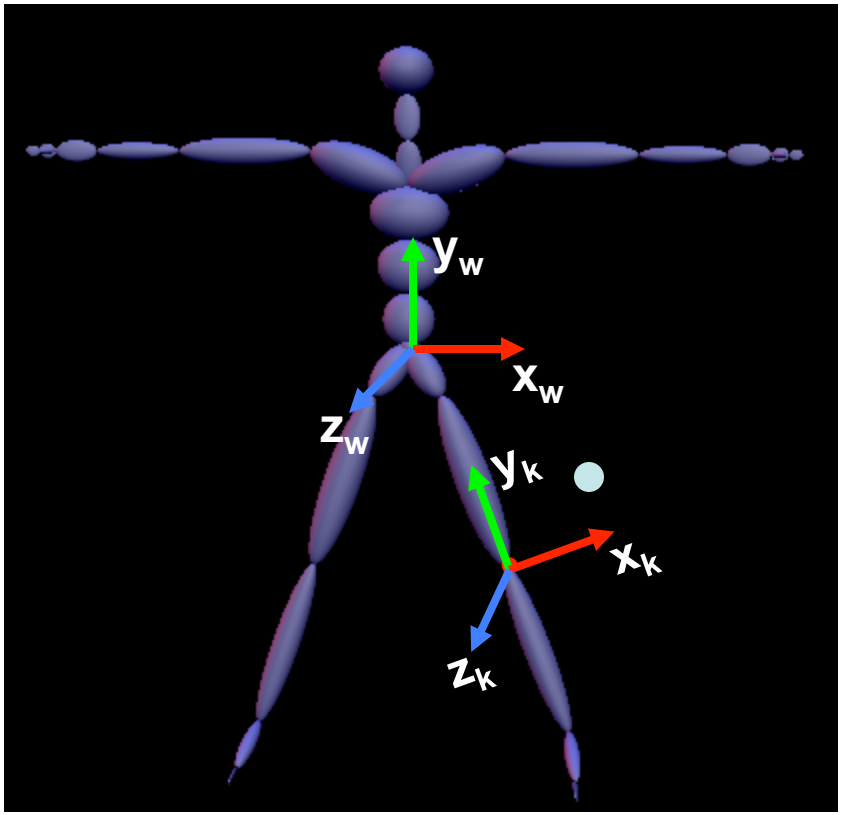


We can compute various transforms in the default pose

- between world coordinate frame and local coordinate
- between parent coordinate frame and child coordinate frame

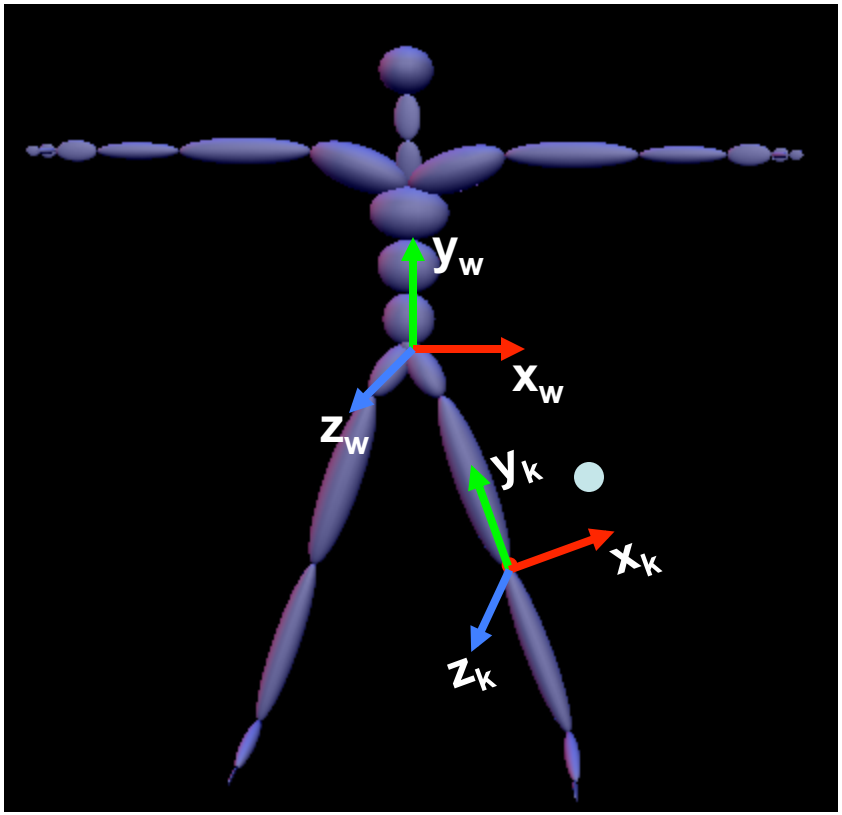
From Local Coordinate to World Coordinate

$$X_w = R_k X_k + T_k$$



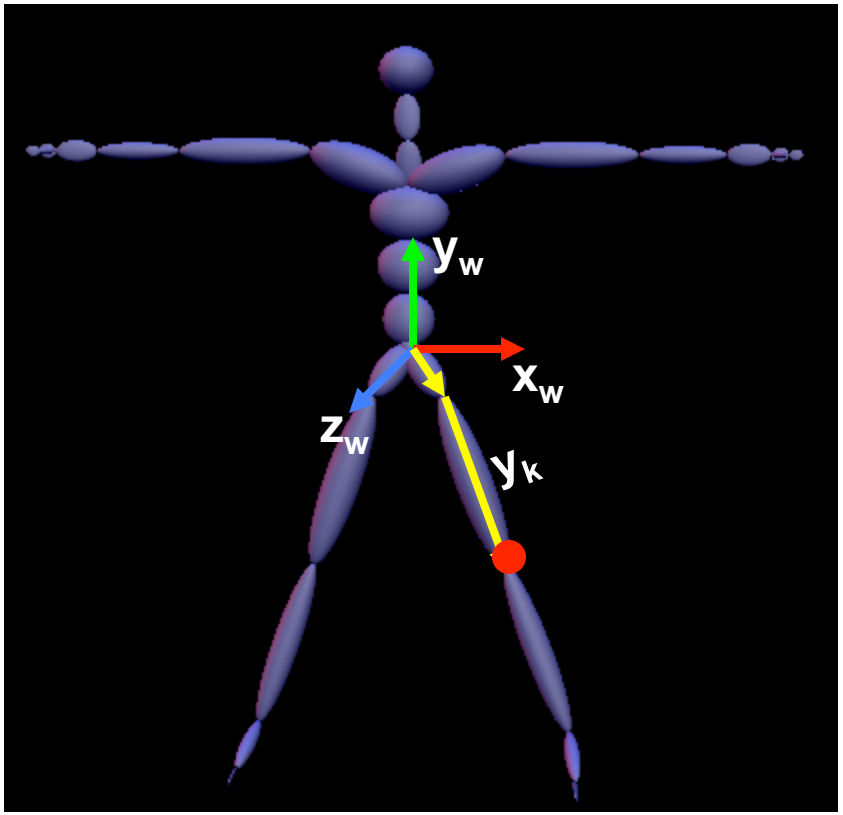
From Local Coordinate to World Coordinate

$$X_w = R_k X_k + T_k$$



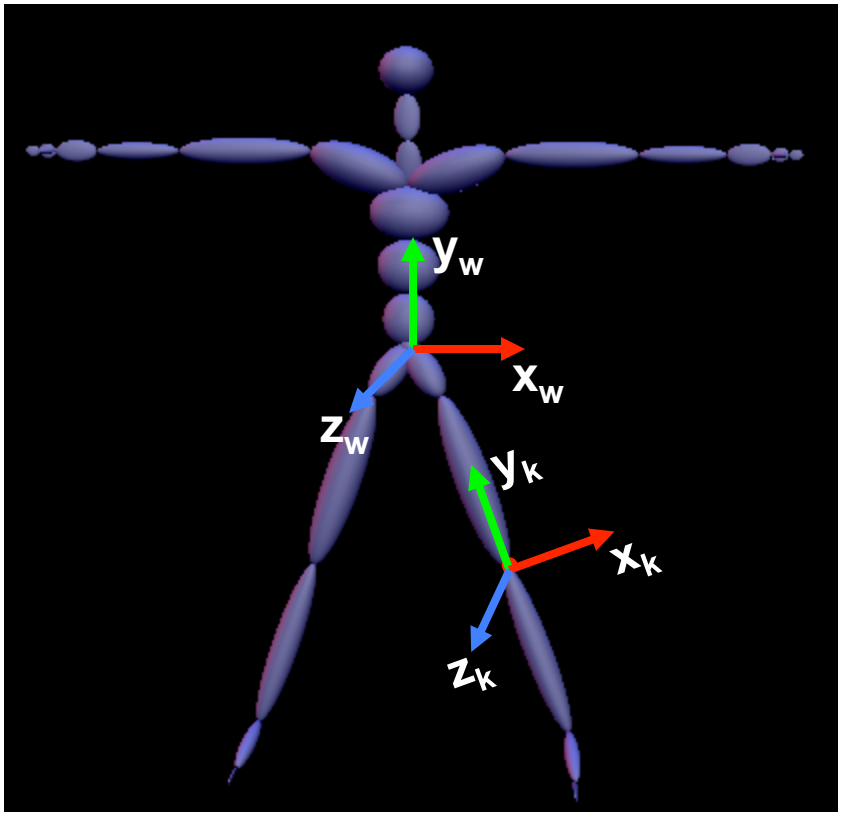
From Local Coordinate to World Coordinate

$$X_w = R_k X_k + T_k$$



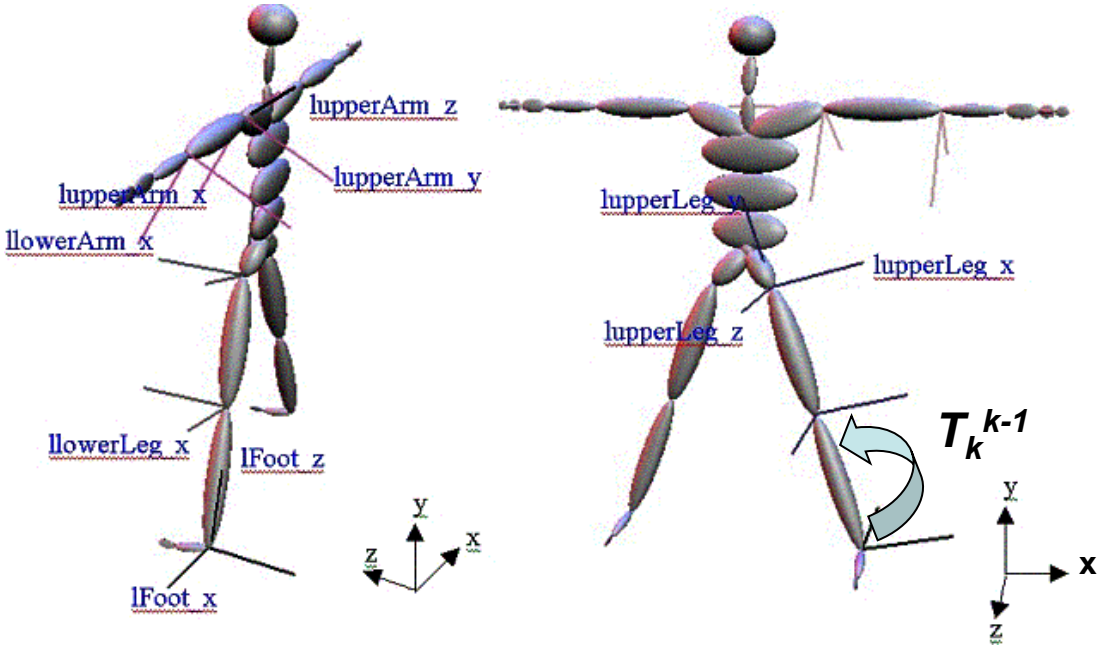
From Local Coordinate to World Coordinate

$$X_w = R_k X_k + T_k$$

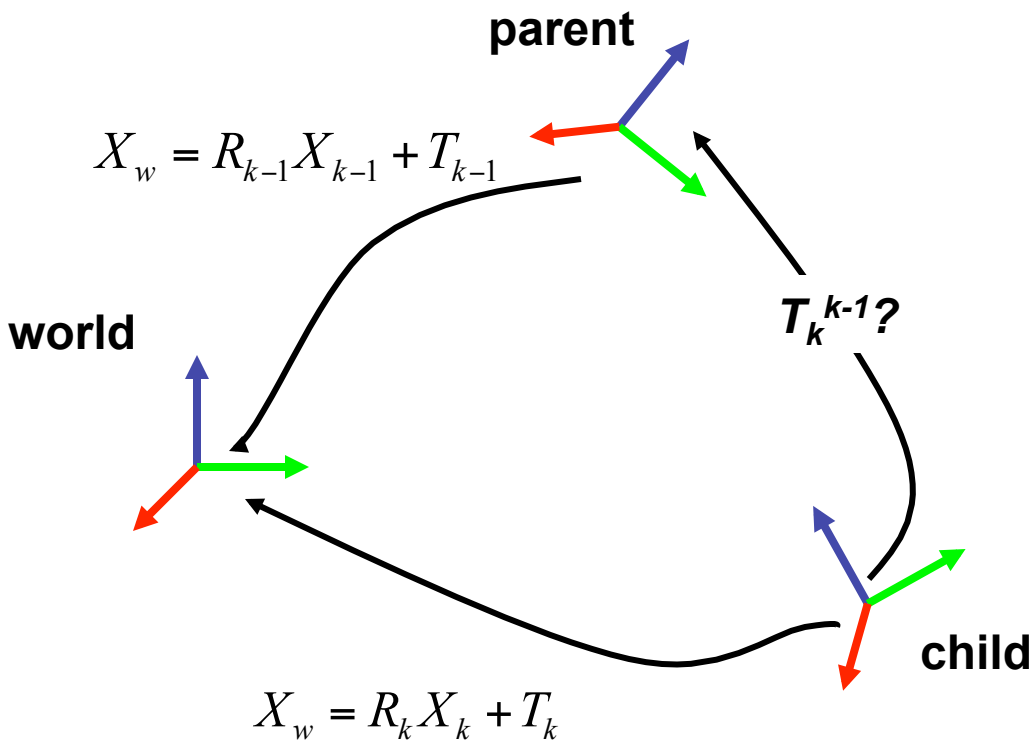


From Child to Parent Node

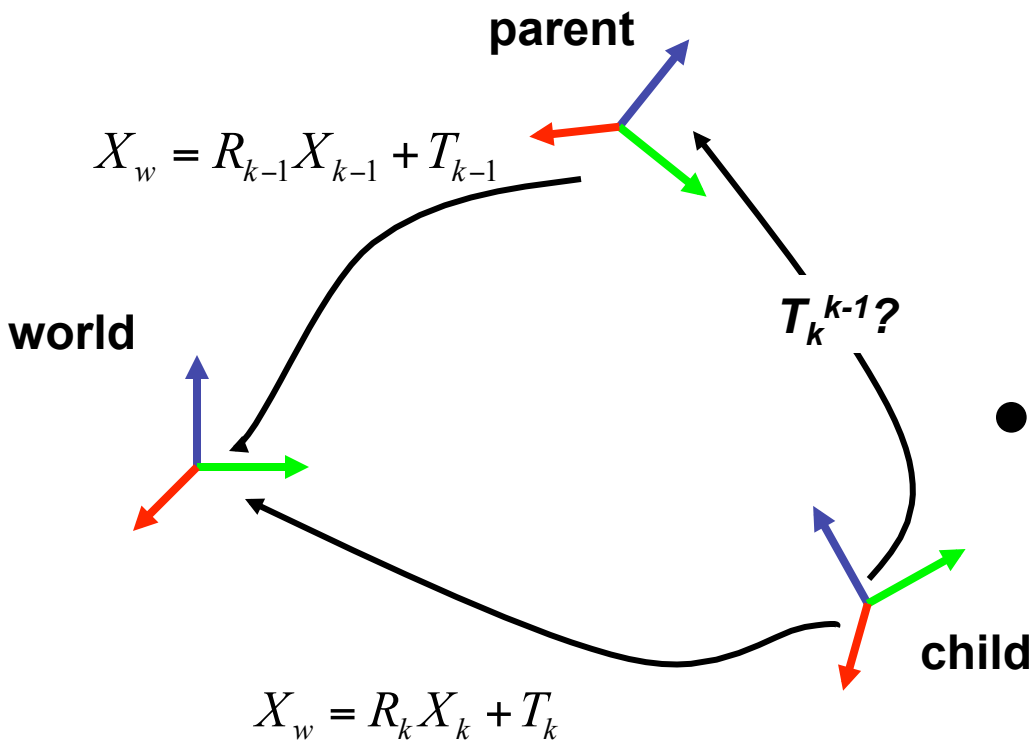
How to Compute the transformation T_k^{k-1} from a child local coordinate frame to its parent local coordinate frame



Bone Transform

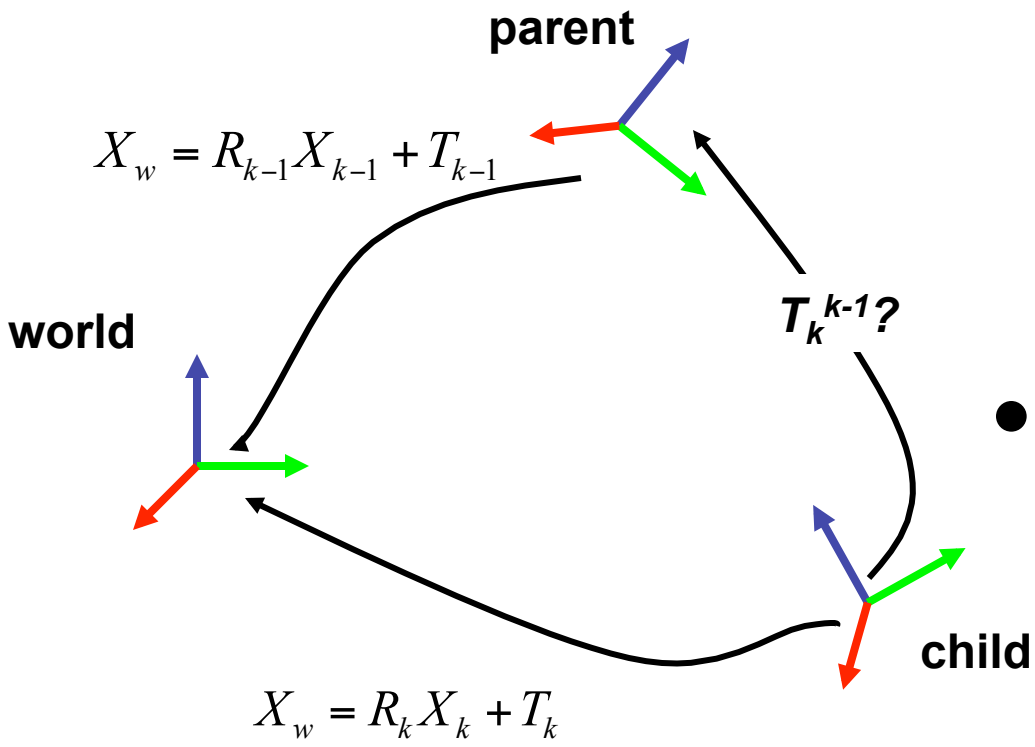


Bone Transform



$$X_w = R_k X_k + T_k = R_{k-1}X_{k-1} + T_{k-1}$$

Bone Transform



$$X_w = R_k X_k + T_k = R_{k-1}X_{k-1} + T_{k-1}$$

$$X_{k-1} = R_{k-1}^{-1}R_k X_k + R_{k-1}^{-1}(T_k - T_{k-1})$$

$$T_k^{k-1} = \begin{bmatrix} R_{k-1}^{-1}R_k & -R_{k-1}^{-1}(T_{k-1} - T_k) \\ 0 & 1 \end{bmatrix}$$

Motion Data File (.amc)

```
i                                                                    // frame number
root 2.36756 16.4521 12.3335 -165.118 31.188 -179.889 // root position and orientation
lowerback -17.2981 -0.243065 -1.41128                 // joint angles for lowerback joint
upperback 0.421503 -0.161394 2.20925                  // joint angles for thorax joint
thorax 10.2185 -0.176777 3.1832
lowerneck -15.0172 -5.84786 -7.55529
upperneck 30.0554 -3.19622 -4.68899
head 12.6247 -2.35554 -0.876544
rclavicle 4.77083e-014 -3.02153e-014
rhumerus -23.3927 30.8588 -91.7324
rradius 108.098
rwrist -35.4375
rhand -5.30059 11.2226
rfingers 7.12502
rthumb 20.5046 -17.7147
lclavicle 4.77083e-014 -3.02153e-014
lhumerus -35.2156 -19.5059 100.612
```

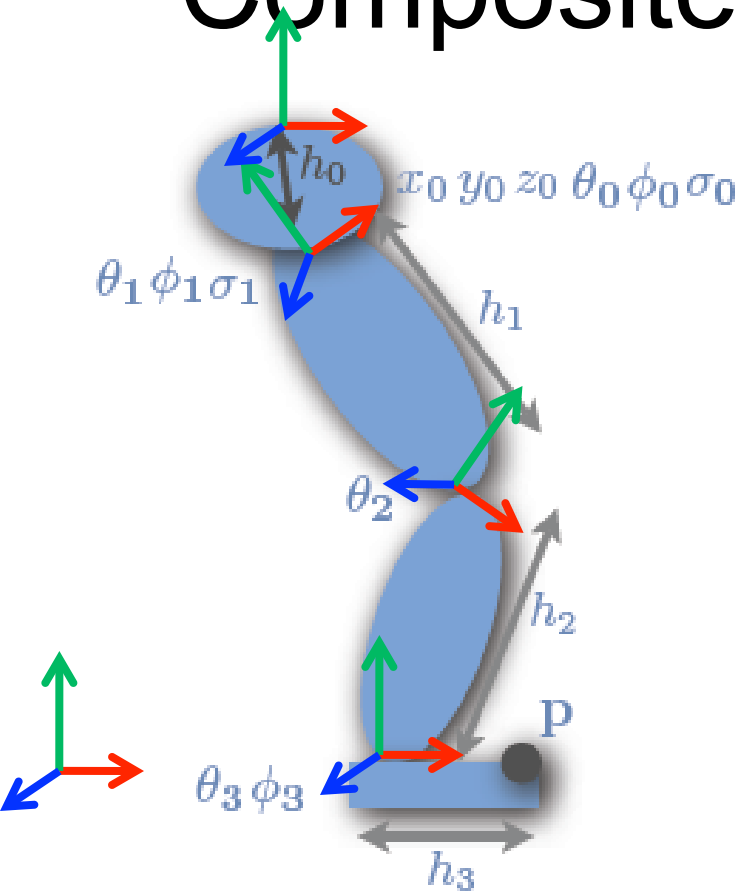
Motion Data File (.amc)

```
i
      // frame number
root 2.36756 16.4521 12.3335 -165.118 31.188
      -179.889 // root position and orientation
lowerback -17.2981 -0.243065
      -1.41128                    // joint angles for
      lowerback joint
upperback 0.421503 -0.161394
      2.20925                     // joint angles for
      thorax joint
thorax 10.2185 -0.176777 3.1832
lowerneck -15.0172 -5.84786 -7.55529
upperneck 30.0554 -3.19622 -4.68899
head 12.6247 -2.35554 -0.876544
rclavicle 4.77083e-014 -3.02153e-014
rhumerus -23.3927 30.8588 -91.7324
rradius 108.098
rwrist -35.4375
rhand -5.30059 11.2226
rfingers 7.12502
rthumb 20.5046 -17.7147
lclavicle 4.77083e-014 -3.02153e-014
lhumerus -35.2156 -19.5059 100.612
```

- Rotation described in local coordinate frame
- Euler angle representation x-y-z

$$R_z(\delta)R_y(\phi)R_x(\theta)$$

Composite 3D Transformation



A series of transformations on an object can be applied as a series of matrix multiplications

p : position in the global coordinate

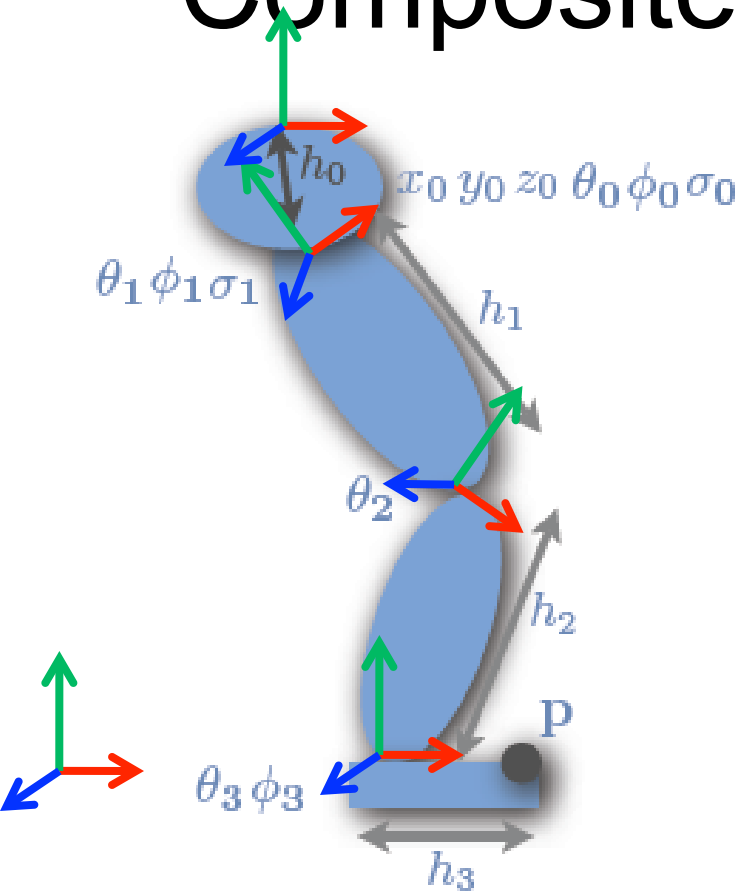
x : position in the local coordinate

$(h_3, 0, 0)$

$$p = T(x_0, y_0, z_0)R(\theta_0)R(\varphi_0)R(\delta_0)T_1^0 R(\theta_1)R(\varphi_1)R(\delta_1)T_2^1 R(\theta_2)T_3^2 R(\theta_3)R(\varphi_3)x$$

From .asf file

Composite 3D Transformation



A series of transformations on an object can be applied as a series of matrix multiplications

p : position in the global coordinate

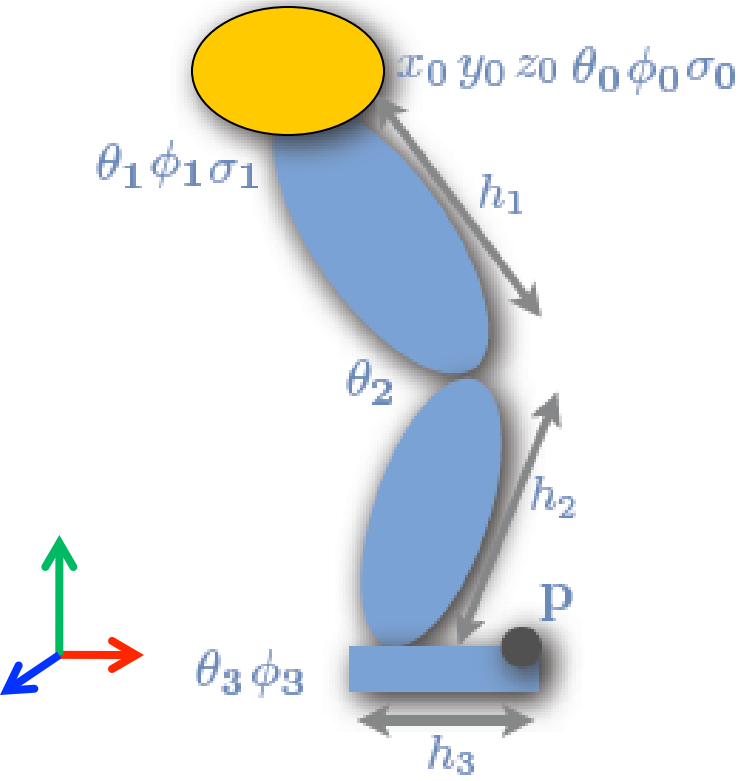
x : position in the local coordinate

$(h_3, 0, 0)$

$$p = T(x_0, y_0, z_0)R(\theta_0)R(\varphi_0)R(\delta_0)T_1^0 R(\theta_1)R(\varphi_1)R(\delta_1)T_2^1 R(\theta_2)T_3^2 R(\theta_3)R(\varphi_3)x$$

From .amc file

Composite 3D Transformation



A series of transformations on an object can be applied as a series of matrix multiplications

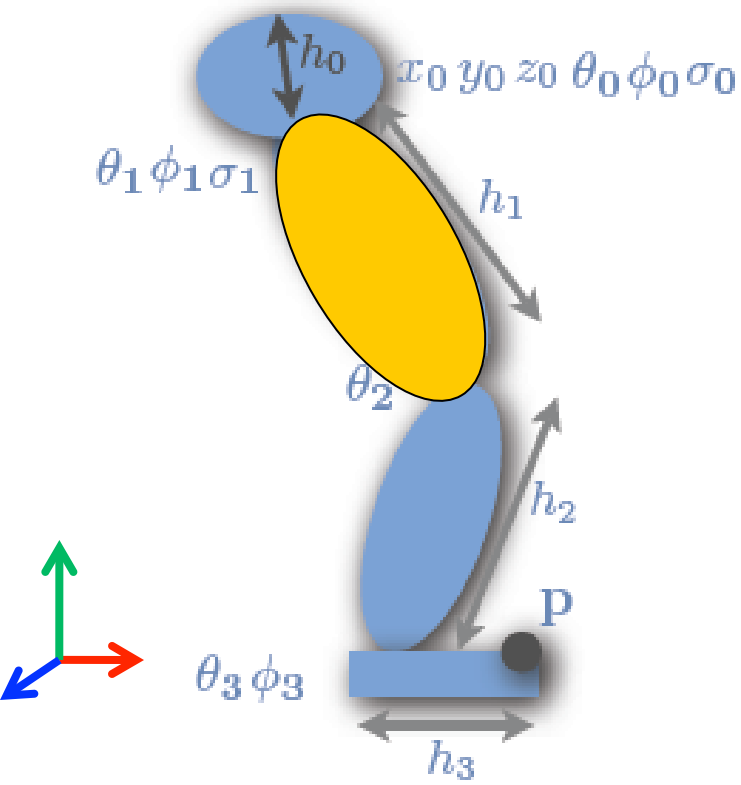
p : position in the global coordinate

x : position in the local coordinate

$(h_3, 0, 0)$

$$p = T(x_0, y_0, z_0)R(\theta_0)R(\varphi_0)R(\delta_0)T_1^0R(\theta_1)R(\varphi_1)R(\delta_1)T_2^1R(\theta_2)T_3^2R(\theta_3)R(\varphi_3)x$$

Composite 3D Transformation



A series of transformations on an object can be applied as a series of matrix multiplications

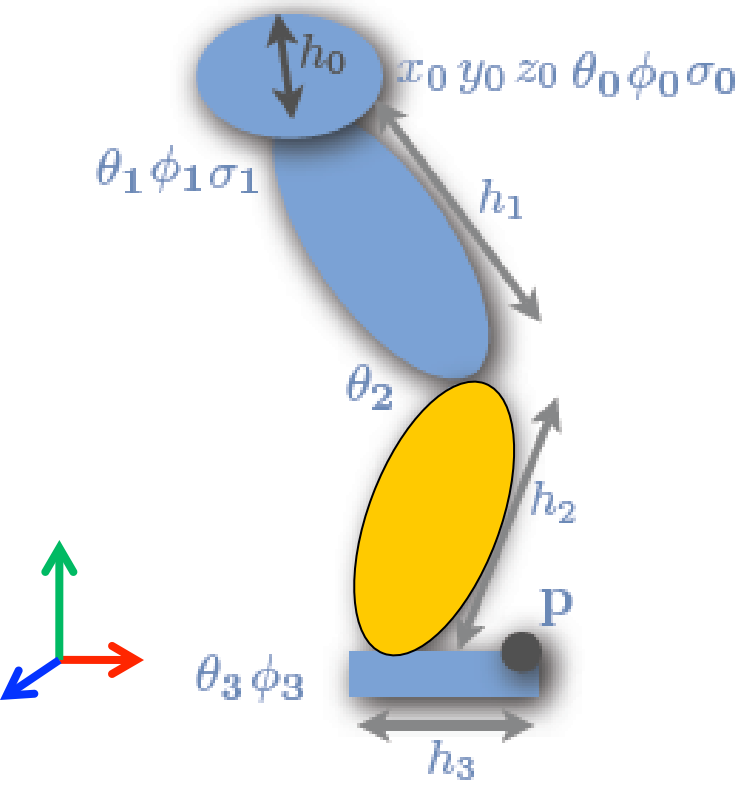
p : position in the global coordinate

x : position in the local coordinate

$(h_3, 0, 0)$

$$p = T(x_0, y_0, z_0)R(\theta_0)R(\varphi_0)R(\delta_0)T_1^0 R(\theta_1)R(\varphi_1)R(\delta_1)T_2^1 R(\theta_2)T_3^2 R(\theta_3)R(\varphi_3)x$$

Composite 3D Transformation



A series of transformations on an object can be applied as a series of matrix multiplications

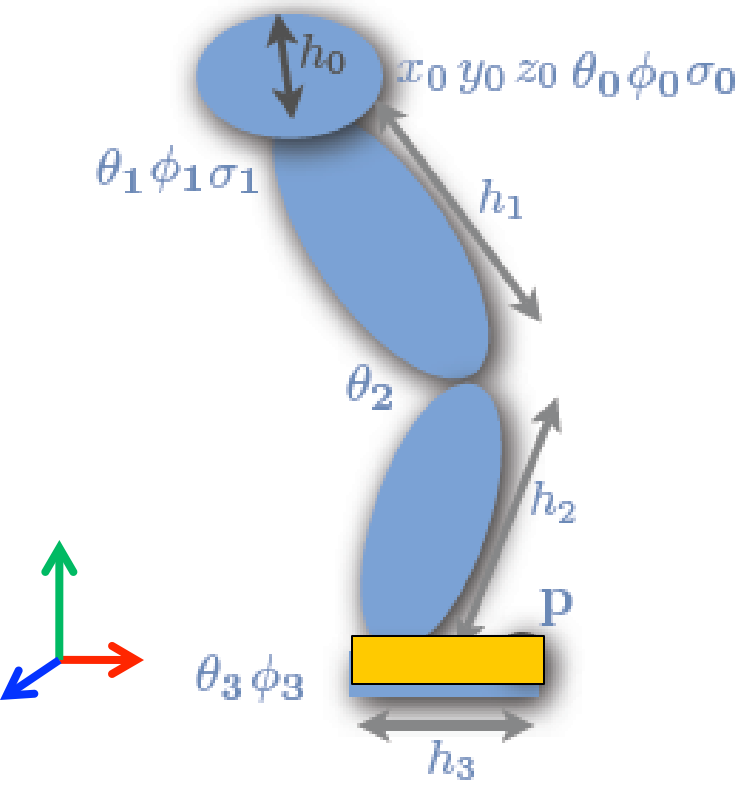
p : position in the global coordinate

x : position in the local coordinate

$(h_3, 0, 0)$

$$p = T(x_0, y_0, z_0)R(\theta_0)R(\varphi_0)R(\delta_0)T_1^0 R(\theta_1)R(\varphi_1)R(\delta_1)T_2^1 R(\theta_2)T_3^2 R(\theta_3)R(\varphi_3)x$$

Composite 3D Transformation



A series of transformations on an object can be applied as a series of matrix multiplications

p : position in the global coordinate

x : position in the local coordinate

$(h_3, 0, 0)$

$$p = T(x_0, y_0, z_0)R(\theta_0)R(\varphi_0)R(\delta_0)T_1^0 R(\theta_1)R(\varphi_1)R(\delta_1)T_2^1 R(\theta_2)T_3^2 R(\theta_3)R(\varphi_3)x$$

Online Motion Capture Database

- http://mocap.cs.cmu.edu/

Mocap for performance-based visualization and interaction

- Universal everything (youtube video)
- Motion meets music (youtube video)
- Another example (youtube video)
- Or (youtube video)
- Dancing movement visualization and analysis (youtube video)

Try Vicon/Kinect

- Online/offline visualization/interaction using Vicon

- Online visualization/interaction using Kinect cameras

  - improve accuracy using prerecorded mocap data from Vicon