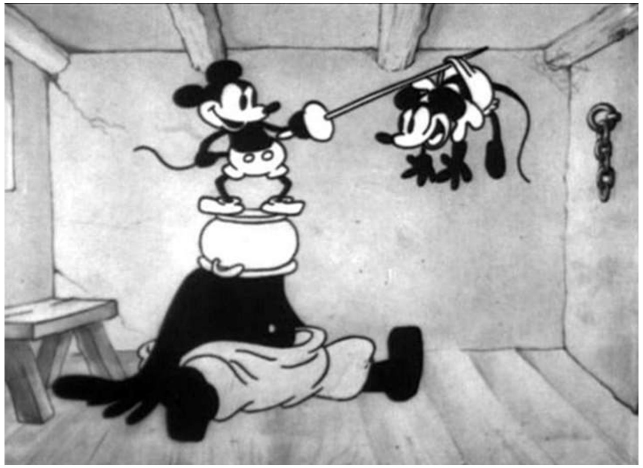
Animation & Simulation

He Wang (王鹤)

Motion Capture

- Sensing, digitalising and recording motions
- Why? The paramount question: motion naturalness
  - Physical naturalness <-> perceived naturalness
  - Artistic naturalness <-> physical naturalness

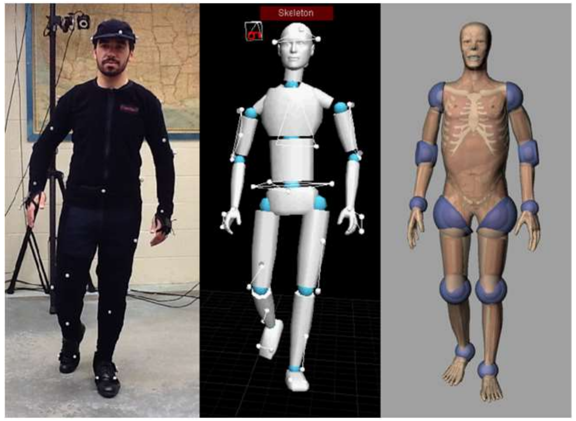


Motion Capture

- Sensing, digitalising and recording motions
- Why? The paramount question: motion naturalness
- Capture what : Anything we can

Motion Capture

- Sensing, digitalising and recording motions
- Why? The paramount question: motion naturalness
- Capture what

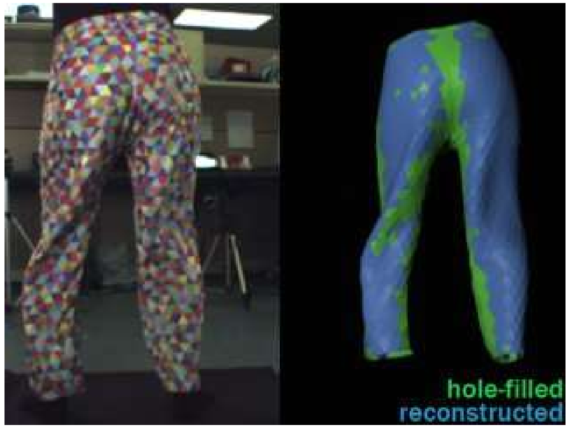


Motion Capture

- Sensing, digitalising and recording motions
- Why? The paramount question: motion naturalness
- Capture what

Motion Capture

- Sensing, digitalising and recording motions
- Why? The paramount question: motion naturalness
- Capture what

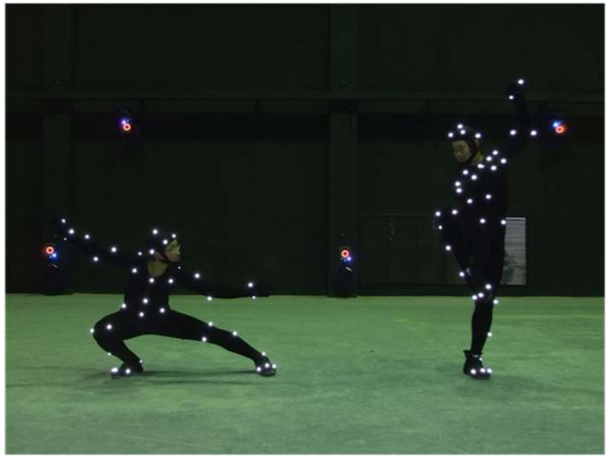


Motion Capture

- Sensing
  - Optical
  - Electro-magnetic
  - Inertial/accelerometer

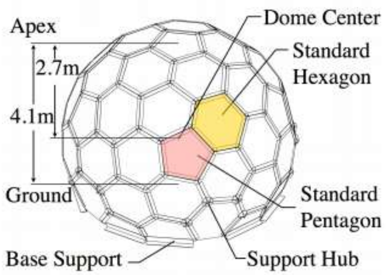
Motion Capture

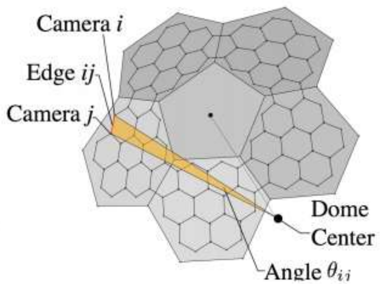
- Sensing
  - Optical (tracking or imaging)
    - Pros: more accurate, lightweight on the subject, small error accumulations
    - Cons: occlusions, errors caused by moving sensors, slow post-processing
  - Electro-magnetic
  - Inertia



Motion Capture

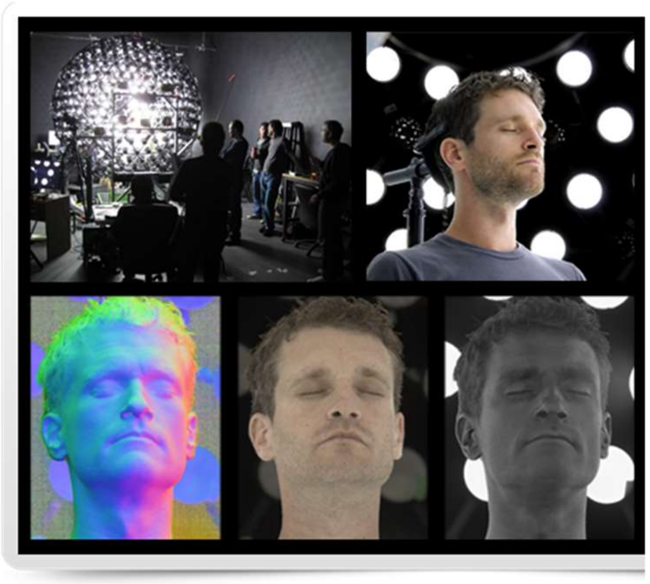
- Sensing
  - Optical (tracking or imaging)
    - Pros: more accurate, lightweight on the subject, small error accumulations
    - Cons: occlusions, errors caused by moving sensors, slow post-processing
  - Electro-magnetic
  - Inertia

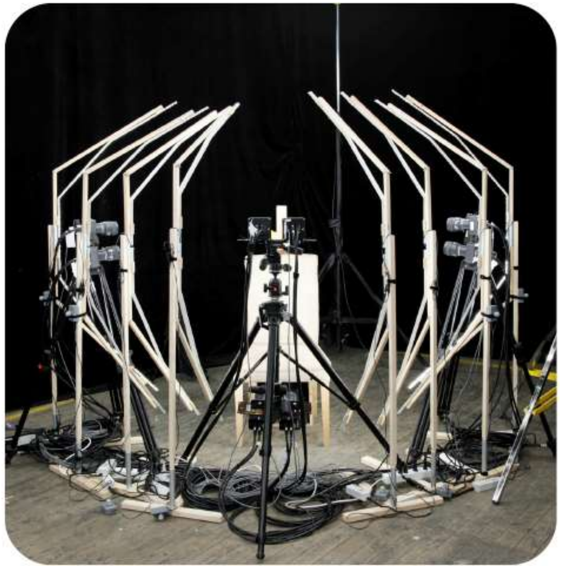





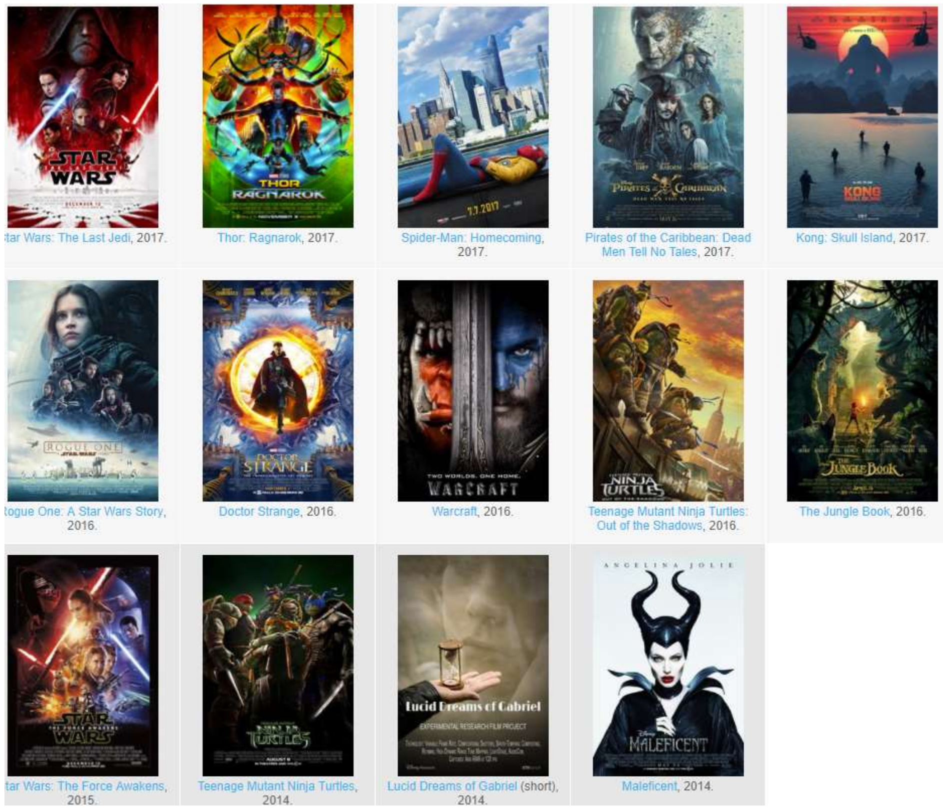

Motion Capture

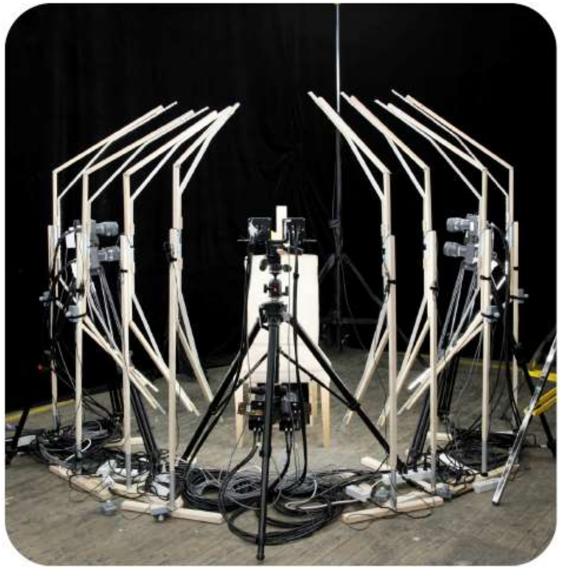
- Sensing
  - Optical (tracking or imaging)
    - Pros: more accurate, lightweight on the subject, small error accumulations
    - Cons: occlusions, errors caused by moving sensors, slow post-processing
  - Electro-magnetic
  - Inertia





Motion Capture



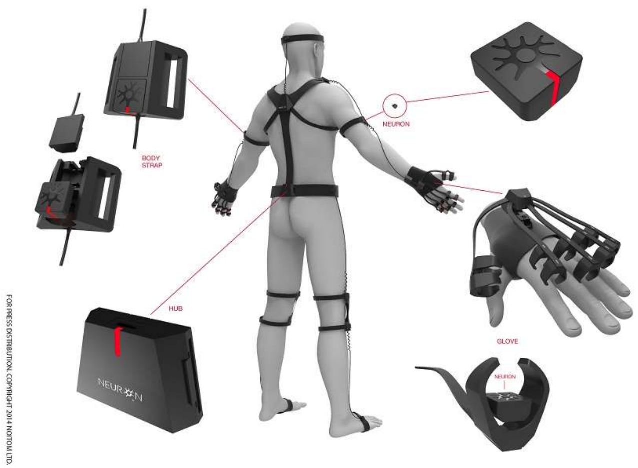


Motion Capture

- Sensing
  - Optical (tracking or imaging)
  - Electro-magnetic
    - Pros: no occlusion, faster
    - Cons: bigger errors in general, devoid of mag field distortions, cumbersome
  - Inertia

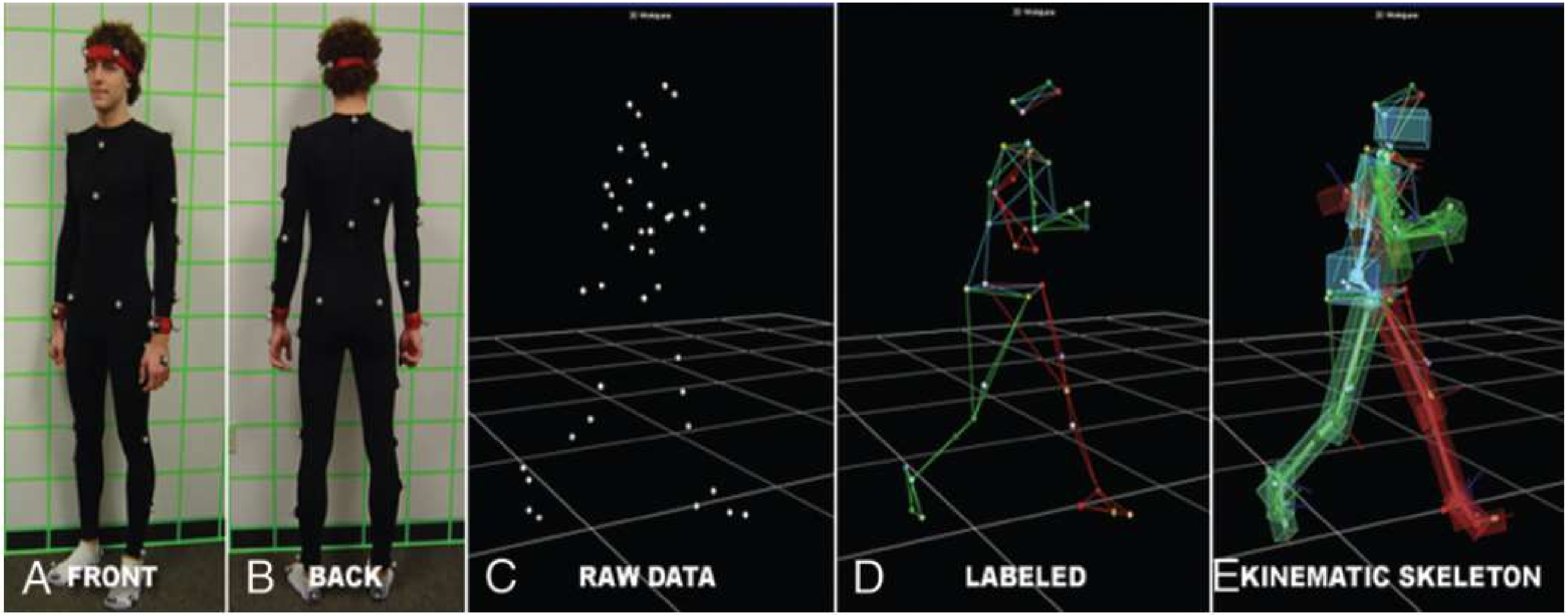
Motion Capture

- Sensing
  - Optical (tracking or imaging)
  - Electro-magnetic
  - Inertia
    - Pros: lightweight, fast
    - Cons: accumulative errors



Motion Capture

- Optical tracking – Processing images



Motion Capture

- Optical tracking – Processing images
  - Locate markers
  - Marker reconstruction
  - Constrained markers

Motion Capture

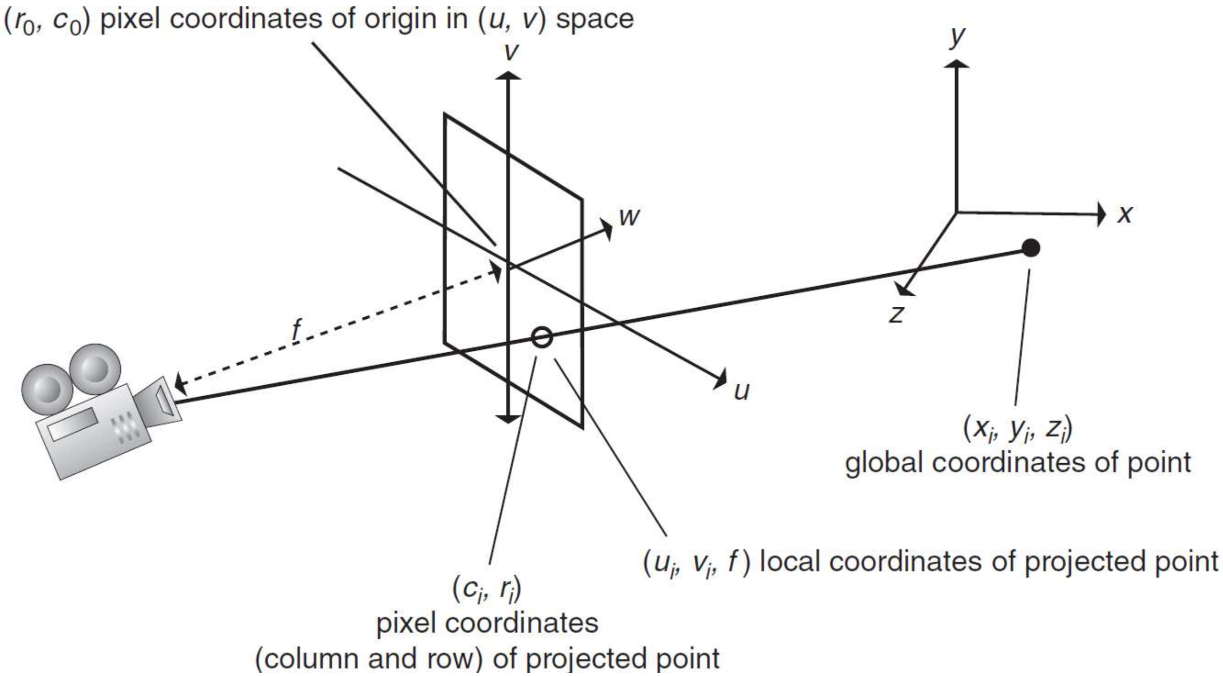
- Optical tracking – Processing images
  - Locate markers
    - Markers put near joints (Source of error: relative displacement; relative movements)
    - Static background, easy to identify one marker, harder when multiple are present
    - Track each marker through time.
      - Cross-frame coherency
      - Still very difficult in general
  - Marker reconstruction
  - Constrained markers

Motion Capture

- Camera Calibration
  - Pinhole method – good enough for computer graphics
    - One world coordinate, multiple cameras
    - Known camera focal lengths, image centre and aspect ratio
    - Need the positions and orientations of the cameras

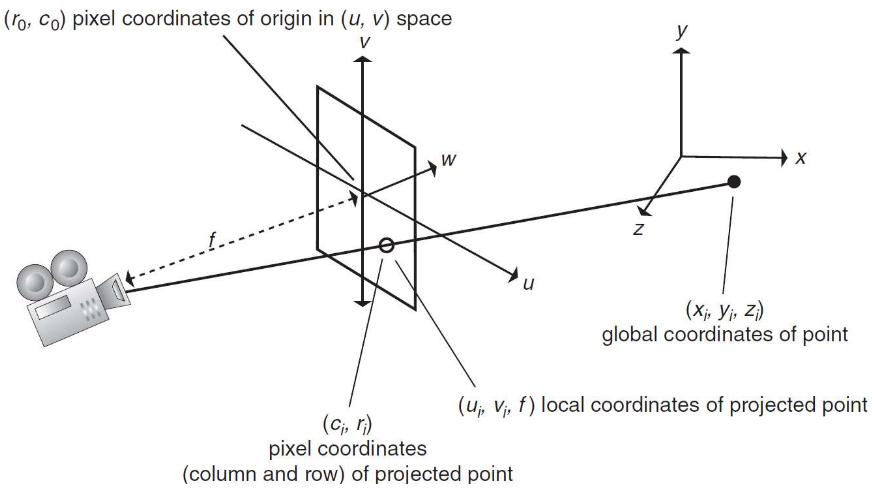
Motion Capture

- Camera Calibration
  - Pinhole method – good enough for computer graphics



Motion Capture

- Camera Calibration
  - Pinhole method – good enough for computer graphics



$$c_i - c_0 = s_u u_i$$
$$r - r_0 = s_v v_i$$

$$\begin{bmatrix} u_i \\ v_i \\ f \end{bmatrix} = R \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} + T$$

$$R = \begin{bmatrix} R_0 \\ R_1 \\ R_2 \end{bmatrix} = \begin{bmatrix} r_{0,0} & r_{0,1} & r_{0,2} \\ r_{1,0} & r_{1,1} & r_{1,2} \\ r_{2,0} & r_{2,1} & r_{2,2} \end{bmatrix}$$

$$T = \begin{bmatrix} t_0 \\ t_1 \\ t_2 \end{bmatrix}$$

Motion Capture

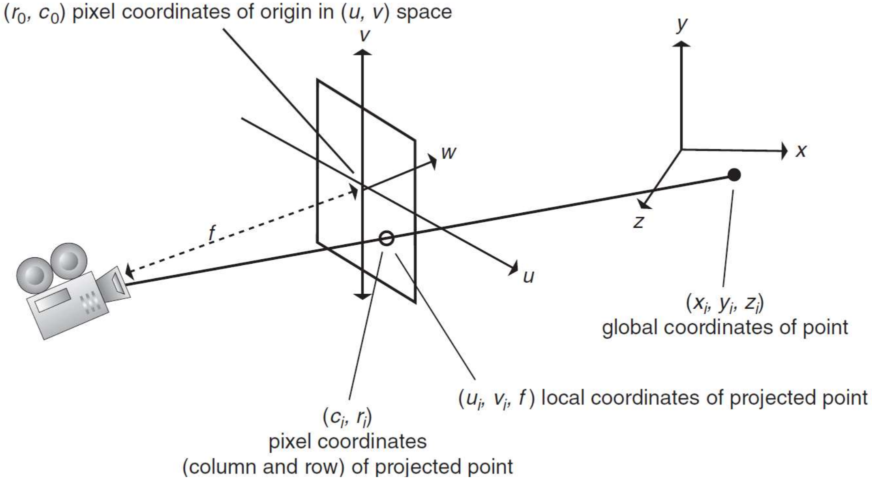
$$c_i - c_0 = s_u u_i$$
$$r - r_0 = s_v v_i$$

$$R = \begin{bmatrix} R_0 \\ R_1 \\ R_2 \end{bmatrix} = \begin{bmatrix} r_{0,0} & r_{0,1} & r_{0,2} \\ r_{1,0} & r_{1,1} & r_{1,2} \\ r_{2,0} & r_{2,1} & r_{2,2} \end{bmatrix}$$

$$\begin{bmatrix} u_i \\ v_i \\ f \end{bmatrix} = R \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} + T \qquad T = \begin{bmatrix} t_0 \\ t_1 \\ t_2 \end{bmatrix}$$

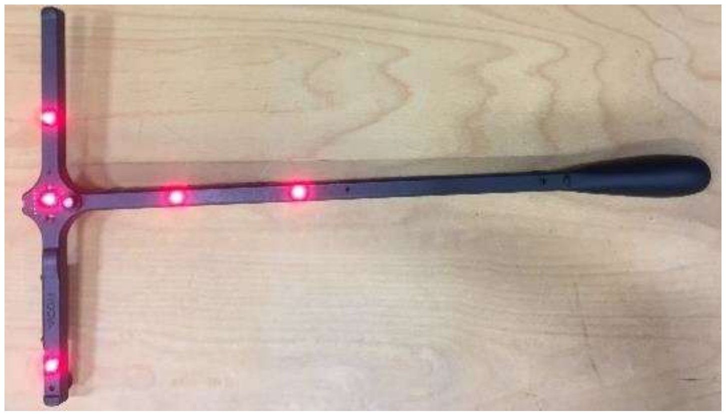
$$\frac{u_i}{f} = \frac{c_i - c_0}{s_u f} = \frac{c_i - c_0}{f_u} = \frac{R_0 \cdot [x_i y_i z_i] + t_0}{R_2 \cdot [x_i y_i z_i] + t_2}$$
$$\frac{v_i}{f} = \frac{r_i - r_0}{s_v f} = \frac{c_i - c_0}{f v} = \frac{R_1 \cdot [x_i y_i z_i] + t_1}{R_2 \cdot [x_i y_i z_i] + t_2}$$

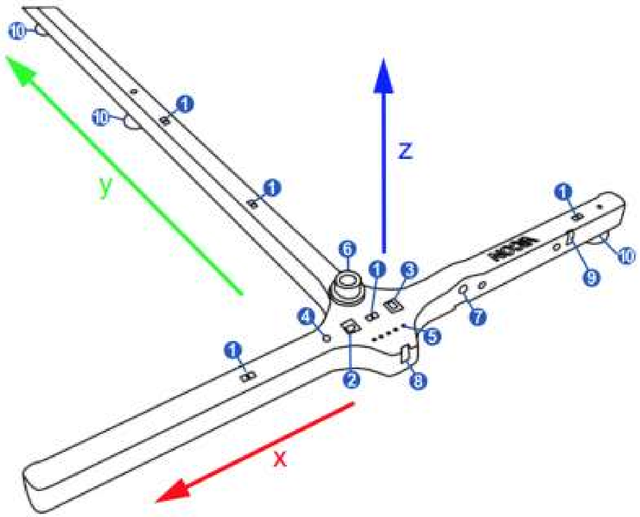
$$(c_i - c_0)(R_2 \cdot [x_i y_i z_i] + t_2) - f_u(R_0 \cdot [x_i y_i z_i] + t_0) = 0$$
$$(r_i - r_0)(R_2 \cdot [x_i y_i z_i] + t_2) - f_v(R_1 \cdot [x_i y_i z_i] + t_1) = 0$$
$$AW = 0$$



Motion Capture

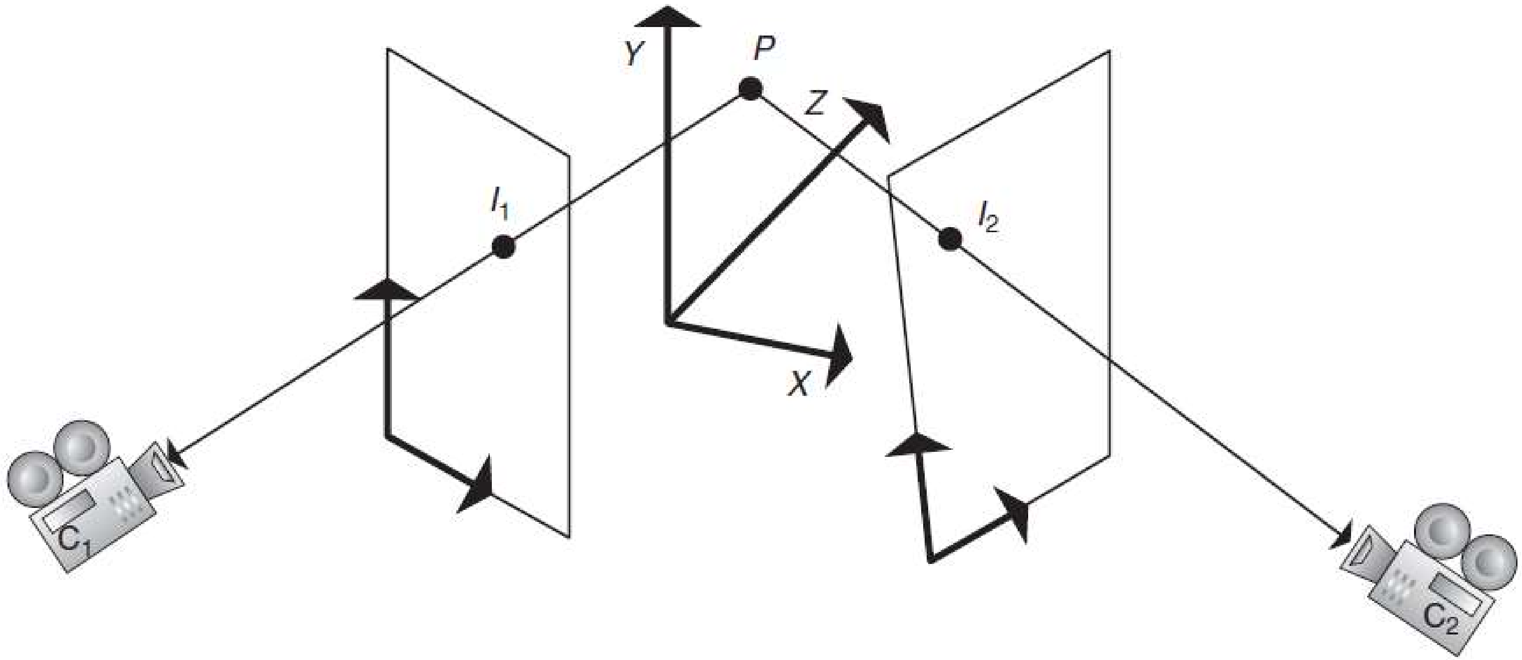
- Camera Calibration
  - Pinhole method – good enough for computer graphics
    - Record points with known world coordinates
    - Solve a least-square problem





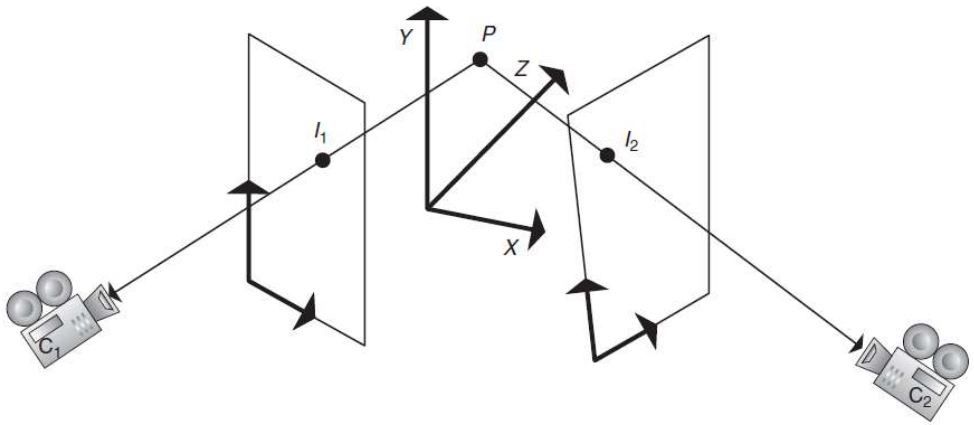
Motion Capture

- 3D position reconstruction



Motion Capture

- 3D position reconstruction



$$C_1 + k_1(I_1 - C_1) = P$$

$$C_2 + k_2(I_2 - C_2) = P$$

$$C_1 + k_1(I_1 - C_1) = C_2 + k_2(I_2 - C_2)$$

Three equations, two unknowns

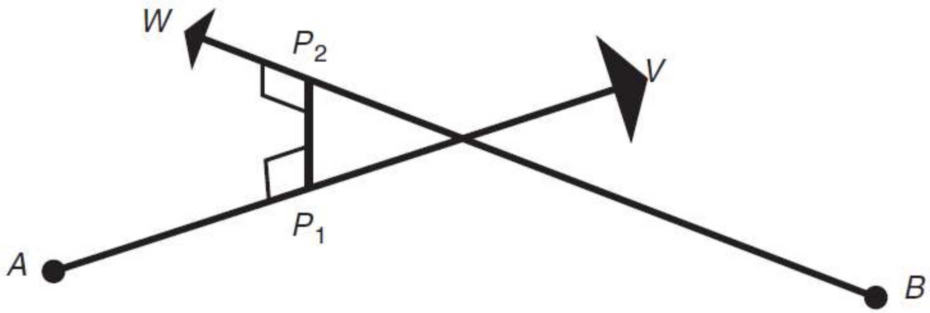
Motion Capture

- 3D position reconstruction

In reality, they never intersect, find the closest point

$$C_1 + k_1(I_1 - C_1) = P$$

$$C_2 + k_2(I_2 - C_2) = P$$



$$P_1 = A + sV$$
$$P_2 = B + tW$$

$$(P_2 - P_1) \cdot (I_1 - C_1) = 0$$

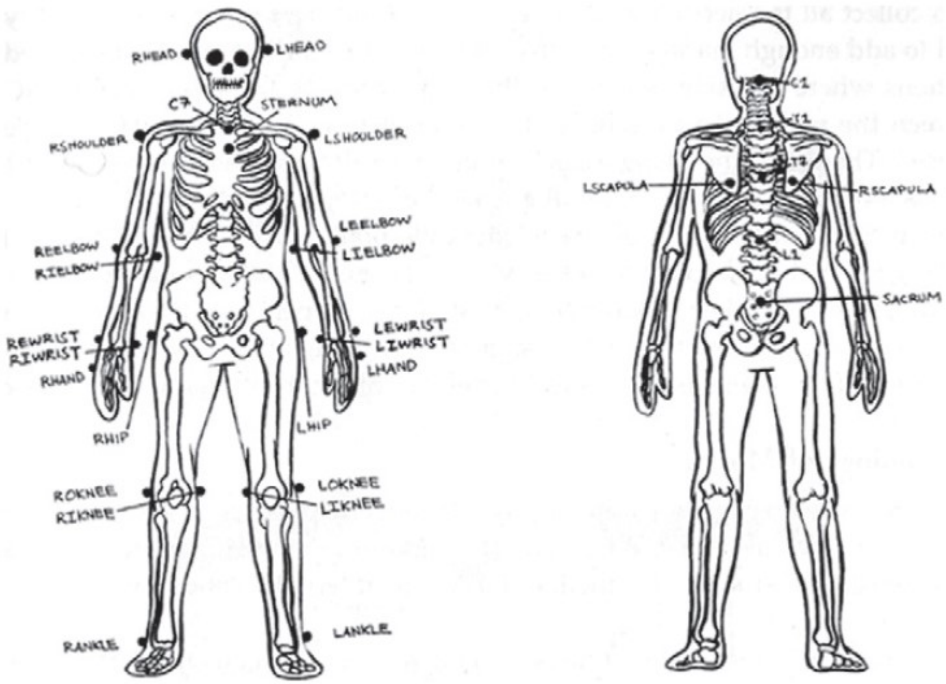
$$(P_2 - P_1) \cdot (I_2 - C_2) = 0$$

$$C_1 + k_1(I_1 - C_1) = C_2 + k_2(I_2 - C_2)$$

Three equations, two unknowns
In the ideal world

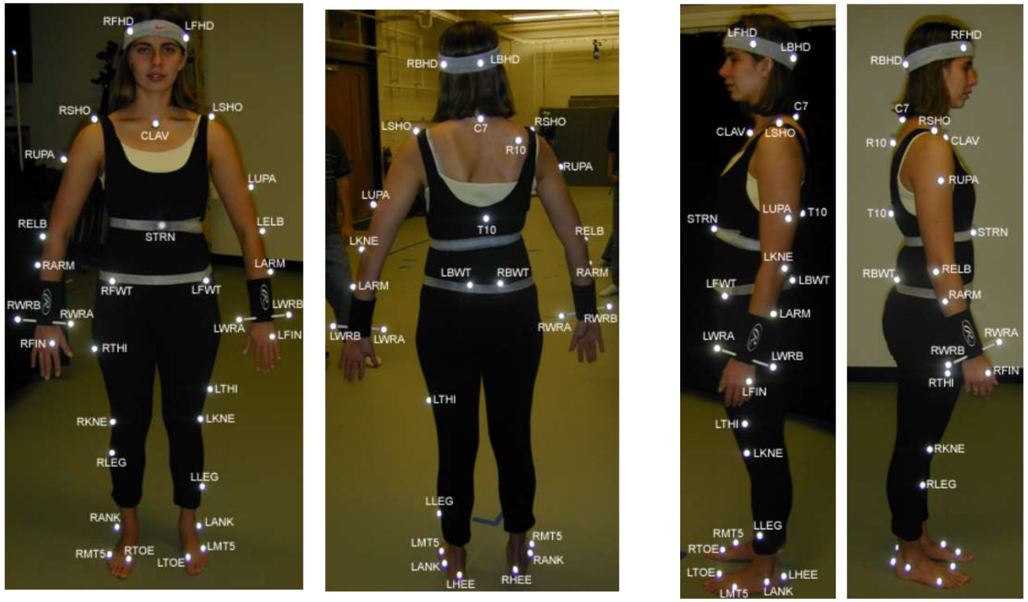
Motion Capture

- Fitting into skeleton



Motion Capture

- Fitting into skeleton
  - Not ideal to use markers' positions
    - Not on joints
    - Relative movements (bone-length not kept)
    - Foot-sliding



Motion Capture

- Fitting into skeleton
  - Not ideal to use markers' positions
  - Using markers to compute joint positions (not accurate)
  - Putting two markers on both sides of a joint and use the mid-point
    - Hips? Shoulders?
  - Three markers to form a plane
    - One on elbow, two on wrist, then the wrist could be computed
    - Not always work, wrist-elbow-shoulder, degenerates when collinear

Motion Capture

- Fitting in
  - Not i
  - Using
  - Putti
    - H
  - Three
    - C
    - N

Motion Capture

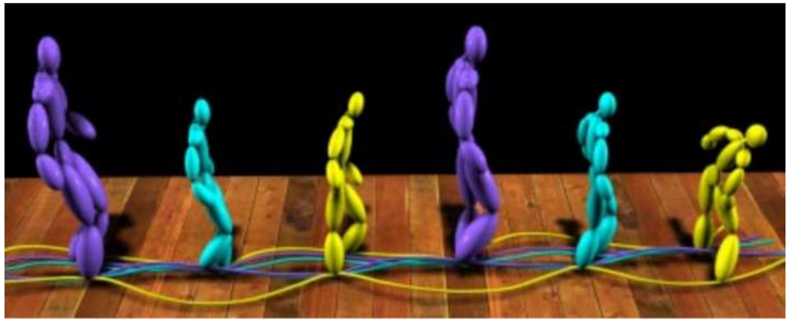
- Fitting into skeleton
  - Not ideal to use markers' positions
  - Using markers to compute joint positions (not accurate)
  - Putting two markers on both sides of a joint and use the mid-point
    - Hips? Shoulders?
  - Three markers to form a plane
    - One on elbow, two on wrist, then the wrist could be computed
    - Not always work, wrist-elbow-shoulder, degenerates when collinear
  - IK, with constraints (foot planting)

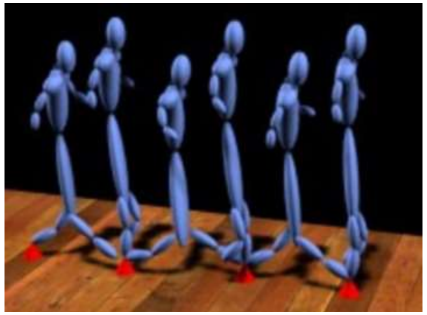
Motion Capture

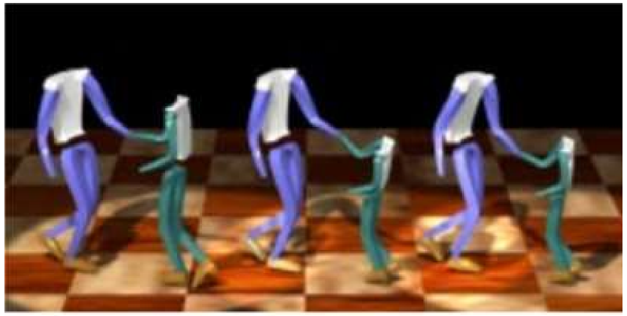
- Manipulating mocap data
  - Error-prone (Assuming you are not super rich)
    - Recapture: Could be very expensive, e.g. hire Angelina Jolie again
    - Post-processing
  - Processing the signals (Assuming you are NOT super rich)
    - Signal processing
      - Each joint trajectory (angle or position) can be seen as a time-series data
      - e.g. low frequency vs high frequency
      - Smoothing
    - Motion Warping
      - To satisfy user-defined key-frame like constraints
      - Postural, positional, velocity constraints, etc.

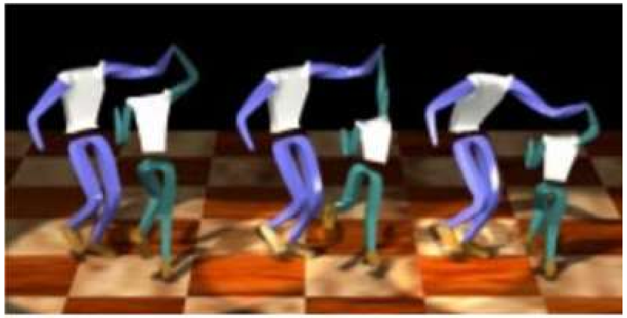
Motion Capture

- Manipulating mocap data (Assuming you are NOT super rich)
  - Retargeting the motion
    - Different models (body proportions, Gleicher 98 Siggraph)









Motion Capture

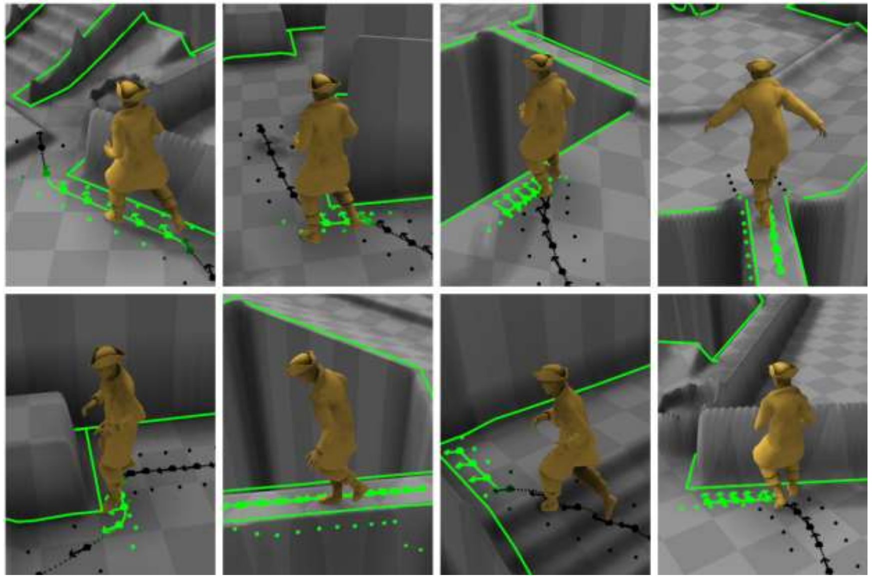
- Manipulating mocap data (Assuming you are not super rich)
  - Retargeting the motion
    - Different models (body proportions)
    - Different environments

(a) Original motion over smooth terrain.

(a) Retargetted motion over rough terrain.

Motion Capture

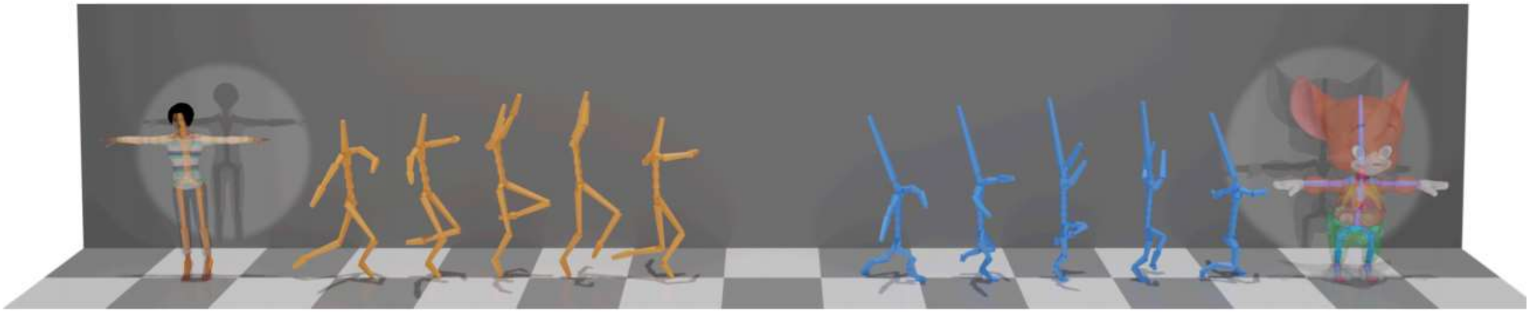
- Manipulating mocap data (Assuming you are not super rich)
  - Retargeting the motion
    - Different models (body proportions)
    - Different environments



Holden et al. "Phase-Functioned Neural Networks for Character Control“ SIGGRAPH 2017

Motion Capture

- Manipulating mocap data (Assuming you are not super rich)
  - Retargeting the motion
    - Different models (body proportions)
    - Different species!
    - Different Morphology!



ABERMAN et al. Skeleton-Aware Networks for Deep Motion Retargeting, SIGGRAPH 2020

Motion Capture

- Manipulating mocap data (Assuming you are not super rich)
  - Retargeting the motion
    - Different environments

STARKE et al. Local Motion Phases for Learning Multi-Contact Character Movements, SIGGRAPH 2020