MOTION CAPTURE TECHNOLOGIES

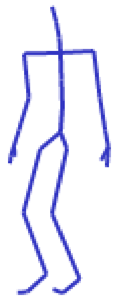
24. 3. 2015

DISA Seminar

Jakub Valcik

Outline

- Human Motion and Digitalization
- Motion Capture Devices
  - Optical
  - Ranging Sensor
  - Inertial and Magnetic
- Comparison



Human Motion

- Brain + Skeleton + Muscles

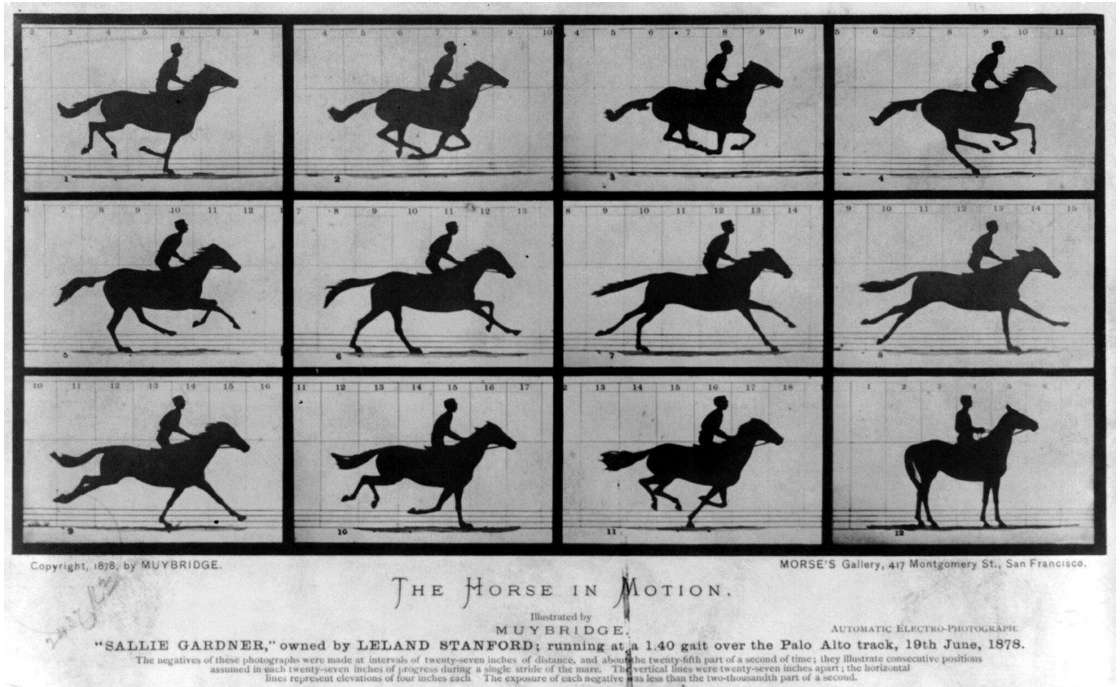
Internal factors	External Factors
Musculature	Shoes
Skeleton	Clothes
Weight	Type of surface
Injuries	Slope of surface
Movement habits	Wind
Pregnancy	Gravity
State of mind	Environment

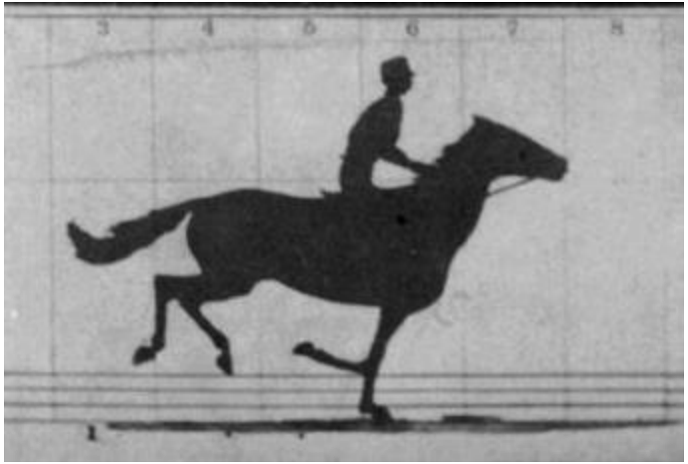
Motion Capture

- Sequence of individual frames $m = (f_k)_{k=1}^{n}$
- Captured information
  - Static body configuration
  - Position in space
  - Orientation of body
- Silhouette vs. Skeleton

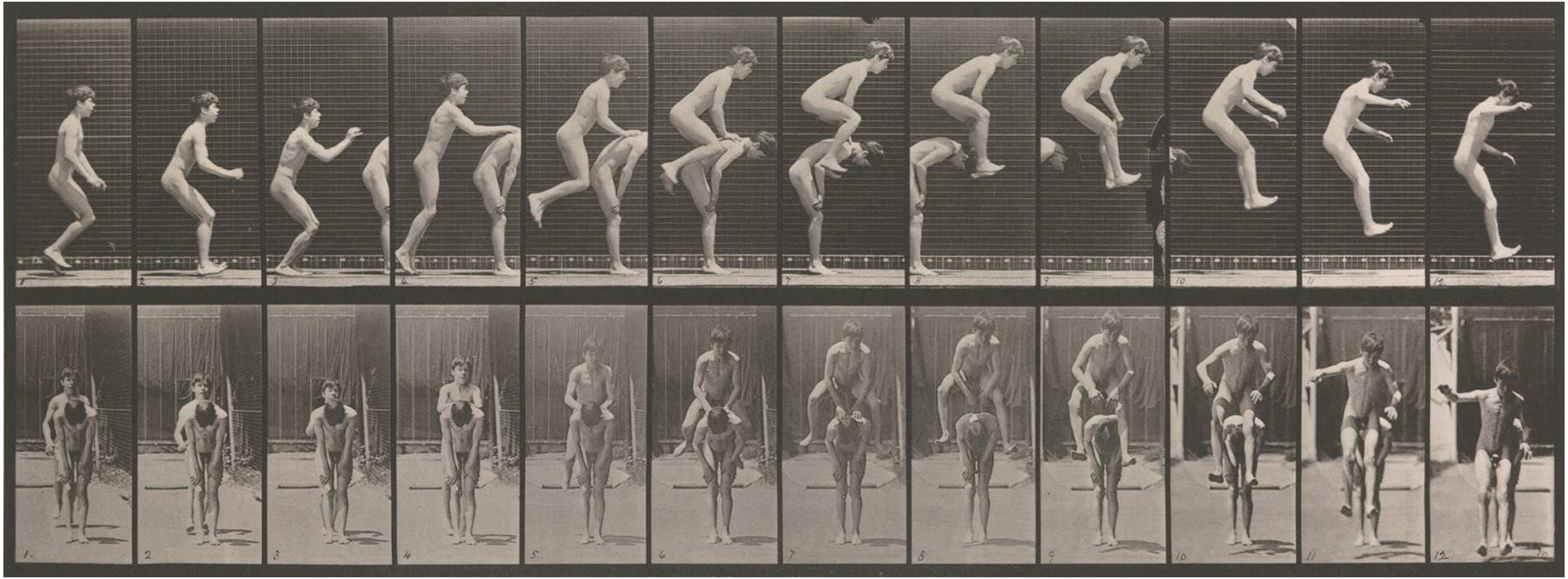
Appearance Based Approaches

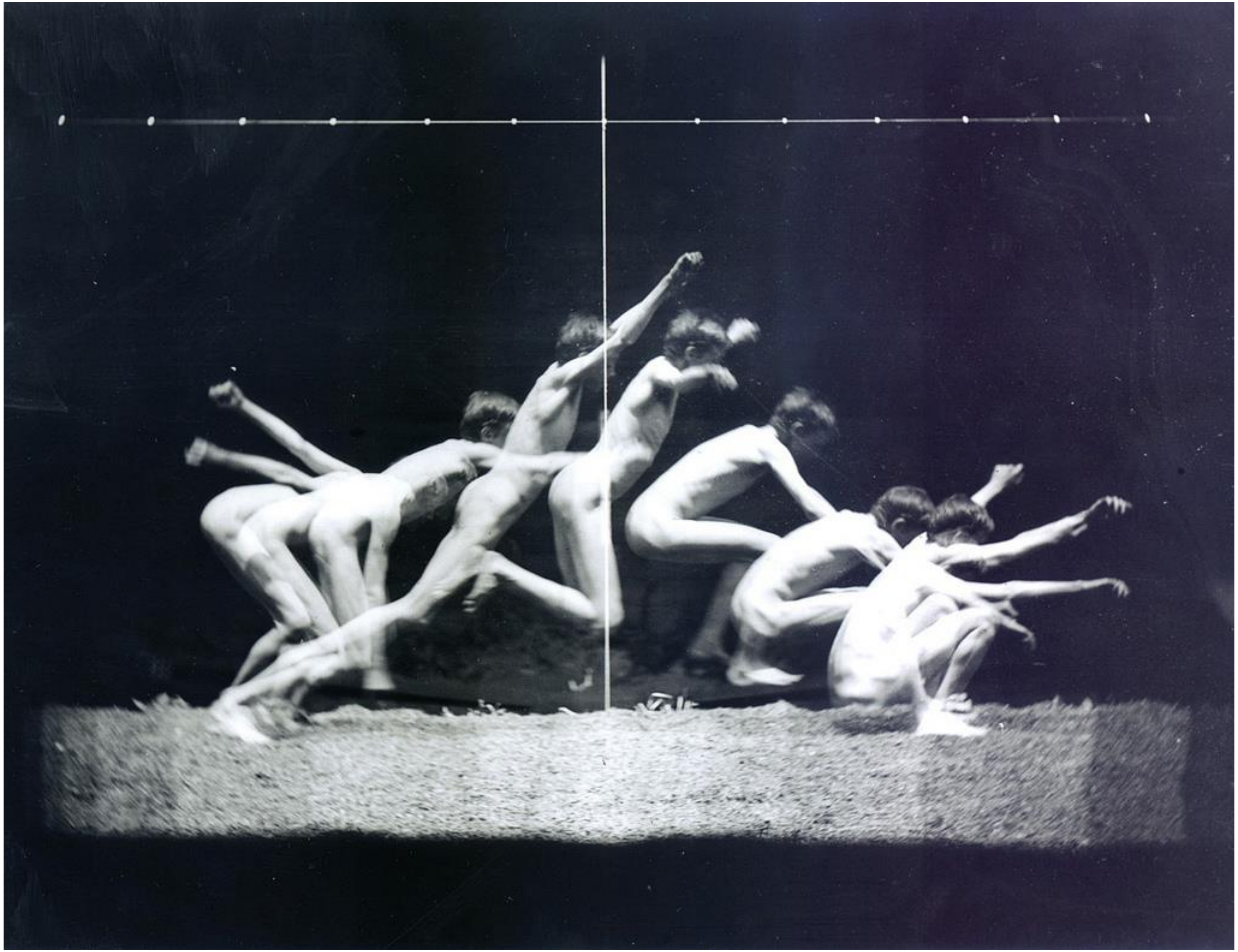
- Eadweard Muybridge, 1878





Boys Playing Leapfrog, printed 1887





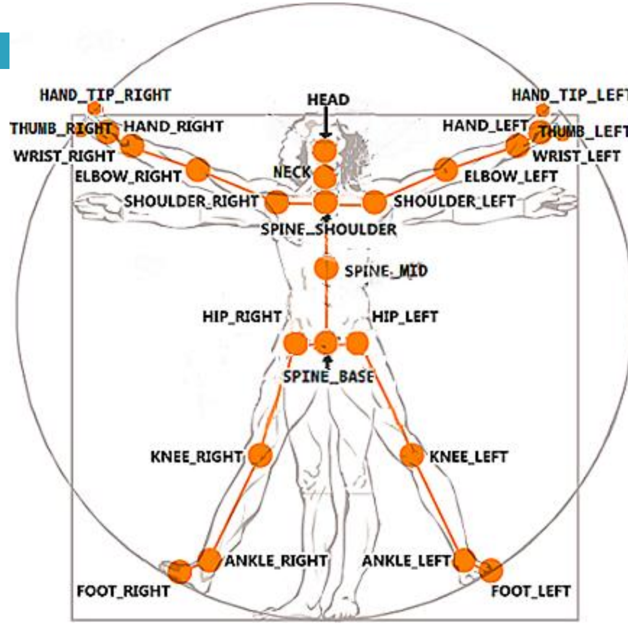
Appearance Base Approaches

- Originally only regular video cameras, CCTV
- Silhouette oriented $f_k := B_k(x, y) \in \{0, 1\}$

Silhouette extraction is a common bottleneck

Model Based Approaches

- Additional abstraction – 2D, 3D
  - Stick figure, volumetric model, …
- Skeleton
  - Joint (or end-effector), Bone
  - Undirected acyclic graph J – tree
    - Joint ~ Vertex, Bone ~ Edge
  - Static configuration ~ Pose $p \in \mathbb{R}^{3\times|V(J)|}$
  - $m = (f_k)_{k=1}^{n} = (p_k)_{k=1}^{n}$



Motion Capture Devices

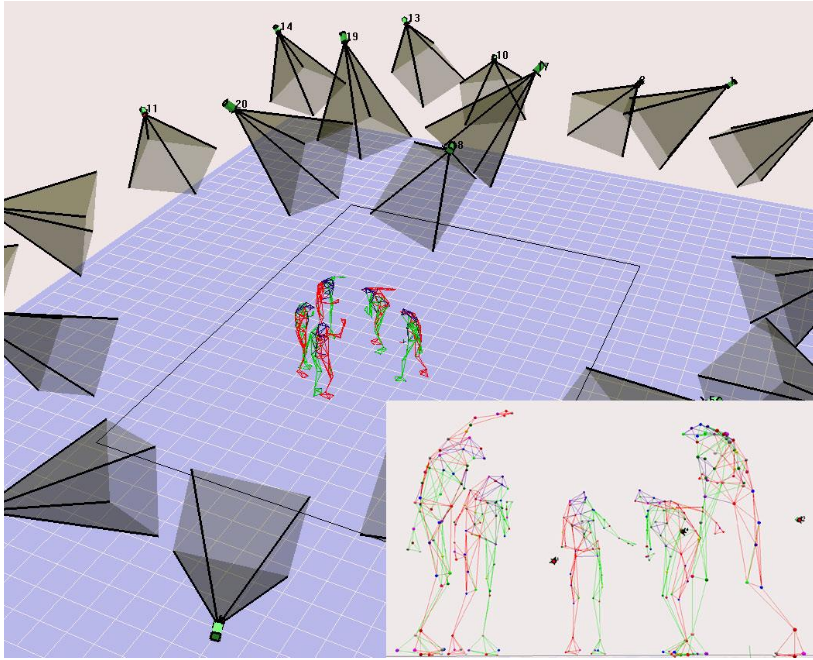
- Optical
  - Markerless
  - Invasive
- Inertial
- Magnetic
- Mechanical
- Radio frequency

Both, appearance and model based approaches

Only model based approaches

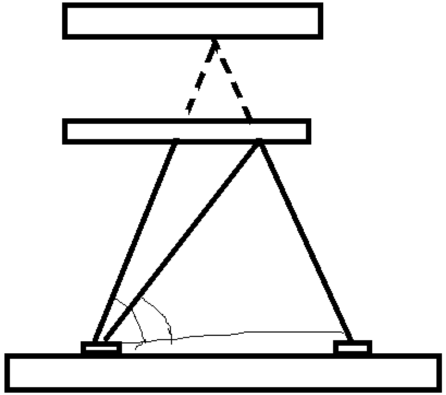
Optical Markerless MoCap Devices

- 3D scene reconstruction
  - Multiple views
  - Depth sensors
- RGB stereoscopic cameras
- Multiple synchronized cameras
- Additional sensors
  - Silhouette extraction
  - Depth sensing
  - IR camera, ranging sensor



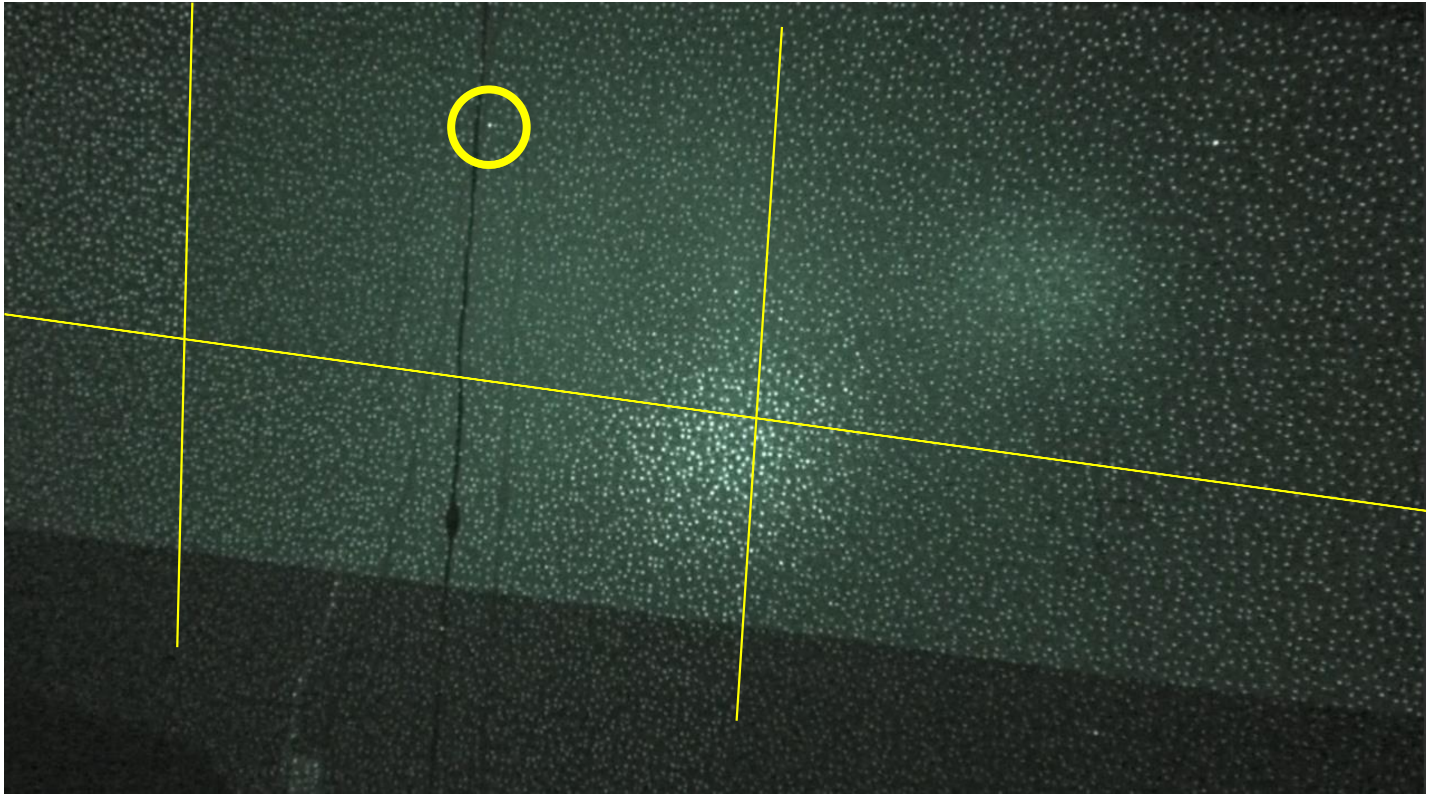
Ranging Sensor - Triangulation

- Laser beam + IR camera
- Projection grid – 'structured light'
- Known variables
  - Emitter-camera distance
  - Dot distribution in grid
- Depth ~ dot translation
- Grid resolution ↑↓ Object distance
- PrimeSense (now Apple)
  - Project Natal => Kinect

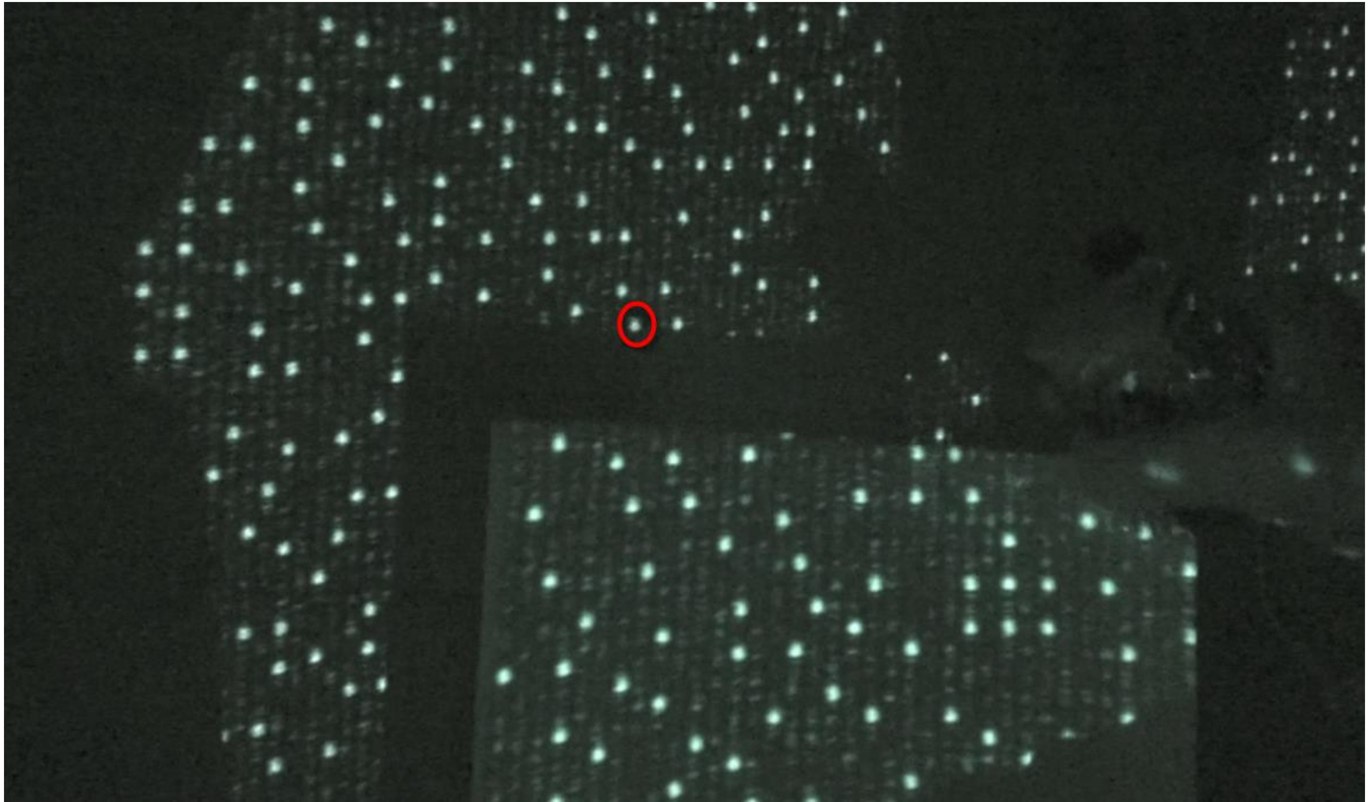


Kinect v1 IR Structured Light

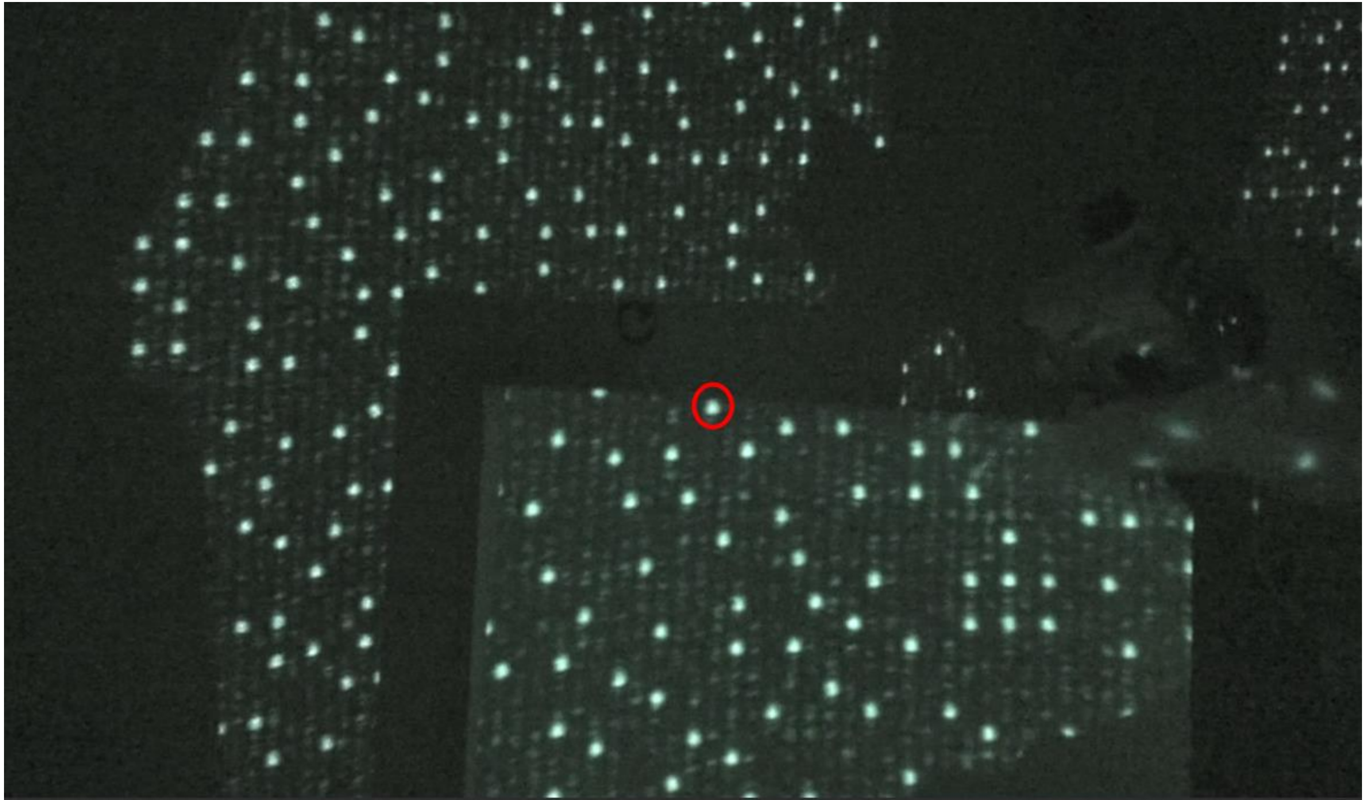
Kinect v1 – Reference Points



Kinect v1 – Point Shift (1)



Kinect v1 – Point Shift (2)



Ranging Sensor – Time of Flight

- Light speed $c \doteq 300\,000\ {}^{km}/_{s} = 30\ {}^{cm}/_{ns}$
- Principles
  - RF modulated – phase shift
    - Swiss Ranger 4000
    - PMD CamCube 3.0
    - Canesta Vision (now MS)
  - Range gated
    - Zcam by 3DV (now MS)
    - TriDiCam
  - Direct ToF – 3D flash LIDAR
    - Advanced Scientific Concepts, Inc.

Ranging Sensor Comparison

Depth Sensor	Maximal Range	Resolution	Field of View [°]	Repeatability [mm](1 Sigma)
Kinect v1	10m	640x480	57.8 x 43.3	7.6@2m, 27.5@4m
Swiss Ranger 4000	8m	176x144	43 x 34 / 69 x 56	4/6
PMD CamCube 3.0	7m	200x200	40 x 40	3@4m

Camera Types Comparison

- Interference comparison

Camera Type	Complex Background	Heat Source	Other Camera	Clothes
RGB (B/W)	○○	●●	●●	○○
IR	●○	○○	●●	●●
Ranging	●●	●○	○○	●○

Optical Markerless MoCap Devices

- Stereoscopic video cameras
  - Sony Playstation Eye
- Video camera + ranging sensor – triangulation
  - MS Kinect v1, Asus Xtion live, Structure Sensor, PrimeSense Carmine 1.08
- Video camera + ranging sensor – ToF
  - MS Kinect v2
- 360° video cameras
  - Organic Motion

PrimeSens Carmine

Asus Xtion Live

Structured Sensor

Microsoft Kinect v1

Microsoft Kinect v2

Price $300

Fun fact #1: Body part estimation based on ML,
Learning phase take 24,000 CPU hours

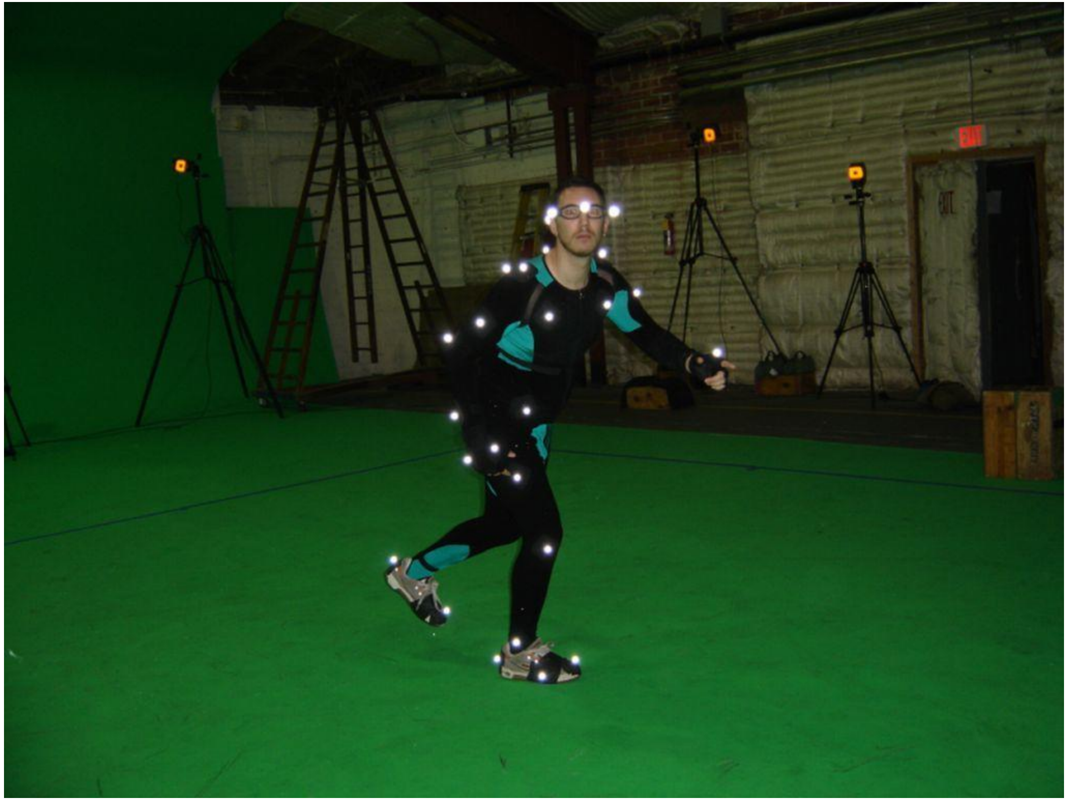
Organic Motion Openstage2

Starting price $40,000

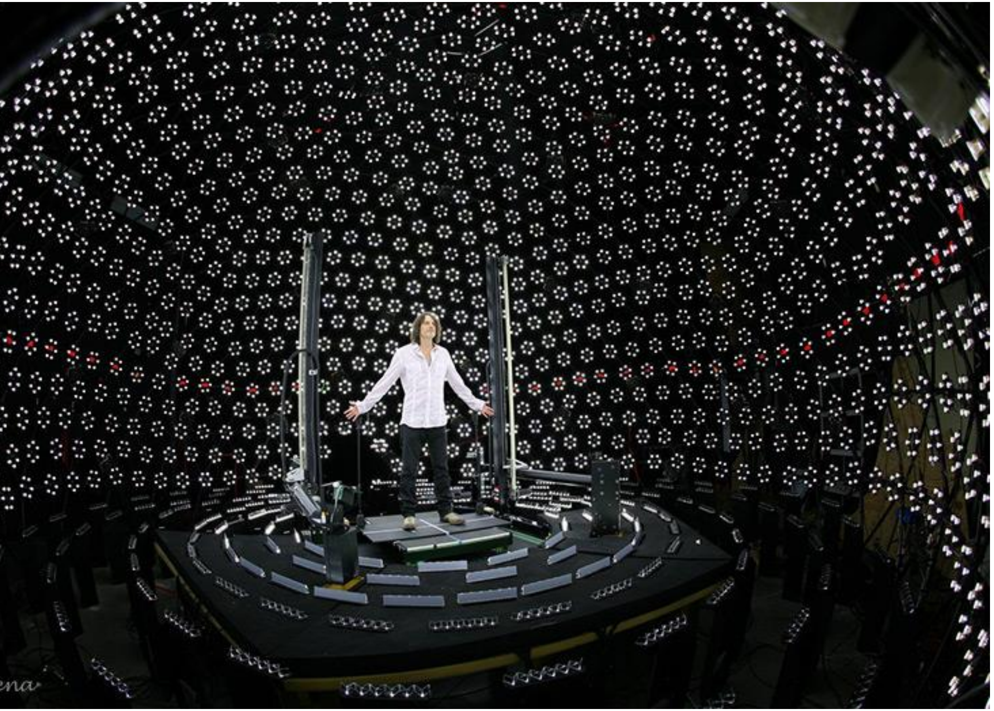
Invasive Optical MoCap Devices

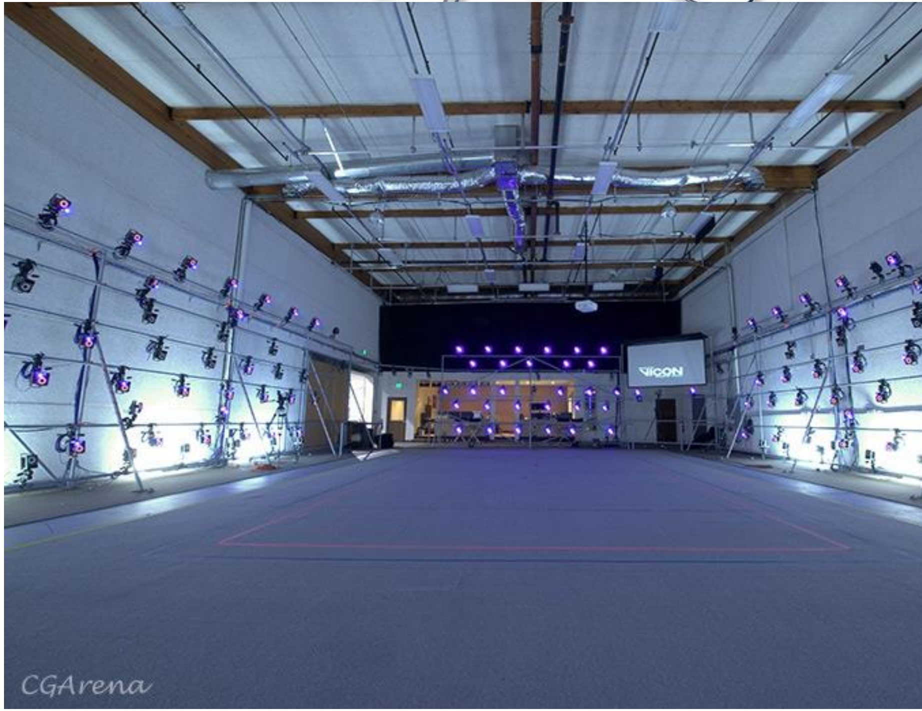
- Active vs. Passive
- Multiple RGB & IR cameras
- Precise, fast
- Price, markers, additional electric source (active)
- Problems
  - Marker swapping

MoCap Suite



Vicon MX





Starting price $100,000 ?

Inertial & Magnetic MoCap Devices

- Acceleration, magnetic flux
- No global position nor orientation
- Pose initialization
- Accumulation of error
- Gravity, wiring, reinforced concrete

MVN XSens

Feature	MVN Awinda	MVN Link
Trackers	17 Wireless	17 Wired
Latency	30ms	20ms
Wireless range	20m/50m	50m/150m
Output rate	60Hz	240Hz

Starting price from $12,000

Other MoCap Devices

- Mechanical system
  - Exoskeleton directly measures joint angle rotations
- Radio Frequency Positioning
  - RADAR working on high frequencies >50GHz
  - Inaccurate, large areas (hundreds of meters^2)

MoCap Devices Comparison

Sensor name	Range	Resolution	Framerate	Sensor Type	Invasive	Field of View
Kinect v1	0.8-4m	640x480	30Hz	RGB, IR-triangulation	-	57°
Kinect v2	0.5-4.5m	512x424	30Hz	RGB(1080p), ToF, IR camera	-	70°
Asus Xtion Pro Live	0.8-3.5m	640x480	30Hz	RGB, IR-triangulation	-	58°
PrimeSense Carmine 1.08	0.8-3.5m	640x480	30Hz	RGB, IR-triangulation	-	57.5°
Structure Sensor	0.4-3.5	640x480	30Hz	RGB, IR-triangulation	-	58°
Playstation Eye Camera	30cm-inf	1280x800	120Hz	2RGB	-	85°
Vicon MX40+	space 7x7m	2352x1728	120Hz	8-40 cameras (IR & RGB combination)	markers	360°
Xsens MVN	-	0.05 deg	120Hz	17 Inertial and magnetic sensors	on body sensors	-
Organic Motion BioStage	space 4.3x3.8m	640x480	120Hz	12-24 cameras (B/W)	-	360°

MoCap Software Comparison

Sensor	Software	Tracked Subjects	Morphology Stability	Positional Accuracy [mm]	Rotational Accuracy [°]	Landmarks
Kinect v1	Kinect SDK v1.*	2	No	$50 - 150$	?	20
Kinect v1	OpenNi+Nite 1.3	2	No	$50 - 150$	?	15
Kinect v2	Kinect SDK v2	6	No	?	$1 - 3$	25
Vicon MX40	Vicon-Workstation V4.6/142	?	Yes	0.063	?	32 - variable
Xsens MVN	MVN Human model	1	Yes	-	$0.5 - 1$	22
Organic Motion	OpenStage2	5	Yes	1	$1 - 2$	22

Joint Tracking Error
Kinect v1

Joint	System	Walk	360	Hide	Box	Occ	Sit
Head	MSSDK	15.8	17.2	16.1	13.2	**32.5**	14.8
	NITE	**10.6**	**11.8**	**13.3**	**12.2**	76.2	**11.0**
Neck	MSSDK	11.2	14.5	10.9	**8.5**	**31.8**	7.7
	NITE	**4.6**	**4.9**	**3.2**	10.5	76.6	**5.7**
Torso	MSSDK	**4.4**	**5.9**	**3.9**	**10.1**	**30.7**	**7.5**
	NITE	6.7	6.7	6.8	15.5	82.0	12.1
Shoulders	MSSDK	7.8	**16.8**	**5.8**	9.3	**34.6**	**7.9**
	NITE	**5.6**	18.6	7.1	**8.7**	82.4	9.2
Elbows	MSSDK	9.6	**28.6**	7.6	**6.4**	**42.5**	**8.7**
	NITE	**9.0**	32.0	**7.4**	9.1	78.5	11.1
Hands	MSSDK	14.8	**47.3**	15.6	**12.2**	**52.9**	**14.1**
	NITE	14.8	50.2	**11.0**	15.9	84.7	14.2

Cosgun, A., Bünger, M., & Christensen, H. I. (2013). *Accuracy Analysis of Skeleton Trackers for Safety in HRI* (p. 1).

Percentage of Tracked Frames Kinect v1

Joint	System	Walk	360	Hide	Box	Occ	Sit
Head	MSSDK	100	100	100	100	66	100
	NITE	91	99	93	82	83	68
Neck	MSSDK	100	100	100	100	67	100
	NITE	100	100	100	100	83	99
Torso	MSSDK	100	100	100	100	67	100
	NITE	100	100	100	100	83	99
Shoulders	MSSDK	100	73	100	100	60	99
	NITE	100	100	100	100	83	99
Elbows	MSSDK	100	75	100	96	57	74
	NITE	99	90	100	71	59	95
Hands	MSSDK	100	76	58	48	47	86
	NITE	99	88	100	70	58	94

Cosgun, A., Bünger, M., & Christensen, H. I. (2013). *Accuracy Analysis of Skeleton Trackers for Safety in HRI* (p. 1).

Thank you for your attention

Resources

- http://nongenre.blogspot.cz/2010/12/how-kinect-senses-depth.html
- http://www.freepatentsonline.com/20100118123.pdf
- http://users.dickinson.edu/~jmac/selected-talks/kinect.pdf
- http://www.nimbocg.com.br/wp-content/uploads/2013/02/mocap11.jpg
- Other sources cited in thesis