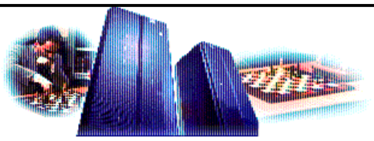
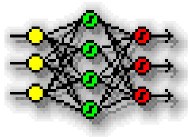
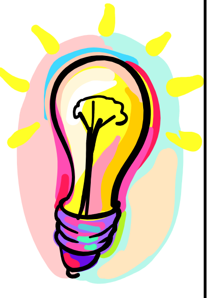




Introduction to Virtual Reality
Parte I

Prof. Alberto Borghese
Applied Intelligent Systems Laboratory (AIS-Lab)
Department of Computer Science
University of Milano



Obbiettivo del corso

- Fornire i fondamenti per capire cosa succede dentro ad un sistema di Realtà Virtuale (trasformazioni, proiezioni, animazione di scheletri).
- Esperienza pratica estesa in laboratorio con i dispositivi di VR di utilizzo corrente (Oculus-rift, Hololens, Google card, Kinect, Leap, MoCap,...).
- Modalità d'esame: progetto + discussione teoria
  - Il progetto può essere associato a altri corsi e/o alla tesi.
  - La valutazione della parte di teoria [pass / fail
  - La valutazione della parte di laboratorio sarò in 30esimi e sarà composta dalla valutazione di micro-progetti in itinere (1/3 della valutazione) e del progetto finale (2/3 della valutazione).

Realtà Virtuale – 6 CFU

Sito principale:
http://borghese.di.unimi.it/Teaching/VR/VR.html

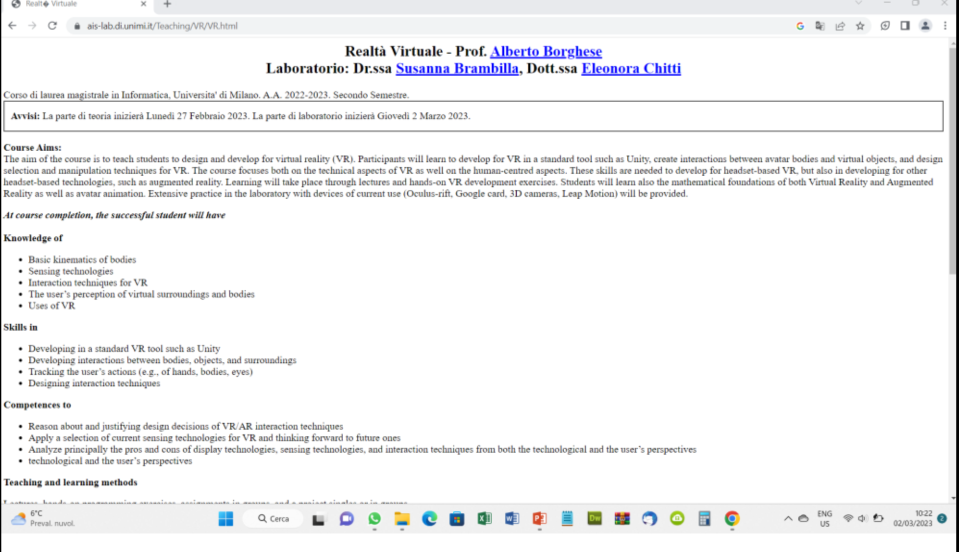
Programma:
http://borghese.di.unimi.it/Teaching/VR/Programma_2022-2023.html

Let's try to keep the course interactive

Orario:
Lunedì Ore 08.30-10.30 – ab. LM, 3o piano, via Celoria 18 - Teoria
Giovedì Ore 9.30-12.30 – ab. LM, 3o piano, via Celoria 18 - Laboratorio

Strumento principale di contatto: email (alberto.borghese@unimi.it)
Ricevimento su appuntamento

Home page del corso


Realt◆ Virtuale
ais-lab.di.unimi.it/Teaching/VR/VR.html

Realtà Virtuale - Prof. Alberto Borghese
Laboratorio: Dr.ssa Susanna Brambilla, Dott.ssa Eleonora Chitti

Corso di laurea magistrale in Informatica, Universita' di Milano. A.A. 2022-2023. Secondo Semestre.

Avvisi: La parte di teoria inizierà Lunedì 27 Febbraio 2023. La parte di laboratorio inizierà Giovedì 2 Marzo 2023.

Course Aims:
The aim of the course is to teach students to design and develop for virtual reality (VR). Participants will learn to develop for VR in a standard tool such as Unity, create interactions between avatar bodies and virtual objects, and design selection and manipulation techniques for VR. The course focuses both on the technical aspects of VR as well on the human-centred aspects. These skills are needed to develop for headset-based VR, but also in developing for other headset-based technologies, such as augmented reality. Learning will take place through lectures and hands-on VR development exercises. Students will learn also the mathematical foundations of both Virtual Reality and Augmented Reality as well as avatar animation. Extensive practice in the laboratory with devices of current use (Oculus-rift, Google card, 3D cameras, Leap Motion) will be provided.

At course completion, the successful student will have

Knowledge of

- Basic kinematics of bodies
- Sensing technologies
- Interaction techniques for VR
- The user's perception of virtual surroundings and bodies
- Uses of VR

Skills in

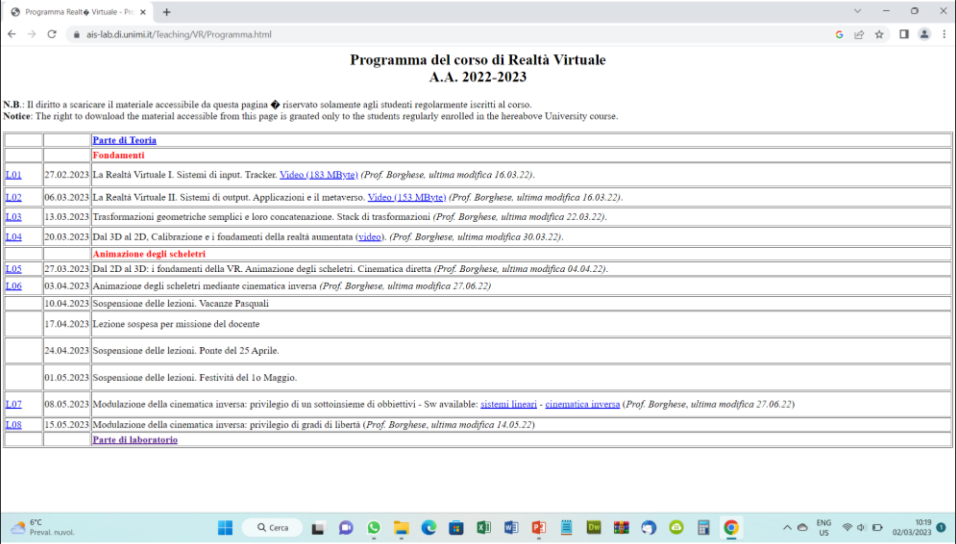
- Developing in a standard VR tool such as Unity
- Developing interactions between bodies, objects, and surroundings
- Tracking the user's actions (e.g., of hands, bodies, eyes)
- Designing interaction techniques

Competences to

- Reason about and justifying design decisions of VR/AR interaction techniques
- Apply a selection of current sensing technologies for VR and thinking forward to future ones
- Analyze principally the pros and cons of display technologies, sensing technologies, and interaction techniques from both the technological and the user's perspectives
- technological and the user's perspectives

Teaching and learning methods

Programma del corso



Programma del corso di Realtà Virtuale
A.A. 2022-2023

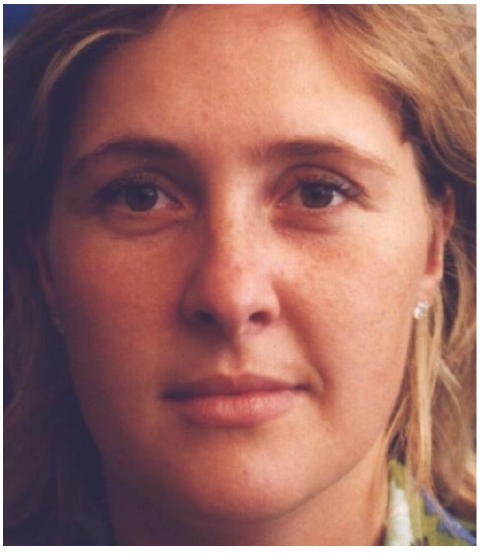
N.B.: Il diritto a scaricare il materiale accessibile da questa pagina � riservato solamente agli studenti regolarmente iscritti al corso.
Notice: The right to download the material accessible from this page is granted only to the students regularly enrolled in the hereabove University course.

		Parte di Teoria
		Fondamenti
L01	27.02.2023	La Realtà Virtuale I. Sistemi di input. Tracker. Video (183 MByte) *(Prof. Borghese, ultima modifica 16.03.22).*
L02	06.03.2023	La Realtà Virtuale II. Sistemi di output. Applicazioni e il metaverso. Video (153 MByte) *(Prof. Borghese, ultima modifica 16.03.22).*
L03	13.03.2023	Trasformazioni geometriche semplici e loro concatenazione. Stack di trasformazioni *(Prof. Borghese, ultima modifica 22.03.22).*
L04	20.03.2023	Dal 3D al 2D, Calibrazione e i fondamenti della realtà aumentata (video). *(Prof. Borghese, ultima modifica 30.03.22).*
		Animazione degli scheletri
L05	27.03.2023	Dal 2D al 3D: i fondamenti della VR. Animazione degli scheletri. Cinematica diretta *(Prof. Borghese, ultima modifica 04.04.22).*
L06	03.04.2023	Animazione degli scheletri mediante cinematica inversa *(Prof. Borghese, ultima modifica 27.06.22)*
	10.04.2023	Sospensione delle lezioni. Vacanze Pasquali
	17.04.2023	Lezione sospesa per missione del docente
	24.04.2023	Sospensione delle lezioni. Ponte del 25 Aprile.
	01.05.2023	Sospensione delle lezioni. Festività del 1o Maggio.
L07	08.05.2023	Modulazione della cinematica inversa: privilegio di un sottoinsieme di obbiettivi - Sw available: sistemi lineari - cinematica inversa *(Prof. Borghese, ultima modifica 27.06.22)*
L08	15.05.2023	Modulazione della cinematica inversa: privilegio di gradi di libertà *(Prof. Borghese, ultima modifica 14.05.22)*
		Parte di laboratorio

Sommario

- Introduzione
- Sistemi di Input
- Generatori di mondi
- Motore di calcolo
- Sistemi di Output
- Conclusioni

Which is real, which is virtual?



Historical Perspective (I)

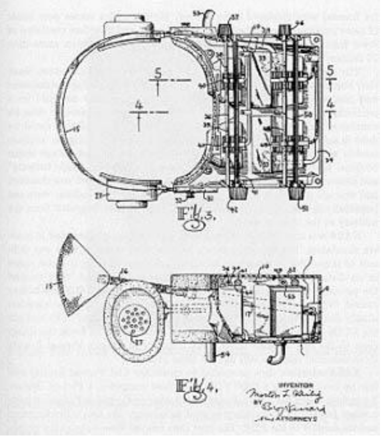
- *The name "Virtual Reality" has been attributed to Jaron Lanier (VPL), 1986.*
- *Virtual Worlds or Synthetic Environments*
- *Philosophical* and *Technological* origin.

Philosophical background

Ontology and *Gnoseology.*

- Plato (world of the ideas) 428-348 a.C.
- Berkeley (sensorial experience is too limited) 1685-1753.
- Hegel ("what is rational is real..") 1770-1831.
- New age.

Historical Perspective (II)



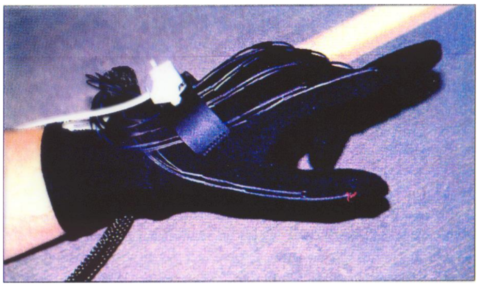
Morton Heilig 1956,
patented in 1961
Non fu mai costruito

projected film,
audio, vibration,
wind, odors.

Historical Perspective (III)

Technological background

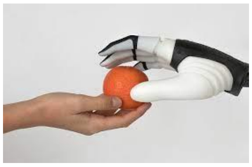
- *Philco HMD, 1961.*
- *"Ultimate display", Sutherland, 1970.*
- *Data Glove, VPL Research, 1988.*



Sutherland, Ivan E. 1968. "A Head-Mounted Three Dimensional Display," pp. 757-764 in Proceedings of the Fall Joint Computer Conference. AFIPS Press, Montvale, N.J.

Virtual Reality Systems

Key characteristics are:
Immersivity.
Interactivity.




VR should be able to stimulate the human sensorial systems
In a coordinated way.

VR output should be able to saturate our sensor systems, congruently.

A typical VR system

VR systems are constituted of:

• *Input systems* (measure the position *in* the environment and force *over* the environment.

• *World generators* (provides a realistic virtual world in which to act. It is a graphical engine).

• *Computational engine* (computes the output, given the input and the virtual world).

• *Output systems* (outputs sensorial stimuli *on* the subject. Vision, sound, force … are generated as if they were provided *by* the virtual environment.

Metaverso

Dispositivi estremamente eterogenei

Nuovi dispositivi sul mercato

E’ possibile definire una inter-operabilita’?

In robotica la risposta è arrivata da ROS

E nella VR? METAVERSO (Neal Stephenson in Snow Crash – 1992).
VR supportata da Internet -> third life?

Nel 2021 Meta Platforms Inc. assume diecimila persone in Europa per creare il metaverso

Facebook cambia il nome in «meta»

Coderblock ha terminato la seconda crowdfunding costruire il metaverso (italiano)

«Internet del 2020»

Sommario

- Introduzione
- Sistemi di Input (trackers)
- Generatori di mondi
- Motore di calcolo
- Sistemi di Output
- Conclusioni

Input systems

Measure human actions on the virtual environment.

•Position. Measure the position of the body segments inside the virtual environment.

• Force. Measure the force exerted by the body segments when in contact with a virtual object.

• Estimate the motor output of the human muscle-skeleton system.

Tracking systems

•Measure the position of the body segments inside the virtual environment.

• Motion capture (batch, complete information on the movement).
• Real-time trackers (real-time position of the body).
• Gloves (specialized for hands).
• Gaze trackers.

<u>Adopted technology</u>

• Optoelectronics (video-camera based)
  •Marker based
  •Computer vision
  •Scanner based.
• Magnetical
• Acoustical
• Mechanical
• Intertial

What is motion capture?

Ensemble of techniques and methodologies to acquire **automatically** the motion of the objects of interest.

Characteristics: sampling rate, accuracy, 2D/3D, real-time, motion amplitude, invasivity,….

Technology: opto-electronical, magnetical, ultrasound, intertial ….

Specific body parts: gloves, gaze trackers….

Applications are increasing (medical applications at the origin, now interest in the enterteinment, robotics, reverse engineering …)

Motion Capture and Synthesis

Reproduce digitally the motion of the body (in real-time in case of tracker).

Time series of the position of the body segments
or
Time series of the motion of the articulations.

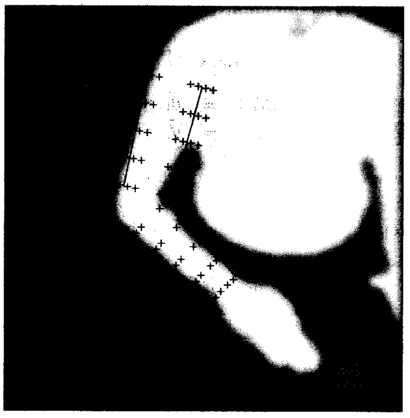
Analysis
Info extraction

Application of the time series to a 3D digital model of the body.

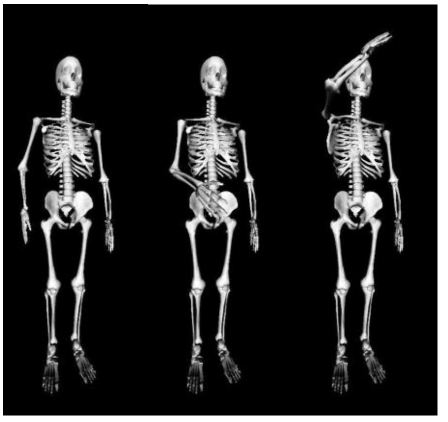
Synthesis
Avatar animation

What is captured?

Silhouette (-> Skeleton)



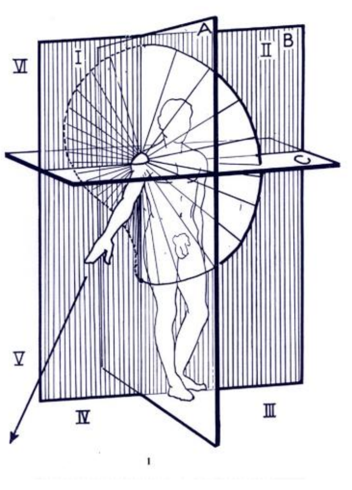
Skeleton



Computer vision techniques
(silhouette, RGB-D cam)

Bony segments or articulations
(marker-based systems, RGB-D cam)

Description of the human skeleton



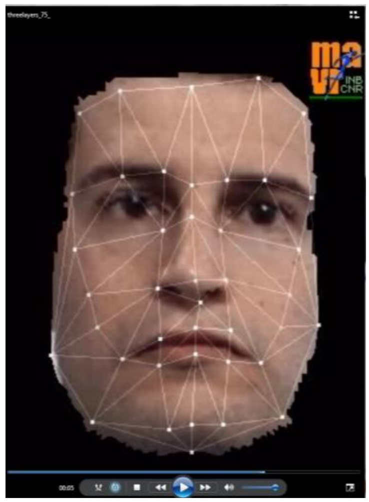
A – Frontal plane
B – Sagittal plane
C – Horizontal plane

Abduction/adduction
Flexion/extension
Axial rotation (V)
Quaternions for 3D rotations

3D position of joint extremes

Definition of the interesting degrees of freedom.

Motion capture through passive markers



https://www.youtube.com/watch?v=uPn26JbRN4g&list=PLxtdgDam3USWUXO7eliIFlg4WJMhJpLUp&index=16

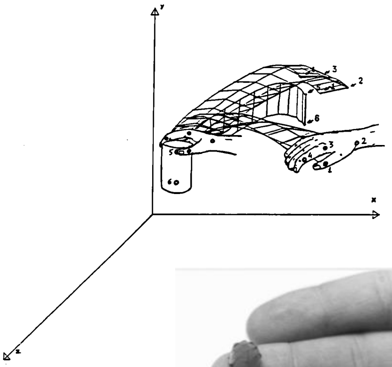
Facial animation is still a difficult task

Emotions provoke very small muscles tension
Muscles tension produces changes on the face:

- Very very small changes
- Concentrated (wrinkles)

Why passive markers?

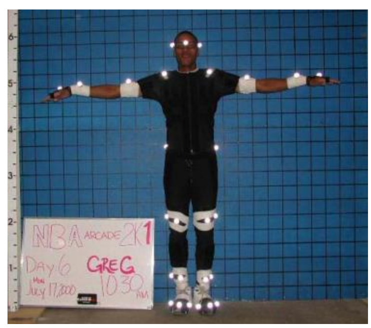
Minimum encoumbrance on the subject: markers do not require any powering and are hardly sensed by the subjects.



No constraint on the dimension of the working volume is prescribed.

How passive markers work?

Passive markers are constituted of a small plastic support covered with retro-reflecting material (3M™). It marks a certain repere point.



Video-cameras are equipped with a co-axial flash.

Markers appear much brighter than the background making their detection, on the video images, easier.

Tracking difficulties

It is a complex problem because:

• Dense set of markers. These may come very close one to the other in certain instants.

• Motion can be easily complex, as it involves rotation and twists of the different body parts (thing at a gymnastic movement).

•Multi-camera information and temporal information is required to achieve a robust tracking.

Tracking difficulties

It is a complex problem because:

• Dense set of markers. These may come very close one to the other in certain instants.

• Motion can be easily complex, as it involves rotation and twists of the different body parts (thing at a gymnastic movement).

•Multi-camera information and temporal information is required to achieve a robust tracking.

Sequential processing

1. Surveying the image of the moving subject on multiple cameras (*frequency & set-up*).
2. Markers extraction from the background scene (*accuracy & reliability*).
3. Computation of the “real” 2D position of the markers (*accuracy <- distortion*).

Low-level Vision

4. **Matching on multiple cameras.**
5. 3D Reconstruction (*accuracy*).

High-level Vision

6. **Model fitting (*labelling, classification*).**

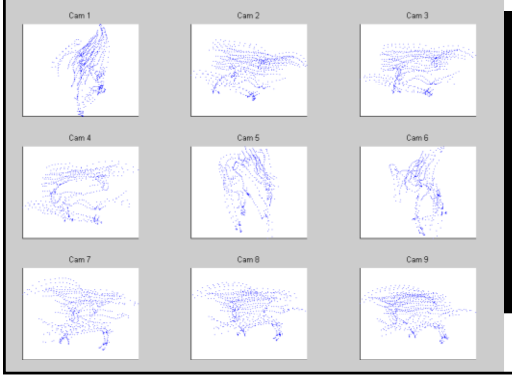
Semantic

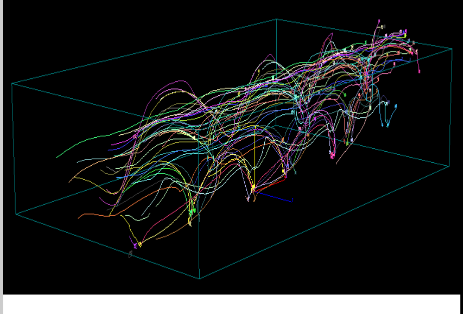
An implicit step is CALIBRATION.

Disadvantages of motion capture systems based on passive markers

When a marker is hidden to the cameras by another body part (e.g. the arm which swings over the hip during gait), the motion capture looses track of it.

The multiple set of 2D data have to be correctly labaled and associated to their corresponding 3D markers.





Tracking difficulties

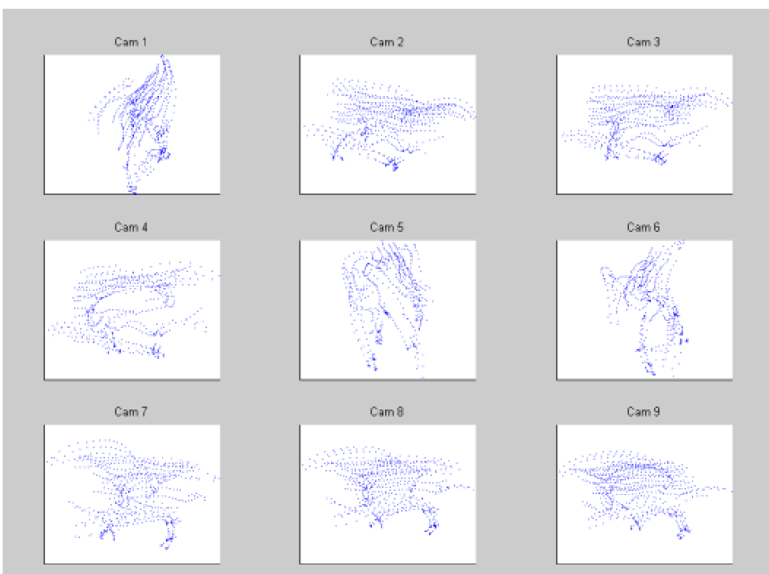
It is a complex problem because:

• Dense set of markers. These may come very close one to the other in certain instants.

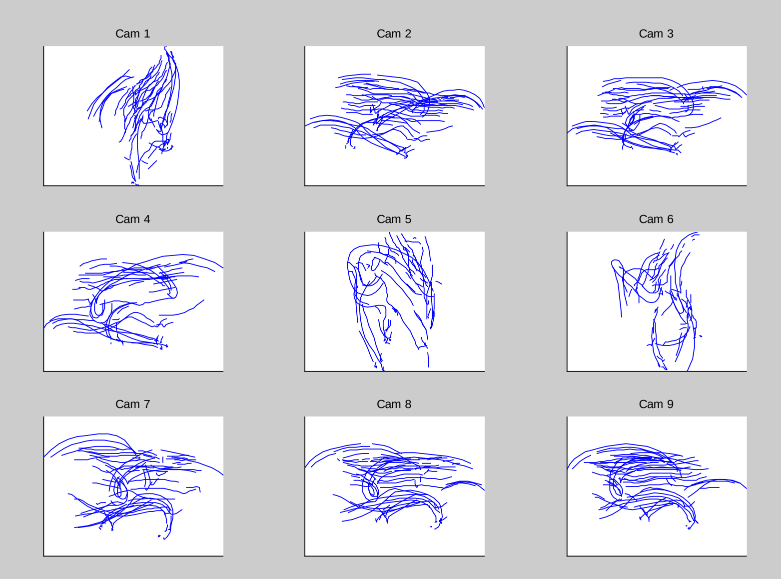
• Motion can be easily complex, as it involves rotation and twists of the different body parts (thing at a gymnastic movement).

•Multi-camera information and temporal information is required to achieve a robust tracking.

2D tracking

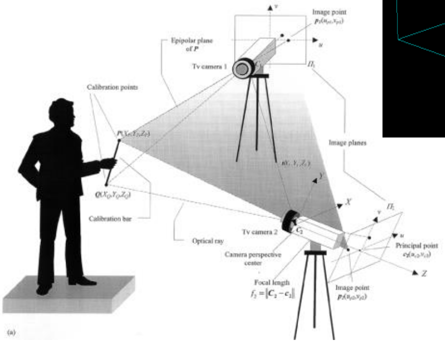


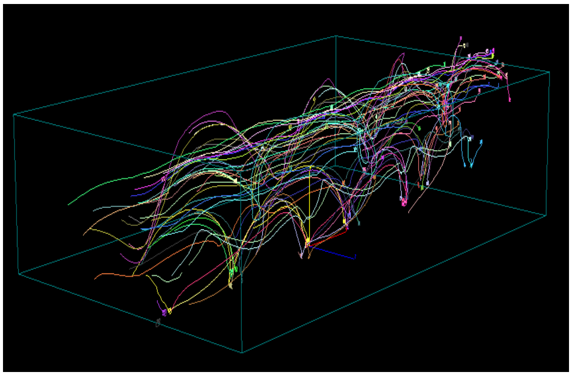
1) Creation of 2D strings



2) Matching 2D strings

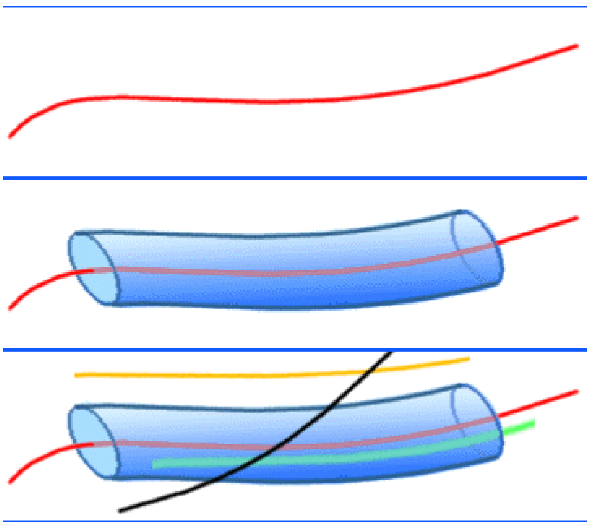
Epipolarity constraint



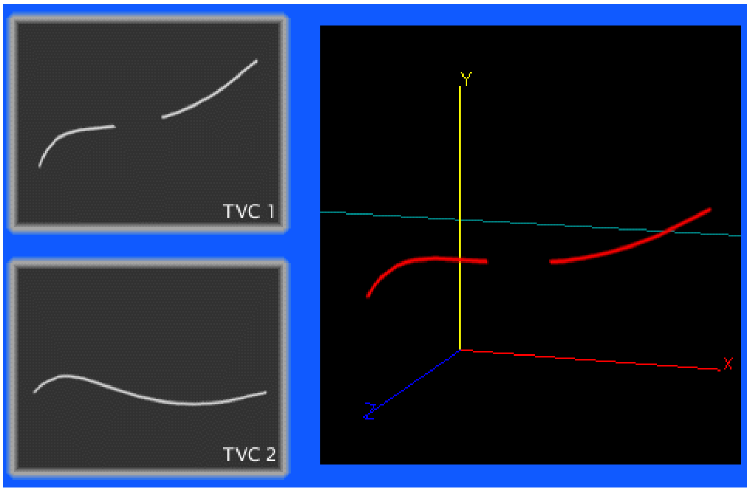


3D strings

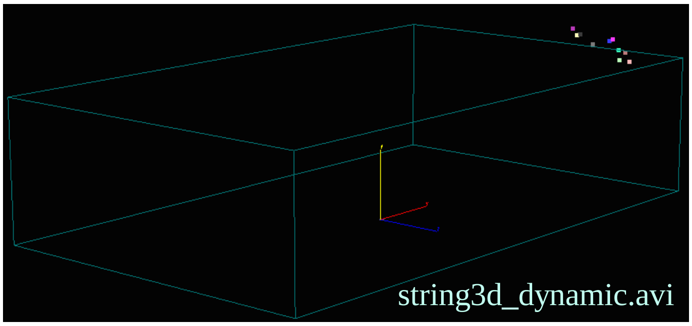
3) Condensation of 3D strings

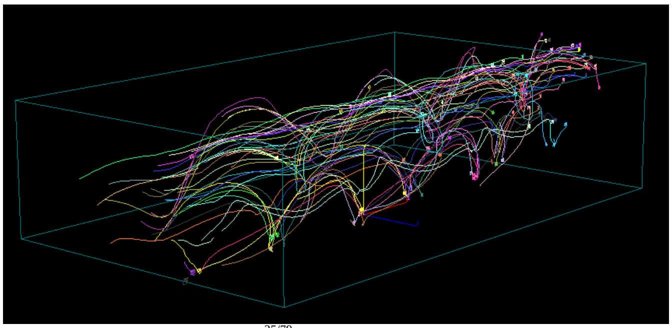


4) Joining 3D strings

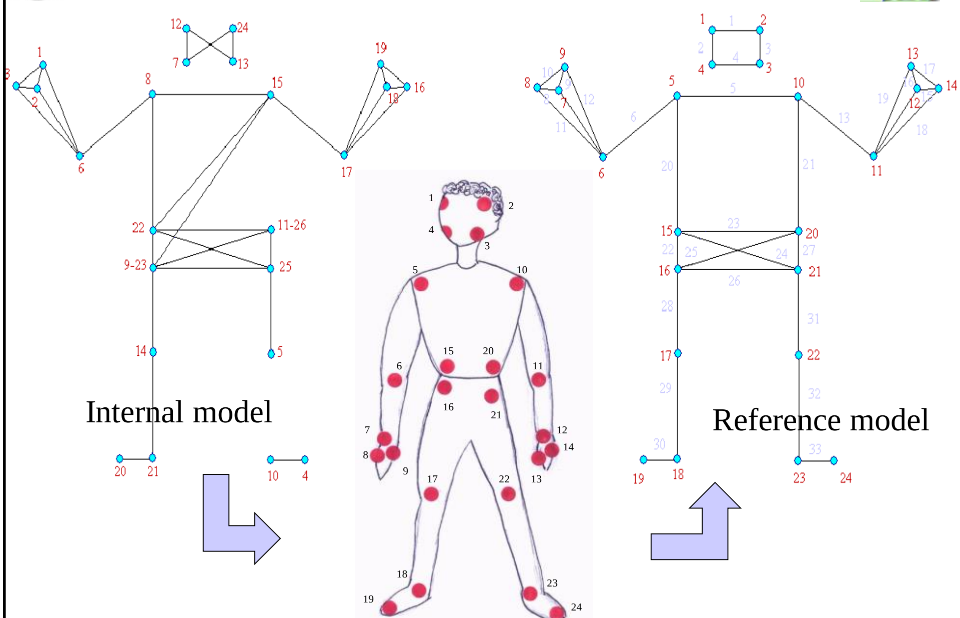


3D strings

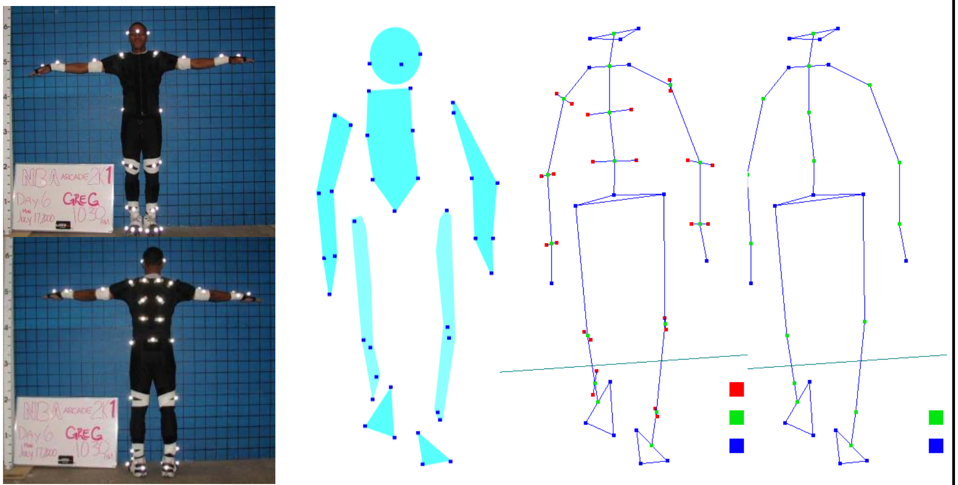




Model fitting



What a model represents?



Markered subject | Modello 3D | Modello a stick | Modello hidden

Problems intrinsic in body tracking

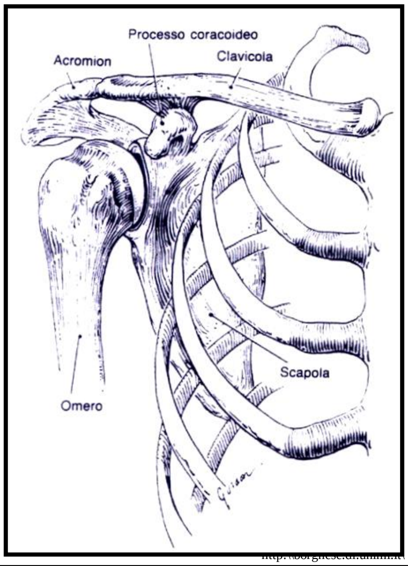
- Joints are points inside the body, markers are attached on the body surface.
- Joint are not fixed points: two adjacent bones rotate <u>and slide</u>.
- Joint are not spherical.
- Joints can be complex (e.g. Shoulder, spine)
- Skin artifacts.

The human skeleton has complex articulations

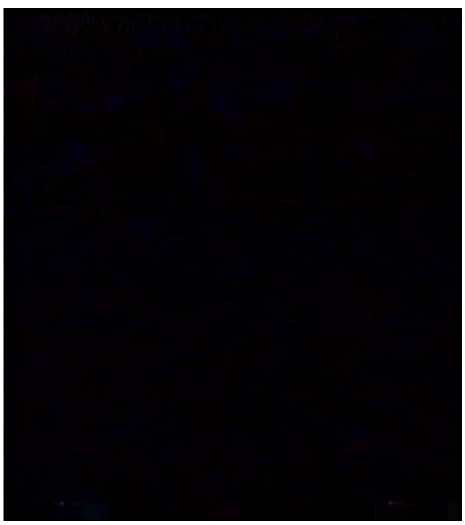
"Rigid" bones connected. Tendons keep the bones in place.

Motion allowed can be very complex (e.g. shoulder, spine).

The reconstruction of the finest details of the motion are beyond reach, simplifying assumptions are made => ***Level of detail*** in motion analysis

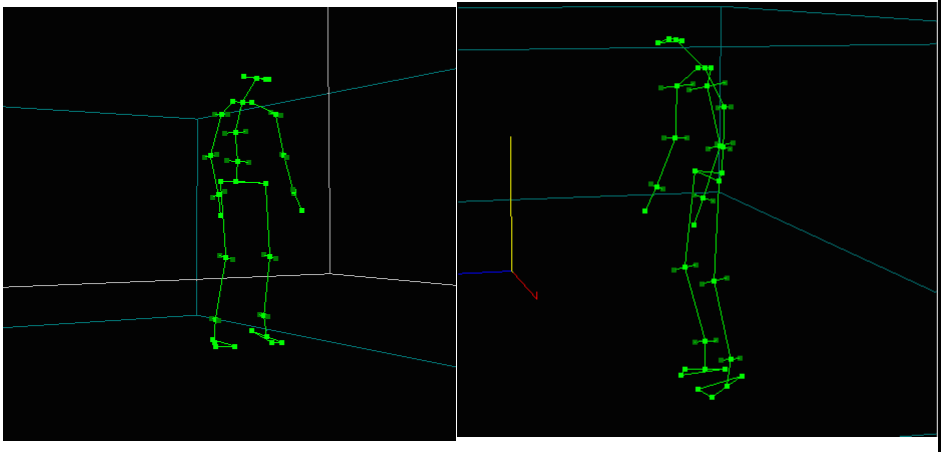


Video by Superfluo

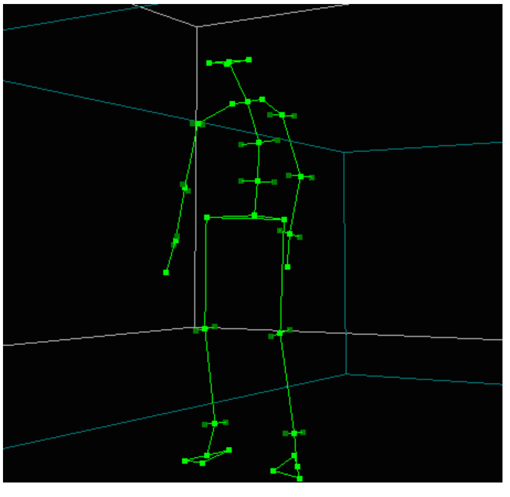


superfluo3.wmv

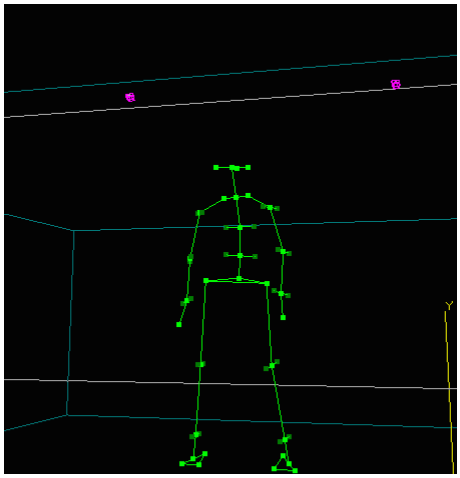
Risultati: escape



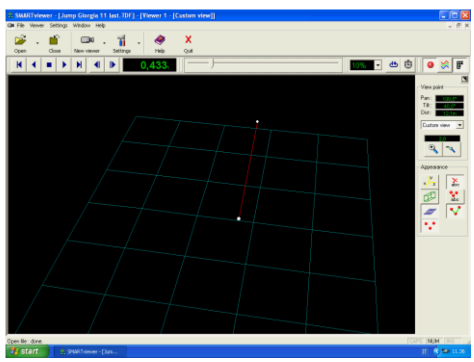
Risultati: fall_run



Risultati: roll

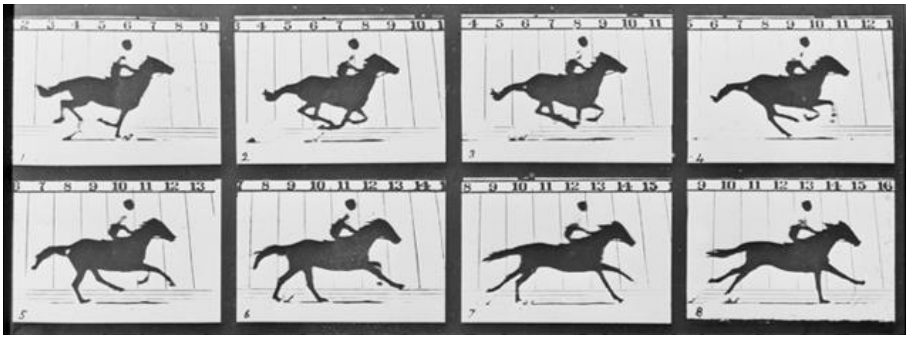


High jump – top athletes



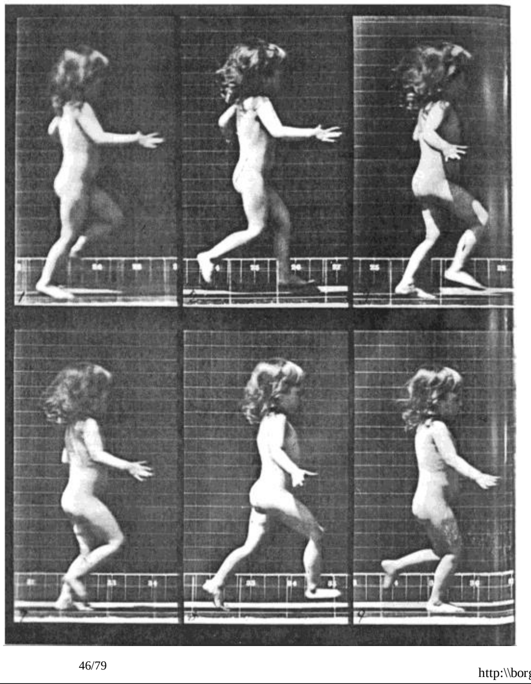
Can we work without markers?

Edward Muybridge 1878-1901



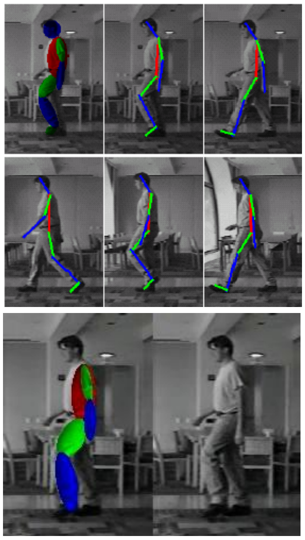
http://www.eadweardmuybridge.co.uk/

Edward Muybridge 1878-1901



Computer vision techniques

Silhouette (-> Skeleton)



Set of difficult problems:

2D Image processing (silhouette identification, optical flow detectors…)

Multi-view invariants.

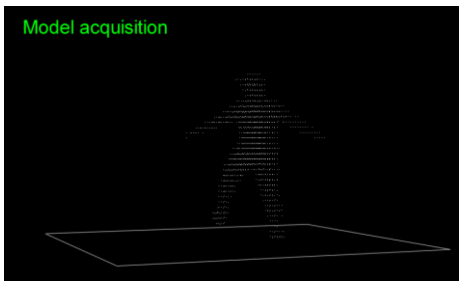
Smooth motion -> temporal filtering.

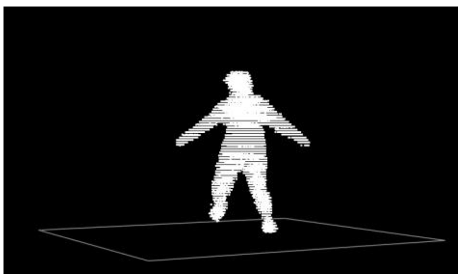
Skeleton fitting (different rigid motion for different segments).

3D cameras help a lot

http://movement.stanford.edu/

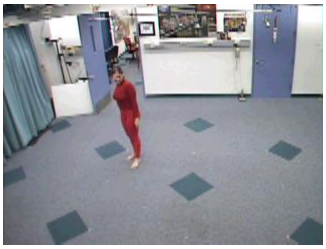
Results: stepping (640 x 480, 10Hz)

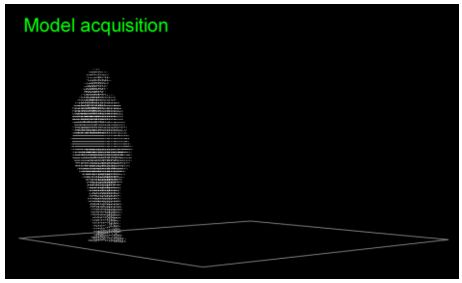


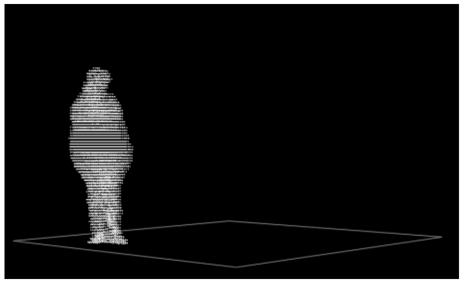


Mikic, Trivedi, Hunter

Results: cartwheel

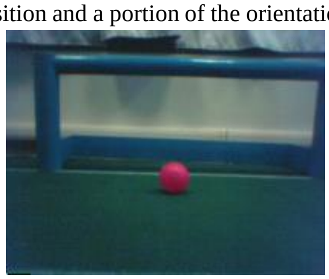


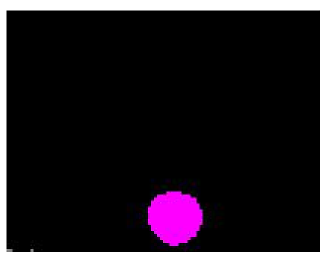




2D color coded tracking

- Players could interact with a 3D scene by moving known brightly saturated colored objects that were visually tracked in PlayStation 2 (EyeToy Webcam). Threshold on color representation.
- Pose recovery can be accomplished robustly for certain shapes of known physical dimensions by measuring the statistical properties of the shape's 2D projection. In this manner, for a sphere the 3D position can be recovered (but no orientation), and for a cylinder, the 3D position and a portion of the orientation can be recovered.





- Multiple objects can be also be combined for complete 3D pose recovery, though occlusion issues arise.
- Perfect recognition in all lighting conditions is difficult.

2D tracking with controlled background

Duck-neglect project **http://borghese.dsi.unimi.it/Research/LinesResearch/Virtual/Virtual.html**

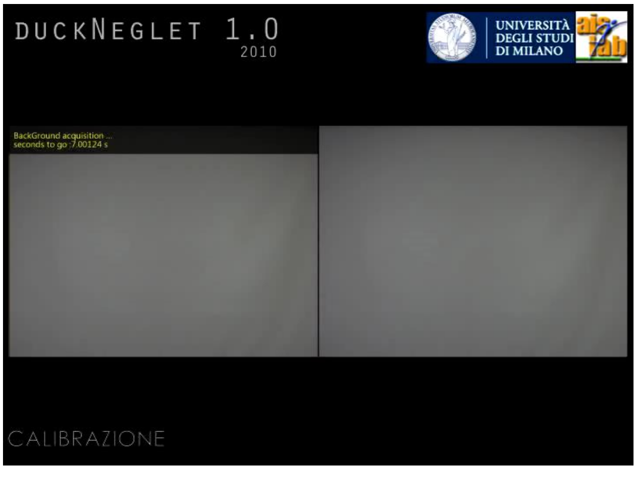
"Magic mirror" paradigm in which video of the player is overlayed with graphics generated by the computer.

Background measurement. Thresholding.

In this case, silhouette is tracked.

Alternative is the difference between consecutive images (glaring and blurring require some filtering).

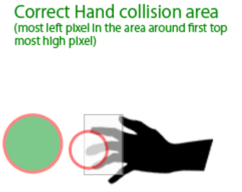
Duckneglect



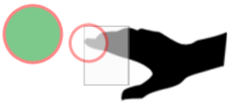
Uniform background subtraction (e.g. green screen)

2D collision detection

- Collision detection with target can be checked by analyzing the overlapping between part of the motion mask only in particular regions.
- Identification of the motion mask as the outermost part of the body. Approximated collision detection defining general shapes.



- Collision with targets gives hit, collision with distractors gives a miss.
- Same principles implemented with Sony EyeToy Webcam (2003).



Markerless optical motion capture

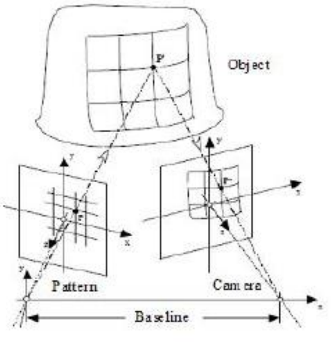
https://www.ideaslab.com/ai-technologies/

2.5D First SDK for Kinect

Primesense drivers, with skeleton tracking: http://www.primesense.comse.com

RGB-D cameras (Kinect)

- 3D scanner with active pattern (Infra Red)
- RGB camera
- Robust background/foreground separation
- Robust skeletal tracking (Kinect)



Used as a Web-cam with advanced silhouette Subtraction for rehabilitation.

Come to the lab to see…

Body tracking (marker-less)

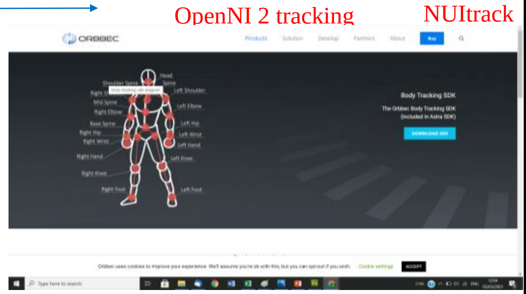
RGB-D Cameras
(Prime Sense - 2000-2013)

RGB-D Cameras (Prime Sense - 2000-2013) → Microsoft (Kinect I e II) → Apple (2014, 360 miliardi) → Integration of gaze Tracking and analysis

OpenNI tracking → X

Occipital → OpenNI 2 tracking

NUItrack



Intel Real Sense cameras (chiusa nel 2021)

Basler
Asus
Ifm
Orbbec
Zvid
….

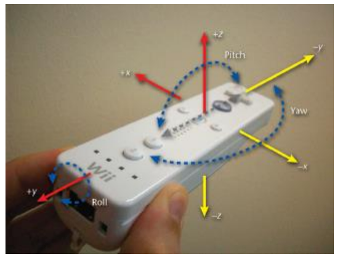
Experience in the lab

Body motion from footage (Structure from Motion)

2 approcci:
- Probabilistico. Stima di un modello parametrizzato e dei parametri di movimento.
- Deterministico. Definisco un modello a-priori e stimo i parametri della camera e del movimento.

Intertial tracking::Wii

$$pitch = \arctan\left(\frac{a_z}{a_y}\right)$$
$$roll = \arctan\left(\frac{a_z}{a_x}\right).$$



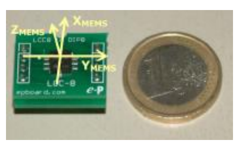
Positional data are obtained through integration.
⇒Instability. A flip of the LSB for one frame generates a rotation at constant speed!!

Other devices are required to stabilize the measurements: Nunchuk (gyroscope), sensor IR-bar

Intertial tracking::Xsens

- Xsens by Moven is a full-body, camera-less inertial motion capture (MoCap) solution. It is flexible motion capture system that can be used indoors or outdoors (on-set). With the short turnaround times MVN is a cost effective system with clean and smooth data.

- Costly



- We have used such system inside the FITREHAB project:
  http://www.innovation4welfare.eu/287/subprojects/fitrehab.html

https://www.xsens.com/products/mvn-animate?hsCtaTracking=0031f976-823a-4074-8cc4-d6f2347422ae%7C584bb7ed-596e-4dd6-992d-245825acf04f

Where are we now (optoelectronic)?

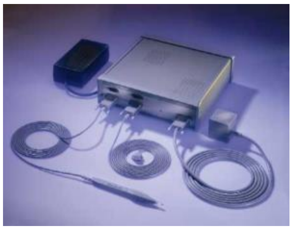
Optotrack, 1991.

LED + cameras

•Measure the position of the joints.
•Time multiplexing for the markers (3 at 450Hz or 750Hz with additional hardware). No-tracking, real-time.
•Power for the LEDs has to be delivered on the subject's body (markers get hot on the skin!!).
•Accuracy 0.1mm (X,Y), 0.15mm (Z, depth).

Where are we now (magnetic)?

***Magnetic technology*: Fastrack & older Polhemus sensors.**



They measure: pitch, yaw and roll; X, Y, Z of the segments.

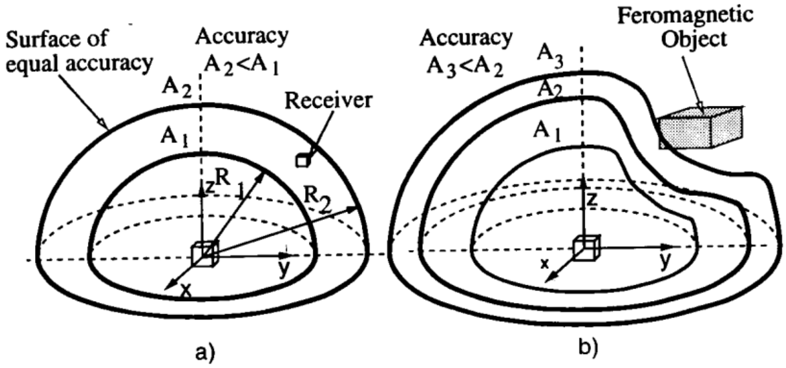
Electro-magnetic induction.

The transmitter is a triad of electromagnetic coils, enclosed in a plastic shell, that emits the magnetic fields. The transmitter is the system's reference frame for receiver measurements.

The receiver is a small triad of electromagnetic coils, enclosed in a plastic shell, that detects the magnetic fields emitted by the transmitter. The receiver is a lightweight cube whose position and orientation are precisely measured as it is moved.

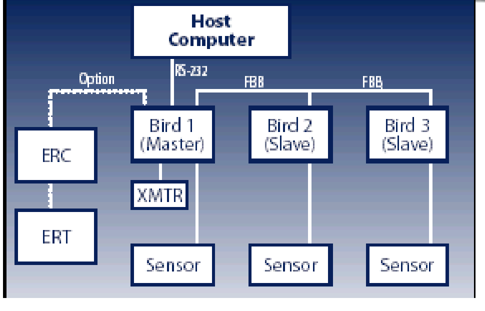
Fast-track Motion Capture

•Higher accuracy through oversampling and DSP signal processing (0,5" and 1.8mm accuracy). Range of 75cm for high accuracy.
•Sensitive to ferromagnetic (metallic) objects.



•Latency: 4msec.
•Sampling rate: 120Hz. Rate drop with multiple receivers because of multiplexing.

Flock of birds Motion Capture



•Each receiver has its own DSP.
•All the DSP are connected with a fast internal bus.
•Latency is increased (8ms).

When more than one transmitter is adopted (exprimental):
larger field (single transmitter at a time)
higher accuracy (time-slicing)

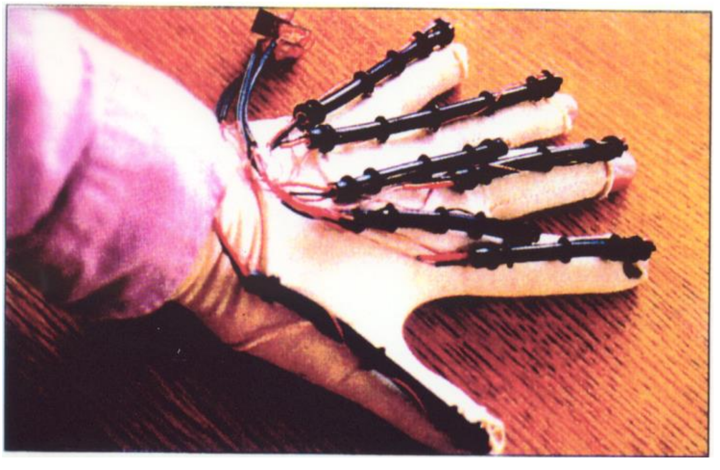
Not really un-obtrusive! Low accuracy. Real-time.

Gloves

Monitor fingers position and force.

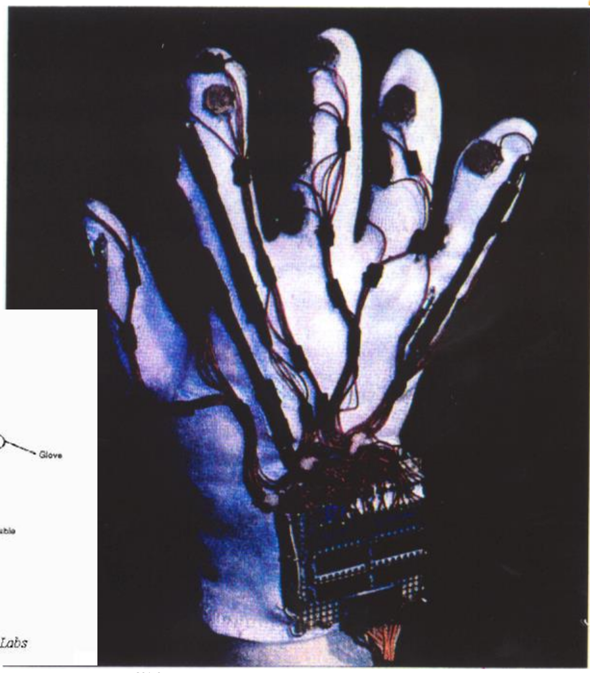
Problems with the motion of the fingers:
- overlap.
- fine movements.
- fast movements.
- rich repertoire.

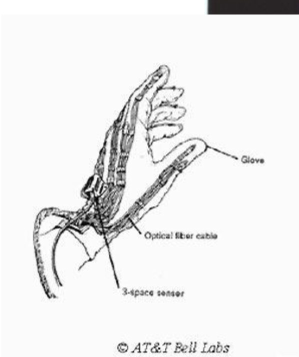
Sayre glove (1976)



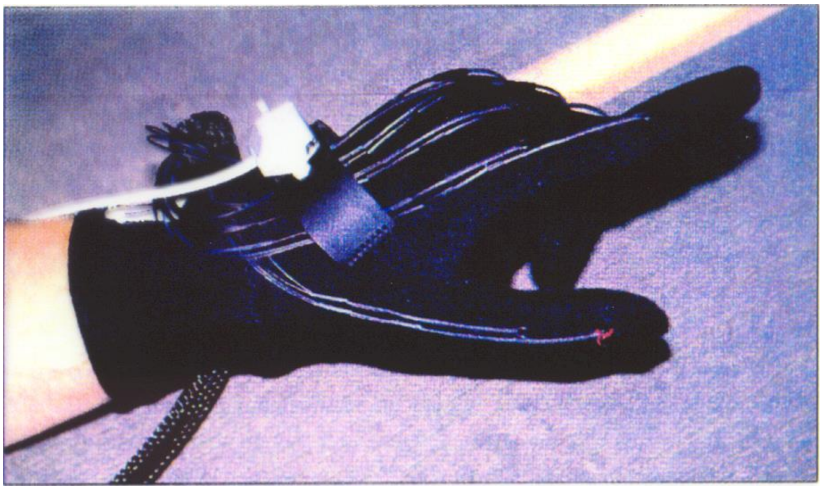
MIT glove (1977)

Digital Data Entry Glove (1983)



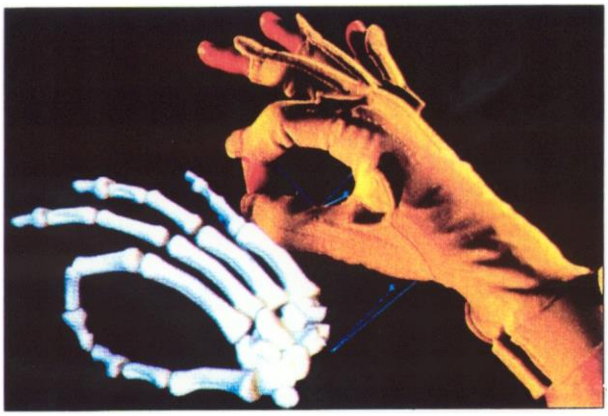


Data Glove (1987)



Power Glove (1990)

Cyber Glove (1995-today)



Cyberglove I

Cyberglove III

AcceleGlove / iGlove (2009)

http://www.anthrotronix.com/index.php?option=com_content&view=article&id=87&Itemid=138

Calibration

Estimate of the geometrical parameters in the transformation operated by the sensors (e.g. the perspective transformation operated by a video-camera).

Estimate of the parameters, which describe distortions introduced by the measurement system.

Measurement of a known pattern. From its distortion, the parameters can be computed.

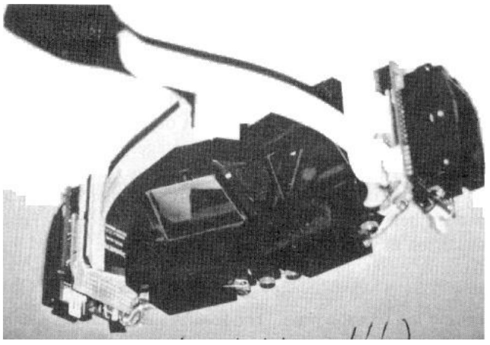
Algorithms adopted: polynomial, local correction (neural networks, fuzzy).

Finger tracking through cameras

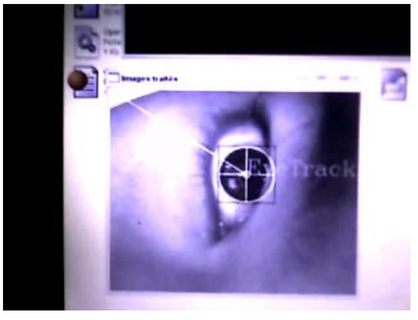
Experience in the lab

Gaze input

- Contact lenses carrying magnetic coils.
- TVcameras aligned with an IR LED source.
- Stereoscopic eye-wear.
- The direction of gaze is decided by measuring the shape of the spot reflected by the frontal portion of the cornea (Ohshima et al., 1996).
- Eye trax http://www.eyetrax.it/en/index.html



Vision based eye trackers

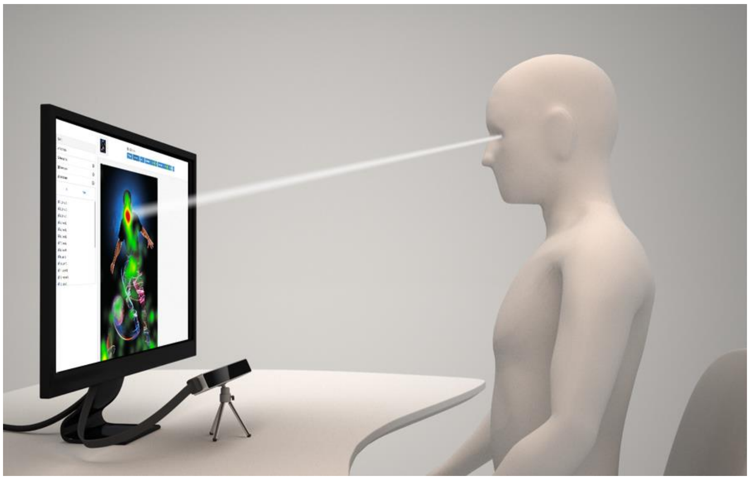


Logitech Quickcam 4000

Color information
Geometry information (circles, relative position...)
Histogram analysis on gray level.
Custom tool for many WEBcams
...

EyeTribe - https://theeyetribe.com/. Passive. 99 US $

Gaze tracking



http://theeyetribe.com/theeyetribe.com/about/index.html
I-Pad

Experience in the lab

History

Video technology (semi-automatic marker detection, slow-motion, 1975)

Optoelecontric active markers: Selspot™ 1977 (Selspot II 1993), Watsmart™ 1985, Optotrack™ 1992, Polaris™ 1998. http://www.ndigital.com/home.html

Automatic video marker detection:
Vicon™ 1981. http://www.oxfordmetrics.com/
Elite™ 1988. http://www.bts.it/
MotionAnalysis™ 1992, Eagle™ 2001. http://www.motionanalysis.com/
Smart™ 2000. http://www.motion-engineering.com/

Magnetic systems:
Sensors: Polhemus 1987, Fastrack 1993. http://www.polhemus.com/
Systems: Flock of birds 1994. http://www.ascension-tech.com/

Intertial systems: Xmoven Xsense 2000, Wii 2008.

Video processing: organicmotion 2010, ideaslab 2020.

3D video systems: RGB-D cameras.

Sommario

- Introduzione
- Sistemi di Input
- Generatori di mondi
- Motore di calcolo
- Sistemi di Output
- Conclusioni