

Introduction to Virtual Reality Parte I

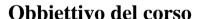
Prof. Alberto Borghese Applied Intelligent Systems Laboratory (AIS-Lab) Department of Computer Science University of Milano





1/75







- 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, Wii balance board, Leap, Smart 3D MoCap, tracker aptico Falcon, solette sensorizzate, eye tribe,...). Causa COVID-19 questa esperienza sarà forzatamente limitata.
- Modalità d'esame: progetto + discussione teoria
 - Il progetto può essere associate a altri corsi e/o alla tesi.

A.A. 2020-2021 2/75

http:\\borghese.di.unimi.it\

1





Realtà Virtuale - 6 CFU

Sito principale:

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

Programma:

http://borghese.di.unimi.it/Teaching/VR/Programma_2020-2021.html

Let's try to keep the course interactive

Orario:

Martedì Ore 10.30-12.30 da remoto (Ariel) - Teoria Giovedì Ore 9.30-12.30 da remoto (Ariel) - Laboratorio

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

Ricevimento su appuntamento

A.A. 2020-2021

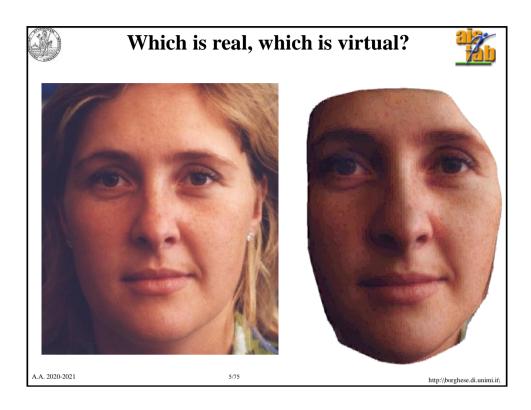




http:\\borghese.di.unimi.it\

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

A.A. 2020-2021 4





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.

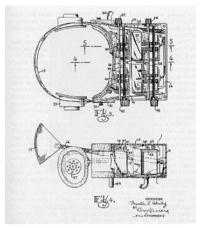
- Plato (world of the ideas) 428-348 a.C.
- Berkeley (sensorial experience is too limited) 1685-1753.
- Hegel ("what is rational is real..") 1770-1831.
- New age.

A.A. 2020-2021 6/75



Historical Perspective (II)





Morton Heilig 1956, patented in 1961 Non fu mai costruito



projected film, audio, vibration, wind, odors.

A.A. 2020-2021

7/75

 $http: \hspace{-0.05cm} \hspace{-$



Historical Perspective (III)

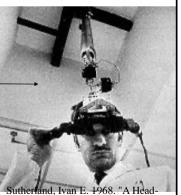


Technological background

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



A. 2020-2021



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.A. 2020-2021

9/75

http:\\borghese.di.unimi.it



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.

A.A. 2020-2021

10/75



Sommario



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



A.A. 2020-2021 11/75 http://borghese.di.jnjin



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.

A.A. 2020-2021 12/75



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.

Adopted technology

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

A.A. 2020-2021

13/7





http:\\borghese.di.unimi.it\



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 ...)

A.A. 2020-2021





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

A.A. 2020-2021 15/75 http://borghese.di.unimi.it/

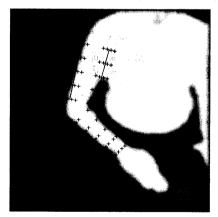


A.A. 2020-2021

What is captured?

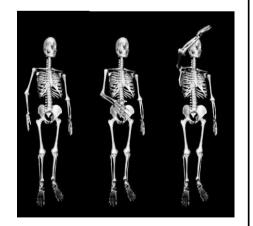


Silhouette (-> Skeleton)

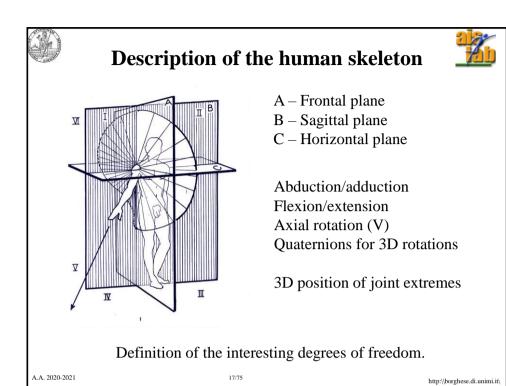


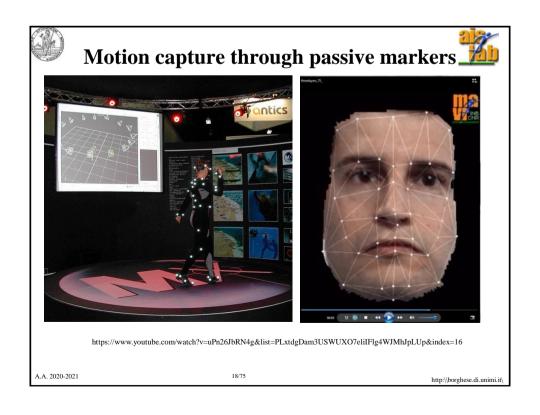
Computer vision techniques (silhouette, RGB-D cam)

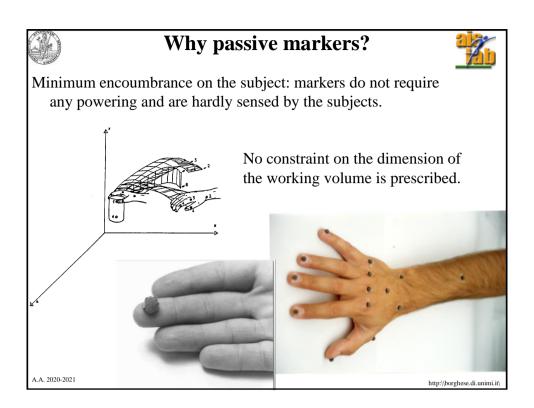
Skeleton



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









How passive markers work?



Passive markers are constituted of a small plastic support covered with retro-reflecting material $(3M^{TM})$. 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.

A.A. 2020-2021

20/75



Tracking difficulties



<u>It is a complex</u> 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.

A.A. 2020-2021

22/75



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.

http:\\borghese.di.unimi.it\

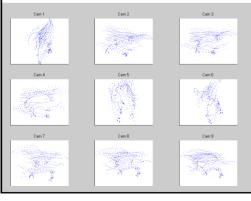


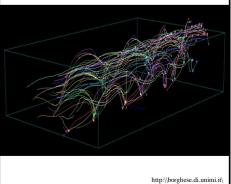
Disadvantages of motion capture systems based on passive markers



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.





12

Tracking difficulties

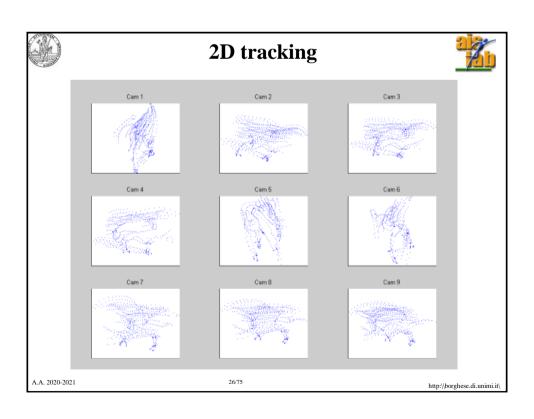


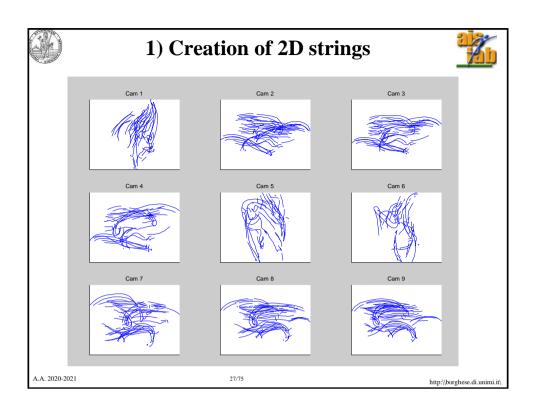
<u>It is a complex</u> problem because:

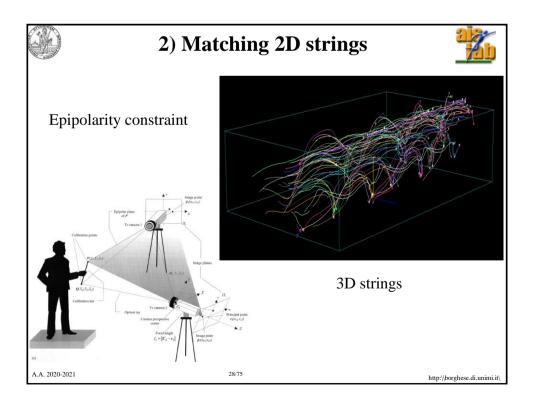
• Dense set of markers. These may come very close one to the other in certain instants.

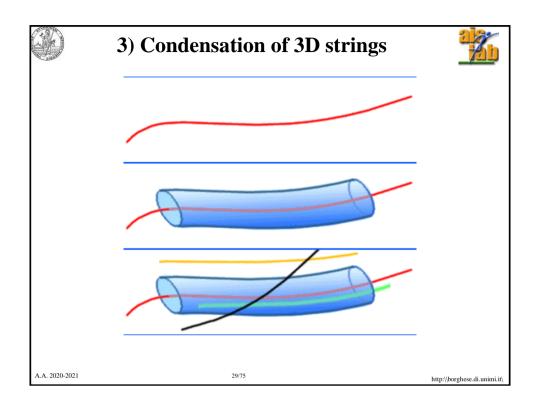


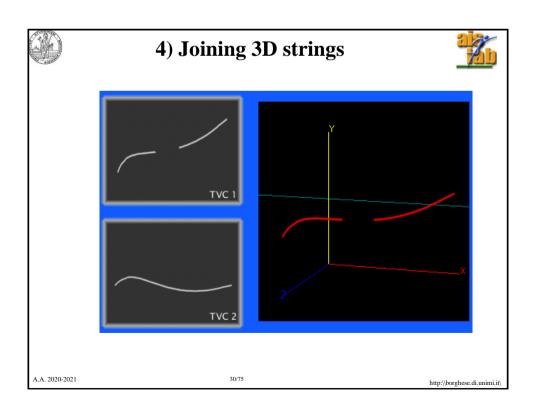
- Motion can be easily complex, as it involves rotation and twists of the different body parts (thing at a gymnastic movement).
- •Multi-camera information and temporal information is required to achieve a robust tracking.

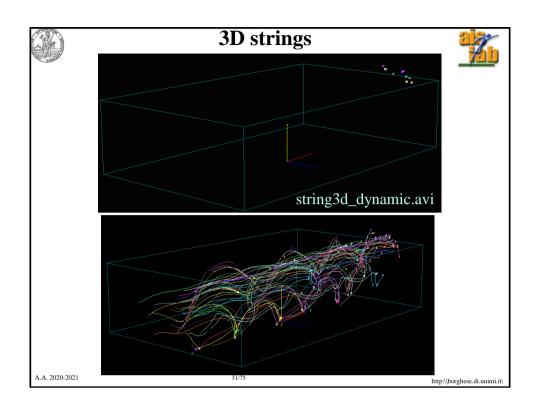


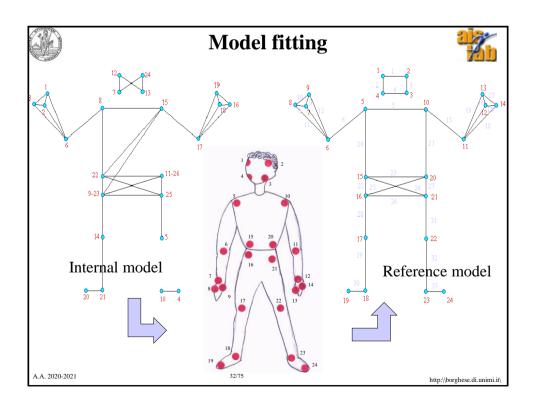


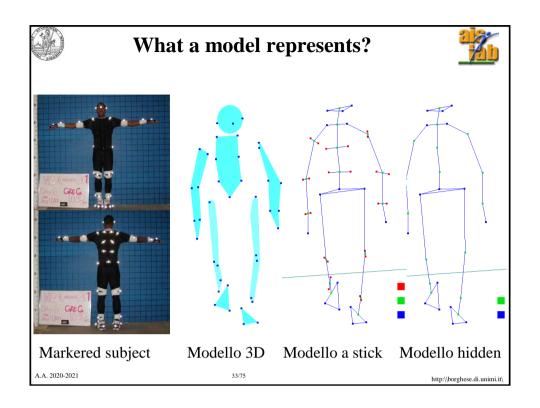














Problems intrinsic in body tracking



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

A.A. 2020-2021 34/75

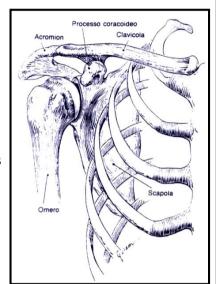
 $http: \hspace{-0.05cm} \hspace{-0.05cm} \hspace{-0.05cm} \hspace{-0.05cm} \hspace{-0.05cm} http: \hspace{-0.05cm} \hspace{-0.05c$

The human skeleton has complex articulations

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

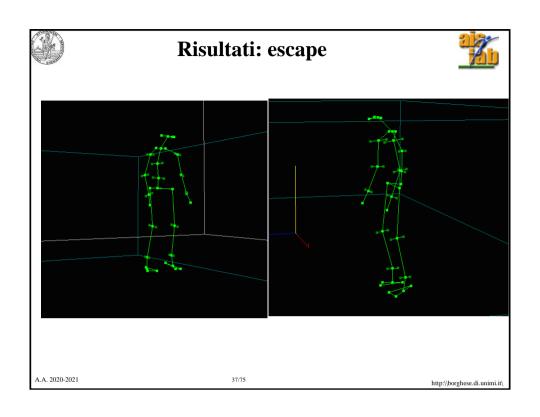
Motion allowed can be very complex (e.g. shoulder, spine).

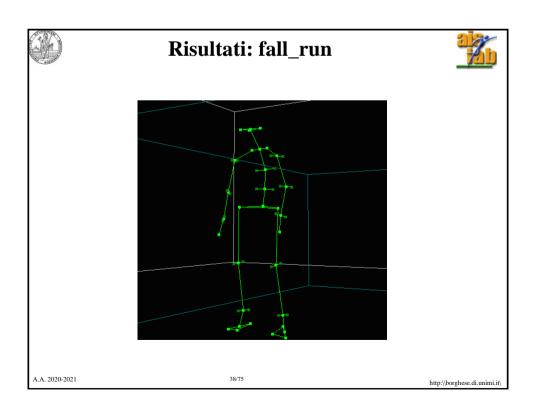
The reconstruction of the finest details of the motion are beyond reach, simplifying assumptions are made => *Level of detail* in motion analysis

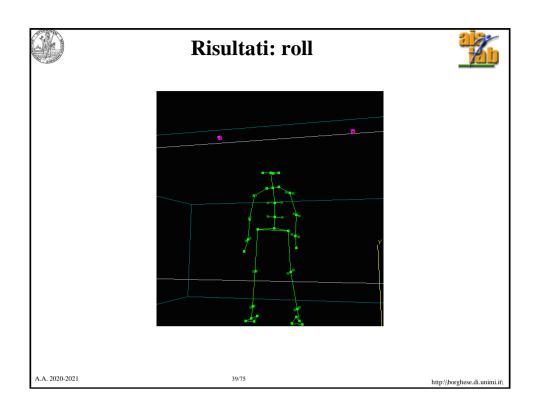


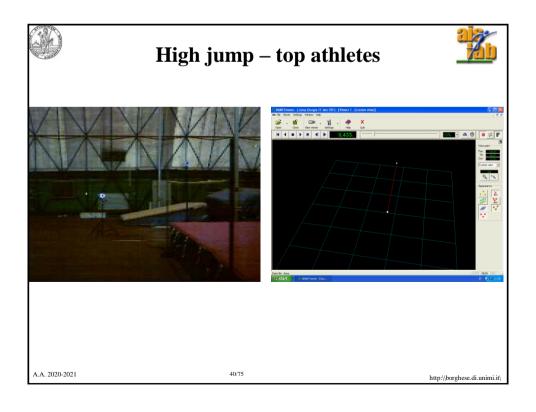
A.A. 2020-2021 35



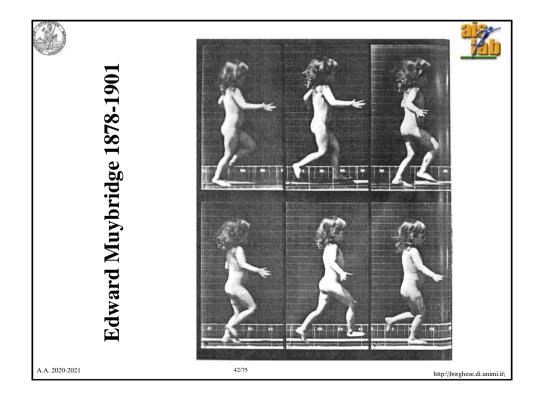


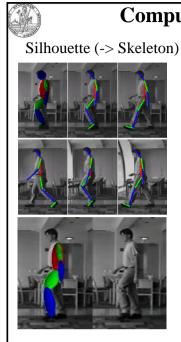












A.A. 2020-2021

Computer vision techniques



Set of difficult problems:

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

Multi-view invariants.

Smooth motion -> temporal filtering.

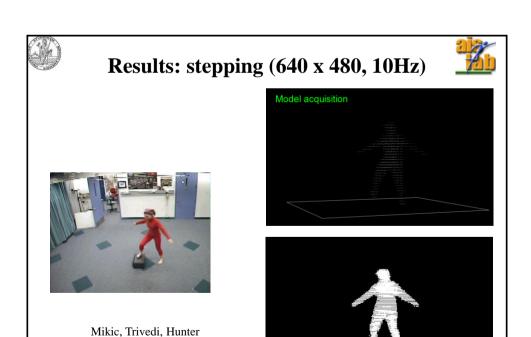
Skeleton fitting (different rigid motion for different segments).

3D cameras help a lot

http://movement.stanford.edu/

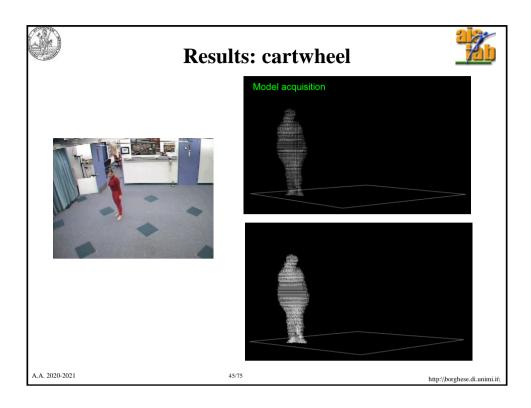
http:\\borghese.di.unimi.it\

http:\\borghese.di.unimi.it\



A.A. 2020-2021

22





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.

A.A. 2020-2021



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).

47/75

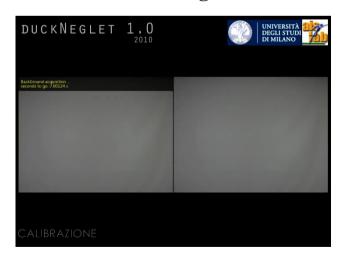
A.A. 2020-2021

 $http: \hspace{-0.05cm} \hspace{-$



Duckneglect





Uniform background subtraction (e.g. green screen)

A.A. 2020-2021

48/75





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.

Correct Hand collision area (most left pixel in the area around first top most high pixel)





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



A.A. 2020-2021

di unimi it



Markerless optical motion capture



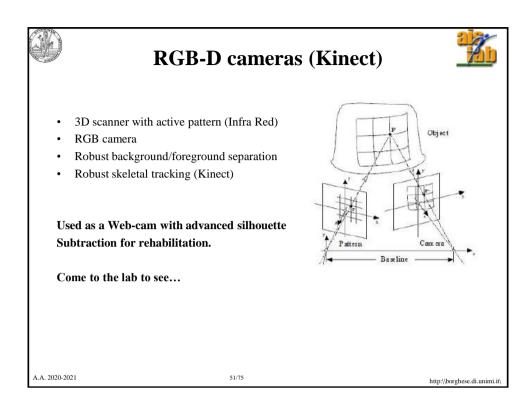
https://www.ideaslab.com/ai-technologies/

49/75

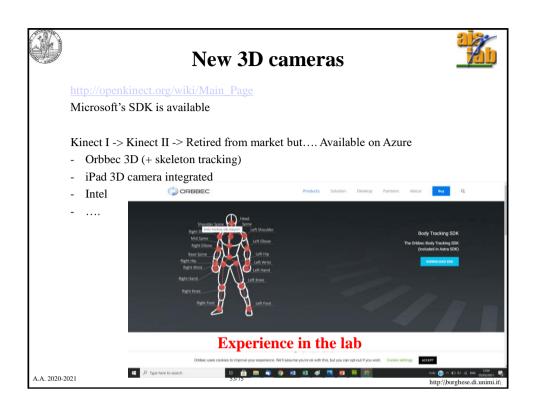


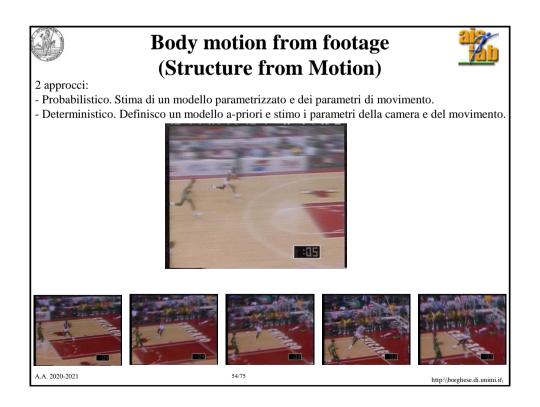
A.A. 2020-2021

50/75









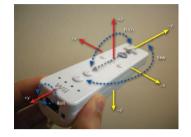




Intertial tracking::Wii

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

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



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



A.A. 2020-2021

ttn:\\borghese di unimi it\

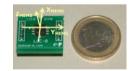


Intertial tracking::Xsens



- 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

A.A. 2020-2021







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).

A.A. 2020-2021

57/7

http:\\borghese.di.unimi.it\



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.

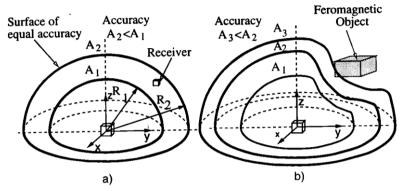
A.A. 2020-2021

58/75

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.

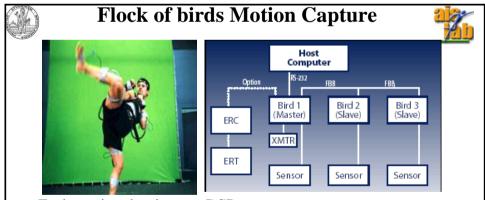


- •Latency: 4msec.
- •Sampling rate: 120Hz. Rate drop with multiple receivers because

of multiplexing.

59/75

 $http: \hspace{-0.05cm} \hspace{-$



- •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.

http://borohosa.di unimi i

A.A. 2020-2021



Gloves



 $http: \hspace{-0.05cm} \hspace{-0.05cm} \hspace{-0.05cm} \hspace{-0.05cm} \hspace{-0.05cm} http: \hspace{-0.05cm} \hspace{-0.05c$

Monitor fingers position and force.

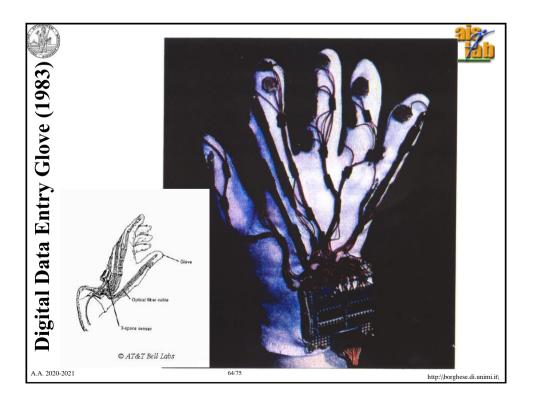
Problems with the motion of the fingers:

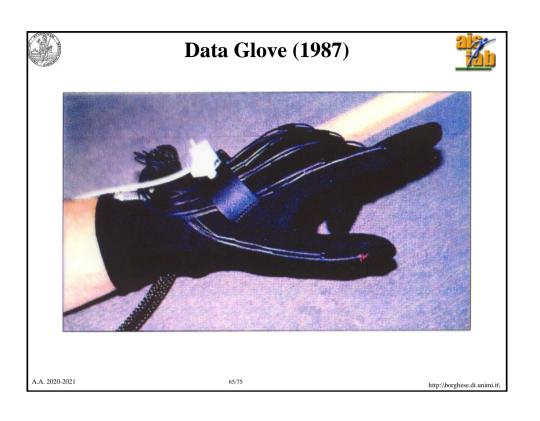
- overlap.
- fine movements.
- fast movements.
- rich repertoire.

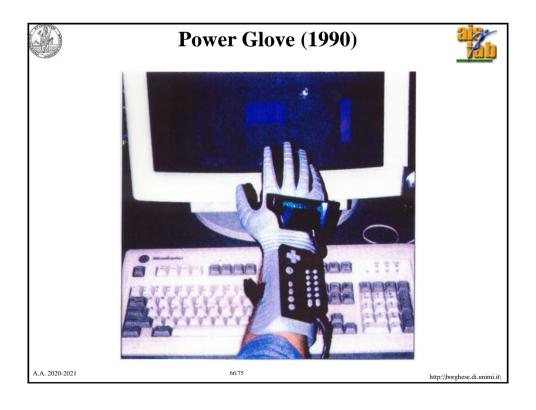
A.A. 2020-2021 61

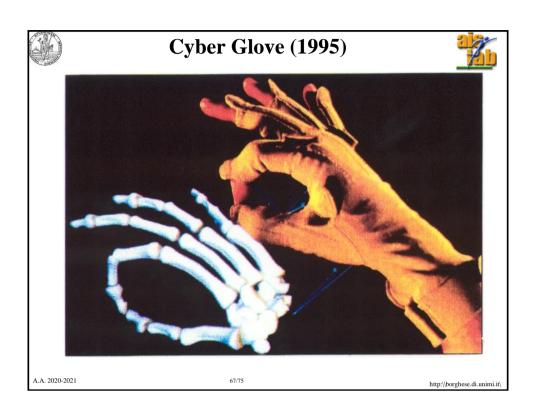
















Calibration



http:\\borghese.di.unimi.it\

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).

A.A. 2020-2021

Finger tracking through cameras

LEAP

PRODUCT NO APPS DEVELORE SOLUTIONS COMMANTY STEP WHAT

HE Leap Motion Keyboard

The left great leag Motion tetraheling
Hey une a compactive life device and transform
the way you regerence your digit device.

Experience in the lab

A.A. 2020-2021

PRODUCT NO APPS DEVELORE SOLUTIONS COMMANTY STEP WHAT

HE Leap Motion Keyboard

The left great leag Motion tetraheling
Hey une a compactive life device and transform
the way you regerence your digit device.

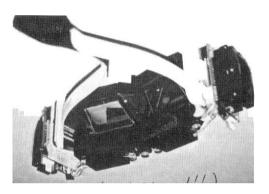
A.A. 2020-2021



Gaze input



- Contact lenses carrying magnetic coils.
- TV cameras 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



A.A. 2020-2021 71/75 http://borghese.di.unimi.it/



Vision based eye trackers





Color information Geometry information (circles, relative position...) Histogram analysis on gray level.

Custom tool for many WEBcams



Logitech Quickcam 4000

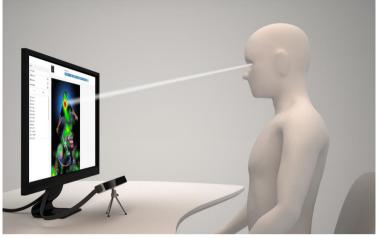
EyeTribe - https://theeyetribe.com/. Passive. 99 US \$

A.A. 2020-2021



Gaze tracking





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

Experience in the lab

A.A. 2020-2021

73/75

http:\\borghese.di.unimi.it\



History



<u>Video technology</u> (semi-automatic marker detection, slow-motion, 1975)

Optoelecontric active markers: SelspotTM 1977 (Selspot II 1993), WatsmartTM 1985, OptotrackTM 1992, PolarisTM 1998. http://www.ndigital.com/home.html

Automatic video marker detection:

ViconTM 1981. http://www.oxfordmetrics.com/

EliteTM 1988. http://www.bts.it/

MotionAnalysisTM 1992, EagleTM 2001. http://www.motionanalysis.com/

SmartTM 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.

A.A. 2020-2021

74/75



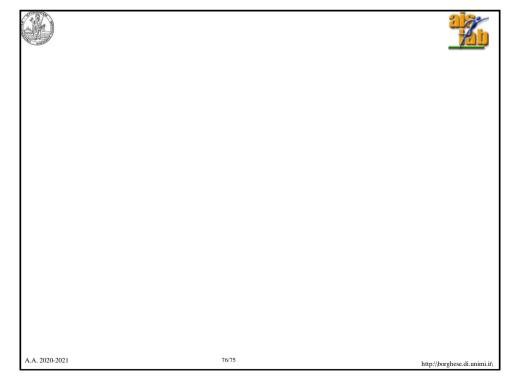
Sommario

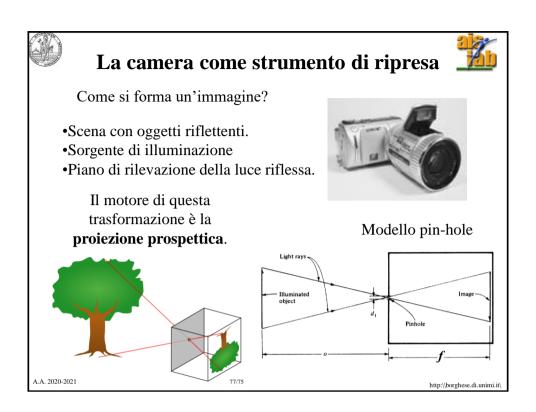


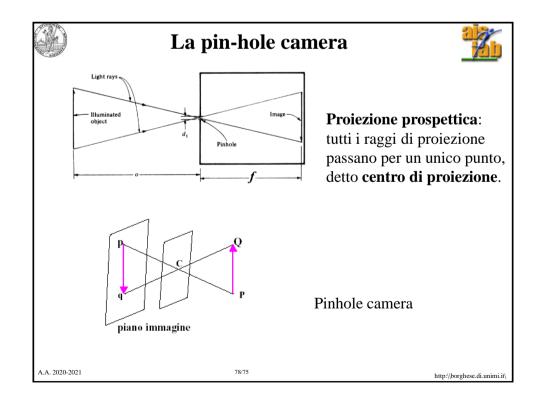
http:\\borghese.di.unimi.it\

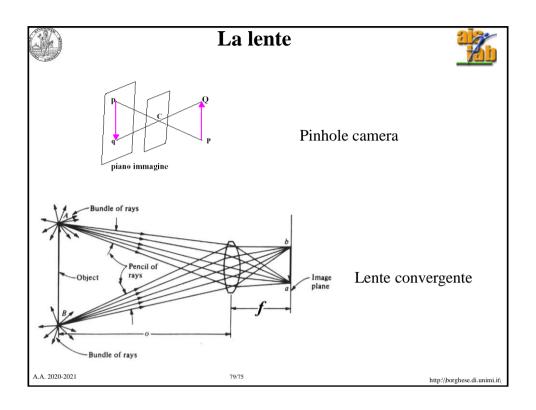
- Introduzione
- Sistemi di Input
- Generatori di mondi
- Motore di calcolo
- Sistemi di Output
- Conclusioni

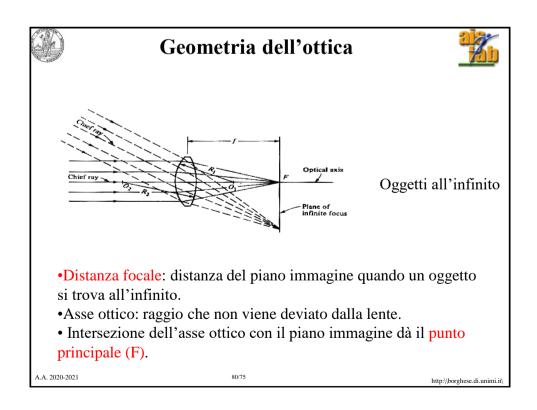
A.A. 2020-2021 75









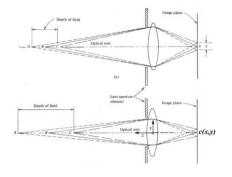




Messa a fuoco



Problema della messa a fuoco

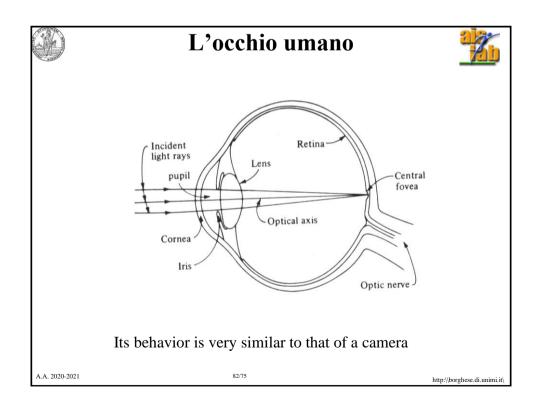


Parametri di camera (o intrinseci):

- •Punto principale c(x,y) + lunghezza focale, f (3 parametri).
- •Occorre conoscere anche il fattore di forma dei pixel nel caso di immagini digitali (è una costante, non un parametro).
- •(Distorsioni).

A.A. 2020-2021

81/75

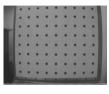






Calibration is a pre-requisite

Camera calibration











Set-up calibration

Excellent for special effects, not so good for measurements....

Cameras are not metric.

A.A. 2020-2021

83/75