

Predictive Vision in Human Robot Collaboration

Giulio Sandini
IIT - Robotics, Brain and Cognitive Sciences
and
University of Genova

Active Vision and Perception in Human(-Robot) Collaboration (AVHRC 2020) - ROMAN-2020

Goal of talk

To revisit some of our past work on Cognitive Robotics and to propose a few research topics and methodologies which I consider important for the advancement of Social Cognition (in humans and robots)...

...no solutions...just aspects to think about and, perhaps, address more specifically in the future...

...from a biased point of view...

© Giulio Sandini

...use artificial systems to study humans...

Artificial Systems and Neuroscience

- Use artificial systems to study human visual perception and sensorimotor control
- Propose new technologies

Background

- Biological and Artificial Vision
- Robotics

Sandini, G. et. al., (1997) Human Sensorimotor Development and Artificial Systems. Symposium on Artificial Intelligence, Robotics and Intellectual Human Activity Support (AIR&IHAS '97, November 19-21, 1997 - Riken - Japan)

Outline of Talk:

- How I got here... (a bit of history)
- Cognition: "Beyond Real-time" (internal models...)
- Timeframes of adaptation (a role for Deep Learning?)
- Social Interaction and HRI (Few examples)
- Interdisciplinarity (...and the ethics of science)

© Giulio Sandini

...use artificial systems to study humans...

...the target is a more humane robot(*)...

A humane robot is not necessarily a humanoid or a realistic copy of a human but should stimulate our Anthropomorphic Imagination...

AI ...our imagination cannot be anything But anthropomorphic.

Halo Cavrino: The Mentor for the next millennium: Visibility (written 1945 - published 1988)

(*) Humane (def.): someone or something gentler, kinder, or more appealing to people. (Merriam-Webster Dictionary)

Sandini, G. et. al., (1997) Human Sensorimotor Development and Artificial Systems. Symposium on Artificial Intelligence, Robotics and Intellectual Human Activity Support (AIR&IHAS '97, November 19-21, 1997 - Riken - Japan)

Outline of Talk:

- How I got here... (a bit of history)
- Cognition: "Beyond Real-time" (internal models...)
- Timeframes of adaptation (A Role for Deep Learning?)
- Social Interaction and HRI (Few examples)
- Interdisciplinarity (...and the ethics of science)

© Giulio Sandini

iit

June 1996 EIBA Lab

Artificial Systems and Neuroscience

- Use artificial systems to study human visual perception and sensorimotor control
- Propose new technologies

Background

- Biological and Artificial Vision
- Robotics

**How I got here...
(a brief history of the parallel between neuroscience and the study of artificial systems)**

Sandini, G. et. al., (1997) Human Sensorimotor Development and Artificial Systems. Symposium on Artificial Intelligence, Robotics and Intellectual Human Activity Support (AIR&IHAS '97, November 19-21, 1997 - Riken - Japan

© Giulio Sandini

iit

...master thesis (bioengineering) on modelling eye-head coordination (a hot topic at the time)

...visual perception and eye movements...

From neuroscience: Research on "eye movements and attention": Yarbus (1967), Emilio Bizzi (1973), Larry Stark

...but gaze is not only "control" and...

...I fell in love with "vision".

...but gaze is not only a "control"...

Morasso P., Sandini G., Tagliasco V., Zaccaria R., (1977) "Control Strategies in the Eye-Head Coordination System". IEEE Trans. Systems, Man and Cybernetics, SMC-7

© Giulio Sandini

iit

...a PhD in neuroscience ...at the time of local features

The early years of brain vision ('70-'80)

In the same years computer vision also focused on local features...

Objects are uniquely identified through contours

David Marr (1978) Representing Visual Information: A Computational Approach, COGTRG-80

...1985 the birth of computational neuroscience

Edges

Sandini G., Torre V., (1983): "Thresholding Techniques for Zero Crossings". Proc. Optical Society of America

© Giulio Sandini

iit

Active Vision and Action Vision

...the Active Vision Revolution

Motion can be exploited to extract visual (perceptual) information

Vision during action (sensory-motor strategies for object understanding)

From "active observer" to "actor"...

...anticipating the effect of actions...

G. SANDINI, F. GANDOLFO, E. GROSSO, M. TISTARELLI, "VISION DURING ACTION", IN ACTIVE PERCEPTION, Y. ALOMONOS, Ed., LAWRENCE Erlbaum ASSOCIATES, 1993, pp. 151-190.

© Giulio Sandini

iit From eye movements and space-variant retina

Alfred Yarbus studies of scanpaths in the '50 and '60

Need to move the eyes to explore a scene

Retino-cortical mapping and the cortical magnification factor

G. Sandini, V. Tagliaco (1980). An Anthropomorphic Retina-like Structure for Scene Analysis, Computer Graphics and Image Processing, Vol.14 No.3, pp.365-372.

© Giulio Sandini

iit ..from mathematical models to embodied (robotic) implementation (e.g. gaze control and reaching)

...visual perception and eye movements...
...gaze is not only control...

1997

© Giulio Sandini

iit ...from general purpose vision to embodiment and multimodality

...the "what and where" pathways

The end of general purpose vision: there is a stream of visual processing specialized for action execution

Where
What

Disorders in the where pathway (optic ataxia) impair, for example, manual reaching and grasping under visual control

...neuroscience research evolved from the study of isolated functions to the study of functional integration through actions...

© Giulio Sandini

iit From Babybot to iCub to investigate the development of intelligent behaviour...

Babybot (1998)

James (2002)

Robotcub (2004)

iCub The Robotcub Project
Open Software, Open Hardware, open mind

EU Project
11 Partners
9 M€

@UNIGE → IIT

© Giulio Sandini

RobotCub Project
1 Sept 2004 – 31 Jan 2010

iCub The RobotCub Project
Open Software, Open Hardware, open mind

A multidisciplinary approach

Consortium

Name	Contact
University of Genova - LINA-Lab, Department of Informatics, Electronics & Telecommunications, Genova - Italy	Giulio Sandini, Giorgio Baldi, Luca Natale
Univ. Roma Tor Vergata - Department of Computer Science, Rome - Italy	Paolo Dario
University of Pisa - Artificial Intelligence Lab, INFN Lab - Pisa - Italy	Paolo Fierro
University of Padova - Department of Psychology, Padova - Italy	Stefano Iacono
University of Ferrara - Department of Biomedical Sciences, Ferrara - Italy	Luca Maria Falga
University of Pisa - Department of Computer Science, Pisa - Italy	Stefano Schiavon
University of Bologna - Department of Computer Science, Bologna - Italy	Stefano Schiavon
IST - Center for Cognitive Vision and Robotics Lab, Lecce - Italy	Stefano Schiavon
University of Sheffield - Sheffield - United Kingdom	Alan Clegg
Robotics Programme - Faculty of Engineering - Queensland University of Technology - Australia	Andreas Koller
University of Bremen - Bremen - Germany	Stefano Schiavon
University of Bremen - Bremen - Germany	Stefano Schiavon

iCub: from hands to facial expressions...

Software Architecture

<https://confis.europa.eu/project/rcn/71867/factsheet/>

© Giulio Sandini

Since then ... a lot of learning...

- Learning to Control Gaze**
- Learning gravity...**
- Learning to coordinate sensory...**
- The system has learned to direct gaze using arm's proprioception**
- Learning the effects of actions**

Babydoll point of view

© Giulio Sandini

...a lot of "individual" learning...

- Learning to track faces**
- Learning to repeat actions**
- Learning to name objects**
- Zero force control**
- Learning object's affordance**
- Learning to sit**

© Giulio Sandini

...a lot of friends...

S/N 41 by end '19

...a lot of science

...a lot of Software ...and counting..
5 million lines of code (excluding comments...)

A Roadmap for Cognitive Development in Humanoid Robots

© Giulio Sandini

Robot Bodies Have Evolved...

1996: Babybot 2001: James 2008: iCub



Babybot trying to reach the red object

© Giulio Sandini

.....and still do...

Better set-ups



Better Control



Better Sensors



Daniele Pucci Francesco Nori

© Giulio Sandini

...today we have....

- Better Bodyware**
 - Better and cheaper sensors (touch, audition, vision...)
 - Better and more powerful and accurate motors
 - More powerful processors
- (A bit) More Knowledge about the Mindware**
 - Human sensorimotor coordination and development
 - Machine Learning and Motor Control
 - Embodiment and Cognitive Architectures

...keeping in mind that there are still important differences...

While physical growth and morphological change is an essential ingredient of nature's smart solutions

$r = ae^{b\theta}$

Artificial systems are assembled and then «turned on»

© Giulio Sandini

...I think we have enough *bodyware* technology to advance key aspects of research on cognitive robotics...

...but I also think we need to refocus our research on developmental robotics

© Giulio Sandini

...robots are getting closer to us...

Robotic Technologies

time

Augmentation tool

manufacturing tool

...a guide

...collaborator?

...interviewer

...mannequin

...team mate

...actor

© Giulio Sandini

Sometimes it looks like we are (close to) delivering...

Personal Robots

Self driving cars

...and the scientific communities have contributed to create the illusion that robots will soon be among us in the form of autonomous cars, personal robots and robot companions.

© Giulio Sandini

...the reality: failure of "personal robots" ... and cobots as well...

Working next to humans

vs.

Working with humans

... Collaboration ... is the close proximity at which our collaborative robot (cobot) work hand-in-hand with operators...."

© Giulio Sandini

Athletic abilities of robots have evolved faster than their ability to understand...

...robot bodies have progressed a lot...

...we communicate with technology using technology's own language ...

...exchanging signals without understanding...

hieroglyphs of the third millennium

...ancient gestures assume new meaning

Still very much exploiting the adaptability of human mind

© Giulio Sandini

iit **There is a mismatch between *acting* and *understanding***



...due to the limited robot's ability to read, interpret and adapt to "human skills"...

...the limited ability to understand intentions and internal states of humans

Most failures of "personal robots" are due to exaggerated expectations about robot "understanding" not about robot mobility and/or sensing

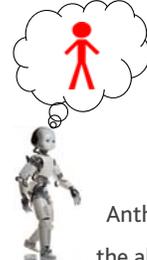
© Giulio Sandini

iit **...we need to make explicit the relationship between the body of the actor and the mind of the observer (and viceversa...)**

The central question is how to map visual input into "internal" motor representations for perception and from intention to motor representation in the case of production.

© Giulio Sandini

iit **...need to reduce the asymmetry between the two sides of anthropomorphism**

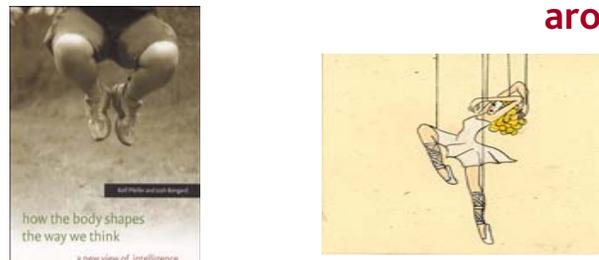


Anthropomorphism of the mind
...the ability to relate to an anthropomorphic body (to *think* like a human)
Hidden/Covert humanoid

Anthropomorphism of the body
... the ability to sense and to move like a human
Explicit/Overt humanoid

© Giulio Sandini

iit **...exploiting the strings between our body and how we interpret the world around us**



Embodiment!!

Rolf Pfeifer and Josh Bongard (2006)

© Giulio Sandini

We interpret the world using an anthropometric geometry

... to canonical neurons in F5 representing visual properties of objects in motor terms

Active when manipulable objects are presented visually (grasp specific)

It is not the shape of the object that matters but how you can grasp it

Observation only

Observation + action

From: Fadiga, L., Fogassi, V., Gallese, V., Rizzolatti, G. (2000) 'The mirror neuron system: a hypothesis on the basic organization of human action-cognition', *Journal of Cognitive Neuroscience*, 12(1), 179-191.

Abstract geometrical shape does not count, does not have a survival value...

Kinematics of the hand is the metric of shape...

© Giulio Sandini

We interpret the others using the knowledge of our own body

...to mirror neurons where visual actions are coded in motor terms

Active when a monkey does an action and also when another individual is seen performing the same manipulative gestures

Motor Resonance

From: Fadiga, L., Fogassi, V., Gallese, V., Rizzolatti, G. (2000) 'The mirror neuron system: a hypothesis on the basic organization of human action-cognition', *Journal of Cognitive Neuroscience*, 12(1), 179-191.

the better we act the better we understand...

Vision activates the motor system...

© Giulio Sandini

A robot does not need an anthropomorphic body to act on objects

The Jaeger-Lipson coffee balloon gripper.

Even if it may help to use tools designed for humans...

However...

© Giulio Sandini

...an anthropomorphic mind is necessary to imagine the affordances of objects built for humans ... (e.g. to give a meaning to "chairs")

The "concept" of chair is an action not a shape

A chair is "something I can sit on"...

© Giulio Sandini

...an anthropomorphic mind is necessary to imagine the affordances of humans... (to give a meaning to "human")

The "concept" of human is a relationship...




A human is "something like me" ...

© Giulio Sandini

The internal model of ourselves is used as an inverse model to "understand" others

We interpret the others using the knowledge of our own body

...to mirror neurons where visual actions are coded in motor terms

the better we act the better we understand...



...the advantage of being conspecific...

...without shared internal models would be like living in a world populated by aliens...

(*) e.g.

- Kanakogi Y, Itakura S. (2011) Developmental correspondence between action prediction and motor ability in early infancy. *Nat Commun* 2: 341.
- Falck-Ytter T, Gredebäck G, von Hofsten C. (2006) Infants predict other people's action goals. *Nat Neurosci* 9: 878-879.

© Giulio Sandini

...main target is to build a causal relationship between robots and humans...



...to be able to understand humans, to anticipate their intentions and to communicate robot's.

...in order to *understand humans* shape is not essential as long as the robot has an *anthropomorphic mind*

No model, no prediction

"How would a human behave in this situation?"

© Giulio Sandini

...need to address which anthropomorphic features to embed for the robot to become "personal" ...and how



- Speech
- Body Language
- Emotions & Feelings
- Affective signals
- Cognitive Skills
- ...

Which combination is needed to exchange goals and intentions as well as internal emotional states

- Appearance
- Kinematics
- Sensing
- Actuation



...searching the moderate amount of anthropomorphism for personal use...

© Giulio Sandini

...a focus on the role of social interaction as a tool to *develop internal models*

Vygotsky's view of cognitive development makes explicit two Key Issues...

...the zone of proximal development...

Social Interaction

"adult guidance or in collaboration with more capable peers" ...the key role of social interaction

Prospection

"prospectively" ...It is not what you know but where you can go, from where you are..."

... the internal model is built by looking at others but also by self-adapting our own internal model

The role of cognitive development...

Vygotsky, L. (1978): Interaction between learning and development. From: Mind and Society – pp 79-91 – Harvard Univ. Press

...need to study *Development (Cognition, Intelligence) as a Dynamic and Interaction Process...*

Global Dynamics of Development

How each trajectory?

- Why convergence?
- Understand the global structure

High-dim. state space (around 20 years)

... Yasuo Kuniyoshi

Development is not a sequence of disconnected states...

...social interaction and prospection needs systems with intrinsic motivations (internal rewards)...

"...development is about systems with the urge to act and explore" (Claes von Hofsten)

...beyond the action perception loop...

Exploration ↔ Prediction

Intentions (goals & motivation) ↔ Experience (learning & memory)

Study the role and structure of *Internal Models* to develop robots able to act *prospectively*

R. C. Miall and D.M. Wolpert: Forward Models for Physiological Motor Control (1996)

Exploration ↔ Prediction

From inverse models to mental simulation and beyond...

The Mental Simulation Theory of Marc Jeannerod: real movements (overt generation of actions) are consequences of "covert internal simulation" thus simulation and generation of movements share the same cortical and computational basis

Marc Jeannerod: Neural Simulation of Action: A Unifying Mechanism for Motor Cognition, NeuroImage 14, S103-S109 (2001)

The research stream of Predictive Coding and Active Inference

OPEN ACCESS Freely available online July 2009 **PLoS one**

Reinforcement Learning or Active Inference?
Karl J. Friston*, Jean Daunizeau, Stefan J. Kiebel

Neural Networks
Volume 12, Issues 7-8, October–November 1999, Pages 1131-1141

Learning to perceive the world as articulated: an approach for hierarchical learning in sensory-motor systems
J. Tani^a, S. Nolfi^b

“...from reinforcing desired behaviors towards teaching the succession of sensory states that lead to desired outcomes.”

“...combining information from input-driven bottom-up signals and expectation-driven top-down signals.”

“learning to predict the next sensation”

© Giulio Sandini

A brief recap ...

Where to go from here... which models, which tools, which priorities?

- There is a mismatch between acting and understanding
...due to the limited robot's ability to read, interpret and adapt to "human skills"...
- ...need to address which anthropomorphic features to embed for the robot to become "personal" ...and how
• Speech
• Body Language
• Emotions & Feelings
- ...need to study Development (Cognition, Intelligence) as a Dynamic and Interaction Process...
- Study the role and structure of Internal Models to develop robots able to act prospectively
Exploration → Prediction
From inverse models to mental simulation and beyond...
- ...need to address the key role of social interaction as a tool to develop internal models and their relationship
...Model is built by looking at others but also by adapting our own internal model
The role of cognitive development...
...development is not a sequence of disconnected states...

© Giulio Sandini

...which role for deep learning?

Deep Learning: Intelligence from Big Data
Tuesday, September 16, 2014
<https://www.vlab.org/events/deep-learning/>

? Is Deep Learning a way to achieve intelligent systems?

"I personally don't like opacity, so I won't spend my time on deep learning, but I know that it has a place in the makeup of intelligence"

POSSIBLE MINDS: CHALLENGES FOR CREATING AI
Chapter 2: THE LIMITATIONS OF OPAQUE LEARNING MACHINES
JUDEA PEARL

which model?

© Giulio Sandini

...Deep Learning gives a fragmented view and no causation...

Deep Learning is all correlation and no causation

Your can teach to many system to become "experts" of one thing but putting them together does not make them an adaptive system or a system able to anticipate the effects of actions.

Yoshua Bengio: ... we need to be able to extend "deep learning" to do things like reasoning, learning causality, and exploring the world...

Judea Pearl
The Book of Why
The New Science of Cause and Effect

Judea Pearl

"Causality encodes the invariant elements in the world"

CAUSALITY

"So rather than experimenting with opaque learning machines I am trying to understand their limitation... in the context of causal-reasoning tasks"

© Giulio Sandini

...weak on generalization and transfer...

A Meta-Transfer Objective for Learning to Disentangle Causal Mechanisms
 Yoshua Bengio^{1,2,3}, Tristan Deleu¹, Nasim Rahaman¹, Nan Rosemary Ke¹, Sébastien Lachapelle¹, Oleksa Bilaniuk¹, Aniruth Goyal¹ and Christopher Pal^{1,2}
 arXiv:1901.10912v2 [cs.LG] 4 Feb 2019

"Current machine learning methods seem weak when they are required to generalize beyond the training distribution"

...after all...learning how to use the hands may cover the basis of tool use in general...

"...we would like what has been learned previously to form a rich base from which very fast adaptation to a new but related distribution can take place, i.e., obtain good transfer."




© Giulio Sandini

Swiss Army Knife approach does not scale

Assembling a set of "intelligent" solutions by learning task-specific representations using large-scale data does not produce an intelligent, adaptive system...

...need to refer to a «big picture»



To assemble a puzzle...



...we would like to have the reference picture...



...but we should at least look for a structure...

Need to have an architecture as a reference to develop individual skills (just pick one...)



This is a good idea...

This may be a better one



This does not work....

© Giulio Sandini

...we should consider the different timescales of human adaptation

Timescales of human adaptation: the role of epigenetic processes
 Kuzawa & Thayer
 Epigenomics (2011) 3(2), 221-234

Evolution, development and learning acts at different timescales and systems may need different tools to "acquire experience"

How to Grow a Mind: Statistics, Structure, and Abstraction
 Joshua B. Tenenbaum,¹ Charles Kemp,² Thomas L. Griffiths,² Noah D. Goodman¹

"Hierarchical Bayesian models (HBMs) may answer some questions about the origins of knowledge...but surely some aspects of mental representation are innate"

Cyclic duration		Adaptation	
Years	Mode	Process	
0.00000001	Seconds	Physiologic	Homeostasis
0.0001	Hours		
0.001	Days		
0.1	Months	Developmental	Plasticity
1	Years		
10	Decades		
100	Centuries	Genetic	Natural selection
1000	Millenia		
1,000,000	Millions		

Figure 1. Timescales of human adaptability.

© Giulio Sandini

...the blueprints selected by evolution

21 August 2019

PERSPECTIVE
 A critique of pure learning and what artificial neural networks can learn from animal brains
 Anthony M. Zador¹

"...much of our sensory representations and behavior are largely innate."

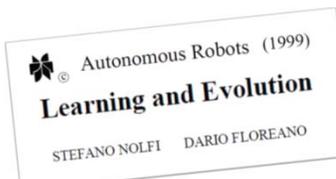
"Thus, it appears that a large component of an animal's behavioral repertoire is not the result of clever learning algorithms — supervised or unsupervised — but rather of behavior programs already present at birth."

"These blueprints have been selected by evolution over hundreds of millions of years, operating on countless quadrillions of individuals. The circuits specified by these blueprints provide the scaffolding for innate behaviors, as well as for any learning that occurs during an animal's lifetime."

the source of big data is not in the experience of a single individual

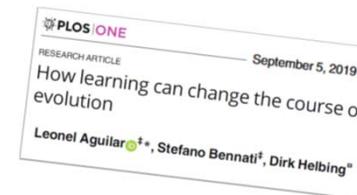
© Giulio Sandini

...revisit the relationship between evolution and learning



Genetically **inherited predisposition to learn** may consist of different things:

1. Presence of **starting conditions at birth** (e.g. initial weights for learning) that canalize learning in the right direction
2. An inherited tendency to behave in such a way that the individual is **exposed to the appropriate learning experiences**



This paper demonstrates that **learning can change the outcome of evolution**, i.e., lead to a genetic configuration that does not correspond to the plastic response.

© Giulio Sandini

Innate functions ... and the shortcuts of evolution...



"The behavioral characteristics of an individual are determined by its genes and by its physical and social environment."

"Not only the individual's early life and current environment is of importance, but also the environment of previous generations"



"Epigenetic mechanisms that support instinct by operating on developmental time scales also support learning by operating on physiological time scales."

"Evolutionary processes can adjust the extent to which behavioral response is learned or instinctive by acting on where, when, and how epigenetic mechanisms operate."

© Giulio Sandini

Deep Learning may be part of the solution (in the framework of developmental robotics)



...**may be** with developmental approach we have been trying to "learn too much" and with the wrong tools...

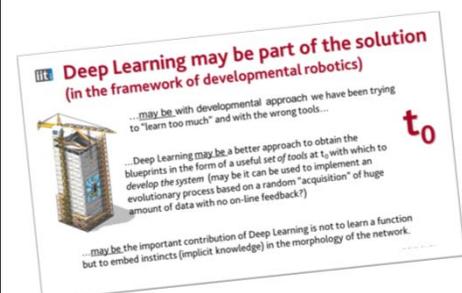
...Deep Learning **may be** a better approach to obtain the blueprints in the form of a useful *set of tools* at t_0 with which to *develop the system* (may be it can be used to implement an evolutionary process based on a random "acquisition" of huge amount of data with no on-line feedback?)

...**may be** the important contribution of Deep Learning is not to learn a function but to embed instincts (implicit knowledge) in the morphology of the network.

t_0

© Giulio Sandini

Need to harmonize the different timescales and tools...



Which tools for evolution?
Which tools for epigenetics?
Which tools for development?
Which tools for learning?

Keeping in mind the limitations of our model but also the fact that we are dealing with an active and social learner

The body (embodiment) is the common reference for *understanding*

© Giulio Sandini

...evolutionary priors

Basic primitives which are coded in the morphology of the brain and body and give a predisposition to learn higher order functions (and are passed from an individual to the next through DNA)

- o Discriminate color (function to distinguish one color from another)
- o Discriminate sounds (function to distinguish different frequencies)
- o Activate muscles synergies
- o **Identify biological motion**
- o **Discriminate faces from other shapes**
- o **Discriminate basic odors**
- o Sensorimotor reflexes...
- o ...

© Giulio Sandini

**Innate Social Skills...
(the bootstrap loader)**

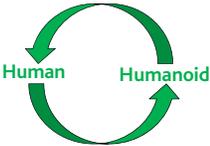
...are represented by a disposition toward "human features" like face-related features, sound of voices, biological motion, emotional states (as basis of a value system)

...and are based on tools such as mouth, face, eyes, hands, voice, body and processed through vision, audition, touch and haptics, proprioception...

→ ...a full human body

© Giulio Sandini

...a humanoid is an ideal tool to study human's internal models and their role in social interaction



Robots used in two ways:

- A model to test theories
- A stimulus to study principles

iCub is a stimulus I can relate with motorically

...use robots to understand humans to build robots

© Giulio Sandini

...study human's internal models to understand social messages to be implemented in robots...



...study human's internal models to understand social messages...

...for Human-Robot Relationship

WHAT IS THE USE OF THE BODY SCHEMA FOR HUMANOID ROBOTS?

INTERFACE

Humanoid after Aristotle: proprioception, shared intention and social interaction through embodied learning operations

Understanding Object Weight from Human and HUMANOID Lifting Actions

Motor Contagion during Human-Human and Human-Robot Interaction

Investigating the ability to read others' intentions using humanoid robots

Revealing the body-schema concept in the context of whole-body postural-focal dynamics

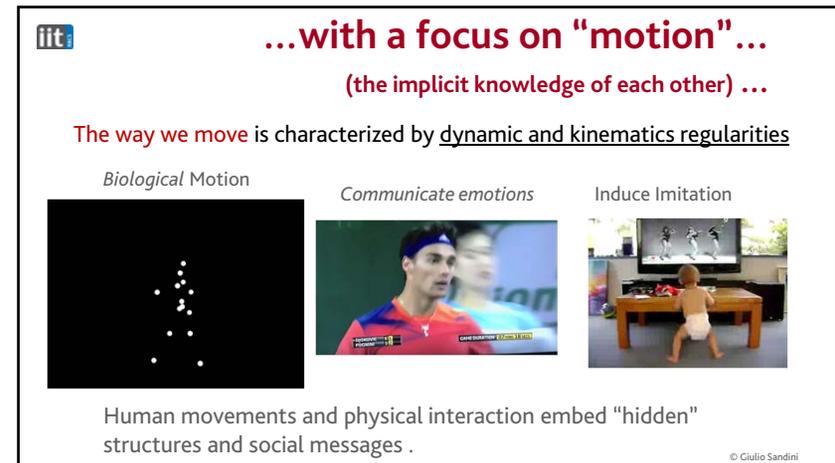
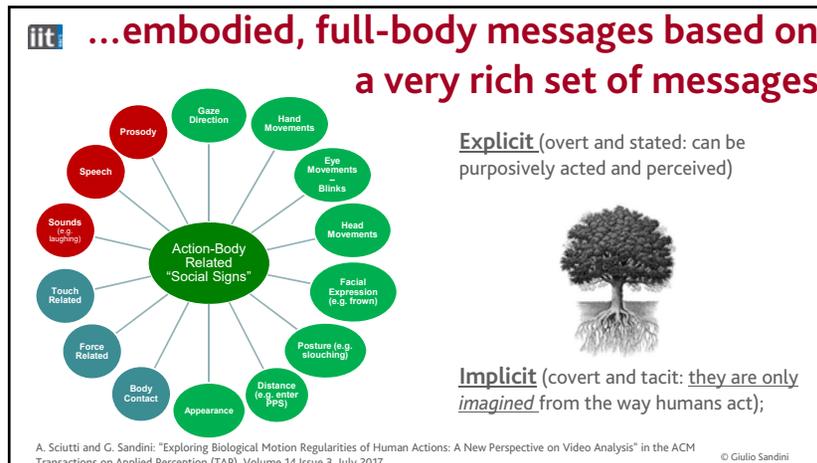
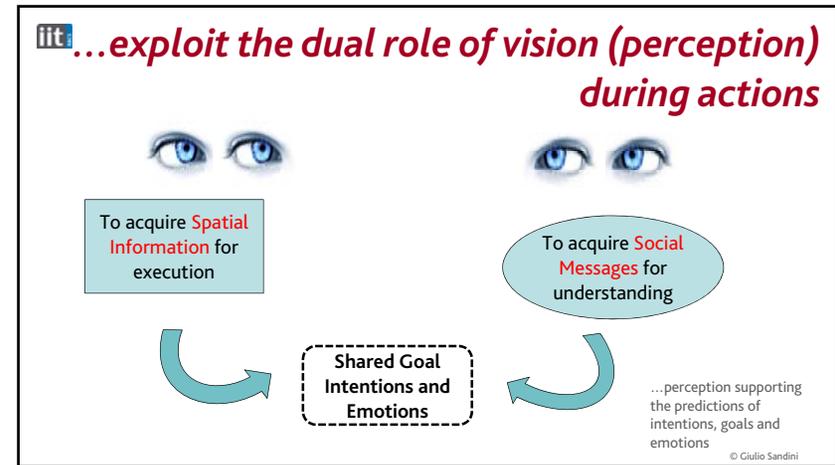
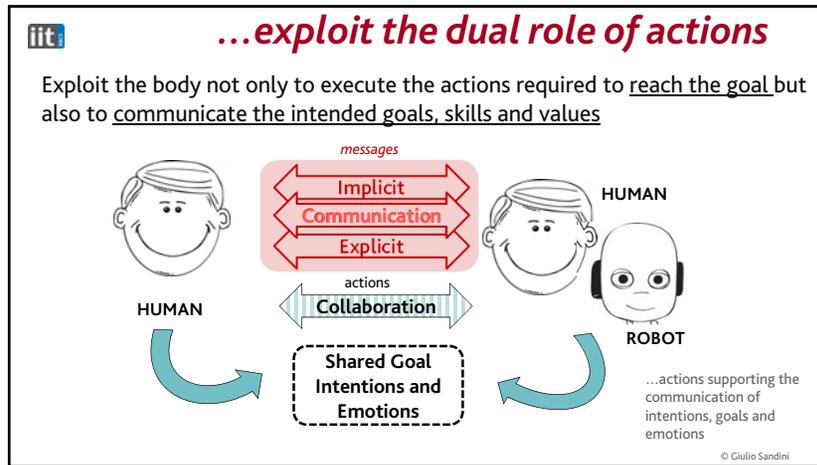
Energy exchanges at contact sensorimotor integration

Development of central dynamical models for humanoid robots

Human ↔ **Humanoid**

!

© Giulio Sandini



How does our "visual brain" transform a stream of images into an object, a "human body", an emotion, an action and a goal to be reached?

...humane interaction has to go beyond the «perception-action» loop ...

Exploration Prediction Experience Motivations

★★★★★

© Giulio Sandini

...it's a team work

Which are the regularities of biological motion affecting the way we communicate with each other

19 females and 18 males!!

How are these regularities mapped into cues perceived during interaction

Alessandra Sciutti

Francesco Rea

Pietro Morasso Gabriel Baud-Bovy Monica Gori Nicoletta Noceti, Francesca Odone, – University of Genoa

© Giulio Sandini

A recent publication with an interesting collection of contribution...

Modelling Human Motion
From Human Perception to Robot Design

Nicoletta Noceti, Alessandra Sciutti, Francesco Rea Editors

Francesco Rea

Alessandra Sciutti

1. Modeling Human Motion: A Task at the Crossroads of Neuroscience, Computer Vision and Robotics
Nicoletta Noceti, Alessandra Sciutti, and Francesco Rea

Part I Motion Perception in Humans

2. The Neurophysiology of Action Perception
Pablo M. Holo, Francisco Castellanos, and Alessandro D'Avella

3. Beyond Automatic Motor Slipping: New Insights into Top-Down Modulation on Action Perception
Alessandra Sciutti, Lucio Soechting, and Cristina Cipriani

4. The Visual Perception of Biological Motion in Adults
Paul Huxner and Yoon Rybaeok

5. The Development of Action Perception
Jenny Christina Siegel

6. The Importance of the Affective Component of Movement in Action Understanding
Giuseppe Di Cesare

Part II Computational Models for Motion Understanding

7. Optical Flow Estimation in the Deep Learning Age
Andrea Flor and Stefan Roth

8. Spatio-Temporal Action Instance Segmentation and Localization
Sumanjaya Ghosh, Gaurav Singh, Michael Sponnagel, Philip H. S. Torr, and Fabio Cuzzolin

9. Vision During Action: Extracting Content and Motion from Manipulation Videos—Event Parsing Human Activity Recognition, Zangpoqiann, Kanika Gopilly, Corina Ferrerillo, and Yannis Alopinos

10. Human Action Recognition and Assessment Via Deep Neural Network Self-Organization
Gertraud I. Faria

11. Movement Expressivity Analysis: From Theory to Computation
Gertraud Varré and Maurizio Mancini

Part III The Robotic Point of View

12. The Promise of Animation in Robotics
Taruja Ribeiro and Ana Paiva

13. Adapting Movement and Behaviour to Foster Communication in Human-Robot Interaction
Karin Lohm, Moritz Inzler, Ahmed, Christian Dandekar, Paolo Arbib, Eric Fahren, and Alessandro Sisti

14. Expressivity Comes First, Movement Follows: Embodied Interaction as Intrinsically Expressive Driver of Robot Behaviour
Carlos Torres-Ferriz and Emilio J. Bevilacqua

15. Gesture in Educational Behavior Coordination: Considering an Executive Robot-Assisted Approach to Didactics
Hajira Lakshmi and Pier Giuseppe Arbib

16. Planning and Timing in Human-Robot Interaction
Allison Langer and Molly Levy-Tzadik

Index

© Giulio Sandini

To establish priorities we can re-visit the taxonomy of action related social sign...

...and take into consideration the contribution of culture, evolution and the laws of physics governing our environment...

Cultural based social signs (acquired through rules, habits and conventions)

Bodily Based Social Signs (embedded in the evolution of human embodiment and in the physics of the environment);

Photoreceptors' Mosaic skeletal and muscle system Gravity and other Physical Laws

© Giulio Sandini

...considering that social messages can be seen often as "side effects" of evolution

The concept of exaptation: the process by which features acquire functions for which they were not originally adapted or selected.

Exaptation-A Missing Term in the Science of Form
 Stephen Jay Gould; Elisabeth S. Vrba
Paleobiology, Vol. 8, No. 1. (Winter, 1982), pp. 4-15.

Exaptations: features that now enhance fitness but were not built by natural selection for their current role

Vs.

Adaptations: features built by selection for their current role

Gould, S., & Vrba, E. (1982). Exaptation-A Missing Term in the Science of Form. *Paleobiology*, 8(1), 4-15. © Giulio Sandini

...the regularities of biological motion deriving from evolution and "physics"

- Shape of the movement's velocity profile;
- Relation between movement's trajectory curvature and velocity;
- Articulated structure of the human body
- Influence of gravity on trajectory;
- ...

Human Movement: Velocity Profile

Morasso Mussa-Ivaldi 1982,

2/3 Power Law

Laquaniti, Terzuolo, Viviani (1983)

© Giulio Sandini

Regularities of biological motion

The movement's velocity profile (a biological motion signature)...
 Bell shaped velocity profile in reaching movements produce perspective invariant features on the retina of the actor and of the observer.

From: P. Morasso - Exp. Brain Res. - 1981

From: Adelman & Hollerbach - J. Neurosci 1985

Also: (Morasso Mussa-Ivaldi 1982, Abend, Bizzi, Morasso, 1982)

Have not evolved to embed "messages" but to optimize performance/energy relationship of biological actuators

...relationship between angular velocity and curvature

...we can "see" the human... with no human in sight

2/3 power law $a(t) = k c(t)^{2/3}$

In rhythmic drawing movements, Laquaniti et. al. noted a power law relationship with proportionality constant k between the angular velocity a(t) of the hand and the curvature of the trajectory path c(t)

From: Laquaniti, Terzuolo, Viviani - Acta Psychologica - 1983

© Giulio Sandini

Biological vs Non-Biological detector

...we can "see" the human with no human in sight

	Lift	Gesticulate	Occlusion	No skin visible
Biological				
Uncertain				
Non biological	Pendulum	Car	Tool Use	shadow

Even in difficult situations

Vignolo, A., N. Noceti, F. Rea, A. Sciutti, F. Odone, and G. Sandini, (2017) Frontiers in Robotics and AI: Detecting Biological Motion for Human-Robot Interaction: A Link between Perception and Action.

© Giulio Sandini

Velocity Profile and Action Reading

velocity profile communicates forces...

HUMAN **ROBOT**

Tuning the velocity profile communicates the weight of lifted objects (we can "see" forces!)

We can see forces!

Action reading since childhood

L. Patané

Lifting speed is correlated with object weight...

... and humans are really good at reading weight through motor observation...

Children Adults

Sciutti, Patané, Nori, Sandini: (2014) IEEE Trans. Autonomous Mental Development

A. Sciutti, L. Patané, G. Sandini, Plos ONE (2019)

© Giulio Sandini

Motor Contagion is affected by motion parameters (but not cultural difference)

Action pace across cultures

Vannucci et al. RO-MAN 2017

Spontaneous motor contagion when working together

Vannucci et al. Paladyn. Journal of Behavioral Robotics (in press)

Collaboration with Y. Nagai & H. Lehmann

© Giulio Sandini

Human Motion Understanding

FIGURE 1 | Examples of dynamic instants detected for a variety of actions (ie. lifting, mixing, and sprinkling the salt) and viewpoints. Dynamic instants are marked in red circles, while blue crosses show the dynamic evolution of the action over time. (A) Lifting, (B) Mixing, (C) Sprinkling salt.

Identification of relevant time instants in an observed action, called **dynamic instants**, informative on the partner's movement timing

MoCA) Multimodal Cooking Actions Dataset

Rea F. et al. : (2019) Front. Robot. AI, Human Motion Understanding for Selecting Action Timing in Collaborative Human-Robot Interaction.

FIGURE 2 | A visual representation of the computation of low-level motion features. Optical flow is computed and thresholded to detect the moving region, where this information is collapsed in a centroid associated with the average velocity magnitude. Such value, analyzed of time, is the signal on which dynamic instants are detected.

© Giulio Sandini

Study which movement's parameters encode the style of actions (communicate internal state)

Can robot actions reveal the how?

F. Vannucci

Vannucci et al. HUMANOIDS 2018

iCub elicits the same subjective and physiological changes if motion parameters are "naturalistic"

Yes!

Subjective evaluation

Neural Response: Insular activation

BOLD Signal Change

HUMAN ROBOT

Collaboration with Giuseppe Di Cesare & Giacomo Rizzolatti, U. Parma

© Giulio Sandini

iit

Social messages mediated by eye movements ...

...social use of gaze is an exapted function...

© Giulio Sandini

iit **Retina is space-variant is to optimize its feasibility and survival functions...**

If human cell res...

Need to move the eyes to explore a scene: it is an "exapted message" of where your interest is...a view of your mental state

© Giulio Sandini

iit **Exploiting Gaze for Interaction**

A gaze-contingent robot
Exploiting mutual gaze in turn taking

Subjects do not know that iCub is monitoring gaze direction: they send messages without knowing it

A gaze-contingent robot
Measuring gaze direction to infer intentions

The eye gaze option worked well most of the time.

Palinko, O., Scutti, A., Schillingmann, L., Rea, F., Nagel, V., Sandini, G.: Gaze Contingency in Turn-Taking for Human Robot Interaction: Advantages and Drawbacks, *RO-HUMAN 2015*

Palinko O., Rea F., Sandini G., Scutti A.: A Robot Reading Human Gaze: Why Eye Tracking is Better Than Head Tracking for Human-Robot Collaboration. *Humanoids 2016*.

iit **...extended to multimodal Joint Attention**

- 1 Integration (Audio + Visual)
- 2 Acyclic extraction of a saliency (hot Point)
- 3 Retinotopic response projection into allocentric spatial representation

...what is salient depends on the task being executed...

Gonzalez Billardon J., Scutti A., Tata M., Sandini G., Rea F., "Audiovisual cognitive architecture for autonomous learning of face localisation by a Humanoid Robot", 2020 International Conference on Robotics and Automation

Tata M., Kothig A., Rea F., "A Bayesian System for Noise Robust Binaural Speaker Counting for Humanoid Robots", IEEE International Conference on Robotics and Automation

UNIVERSITÀ DEGLI STUDI DI GENOVA **Dibris**

iit **Facial Expressions ... and Aerodigestive Behaviors**

BEHAVIORAL AND BRAIN SCIENCES (2017), Page 1 of 55
doi:10.1017/S0140525X16000911_0381

Neonatal imitation in context: Sensorimotor development in the perinatal period

Nazim Keven Kathleen A. Akins

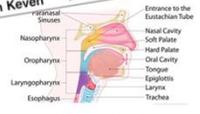


Figure 1. A detailed anatomy of the aerodigestive system.

Beyond aerodigestion: Exaptation of feeding-related mouth movements for social communication in human and nonhuman primates

Lynne Murray,^{a,b,c} Valentina Sclafani,^a Holly Rayson,^a Leonardo De Pascalis,^a Laura Bozicevic,^a and Pier Francesco Ferran^a



"Therefore, while both Tongue Protrusion and Lips Smacking may indeed have aerodigestive origins, they have also been exapted for uniquely social purposes."

"The primary problem of the aerodigestive system is ensuring that air ends up in the lungs and fluids/saliva/masticated food in the stomach"

Keven & Akins, (2017) Behavioral and Brain Sciences

commentary

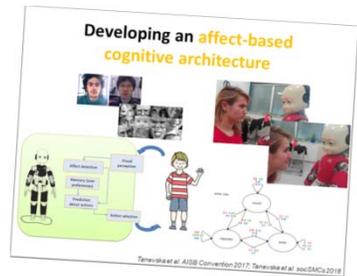
© Giulio Sandini

iit **Reading Expressions and Affect-based Cognitive Architecture**



Reading Expressions

Tanevska et al. AISB Convention 2017: Symposium Computational Modelling of Emotion



Developing an affect-based cognitive architecture

Tanevska et al. AISB Convention 2017; Tanevska et al. SocSciCis 2018

Tanevska et al. AISB Convention 2017: Symposium Computational Modelling of Emotion

© Giulio Sandini

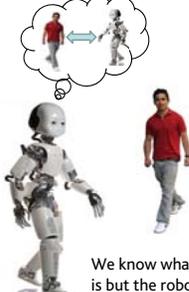
iit **...and so on...to establish mutual understanding and anticipatory behaviours**

...to humanize robots key issue is not acting but understanding (mindware)

...the limit is the lack of mutual understanding...

The limit is robot's ability to read, interpret and complement "human skills"

Understand and model bidirectional intuitive communication in robots



We know what a robot is but the robot does not know who we are...

© Giulio Sandini

iit **Cognitive Interaction Technology is bounded by our lack of knowledge about human's mind ...**



...disruptive technology will be on robots without "threads" behaving and communicating as predictable machines (relating naturally to humans)

© Giulio Sandini

To be fair with the society we have to declare this lack of knowledge and start addressing in parallel how to advance robot technology and how to extend knowledge of humans

..use robots to understand humans to build robots

Human Humanoid

How?

© Giulio Sandini

...to move beyond biomimetic models :
(don't stop at asking how it works...)

Descriptive Model

Explanatory Models Study the Principles

Bio-mimetic

...but study the principles behind... (why it works like that)

...to find different (disruptive?) technologies for artificial systems

© Giulio Sandini

...need to go beyond a superficial level of interdisciplinarity

Scientist Scientific paper Engineer Robot

Doing

Joint Experiments and Robot Design (put the robot in the experiment)

SCIENTISTS ENGINEERS

Need to establish a stronger (real) link between robot design and the behavioural experiment...

...to go from a descriptive to an explanatory role of models (not only the "how" but also the "why")

© Giulio Sandini

...need to join the forces of groups interested in robots able "to understand"!!

...by sharing a "cognitive architecture" not only the bodies and the software architecture

...need to share a «big picture»

To assemble a puzzle... we would like to have the reference picture... but we should at least look for a structure...

Need to have an architecture as a reference to develop individual skills (just pick one...)

© Giulio Sandini

...need to join forces of all actors!!

Each actor involved has to give its own contribution: academia, industries, policy makers, scientific and social media, artists.

© Giulio Sandini

...academia need to nurture the formation of an extended research community...

...to study the central role of anthropomorphism in mediating social interaction and to discuss how we see our society evolving in the direction of a co-existence of humans and artificial systems.

© Giulio Sandini

Summary

Think beyond real time and about the *big picture*

Big picture is missing... beyond real time

Robots to study humans beyond biomimetic models

...move beyond biomimetic models: (don't stop at asking how it works...)

Be fair to society and open to other communities

...academia need to nurture the formation of an extended research community...

...we should consider the different timescales of an adaptive system

...embodied, full-body, conversation based on a very rich set of messages

Timescale of Adaptation and Human-Robot Relationship

Exploration Prediction Experience Intentions

Descriptive Model Explanatory Models Study the Principles

It lo... the scientific communities... contributed to...

...to study the central role of anthropomorphism in mediating social interaction and discuss how we see our society evolving in the direction of a co-existence of humans and artificial systems.

Grazie!

Machines without manuals

Humanizing Human-Robot Interaction