Design and Development of Methodologies, Technologies, and Tools to Support People with Disabilities

Giuseppe AIRÒ FARULLA

Advisor: prof. Paolo PRINETTO

Doctoral Dissertation

Doctoral Program in Computer and Control Engineering

Politecnico di Torino

29th cycle

October the 4th, 2017

The Goal

- Research in the realm of Information Technology for, and **with**, individuals with disabilities
- Enable new interaction paradigms between individuals and machines
  - Empowering and assisting individuals
  - Interaction with Robots

The Means

- Technology can ease and improve everyday life of disabled individuals
- Strict and direct contact with associations and users
- Users actively involved in the overall development process
  - From **end-users** to **begin-users**

My Contributions

	Deafblind Individuals	Rehabilitation Patients	Blind Individuals
Assistive Technologies			
Gesture-based Interfaces			
Optical Character Recognition			

My Contributions - How

- Novel interaction paradigms as Assistive Technology
  - Interfaces to let humans interact with robots in simple ways
- Novel applications of Computer Vision, Image Processing, and Robotics
- Experimental validation

Scientific Outcomes

- Methodologies
  - **Cross-domain**
  - Modular framework for Gesture-based robot teleoperation
  - Non-linear optimization
- Tools
  - Real-time reliable Hand Tracking and Hand Gesture Recognition
  - Segmentation of touching characters in math formulae
  - Released Open Source: https://github.com/guybrush90/OCR
- Publications
  - 4 Journal Papers + 1 to appear
  - 6 Conference Papers

Outline

- Introduction to Disabilities and Inclusive Interfaces
- Hand Gestures-based Interfaces
- Minor Contribution
- Conclusions and Future Research

Outline

- **Introduction to Disabilities and Inclusive Interfaces**
- Hand Gestures-based Interfaces
- Minor Contribution
- Conclusions and Future Research

Individuals with Disabilities

- Have impairments that hinder their full and effective **participation in society on an equal basis with others**
  - UN Convention on the Rights of People with Disabilities
  - World Health Organization
  - European Union

Some Data…

- National level (ISTAT, 2015)
  - **3.1 millions of people** have received a certification of disability (**law n. 104**, 1992)
  - **Only the 3.5% has entered the labor market**

Some Data…

- National level (ISTAT, 2015)
  - **3.1 millions of people** have received a certification of disability (**law n. 104**, 1992)
  - **Only the 3.5% has entered the labor market**
- European level (Eurostat, 2012)
  - **44 millions of people self-reports serious difficulties** in basic daily activities

Some Data…

- National level (ISTAT, 2015)
  - **3.1 millions of people** have received a certification of disability (**law n. 104**, 1992)
  - **Only the 3.5% has entered the labor market**
- European level (Eurostat, 2012)
  - **44 millions of people self-reports serious difficulties** in basic daily activities
- Worldwide level (US Census Bureau, 2010)
  - **Only the 35%** of the working-able individuals with disabilities **has entered the labor market**

Users' Requirements

- Humans eventually need to interact with a machine
  - **Human-Machine Interfaces**
- To fully support the **social inclusion** of individuals with disabilities HMIs should be
  - Dependable

Users' Requirements

- Humans eventually need to interact with a machine
  - **Human-Machine Interfaces**
- To fully support the **social inclusion** of individuals with disabilities HMIs should be
  - Dependable
  - Low-cost

Users' Requirements

- Humans eventually need to interact with a machine
  - **Human-Machine Interfaces**
- To fully support the **social inclusion** of individuals with disabilities HMIs should be
  - Dependable
  - Low-cost
  - Real-time

Users’ Requirements

- Humans eventually need to interact with a machine
  - **Human-Machine Interfaces**
- To fully support the **social inclusion** of individuals with disabilities HMIs should be
  - Dependable
  - Low-cost
  - Real-time
  - Simple to use

Users' Requirements

- Humans eventually need to interact with a machine
  - **Human-Machine Interfaces**
- To fully support the **social inclusion** of individuals with disabilities HMIs should be
  - Dependable
  - Low-cost
  - Real-time
  - Simple to use
  - Designed for Each

Users' Requirements

- Humans eventually need to interact with a machine
  - **Human-Machine Interfaces**
- To fully support the **social inclusion** of individuals with disabilities HMIs should be
  - Dependable
  - Low-cost
  - Real-time
  - Simple to use
  - Designed for Each
  - Designed for All

Design for Each

- **Focus on the individual** to be assisted and empowered
  - Rather than on the solution proposed
- Do blind individuals need a laser cane?
  - (Bolgiano and Meeks, 1967)
  - (Benjamin and Ali, 1974)

Design for All

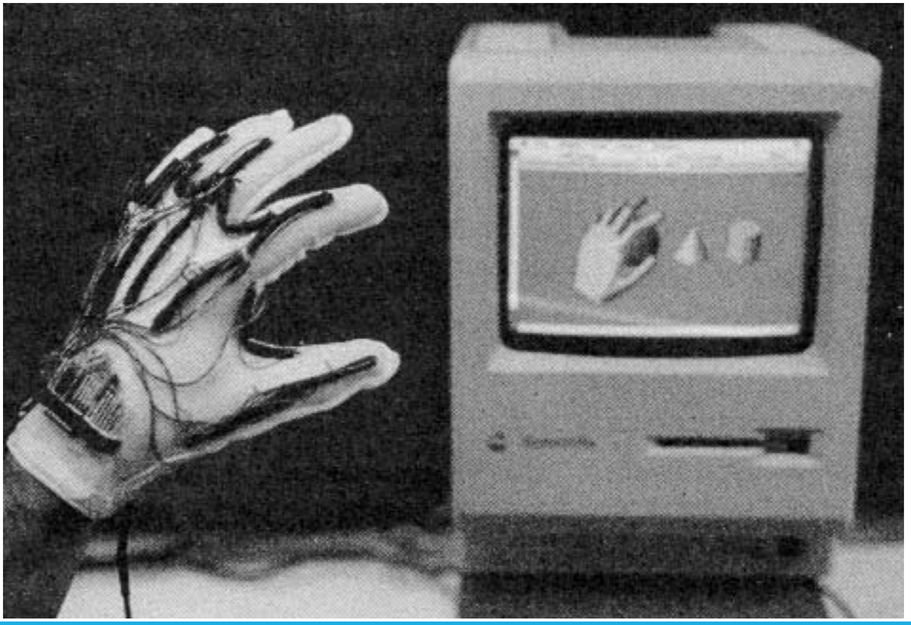
- **Universal design**
  - Concept derived from architecture: barriers-free
- Modern technology is mature to guarantee to **everyone** a **fair interaction with machines**
  - **Gestures** represent a valid alternative to joysticks and keyboards

Outline

- Introduction to Disabilities and Inclusive Interfaces
- **Hand Gestures-based Interfaces**
  - State-of-the-art and Enabling Technology
  - Challenges
  - Proposed Solutions and Their Validation
  - Applications
- Minor Contribution
- Conclusions and Future Research

Gestures

- Voluntary and expressive motion of the body that contain **information** (Kurtenbach and Hulteen, 1990)
- Human individuals find natural to interact via gestures
  - Safety net in foreign contexts
- Does the same hold for **interacting with machines**?
  - Glove to interact with a computer (Zimmerman et al., 1987)



My Research

- Simplify interaction patterns
- Focus on **Hand Tracking**
  - **Simple to use**: No training required
  - **Low-cost** enabling technology

Hand Tracking

Identify *hand pose*

- Set of values to allow unambiguous representation of the global pose a single hand
  - Rotation angles of the hand’s fingers
  - Position and orientation of the wrist with respect to a reference frame

Hand Tracking – The Issues

- 27 degrees of freedom

Hand Tracking – The Issues

- Complex motion compressed in short time
  - Wide movements
  - Movements of small sub-parts
- Self-occlusions

Hand Tracking – Approaches

- *Invasive approaches*
  - Con's: Require training and precise setup
- Enabling technology
  - Electromyography
  - Markers

Hand Tracking – Approaches

- *Non-invasive approaches*
  - Pro's: No constraints on setup
- Classification
  - 2D
  - **3D**
  - Partial tracking
  - **Full tracking**

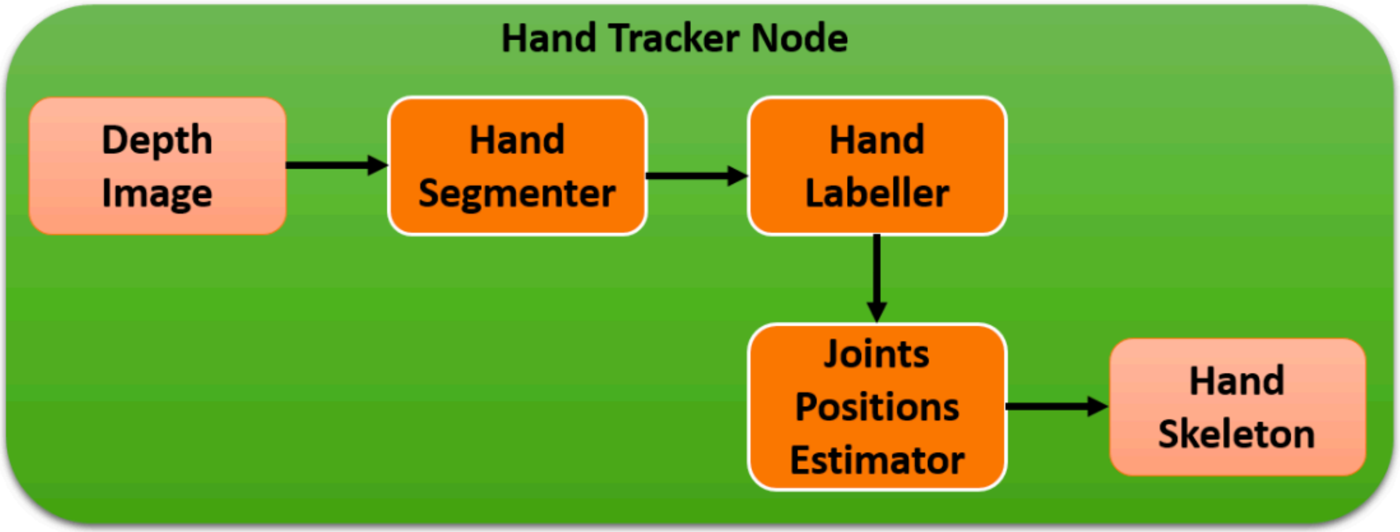
Hand Tracking – The Devices

- Consumer Depth Sensors
  - Capture **depth information** (distance to the sensor) in a per-pixel basis
  - Microsoft Kinect (RGB + D camera)
  - Leap Motion

Hand Tracking – My Contributions

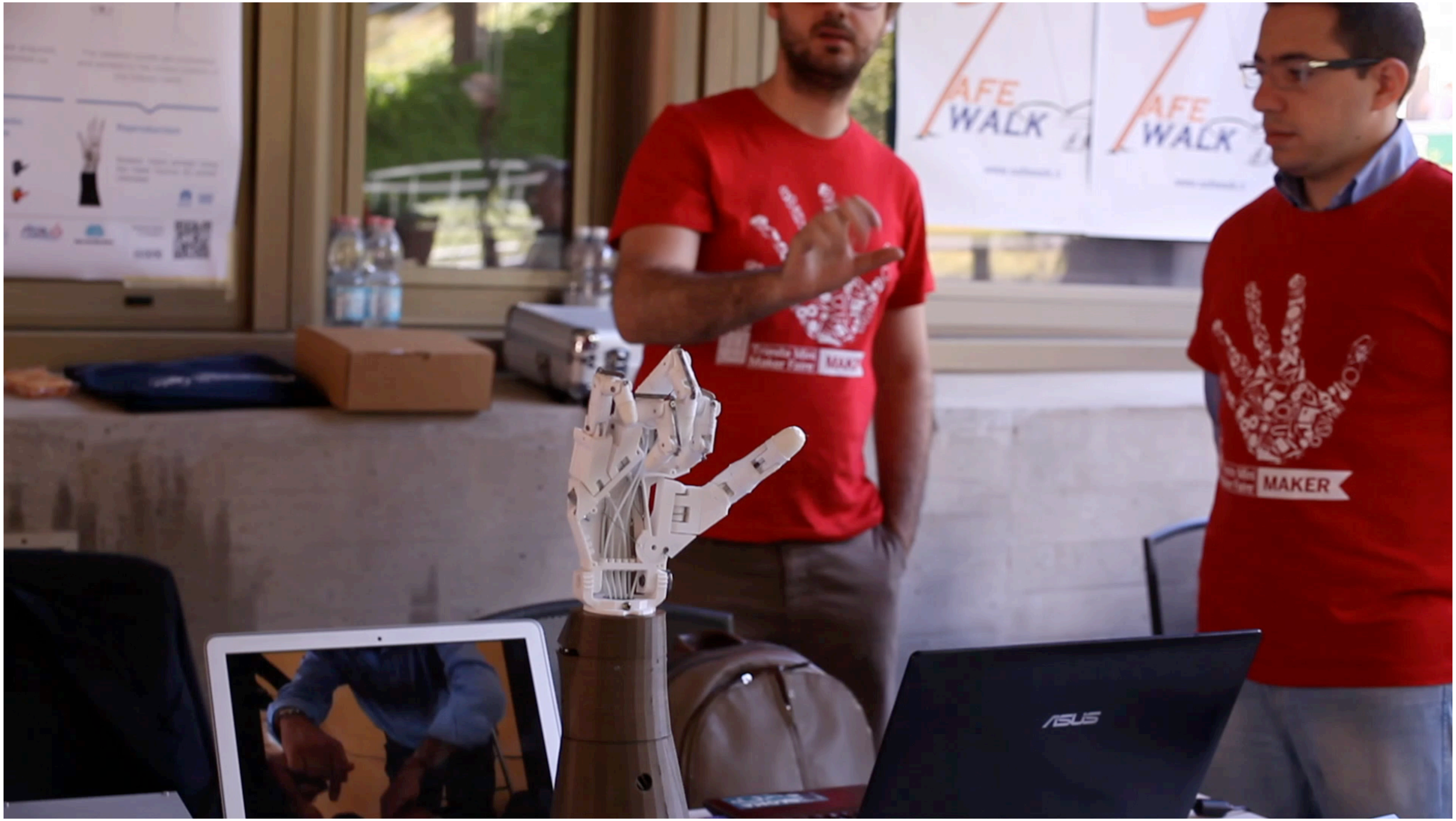
Innovative usage of two techniques

- A *Generative* technique minimizing the difference between the observed hand pose and a 3D hand model
  - **Particle Swarm Optimization** (Oikonomidis et al., 2011)
- A *Discriminative* technique trained for mapping the extracted features to the hand pose
  - **Random Forest** (Shotton et al., 2013) (Keskin et al., 2013)



Hand Tracking – My Contributions

- Enabling **Human-Robot collaboration**

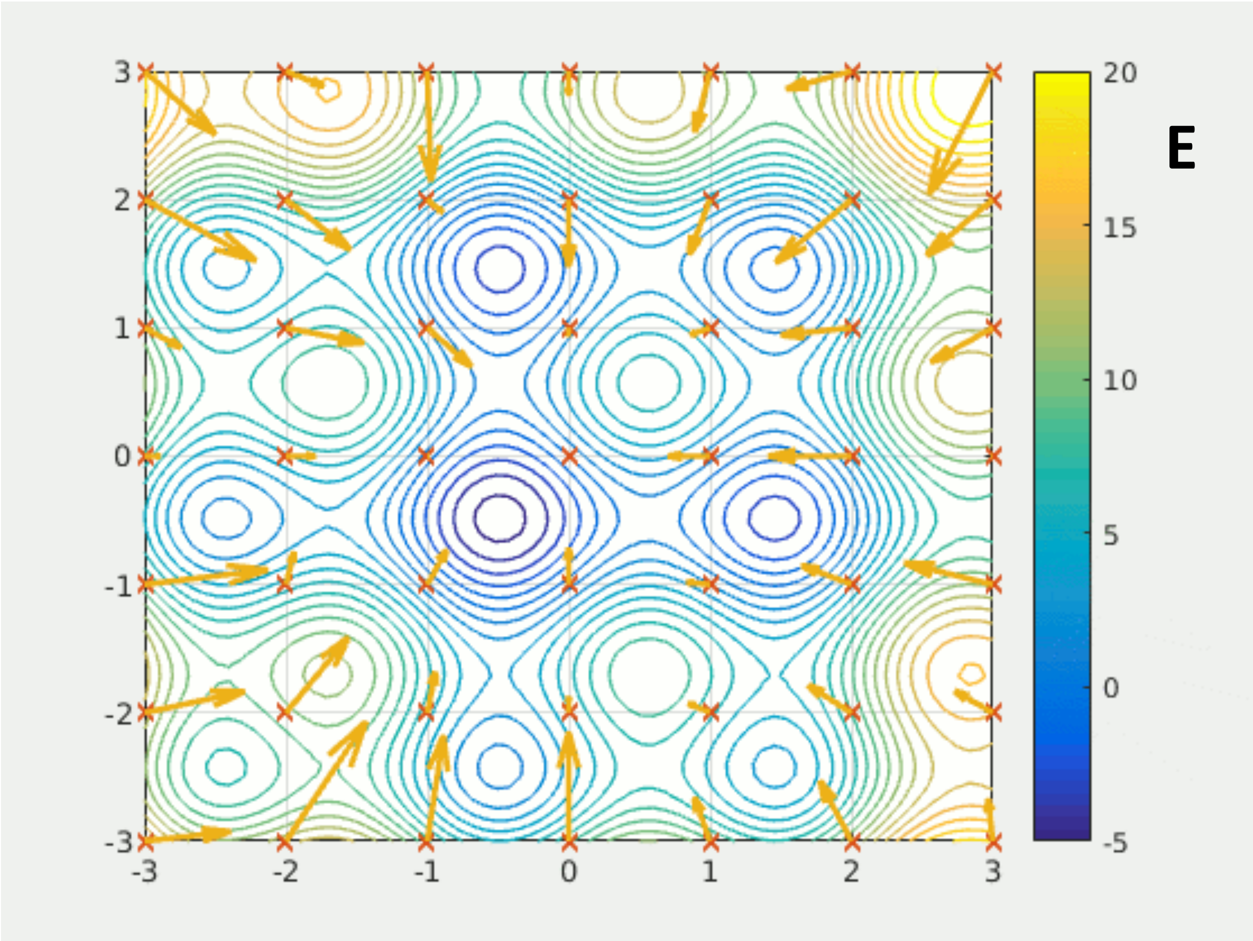


Particle Swarm Optimization

- Swarm of particles with own **positions** and **velocities**
- Positions evolve towards the global minimum of a difference function **E**
- Velocities change to reflect
  - *Cognitive components* private to the particles
  - *Social components* shared among the swarm

Particle Swarm Optimization

- Particles: red crosses
- Velocities: yellow arrows
- 3D Example



Generative Solution – Design

- Input
  - Single markerless observation from depth camera
  - Output of previous iterations
- Output
  - Hand pose
- Core processing in C++
  - Python wrappers for modularity

Generative Solution – Implementation

- 27-dimensional hand model
- Difference function **E** to be minimized between hand model **h** and multiframe depth observations **D**

$$E(h,D) = \sum_{I \in D} G(I,h,C(I)) + \lambda_k \cdot kc(h)$$

$$G(I,h,C) = \frac{\sum o_s(I) \oplus r_s(h,C)}{\sum o_s(I) + \sum r_s(h,C) + \varepsilon} + \lambda \frac{\sum o_d(I) \cdot r_s(h,C)}{\sum r_e(h,C) + \varepsilon}$$

Generative Solution – Validation

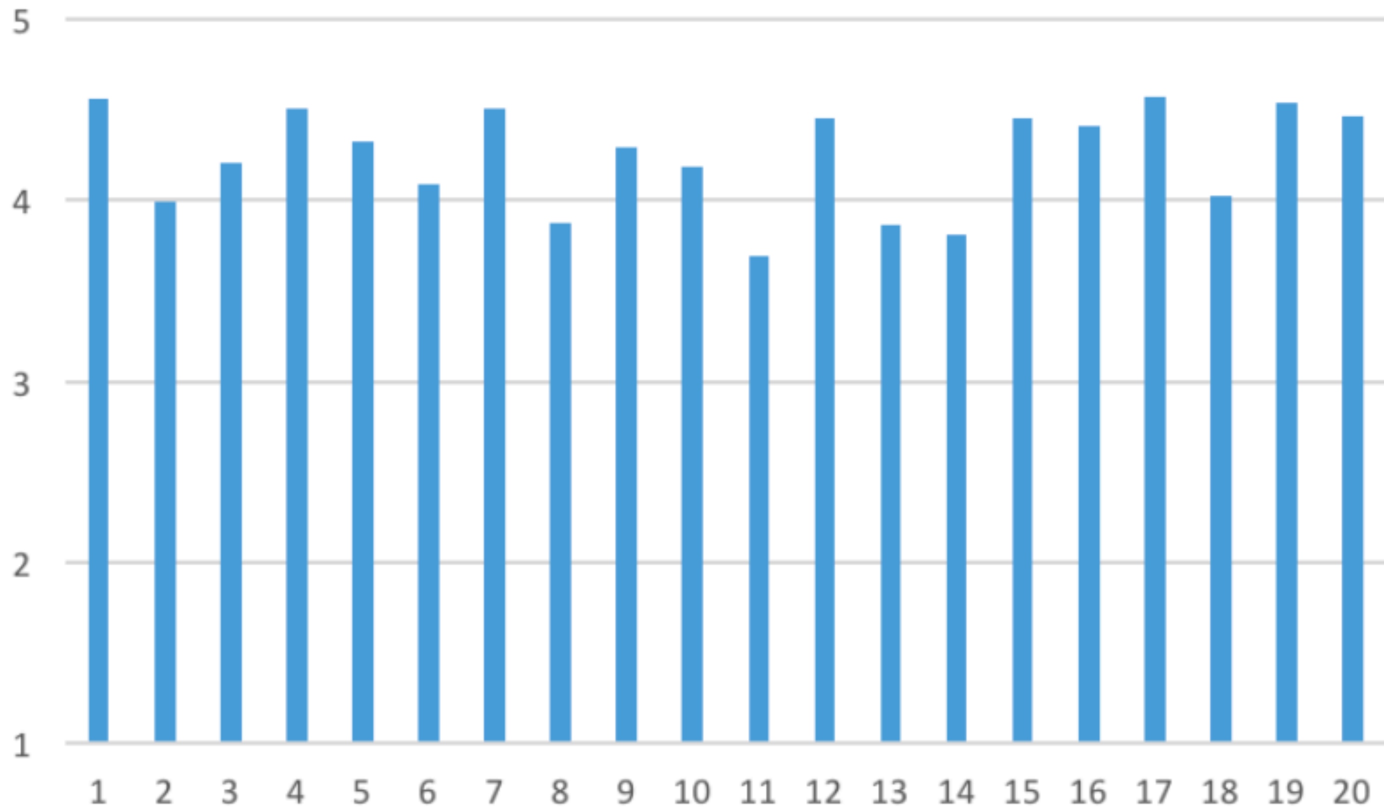
- Qualitative analysis
- Test if the human eye perceives the tracking as proper
  - If so the tracking algorithm is accurate enough

Generative Solution – Validation

- **20 fundamental hand poses** (Brentari, 1998)
- Survey to **rate tracking accuracy** in a 1-to-5 scale
  - Average over **87 anonymous volunteers**
  - No previous knowledge on the system

Very Accurate

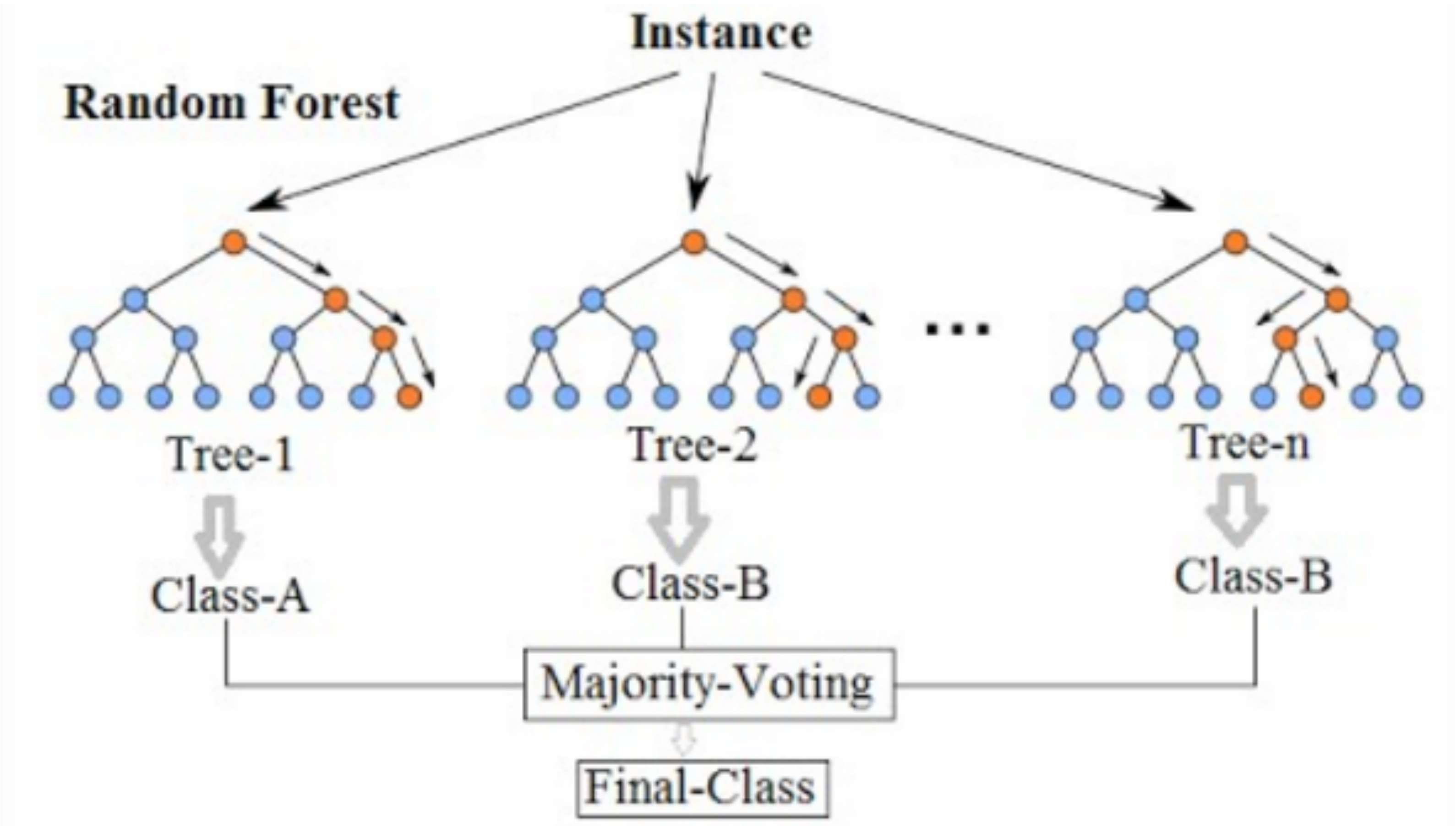
Poor



Generative Solution – Validation

- Pro's
  - Very accurate
  - Does not require to learn hand poses
- Con's
  - Performances affected by wrong swarm initialization
  - Does not match real-time constraints

Random Forest

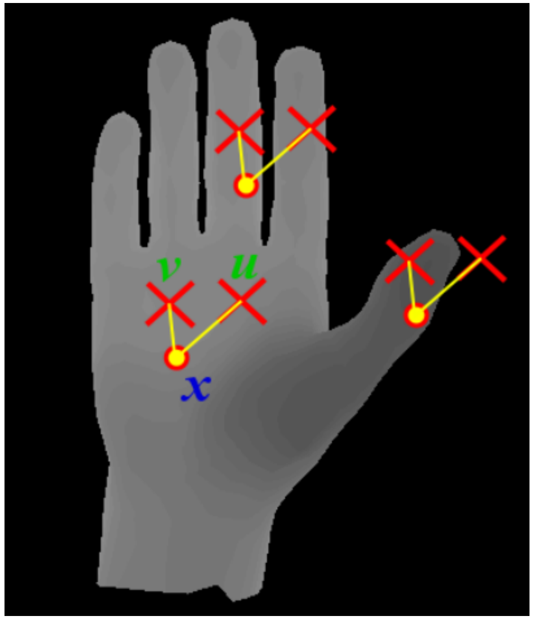


Discriminative Solution – Design

- Input
  - Single markerless static observation from depth camera
- Output
  - Hand pose
  - Recognized gesture
- Core processing in C++
  - Optimization in OpenCL
  - Python wrappers for modularity

Discriminative Solution – Implementation

- The hand is isolated from the background
  - OpenNi hand tracker
- Each pixel **x** is described **on-the-fly** through the **feature**
  - $$\mathscr{F}(\mathbf{x}) = \{F_{\mathbf{u},\mathbf{v}}(D,\mathbf{x}), \|\mathbf{u}\| < R, \|\mathbf{v}\| < R\}\,;\, F_{\mathbf{u},\mathbf{v}} = D\left(x + \frac{\mathbf{u}}{D(\mathbf{x})}\right) - D\left(x + \frac{\mathbf{v}}{D(\mathbf{x})}\right)$$
- Invariant to in-plane and depth translations
  - Not invariant to rotations
  - Large training set

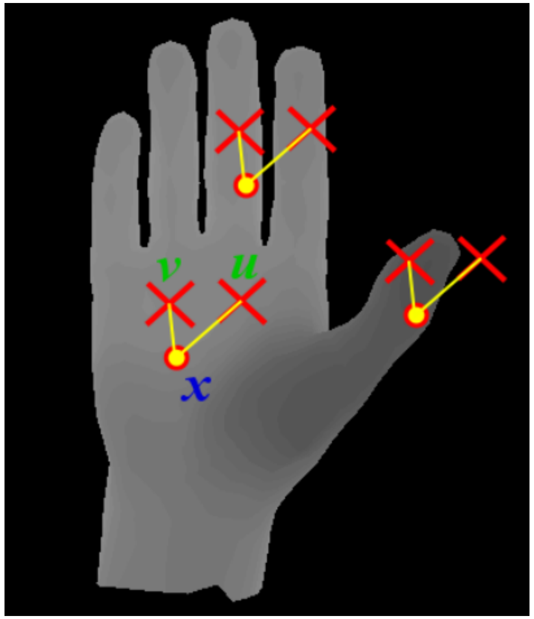


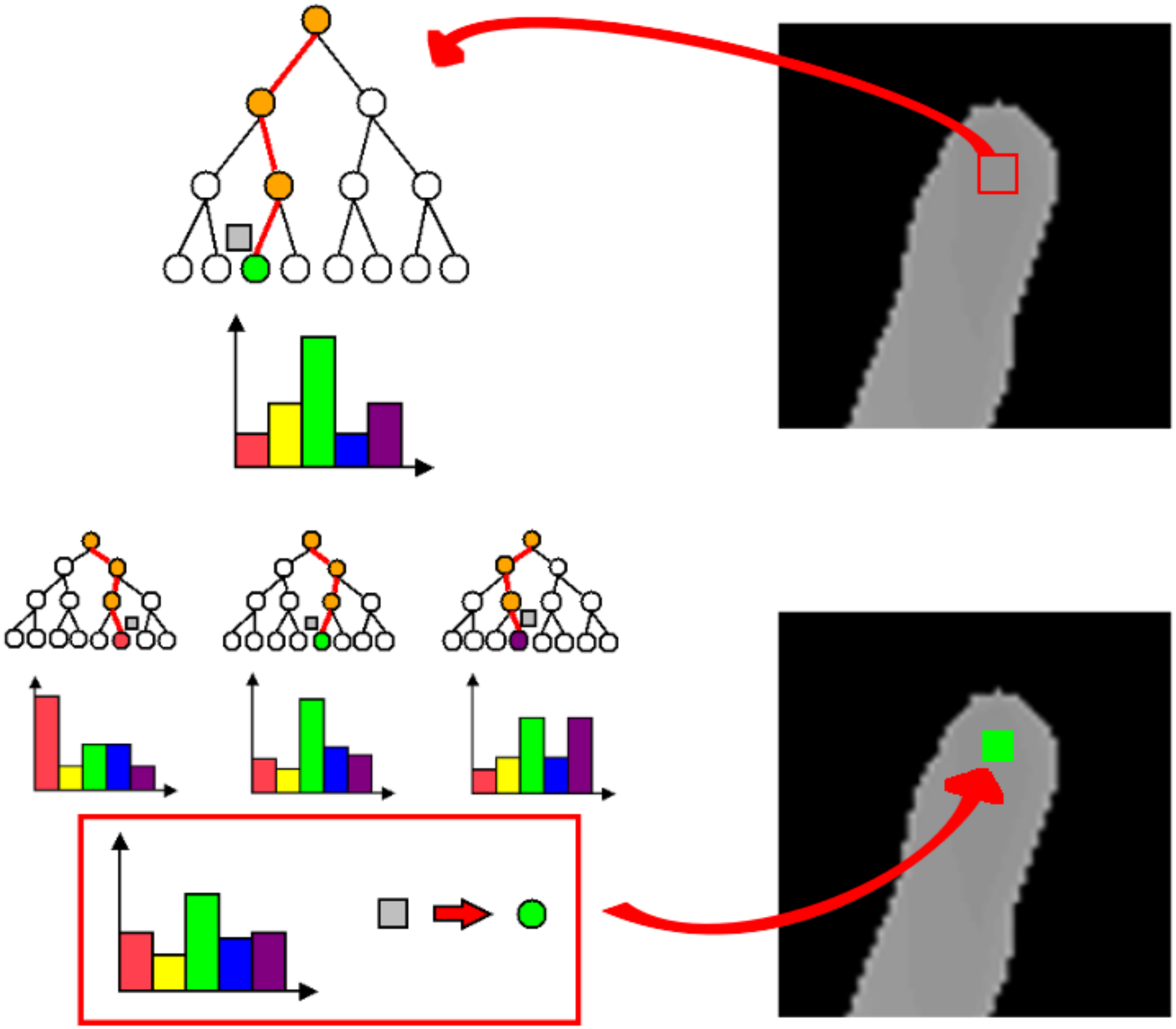
Discriminative Solution – Implementation

- First **Random Forest** classifier
  - Label segmented depthmaps to hand joints
- Mean Shift algorithm
  - Outliers reduction
- Second **Random Forest** classifier
  - Map the hand pose to a gesture
  - Resorts to joint-to-joint Euclidean distances for all the joints pairs

$$\mathscr{P}(\mathscr{S}) = \{d_{k,l} = \|\mathbf{j}_k - \mathbf{j}_l\|, \forall \mathbf{j}_k, \mathbf{j}_l \in \mathscr{S}, k < l\}$$

Discriminative Solution – Classification

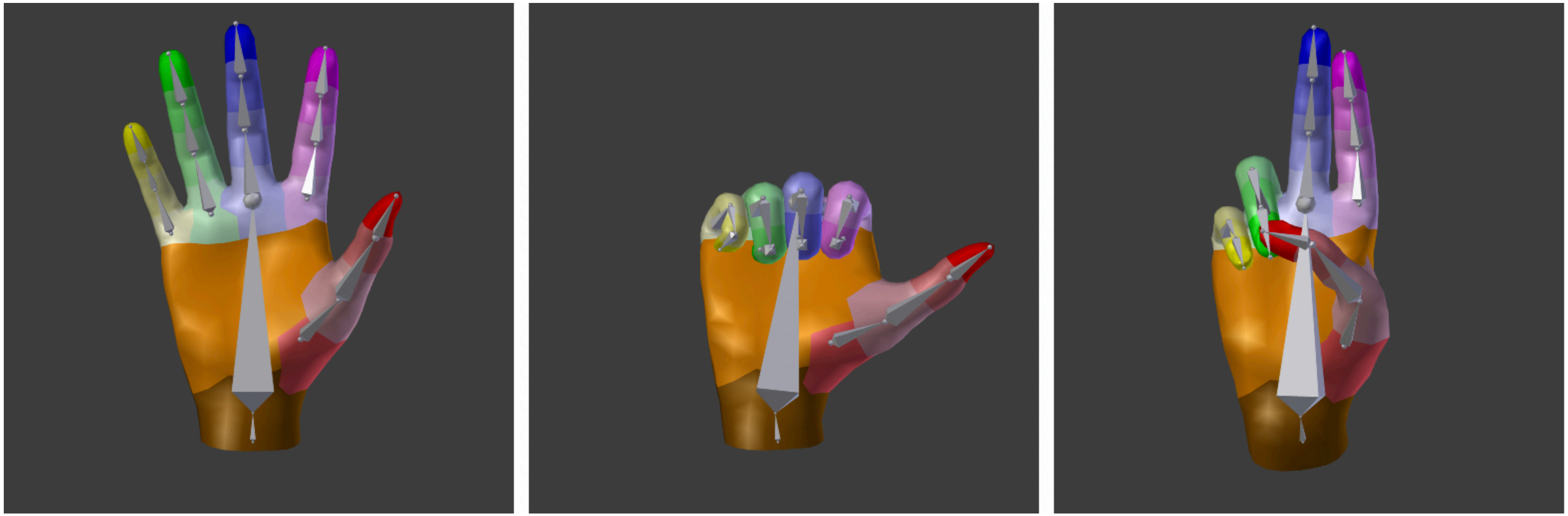




Parameter	Value
R	30 pixel
Threshold τ	10
Sample pixels per image	2,000
Trees in the forest	3
Depth of each tree	18

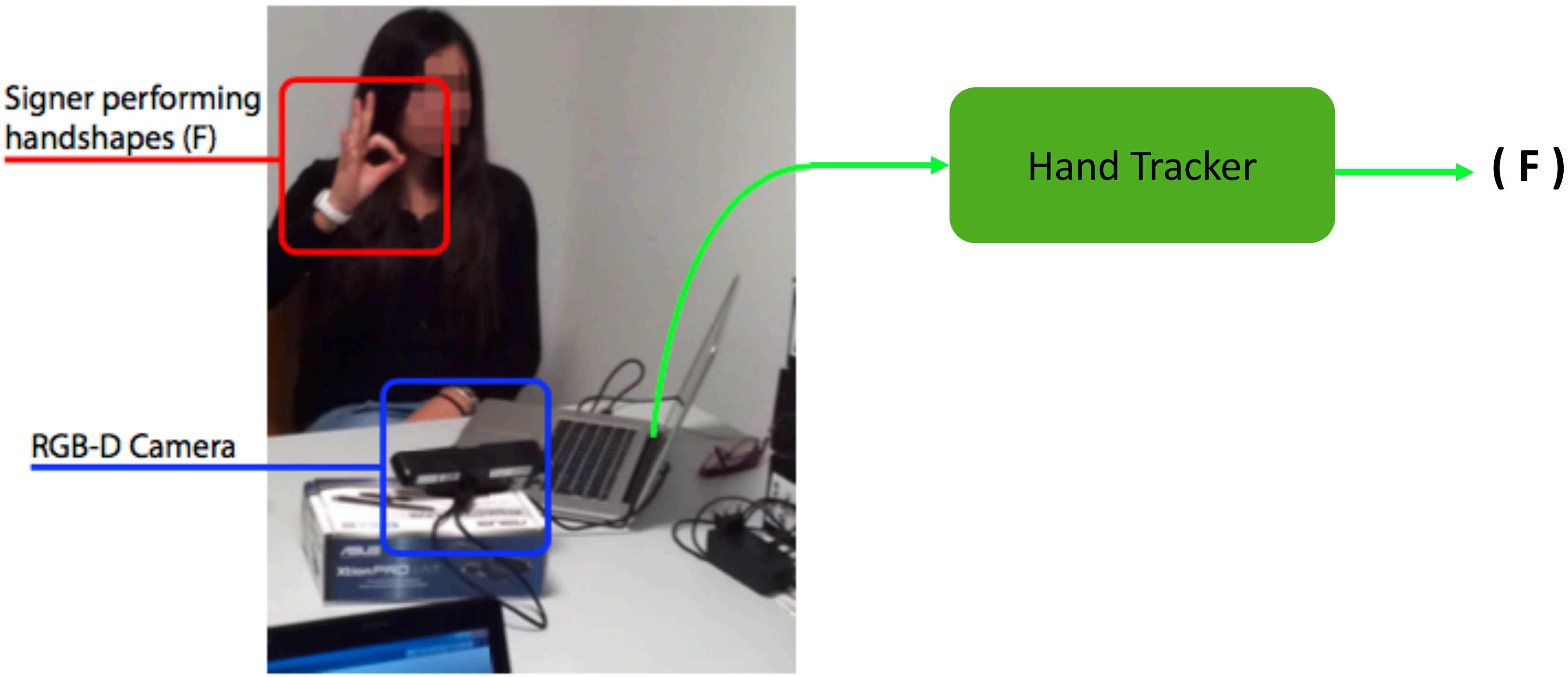
Discriminative Solution – Training

- Datasets of human hand poses
- Synthetic datasets
  - Designed with target applications in mind



Discriminative Solution – Validation

- Quantitative analysis
- Test if the recognized gesture is correct
  - If so the tracking algorithm and its training are accurate enough



Discriminative Solution – Validation

- **16 fundamental hand poses** (Brentari, 1998)
- **16 volunteers**, no previous knowledge on the system
  - 15 used to train the classifiers + 1 used for testing
  - Average over **10 leave-one-out** cross validation

	S	Y	I	A	O	P	K	L	D	C	H	X	B	W	V	F
S	90	4	0	6	0	0	0	0	0	0	0	0	0	0	0	0
Y	21	44	0	32	0	0	0	0	0	0	0	0	0	3	0	0
I	12	3	66	16	0	0	0	0	3	0	0	0	0	0	0	0
A	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0
O	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0
P	0	0	0	75	8	9	0	0	0	0	0	8	0	0	0	0
K	0	0	0	0	0	39	36	17	4	0	0	4	0	0	0	0
L	7	0	0	0	0	0	0	93	0	0	0	0	0	0	0	0
D	0	0	0	3	0	0	3	10	84	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0	3	0	88	0	0	3	0	0	6
H	0	0	0	0	0	0	0	6	6	0	76	6	0	0	6	0
X	5	0	0	0	0	0	0	0	15	0	0	75	0	5	0	0
B	0	0	0	0	0	0	0	0	0	0	0	0	87	13	0	0
W	0	0	0	0	0	4	0	0	0	0	0	0	8	79	9	0
V	0	0	0	0	0	0	3	0	0	0	0	0	0	9	88	0
F	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	97

Generative Solution – Validation

- Pro's
  - Accurate
  - Matches real-time constraints
- Con's
  - Requires to learn hand poses, although off-line

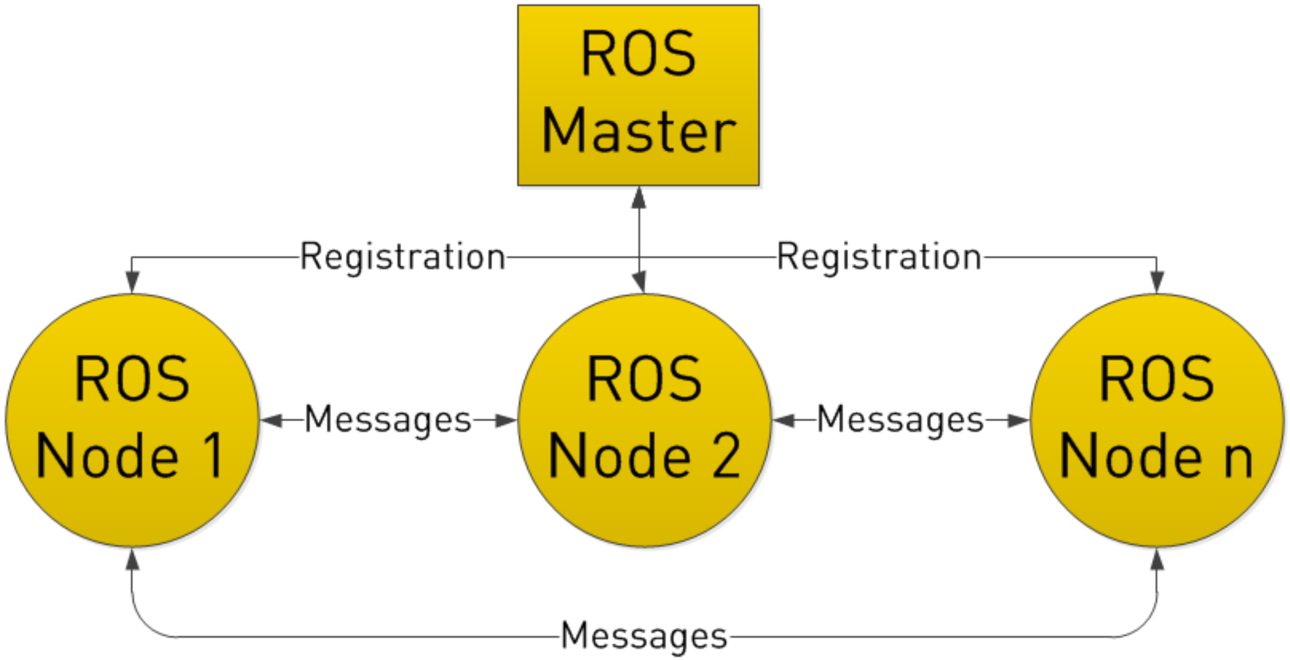
Gestures for Robot Teleoperation

	Deafblind Individuals	Rehabilitation Patients	Blind Individuals
Assistive Technologies			
Gesture-based Interfaces			
Optical Character Recognition			

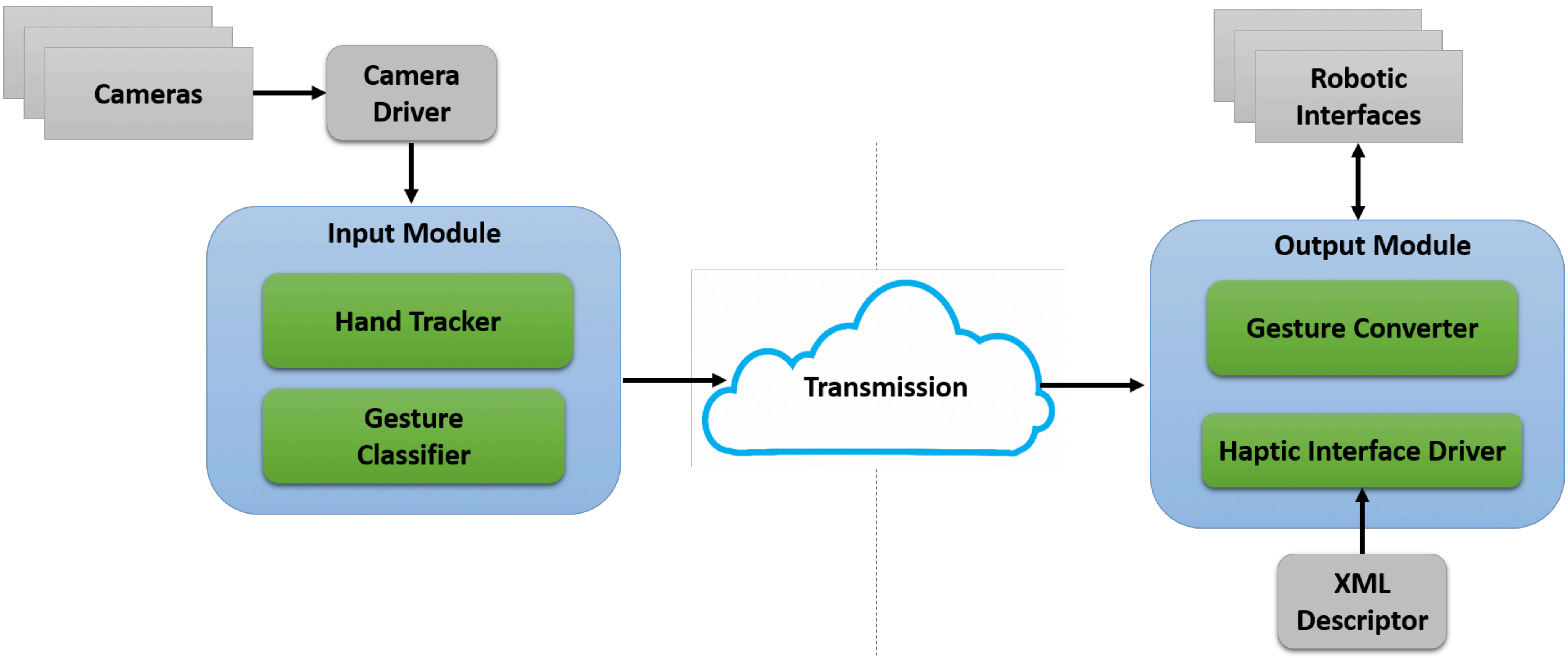
Gestures for Robot Teleoperation

Robot Operating System

- Real-time control of distributed robotic interfaces through hand gestures
- Network infrastructure: **ROS**

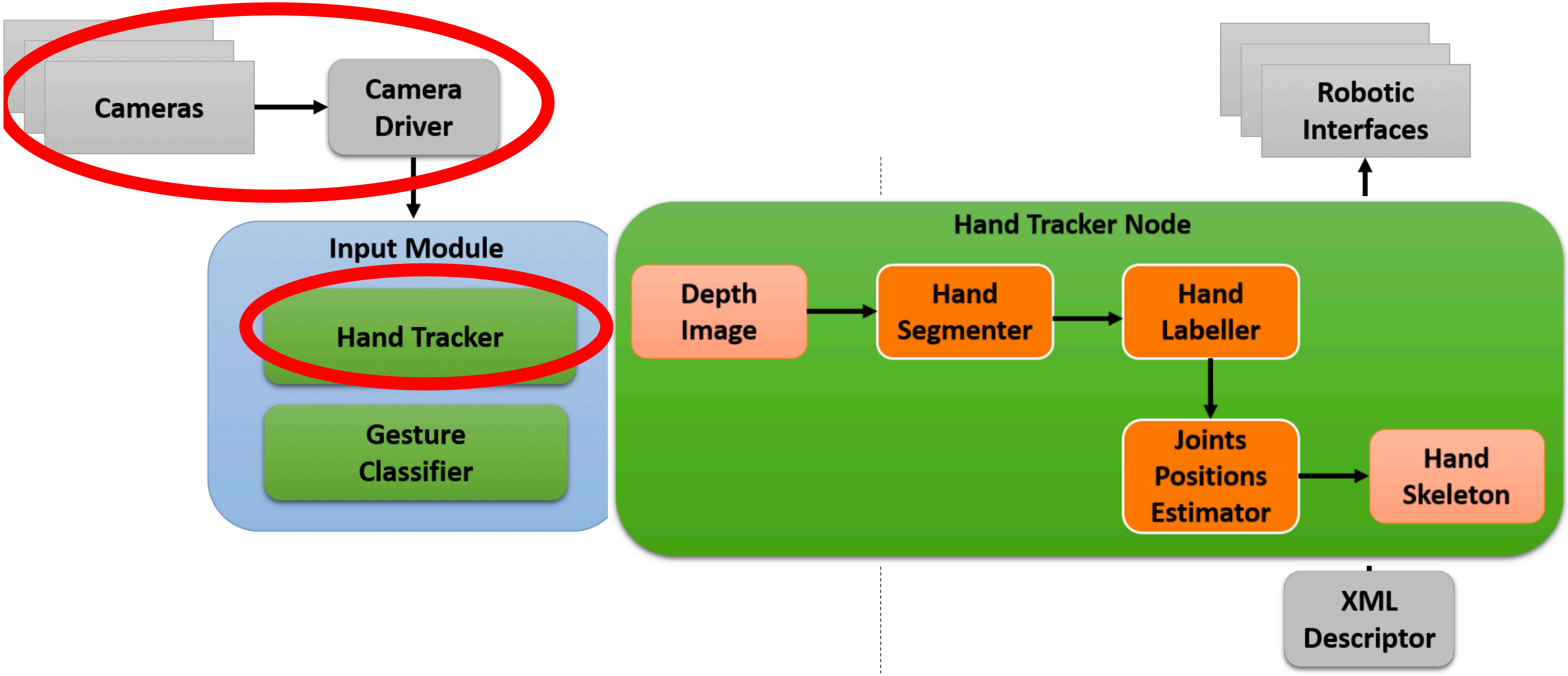


Teleoperation – The Architecture



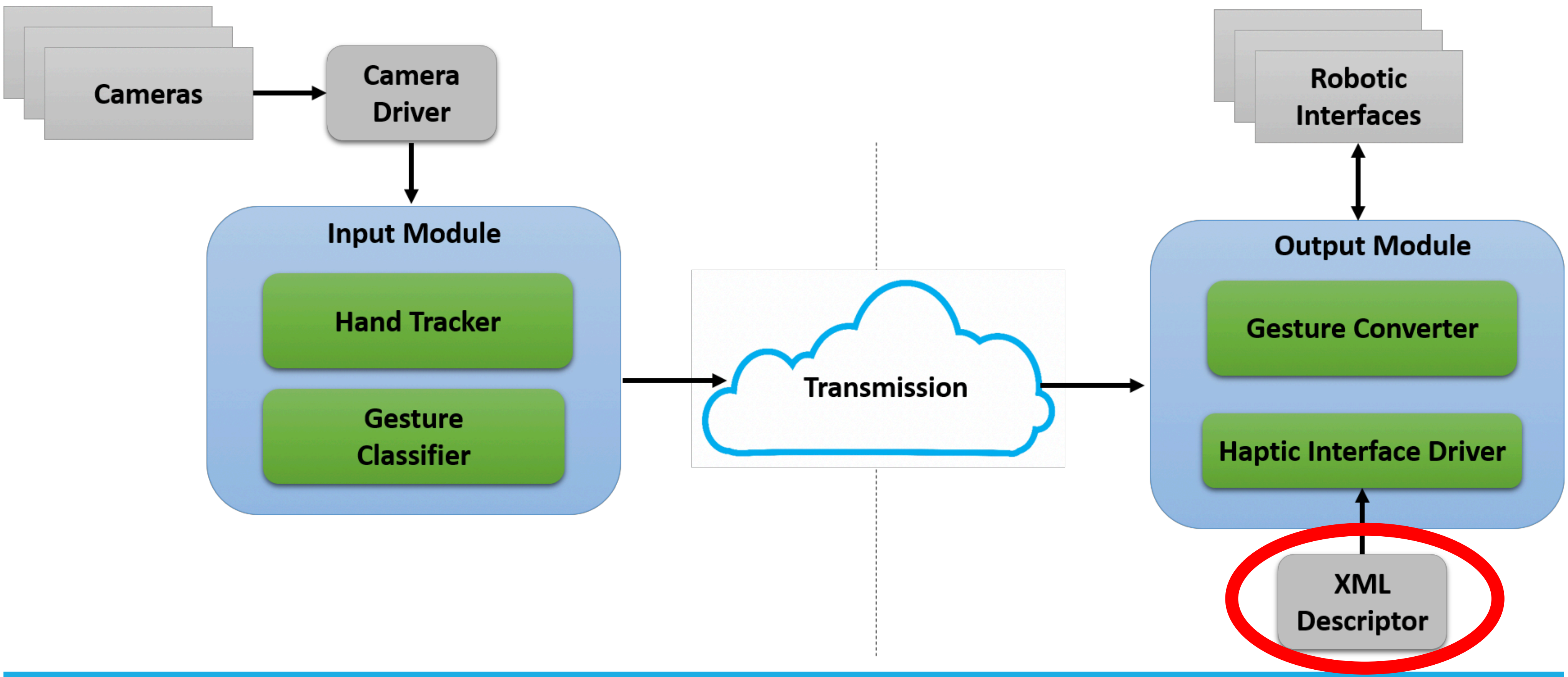
Teleoperation – The Architecture

- Phase I: Input Customization
  - Tuning the Hand Tracker



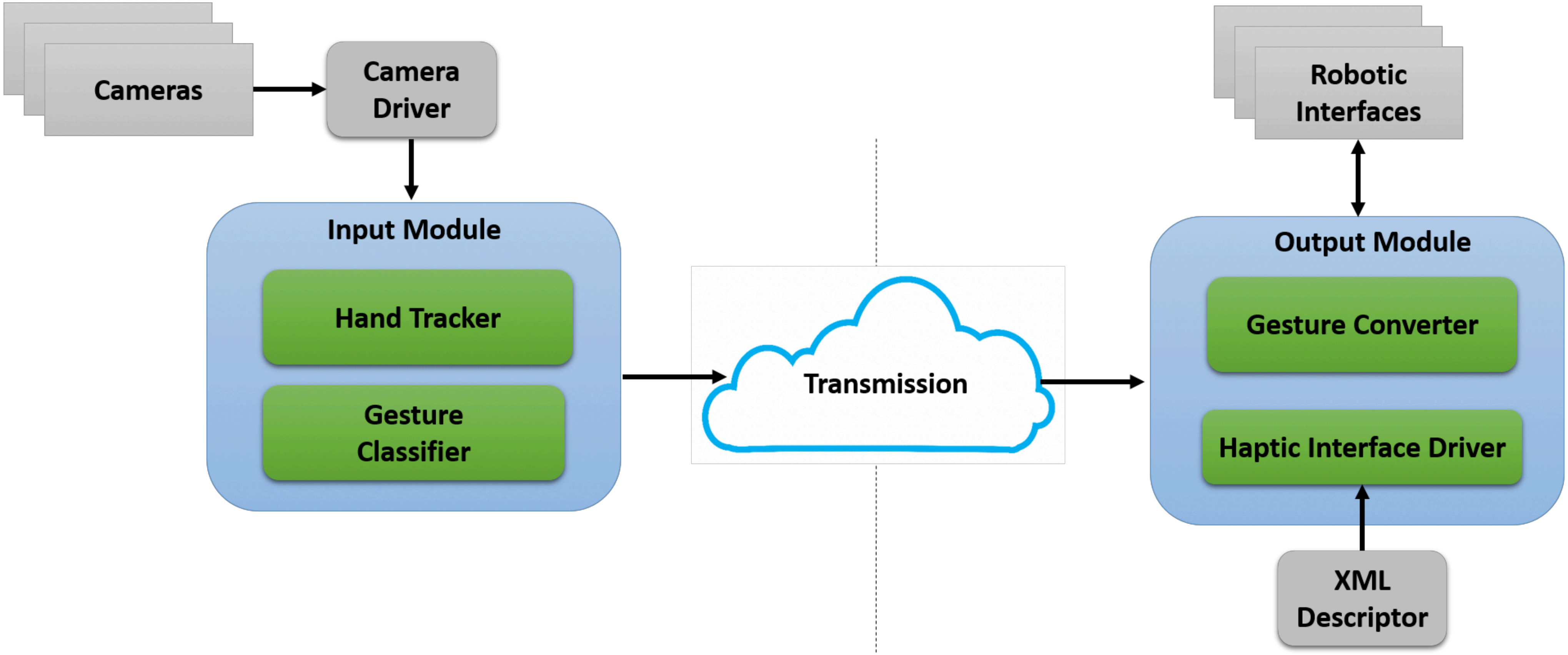
Teleoperation – The Architecture

- Phase II: Output Customization
  - Describing the robotic interface(s)



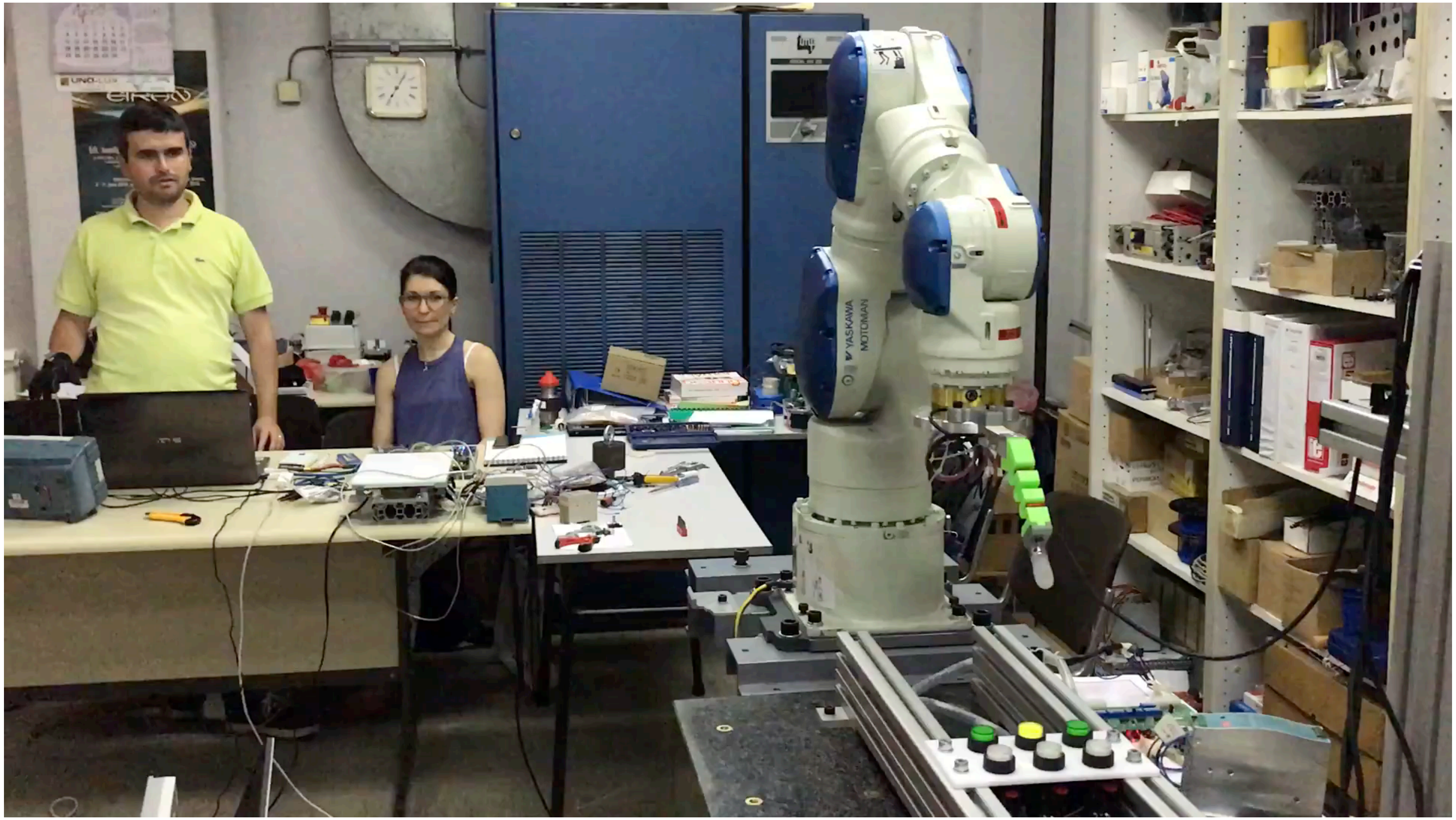
Teleoperation – The Architecture

- Phase III: Deploy



Teleoperation – Publication

- To appear in **Journal of Science Robotics**



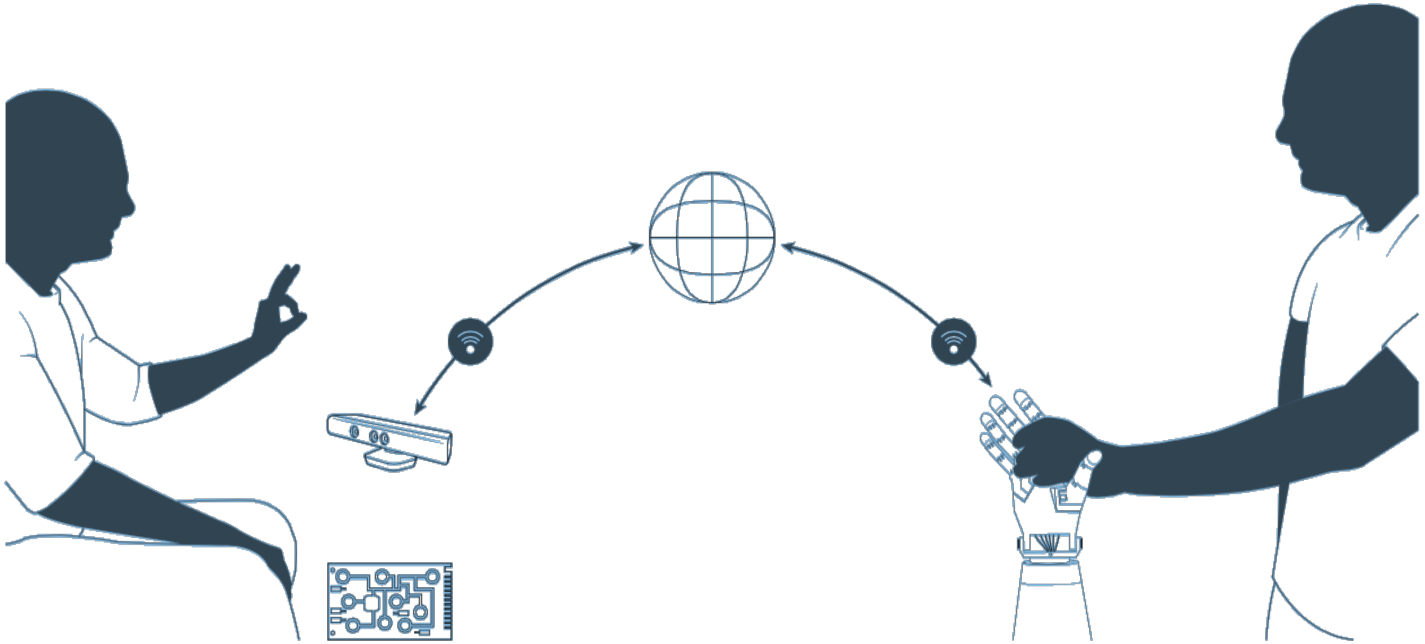
Teleoperation – Application I

	Deafblind Individuals	Rehabilitation Patients	Blind Individuals
Assistive Technologies			
Gesture-based Interfaces			
Optical Character Recognition			

Tactile Sign Languages

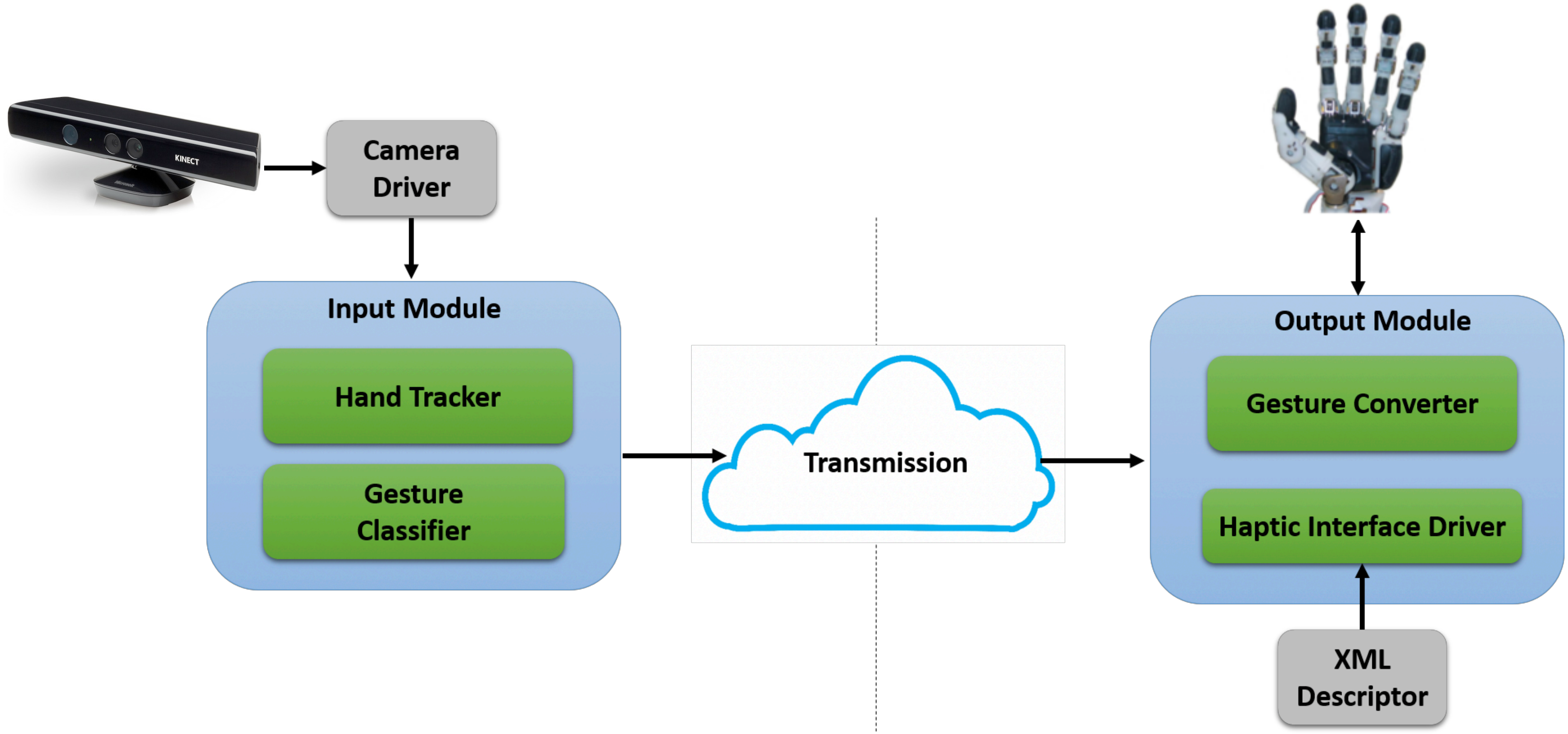
Telephone for Deafblind

- Allowing **remote** and **1-to-many** communication
  - Signer interacts with the Input Hand Gesture Recognition Module
    - Microsoft Kinect
  - Receivers interact with the Output Module
    - 3D-printed anthropomorphic hands
    - Innovative **low-cost** design <1K€ cost



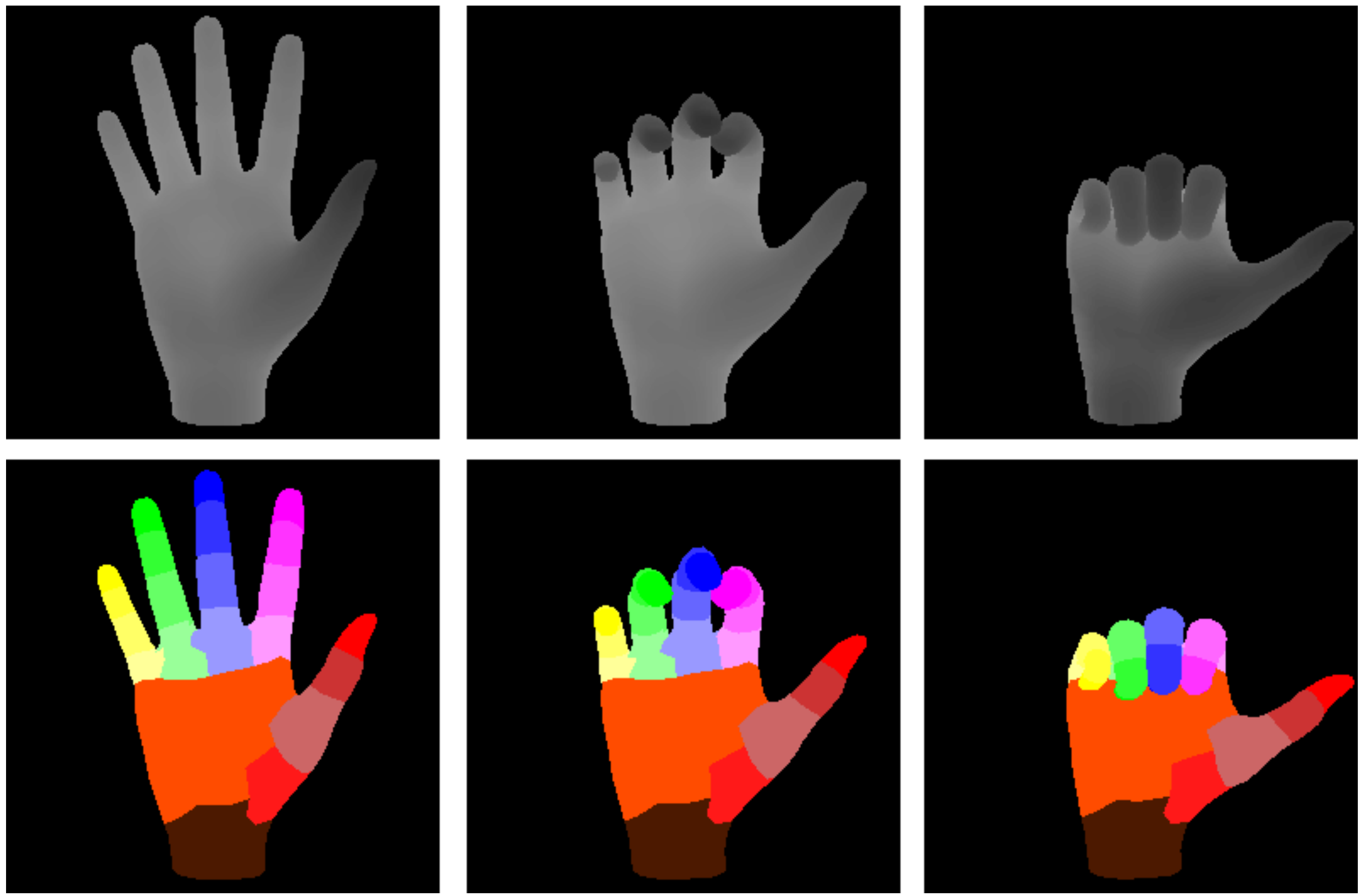
Telephone for Deafblind

- The architecture



Telephone for Deafblind

- Hand Tracker trained with hand poses typical in Sign Languages



Telephone for Deafblind

- Validated with 7 volunteers + 1 proficient signer
  - Reliability of communication pipeline > 73%
- Published in **International Journal of Advanced Robotic Systems**

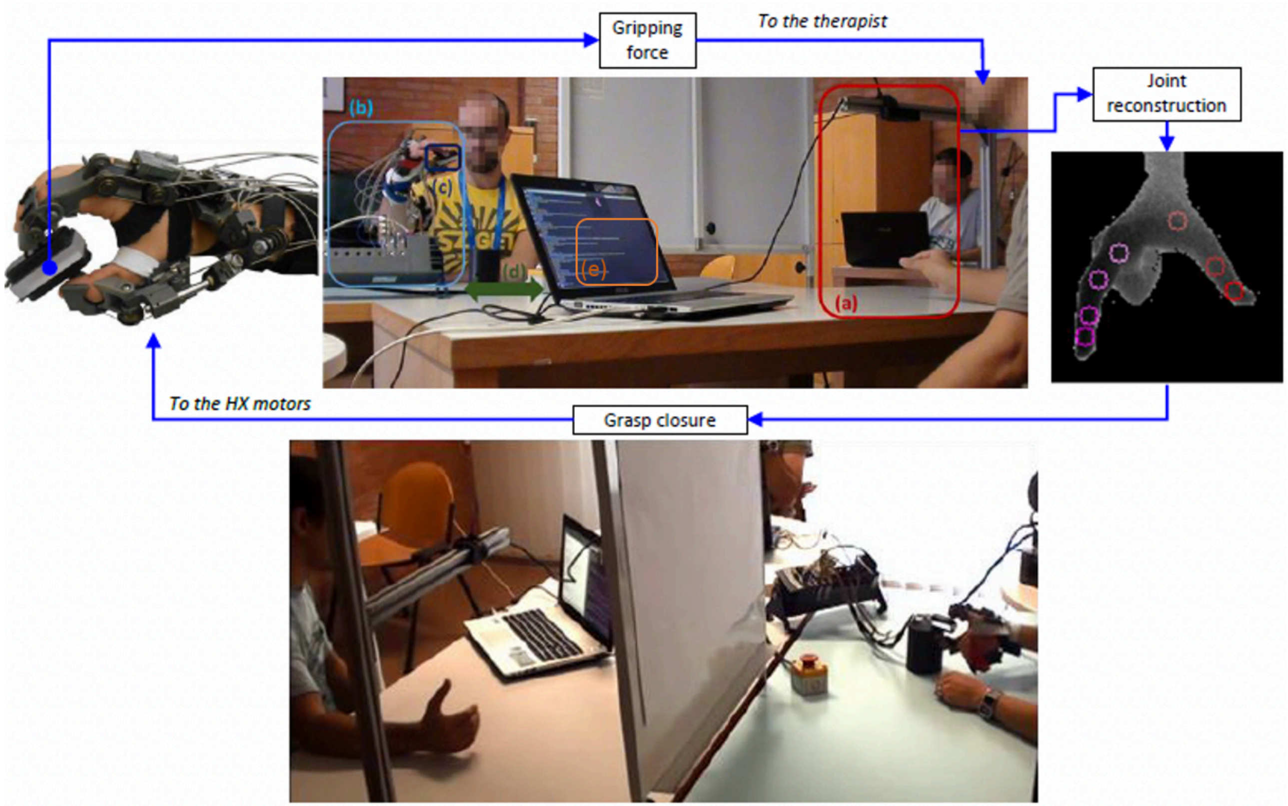
Teleoperation – Application II

	Deafblind Individuals	Rehabilitation Patients	Blind Individuals
Assistive Technologies			
Gesture-based Interfaces			
Optical Character Recognition			

Robot-mediated Rehabilitation

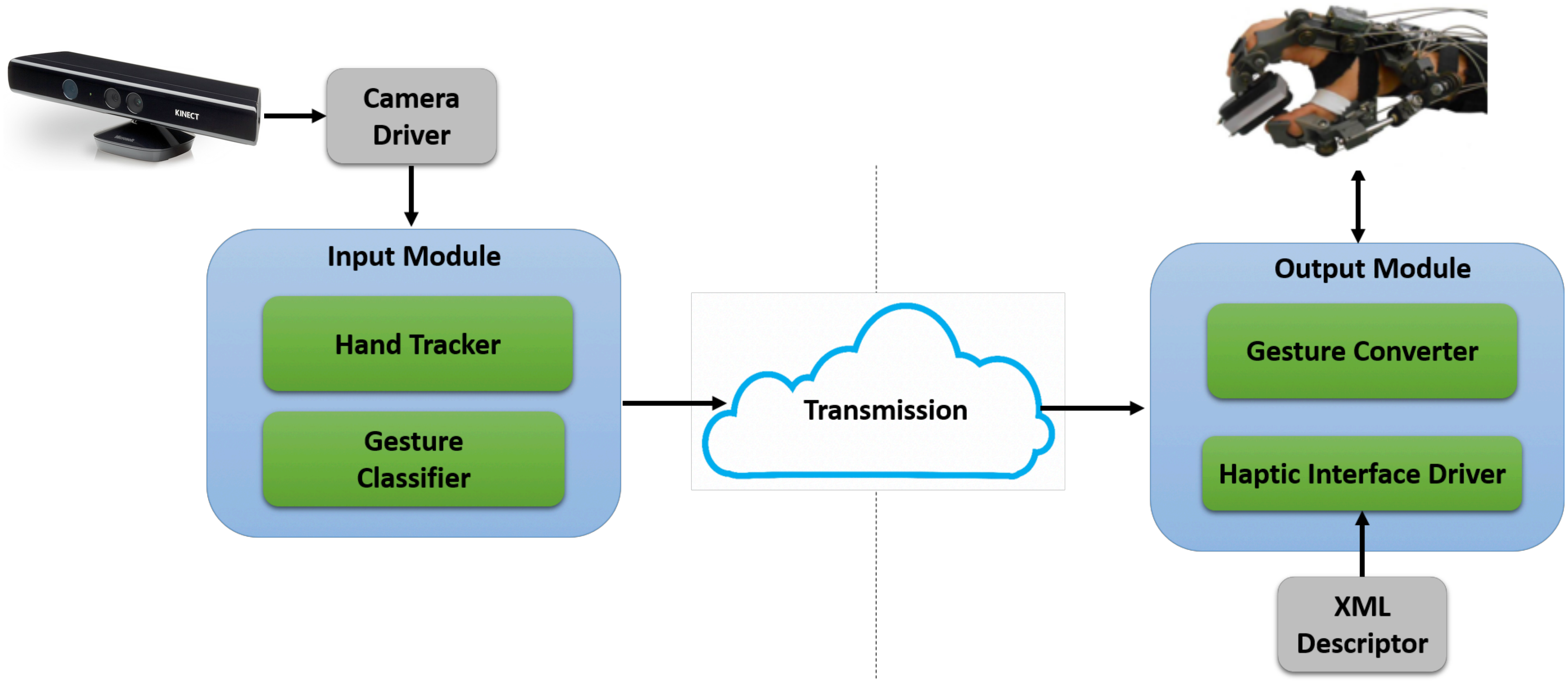
Master-slave Telerehabilitation

- Allowing remote rehabilitative exercises
  - Therapist interacts with the Input Hand Gesture Recognition Module
    - Microsoft Kinect
  - Receivers interact with the Output Module
    - Hand Exoskeleton **HX**



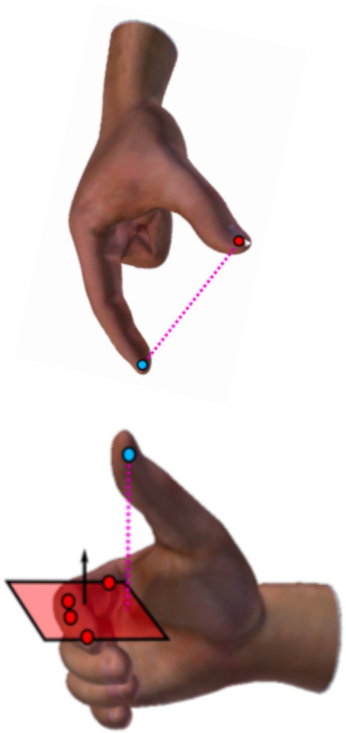
Master-slave Telerehabilitation

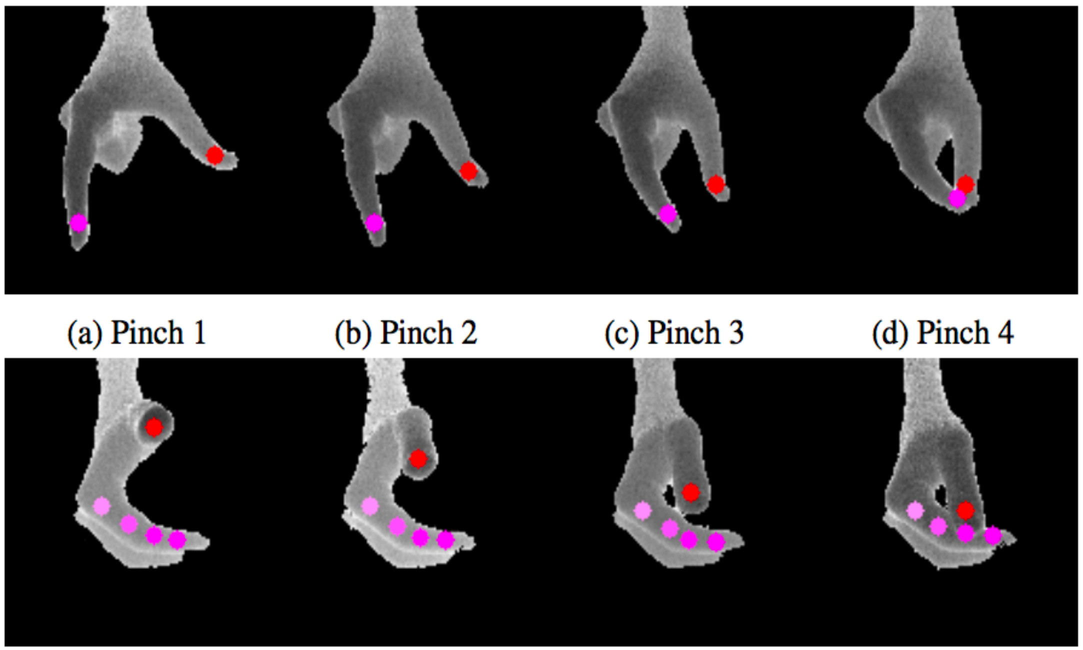
- The architecture



Master-slave Telerehabilitation

- Hand Tracker trained with hand poses typical in rehabilitative exercises





(a) Pinch 1 (b) Pinch 2 (c) Pinch 3 (d) Pinch 4

(e) Lateral 1 (f) Lateral 2 (g) Lateral 3 (h) Lateral 4

Master-slave Telerehabilitation

- Validated with 6 volunteers + 1 therapist
  - Error between the Hand Exoskeleton position and the pose commanded by the therapist < 10°
  - Communication pipeline latency < 200ms
- Published in **Sensors Journal**

A novel tele-rehabilitation system
based on a single depth-camera and
HX

Outline

- Introduction to Disabilities and Inclusive Interfaces
- Hand Gestures-based Interfaces
- **Minor Contribution**
  - Optical Character Recognition for Scientific Texts
- Conclusions and Future Research

Methodologies – Application

	Deafblind Individuals	Rehabilitation Patients	Blind Individuals
Assistive Technologies			
Gesture-based Interfaces			
Optical Character Recognition			

Inclusive Math Texts

- A different domain
- Research on
  - Multi-features segmentation of touching characters through many-valued logic
  - Knowledge from application experts encoded and finely tuned through **Particle Swarm Optimization**
- Outcomes
  - Accuracy with challenging symbols in presence of noise > 80%
  - Published in **Expert Systems with Applications**

Outline

- Introduction to Disabilities and Inclusive Interfaces
- Hand Gestures-based Interfaces
- Minor Contribution
- **Conclusions and Future Research**

Conclusions

- Cross-domain methodologies
  - Valid for several disabilities
  - Considerable use cases enabled
- Tools
  - Real-time reliable Hand Tracking and Hand Gesture Recognition
  - Segmentation of touching characters in math formulae
- Strict and direct collaboration with users

Collaborations

- International Project
  - H2020 Project **SIGN-HUB**
- National Project
  - Smart Cities Under 30 MIUR Project **PARLOMA**
- Associations
  - Lega del Filo d'Oro ONLUS
  - Unione Italiana Ciechi e Ipovedenti ONLUS
    - Istituto di Ricerca e Formazione
- Universities
  - National Laboratory **CINI AsTech**
  - Institute of Electronics, Computer, and Telecommunication Engineering
  - Scuola Sant'Anna di Pisa
  - Università di Torino

Future Research

- **Hand Tracking**
  - In-depth analysis with extensive experimental campaigns
  - Boost performances by means of technological advances
- Alternative applications for the **Teleoperation pipeline**
  - Surgeon robots
  - Industrial robots for manipulating heavy, small, and dangerous items

Thanks for Your Attention