

Tracking hands and hand-object interactions

Antonis Argyros

Computer Science Department, University of Crete (CSD - UoC)

and

Institute of Computer Science, Foundation for Research and Technology – Hellas (FORTH)

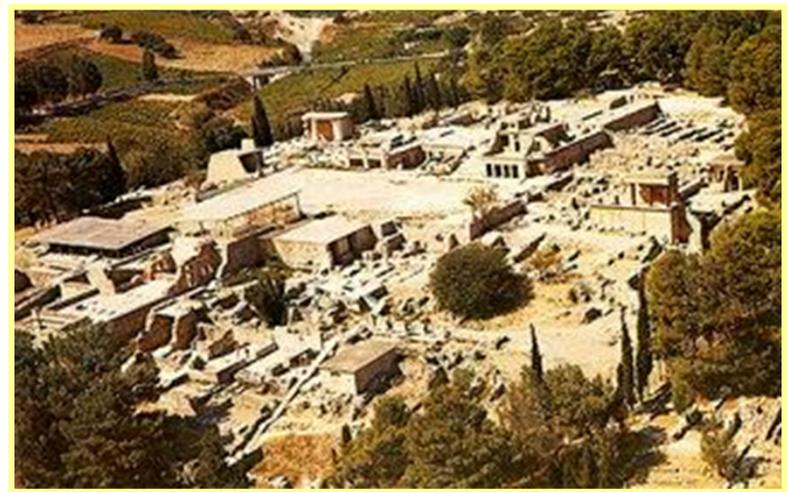
<http://www.ics.forth.gr/~argyros>

Xperience Summer School

Majorca, Spain

October 2013

Archaeological sites



Olive groves, vineyards, beautiful landscapes

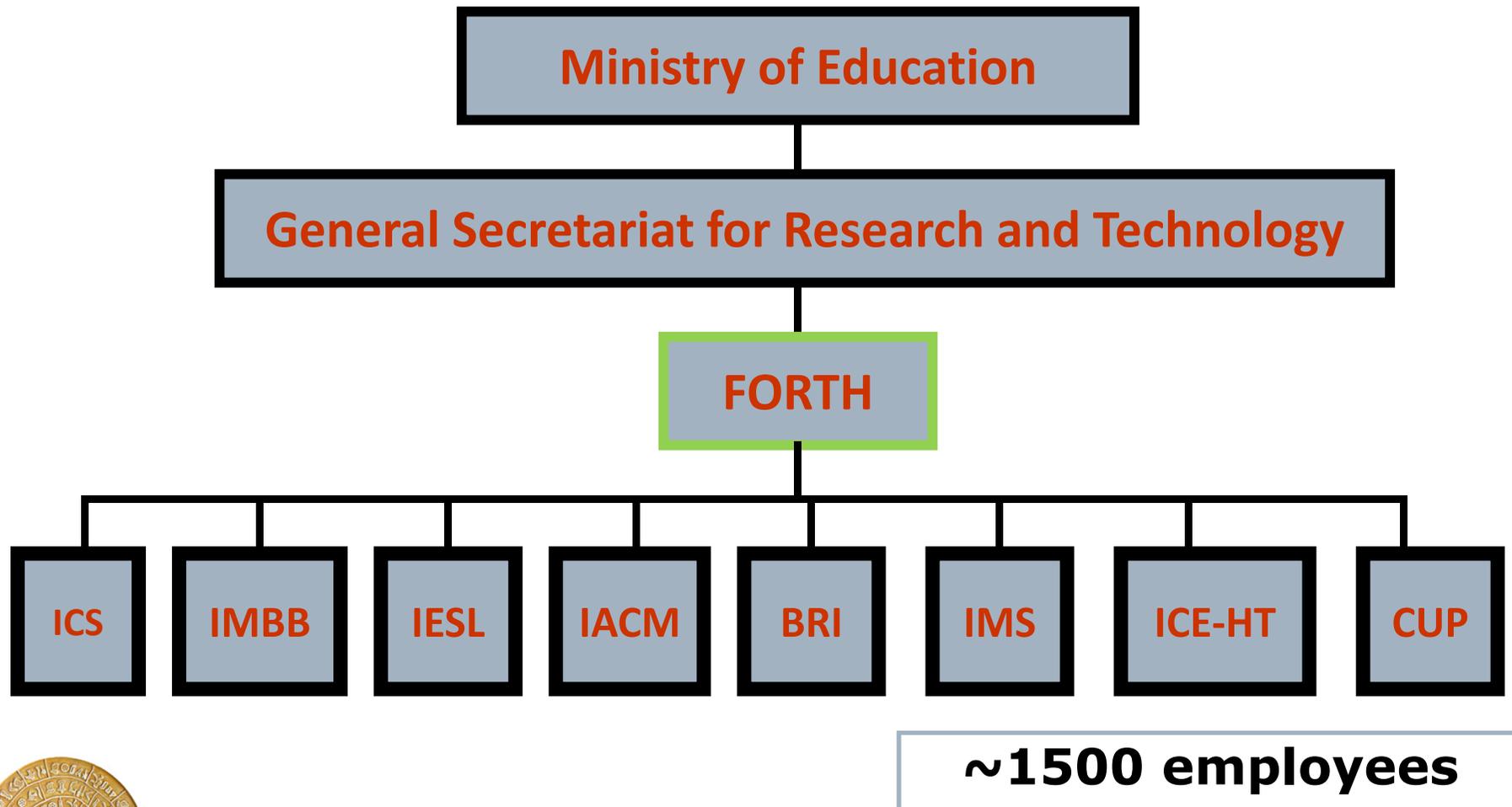


FORTH and University of Crete

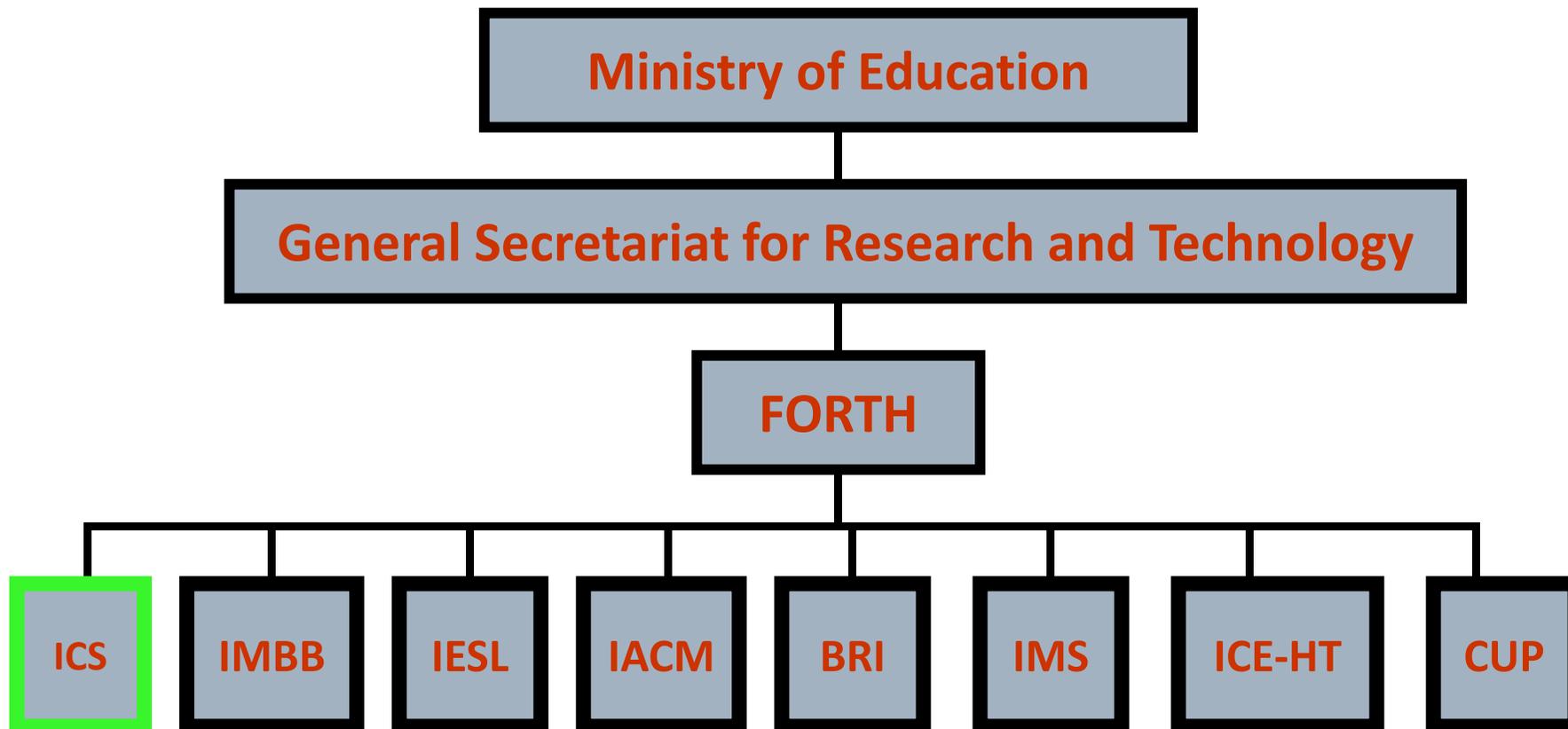


Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>

Foundation for Research and Technology - Hellas (FORTH)



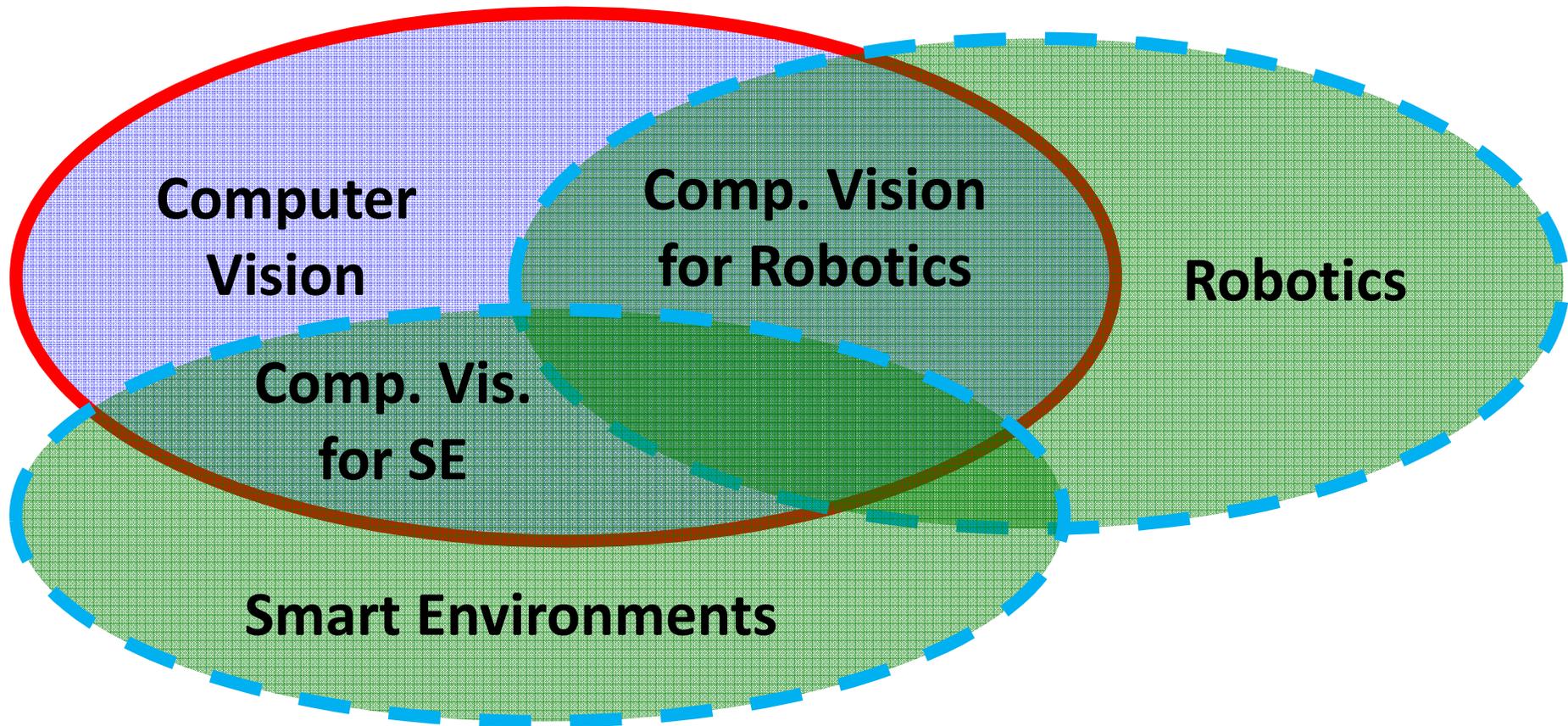
Institute of Computer Science (ICS)



~250 employees



My research interests



Research interests and activities

Core vision problems

- Tracking (2D/3D, rigid/articulated objects, ...)
- Shape representation and matching
- Structure and Motion estimation
(sparse bundle adjustment library)
- ...

Computer vision in applications

- Vision for robotics
 - Human-robot interaction
 - Vision-based robot navigation
- Vision for Ambient Intelligence Environments



Activities...

Real-time Tracking of Multiple Skin-colored Objects with a Possibly Moving Camera

A.A. Argyros, M.I.A. Lourakis
CVRL/ICS/FORTH

Head pose estimation on depth data based on Particle Swarm Optimization

P. Paderleris, X. Zabulis, A. Argyros
Computational Vision and Robotics Laboratory
Institute of Computer Science, FORTH
HAU3D12

Offline Experiment Sequence

Number of Voxels: 15004536
Image Resolution: 1280 x 960

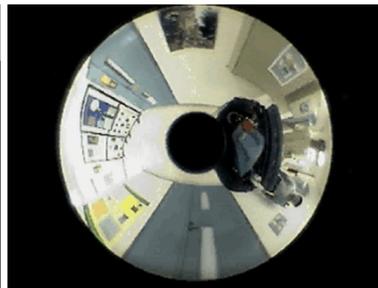
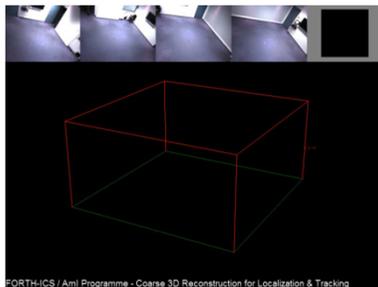
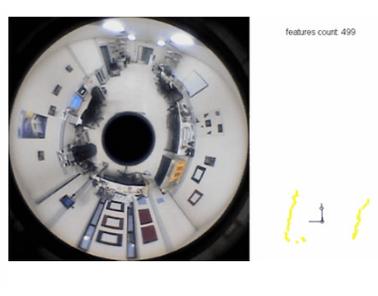
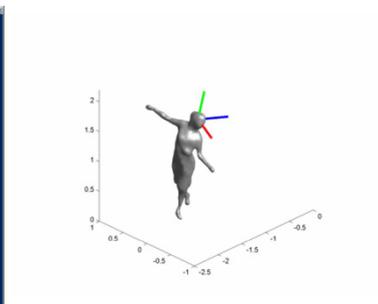
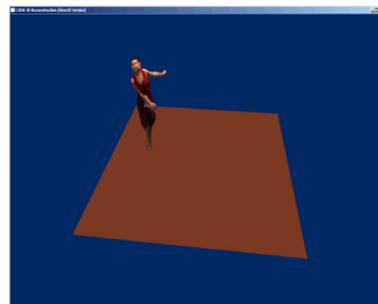


Lumen detection for capsule endoscopy

Xenophon Zabulis, Antonis A. Argyros and Dimitris P. Tsakiris
(zabulis, argyros, tsakiris)@ics.forth.gr
Institute of Computer Science, Foundation for Research & Technology – Hellas
ICS-FORTH

Vision-based interpretation of hand gestures for the remote control of a computer mouse

A.A. Argyros, M.I.A. Lourakis
CVRL/ICS/FORTH



ActIPret

ACIN, GCS, FORTH, PROFACOR, m p



Visual Object Tracking and Segmentation in a Closed Loop

Hand-moving-camera video

Papoutakis K., Argyros A.
ISVC10



Have a look at:

<http://www.ics.forth.gr/~argyros/research.htm>

Problem statement

Given markerless visual observations of a hand-object(s) interaction scenario, track the whole scene in 3D



Problem statement

Given markerless visual observations of a **hand-object(s) interaction scenario**, track the whole scene in 3D

Freely moving hand(s) interact with a number of rigid 3D objects



Problem statement

Multicamera setup, stereo,
RGB-D camera, ...

Given markerless **visual observations** of
a hand-object(s) interaction scenario,
track the whole scene in 3D

Freely moving hand(s)
interact with a number of
rigid 3D objects



Problem statement

Do not interfere with the action!

Multicamera setup, stereo, RGB-D camera, ...

Given **markerless** visual observations of a hand-object(s) interaction scenario, track the whole scene in 3D

Freely moving hand(s) interact with a number of rigid 3D objects



Problem statement

Do not interfere with the action!

Multicamera setup, stereo, RGB-D camera, ...

Given markerless visual observations of a hand-object(s) interaction scenario,
track the whole scene in 3D

- 3D pose of the hand(s) and the object(s)
- Full articulation for the hand(s)

Freely moving hand(s) interact with a number of rigid 3D objects



Is this an interesting problem?

□ Theoretical interest

- Humans solve it, could technical systems solve it, too?
- Solutions can probably prove useful in other, interesting, similar problems

□ Practical interest in supporting the interpretation of human activities

- Understanding grasping and manipulation
- Sign language understanding
- Games, virtual/augmented reality
- Human-Computer Interaction / Human-Robot Interaction
- ...



Is this an easy problem?

- Not really...
- A problem with high dimensionality...
- ... that needs to be solved based on relatively poor observations
 - Chromatically uniform appearance of the hand
 - Severe self occlusions
 - Severe hand/object occlusions
(in case of hand/object interaction)
 - Distant views
 - Rapid hand motions



Our work on this problem...

Emphasis on
“what” and **“why”**
rather than on **“how”**

Detection and tracking of hands in 2D

2D hand detection and tracking

**Real-time Tracking of
Multiple Skin-colored
Objects with a Possibly
Moving Camera**

A.A. Argyros, M.I.A. Lourakis

CVRL/ICS/FORTH



A.A. Argyros, M.I.A. Lourakis, "Real time Tracking of Multiple Skin-Colored Objects with a Possibly Moving Camera", in proceedings of the European Conference on Computer Vision (*ECCV'04*), Springer-Verlag, vol. 3, pp. 368-379, Prague, Czech Republic, May 11-14, 2004.

Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>

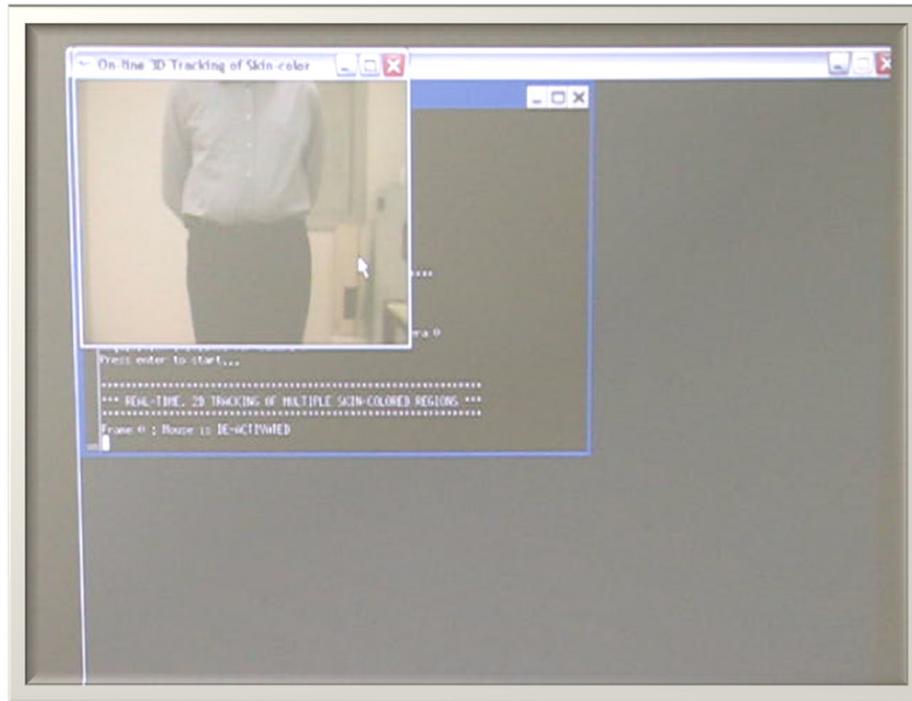
Finger detection



A.A. Argyros, M.I.A. Lourakis, "Vision-based Interpretation of Hand Gestures for Remote Control of a Computer Mouse", in proceedings of the *HCI'06* workshop (in conjunction with *ECCV'06*), LNCS 3979, Springer Verlag, pp.40-51, Graz, Austria, May 13th, 2006. **Recipient of the "Best Paper Award"**.

Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>

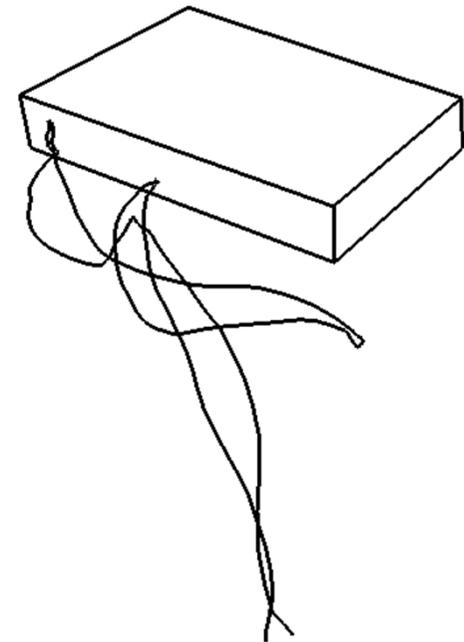
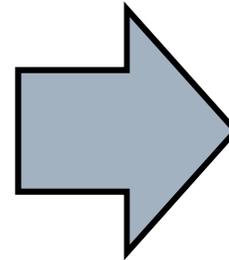
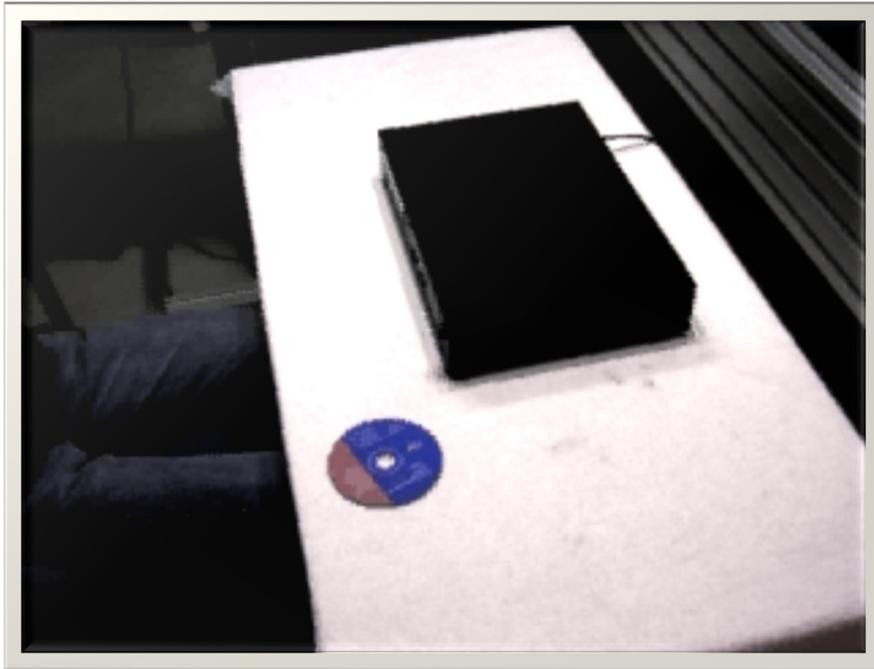
Simple HCI and HRI



A.A. Argyros, M.I.A. Lourakis, "Vision-based Interpretation of Hand Gestures for Remote Control of a Computer Mouse", in proceedings of the *HCI'06* workshop (in conjunction with *ECCV'06*), LNCS 3979, Springer Verlag, pp.40-51, Graz, Austria, May 13th, 2006. **Recipient of the "Best Paper Award"**.

Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>

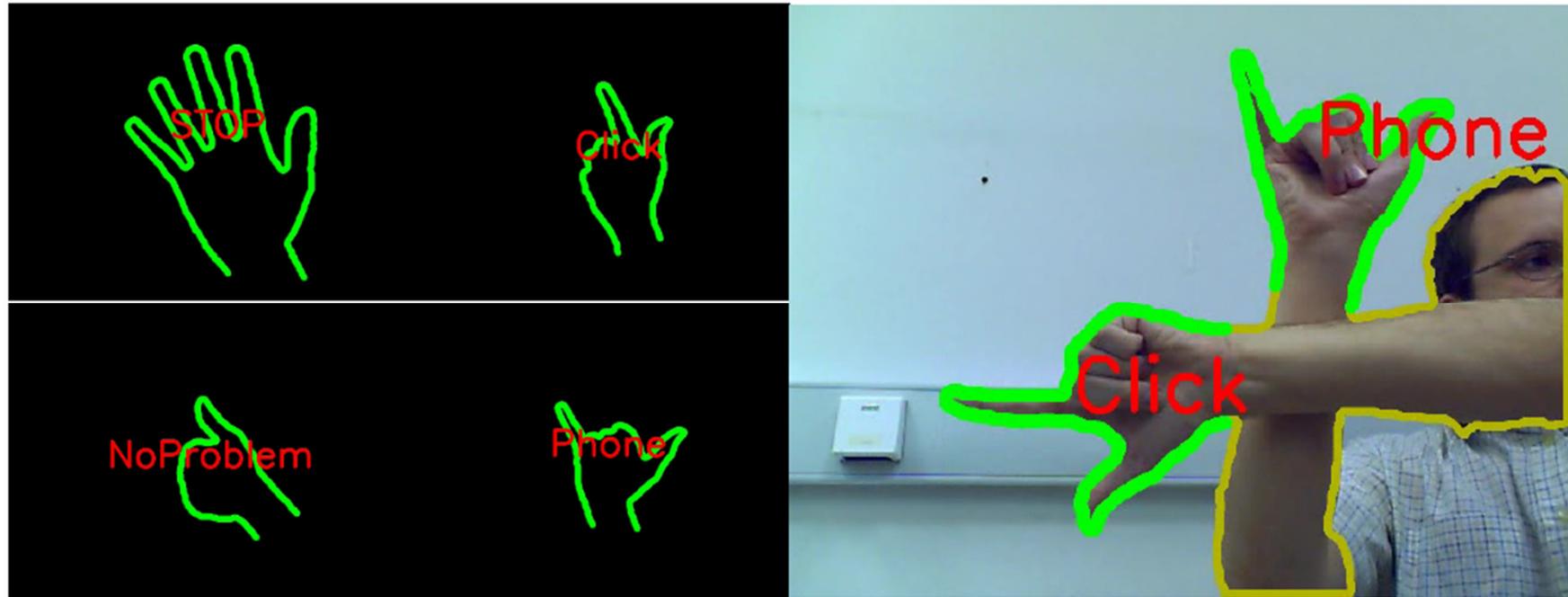
Activity interpretation



A.A. Argyros, M.I.A. Lourakis, "Binocular Hand Tracking and Reconstruction Based on 2D Shape Matching", in proceedings of the International Conference on Pattern Recognition 2006 (*ICPR'06*), Hong Kong, China, 20 – 24, August 2006.

Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>

Deformation tolerant partial shape matching



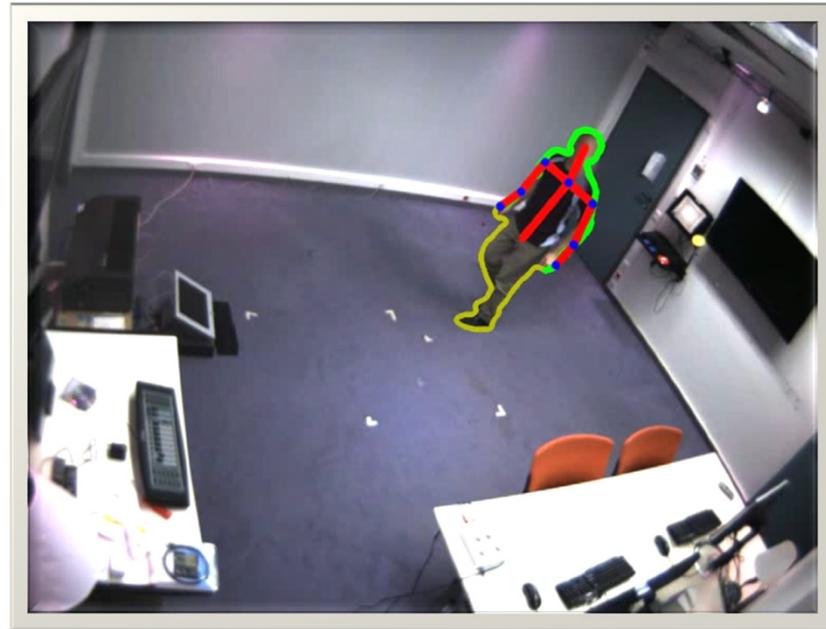
D. Michel, I. Oikonomidis, A.A. Argyros, "Scale invariant and deformation tolerant partial shape matching", in Image and Vision Computing (IVC), Elsevier, vol. 29, issue 7, pp. 459-469, June 2011.

Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>

Deformation tolerant partial shape matching



Interpreting hand postures



Interpreting human body postures



D. Michel, I. Oikonomidis, A.A. Argyros, "Scale invariant and deformation tolerant partial shape matching", in Image and Vision Computing (IVC), Elsevier, vol. 29, issue 7, pp. 459-469, June 2011.

Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>

Tracking in the presence of severe occlusions?



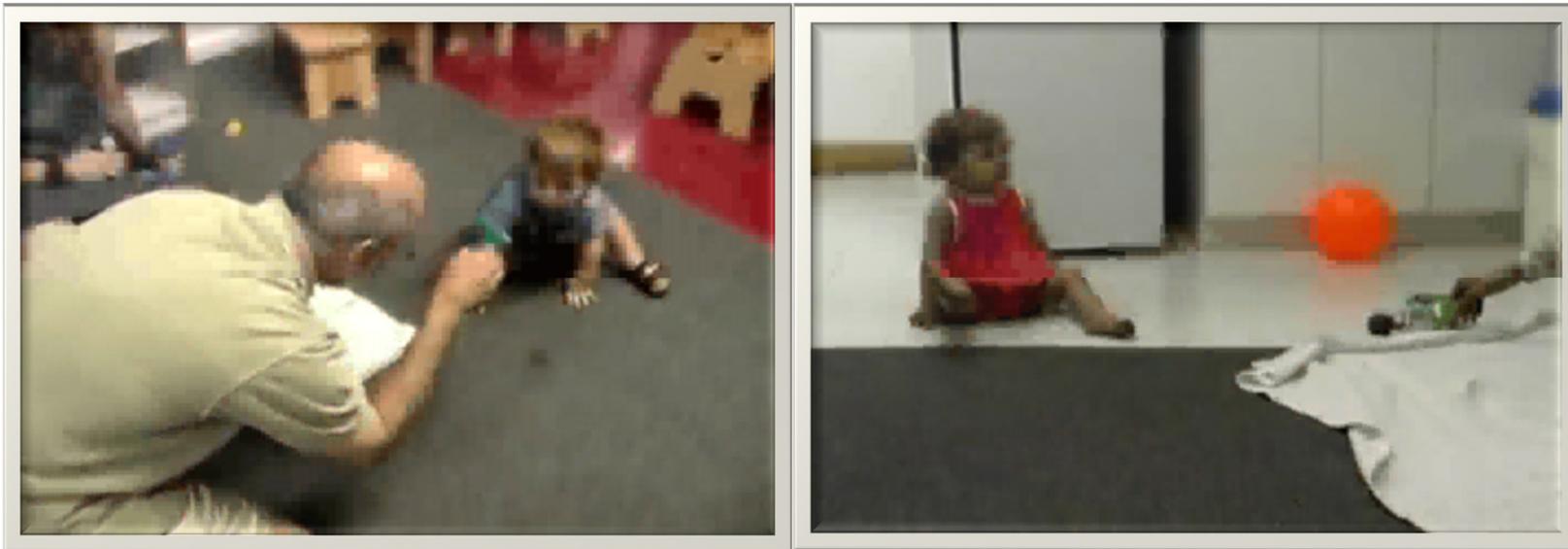
V. Papadourakis, A.A. Argyros, "Multiple Objects Tracking in the Presence of Long-term Occlusions", in Computer Vision and Image Understanding (CVIU), Elsevier, vol. 114, issue 7, pp. 835-846, 2010.

Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>



Tracking in the presence of severe occlusions

Object permanence: the understanding that objects continue to exist even when they cannot be seen, heard, or touched.
Studied in the field of developmental psychology (Jean Piaget)



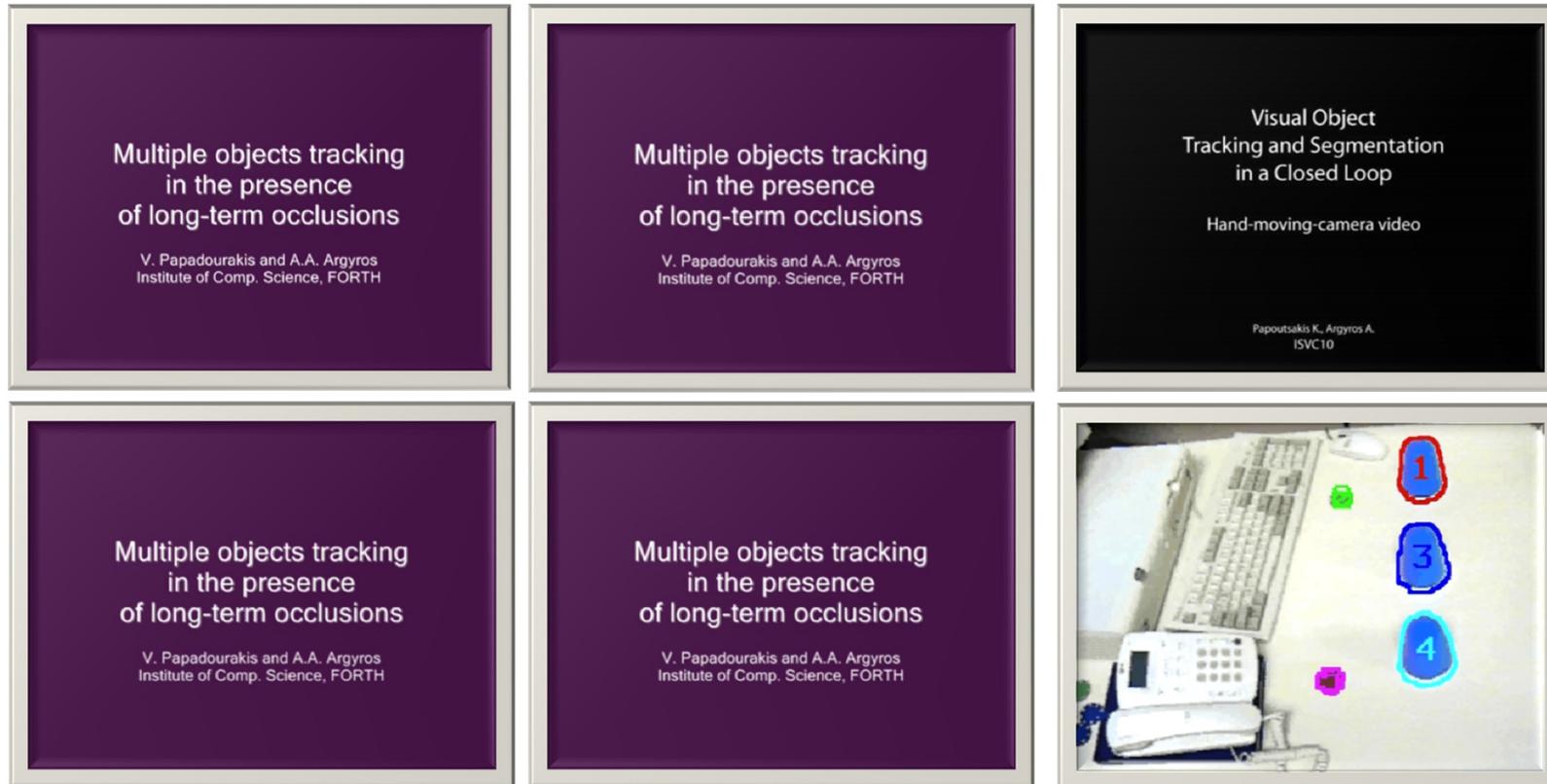
(video source: wikipedia)



V. Papadourakis, A.A. Argyros, "Multiple Objects Tracking in the Presence of Long-term Occlusions", in Computer Vision and Image Understanding (CVIU), Elsevier, vol. 114, issue 7, pp. 835-846, 2010.

Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>

Tracking in the presence of severe occlusions



V. Papadourakis, A.A. Argyros, "Multiple Objects Tracking in the Presence of Long-term Occlusions", in Computer Vision and Image Understanding (CVIU), Elsevier, vol. 114, issue 7, pp. 835-846, 2010.

K. Papoutsakis, A.A. Argyros, "Object tracking and segmentation in a closed loop" to appear in Proceedings of the International Symposium on Visual Computing, ISVC'2010, Las Vegas, USA, Nov 29-Dec 1, 2010.

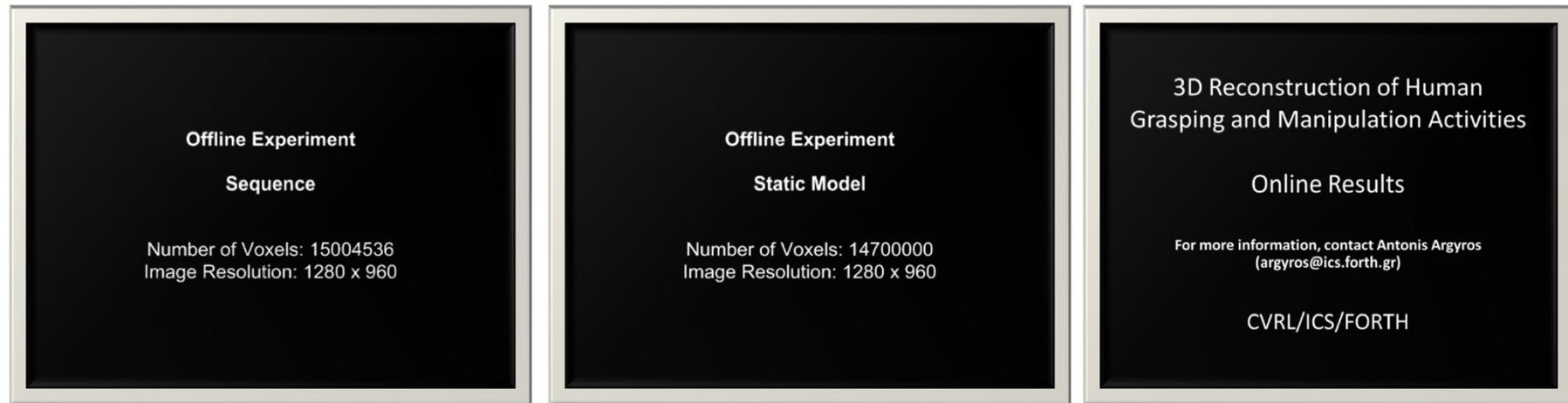


Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>

Detection and tracking of hands in 2D

Detection and tracking of hands in **3D**

From multiple views to textured 3D meshes



K. Tzevanidis, X. Zabulis, T. Sarmis, P. Koutlemanis, N. Kyriazis, A.A. Argyros, "From multiple views to textured 3D meshes: a GPU-powered approach", in Proceedings of the Computer Vision on GPUs Workshop, CVGPU'2010, In conjunction with ECCV'2010, Heraklion, Crete, Greece, 10 September 2010.

Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>

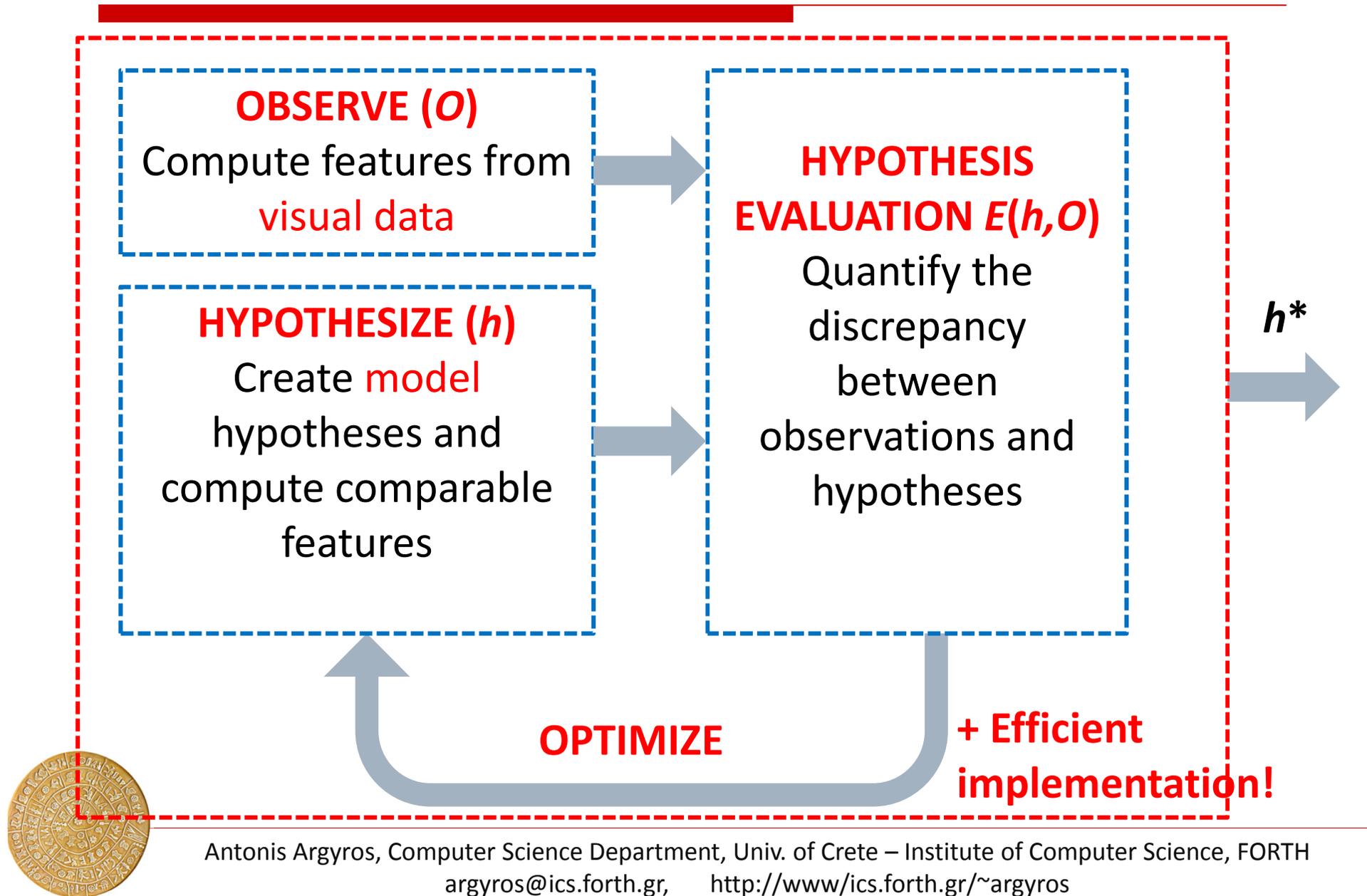
Tracking the articulation of hands

- Estimate the full 3D position, orientation and articulation (i.e., all joint angles of a performing hand)

	Multicamera setup	RGB-D data (Kinect)
A hand in isolation	ACCV'2010	BMVC'2011
A hand interacting with object(s)	ICCV'2011	CVPR'2012 CVPR'2013



Our hypothesize-and-test framework

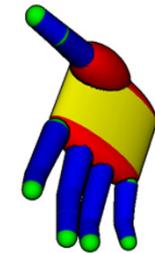
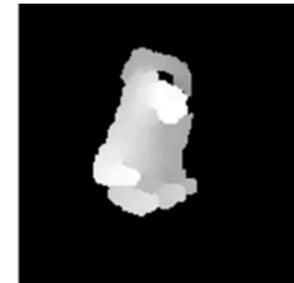


3D hand tracking based on RGB-D images

- **Observations O :** 3D structure of the skin colored regions in each RGBD frame
- **Model:** One hand (37 geometric primitives, 20 intrinsic + 6 extrinsic = **26 DoFs = h**)
- **Hypothesis testing (E):** Rank each hypothesis h based on its compatibility E with the observations O
- Formulate an **optimization problem** to minimize the discrepancies between the hypothesis h and the observations O

$$h^* \stackrel{PSO}{=} \arg \min_h \{E(h, O)\}$$

- **Hypothesis formulation and optimization:** Particle Swarm Optimization on a **27D** parameter space



3D hand tracking based on RGB-D images

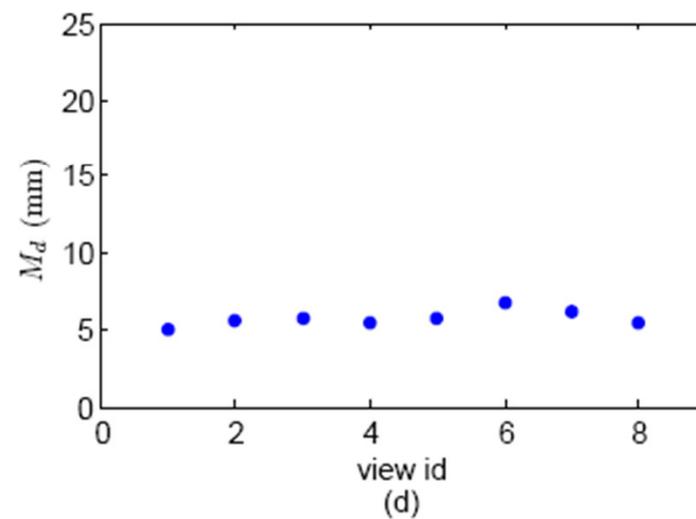
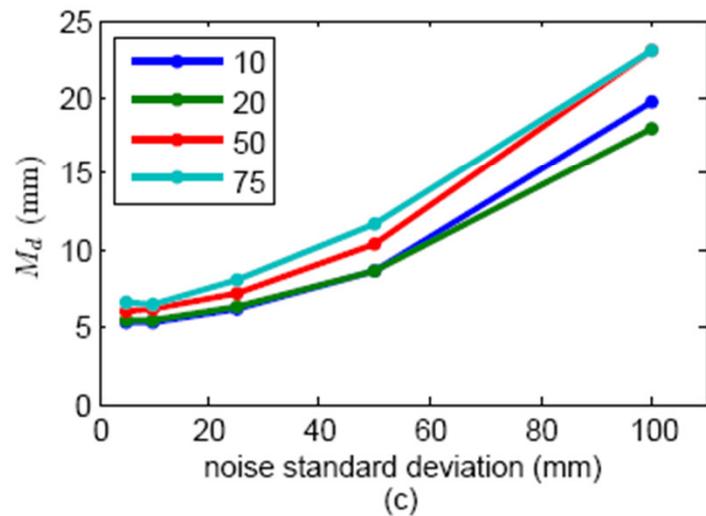
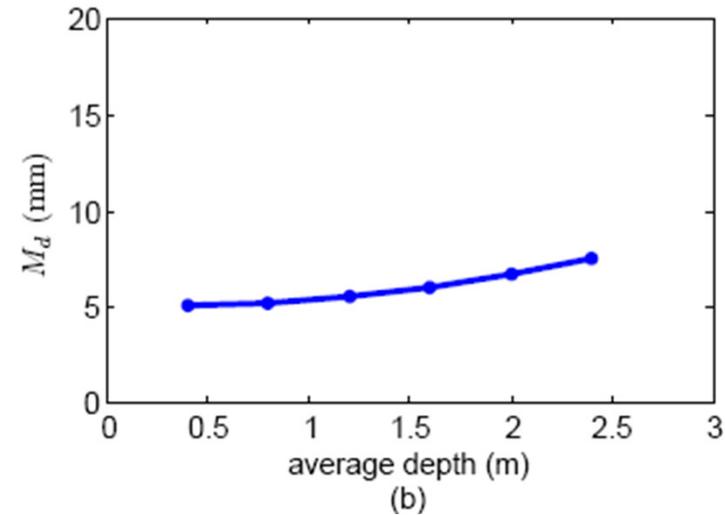
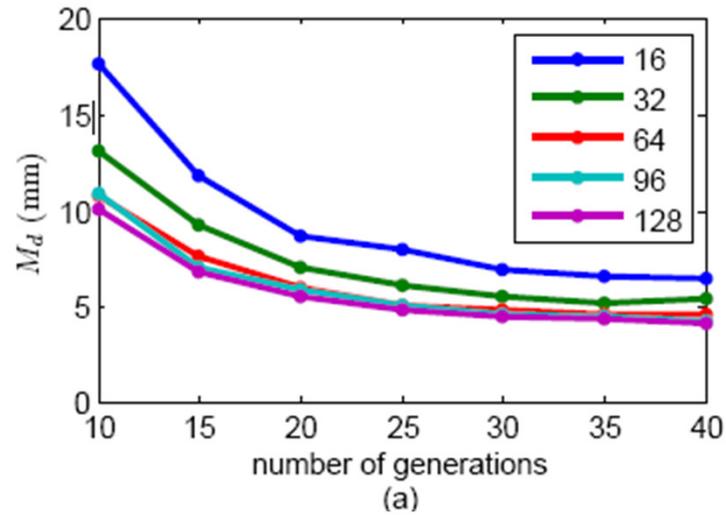
Efficient model-based 3D tracking of
hand articulations using Kinect



I. Oikonomidis, N. Kyriazis, A.A. Argyros, "Efficient model based 3D tracking of hand articulations using Kinect", BMVC 2011, UK, September 2011.

Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>

Quantitative evaluation



Software

- **FORTH 3D Hand Tracking Library:** A [software library](http://www.openni.org/files/3d-hand-tracking-library/) for the 3D tracking of the articulated motion of hands
 - Available with a free license for research purposes at <http://www.openni.org/files/3d-hand-tracking-library/>.
 - 1st place award, CHALEARN Gesture Recognition Demonstration Competition
 - In conjunction with [ICPR 2012](http://www.icpr2012.org/), Tsukuba, Japan, Nov. 2012
 - Sponsored by Microsoft Research, Redmond, USA



BUT...

- ...besides tracking **hands in isolation** we are also interested in tracking **hands interacting** with their environment...
- What does a hand “tell” about the grasped/manipulated objects?
- What do objects “tell” about the grasping/manipulating hand?

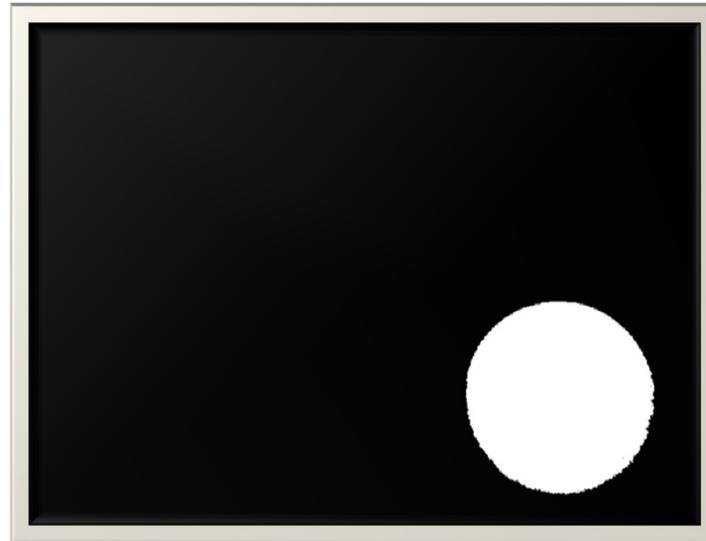


A hand in interaction with an object

A key observation and idea:



Seeing the hand, “only”



Seeing the object, “only”

– Thus

– occlusions due to hand-object interaction is **not a curse to be bypassed** but **a feature to be exploited...**

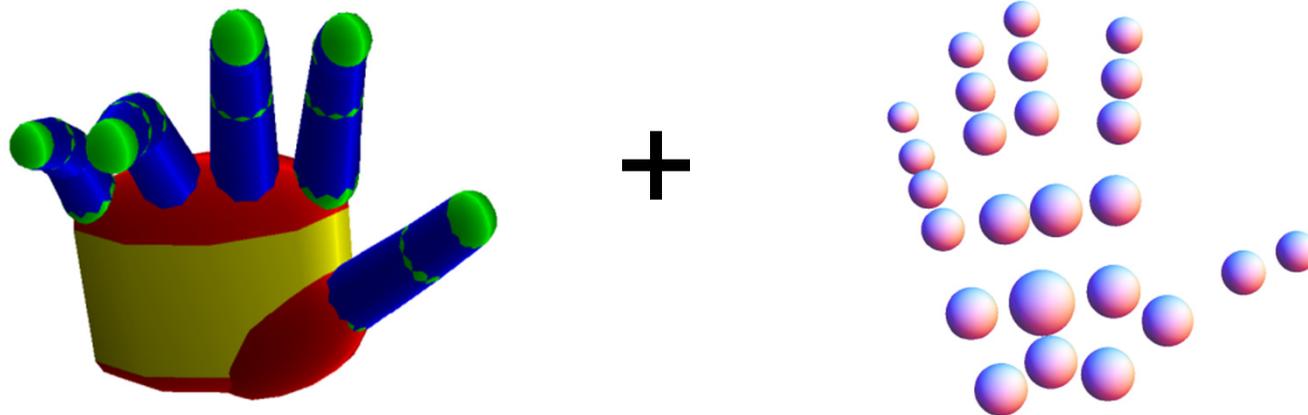
– Joint hand-object modeling



A hand in interaction with an object

Another (obvious, yet important) observation:

- A hand and an object cannot share the same physical space!
- Thus:
 - Penalize physically implausible solutions (i.e., solutions that exhibit hand-object interpenetration)



A hand in interaction with an object

Full DOF tracking of a hand
interacting with an object by modeling
occlusions and physical constraints



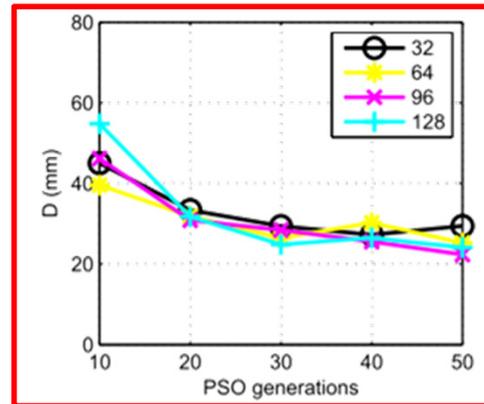
I. Oikonomidis, N. Kyriazis, A.A. Argyros, "Full DOF tracking of a hand interacting with an object by modeling occlusions and physical constraints", ICCV 2011, Barcelona, Spain, November 2011.

Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>

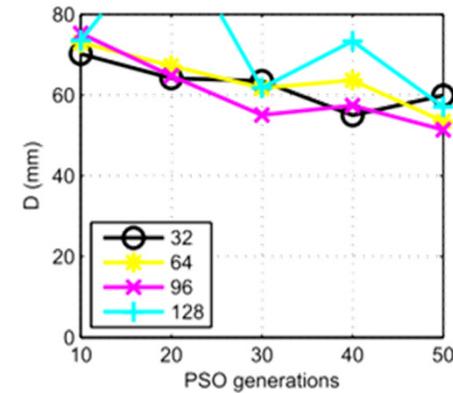
Results - parameter investigation

Input from
2 cameras

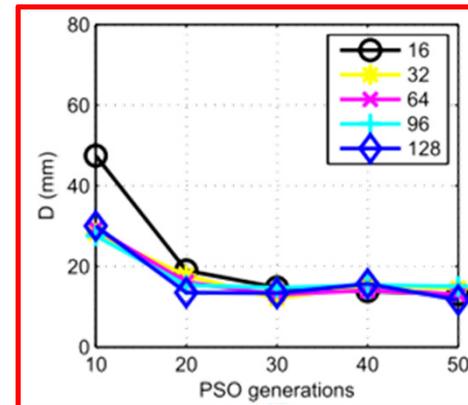
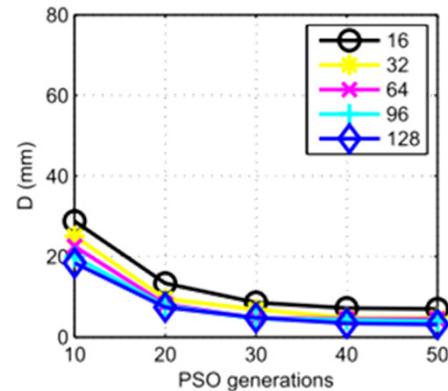
Object
Modeling



No Object
Modeling



Input from
8 cameras



Results - parameter investigation

Object	Estimated/actual parameters (in <i>mm</i>)
Cylinder	Radius: 51/53, Height: 121/131
Ellipsoid	X: 128/116, Y: 128/116, Z: 122/116
Box	X: 66/67, Y 158/150, Z: 84/93

Accuracy in estimating the object parameters



A hand in interaction with multiple objects (RGBD)



Tracking two strongly interacting hands



I. Oikonomidis, N. Kyriazis, A.A. Argyros, "Tracking the articulated motion of two strongly interacting hands", CVPR 2012, Rhode Island, USA, June 2012.

Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>

Extending the existing framework...

□ Why?

- **Computational complexity** increases fast with the number of the objects to be tracked
 - One hand interacting with N rigid objects \rightarrow $(27+7N)$ -Dimensional parameter space...
- Physical constraints can be taken into account, but one needs to consider them one-by-one...

□ How?

- Capitalize on **physics** and on the **single actor hypothesis**:
In a hand-object interaction scenario, state changes of **passive objects** are due to the motion of the **active hand(s)**



Physically plausible 3D scene tracking: The single actor hypothesis

- **Proposed approach:** Track the scene by searching for the **hand motion**, that, in a **physics based simulation** environment, results in hand-object configurations that are as similar as possible to actual, RGBD-camera-based observations
- ... thus, given physics-based simulation of hand motion \mathbf{x} , physics-based simulator \mathbf{S} , observations \mathbf{O} and objective function \mathbf{E} scene tracking amounts to solving:

$$x^* \stackrel{PSO}{=} \arg \min_x \{E(O, S(x))\}$$



Tracking a scene by tracking the actor



N. Kyriazis, A. A. Argyros, "Physically plausible 3D scene tracking: The single actor hypothesis", CVPR 2013, oral presentation, Portland, Oregon, USA, June 2013.

Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>

**Up to now:
from images to 3D info**

**How about moving
from 3D info to “higher level concepts”...**

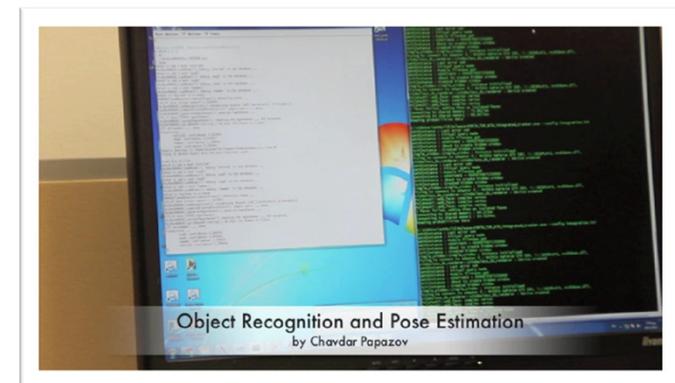
Higher level concepts

- Recognize **action primitives** and **actions**
 - Example: Activity of pouring water from a mug



M. Patel, C.H. Ek, N. Kyriazis, A.A. Argyros, J.V. Miro, D. Kragic, "Language for Learning Complex Human-Object Interactions", in Proceedings of the IEEE Int'l Conference on Robotics and Automation (ICRA 2013), Karlsruhe, Germany, May 6-10, 2013.

- **Infer human intention** to facilitate robot learning by demonstration
 - Example: grasp for tool use, grasp for transfer



D. Song, N. Kyriazis, I. Oikonomidis, C. Papazov, A. Argyros, D. Burschka, D. Kragic, "Predicting Human Intention in Visual Observations of Hand/Object Interactions", in Proceedings of the IEEE ICRA 2013, Karlsruhe, Germany, May 6-10, 2013.



**Different problems,
similar hypothesize-and-test
computational framework**

Head pose estimation

Head pose estimation on depth data based on Particle Swarm Optimization

P. Padeleris, X. Zabulis, A. Argyros
Computational Vision and Robotics Laboratory
Institute of Computer Science, FORTH
HAU3D12



P. Padeleris, X. Zabulis and A.A. Argyros, "Head pose estimation on depth data based on Particle Swarm Optimization", in HAU3D'2012 (CVPR 2012 workshop)

Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>

Human body articulation tracking and beat synchronous dance animation

3D body articulation tracking



... and animation of characters



C. Panagiotakis, A. Holzappel, D. Michel, A.A. Argyros, "Beat Synchronous Dance Animation based on Visual Analysis of Human Motion and Audio Analysis of Music Tempo", to appear, ISVC'2013

Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>

Shape from Interaction



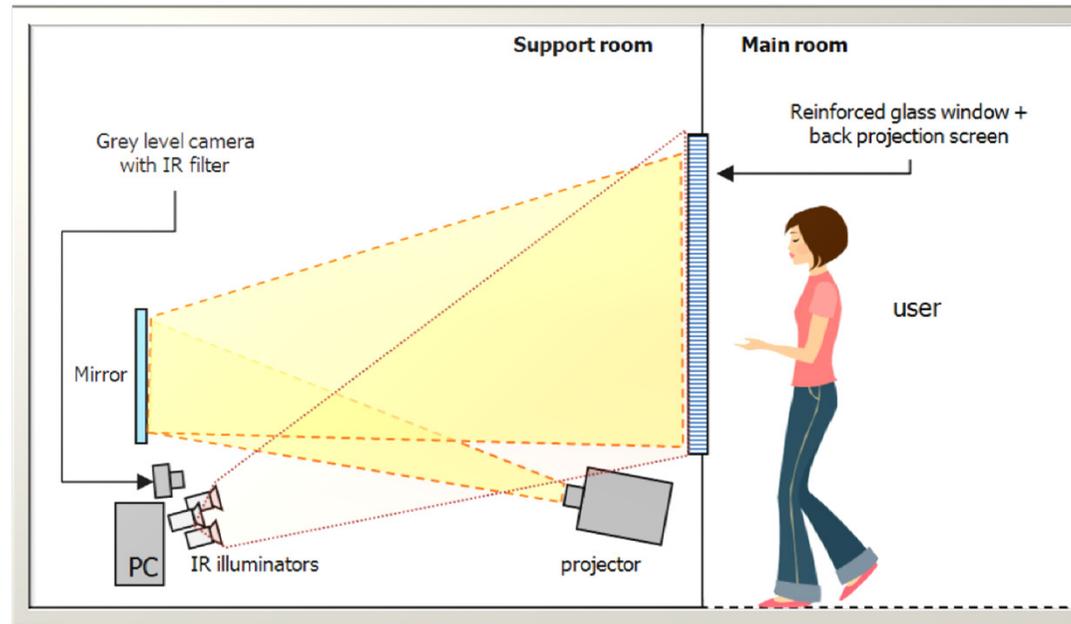
D. Michel, X. Zabulis, A. Argyros, "Shape from Interaction", MVA journal, under review

Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>

**Interesting side effects:
Tracking hands at the service of
building Smart Environments**

Interactive exhibits: Polyapton

- POLYAPTON

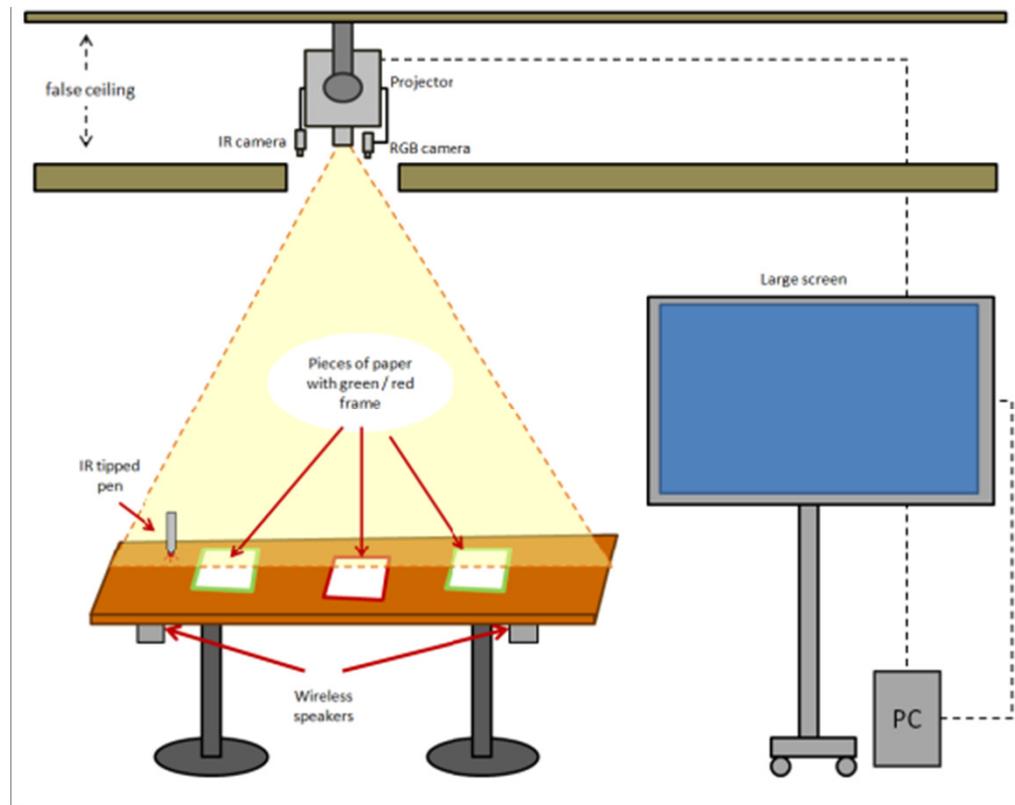


D. Michel, A.A. Argyros, D. Grammenos, X. Zabulis, T. Sarmis, "Building a multi-touch display based on Computer vision techniques", in proceedings of the IAPR Conference on Machine Vision and Applications (MVA'09), pp. 74-77, Hiyoshi Campus, Keio University, Japan, May 20-22, 2009.

Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>

Interactive exhibits: Macedonia Map

- MACEDONIA MAP



D. Grammenos, D. Michel, X. Zabulis, A.A. Argyros, "PaperView: Augmenting Physical Surfaces with Location-Aware Digital Information" in Proceedings of the ACM SIGCHI Conference on Tangible Embedded, Embodied Interaction, TEI'2011, Funchal, Portugal, Jan 23-26, 2011.

Antonis Argyros, Computer Science Department, Univ. of Crete – Institute of Computer Science, FORTH
argyros@ics.forth.gr, <http://www/ics.forth.gr/~argyros>



Acknowledgments

- Dimitris Grammenos
- Cedric Groyer
- Panagiotis Koutlemanis
- Nikolaos Kyriazis
- Manolis Lourakis
- Damien Michel
- Iason Oikonomidis
- Pashalis Padelaris
- Vassilis Papadourakis
- Kostas Papoutsakis
- Thomas Sarmis
- Kostas Tzevanidis
- Xenophon Zabulis

Work supported by the EU projects:

- GRASP



- ActIPret



- Work supported by the ICS FORTH internal Research Programme on Ambient Intelligence Environments



Current work

- EU project **robohow.cog**:

<http://robohow.eu/project>

Cognitive Robots that Learn Complex
Everyday Manipulation Tasks



- EU project **WEARHAP**:
Wearable Haptics for Humans and Robots



- EU project **DALi**:
“Devices for Assisted Living”

<http://www.ict-dali.eu/dali>



- EU project **HOBbit**:
“The Mutual Care Robot”
<http://www.hobbit-project.eu/>



Final remark...

We are not only looking **at** people,
... we are also looking **for** people!!!

argyros@ics.forth.gr

<http://users.ics.forth.gr/~argyros>



**Thanks for your
attention!**

