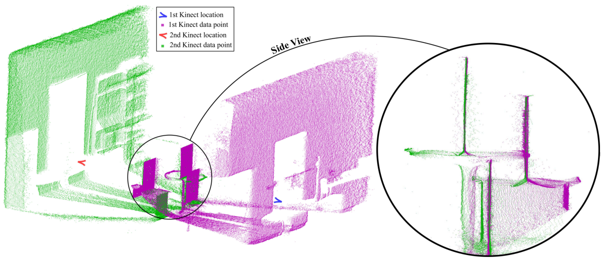
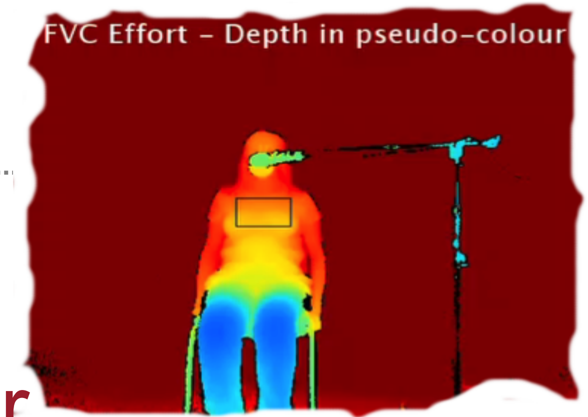
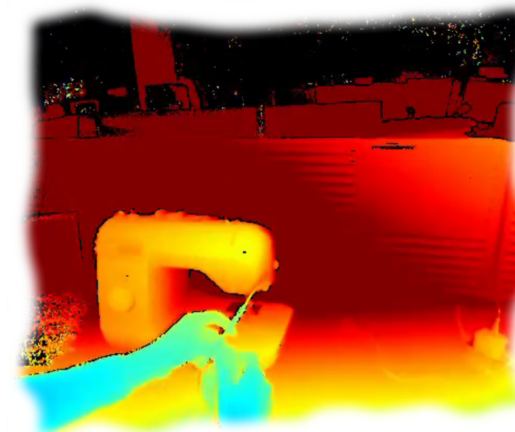




Challenges and Opportunities for Action and Activity Recognition using RGBD Data

BMVA Symposium on Analysis and Processing of RGBD Data



University of BRISTOL

Dima Damen

22 March 2017

Activity Recognition Hierarchy

Visual Sensing – the landscape

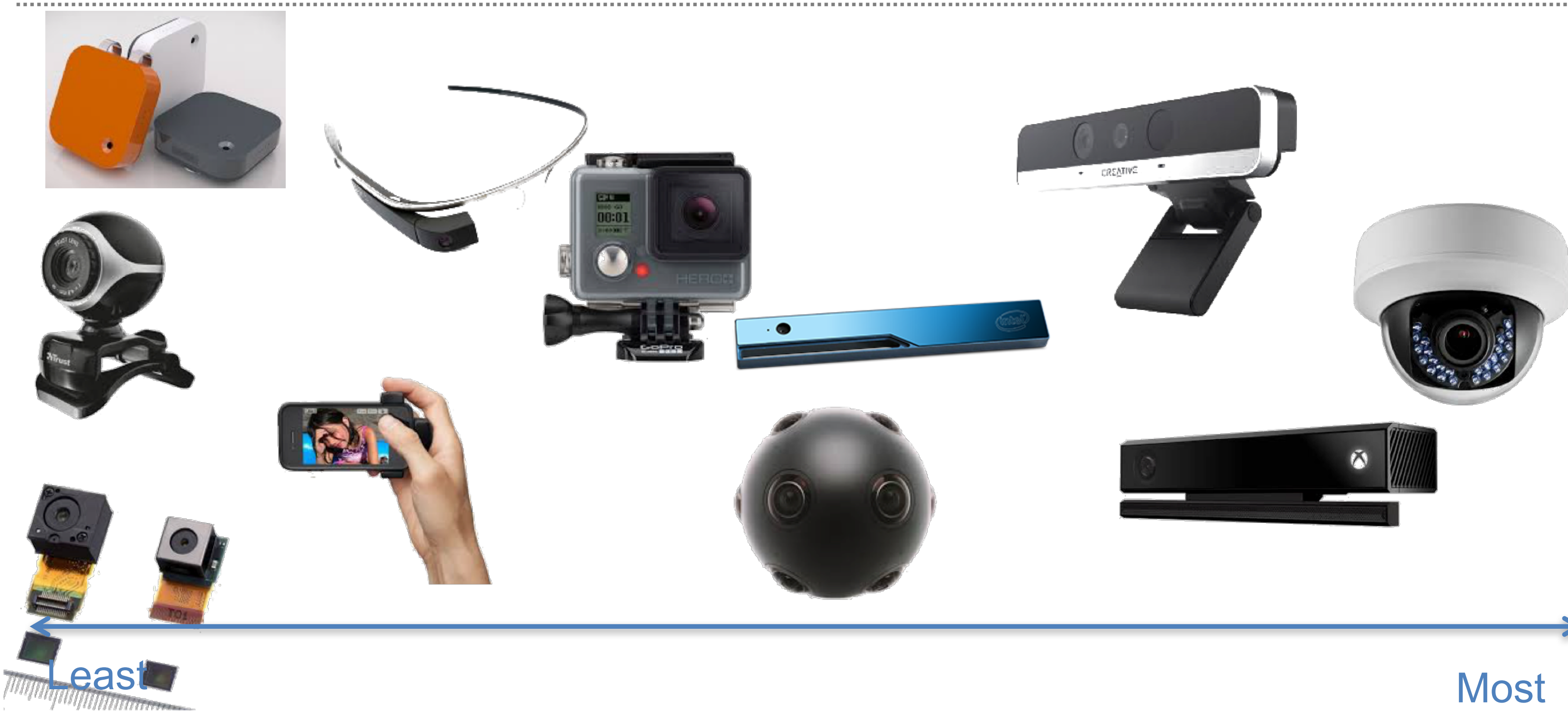
Visual Sensing – the landscape

Least

Most

Expensive

Visual Sensing – the landscape

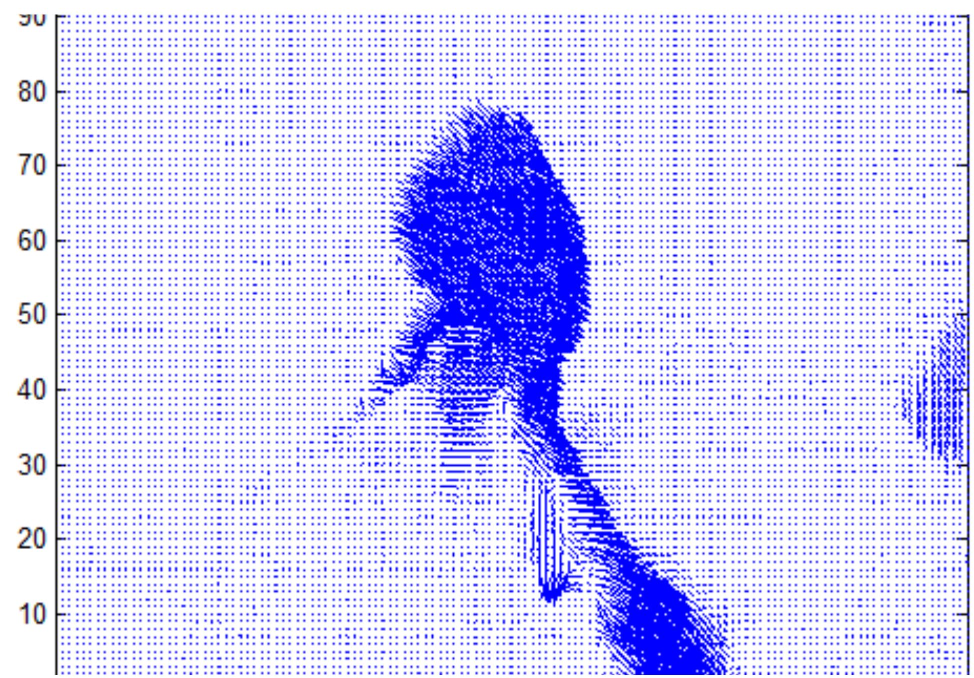


Wearable/Moveable

Visual Sensing – the landscape

- Current affordable RGBD sensors calculate depth on per-frame basis
- They make little usage of the temporal aspect
- Not ideal for action and activity recognition

Visual Sensing – the landscape



Usage of RGBD data for Action &Activity

Three main usages of RGBD sensors in action and activity recognition

1. Separation of Objects at various depths
   - Foreground or Occluder Subtraction
2. Pose Estimation
   - Accurate positioning of body joints
3. Depth from sensor measurements
   - Applications that require accurate depth estimation

Usage of RGBD data for Action &Activity

Three main usages of RGBD sensors in action and activity recognition

1. Separation of Objects at various depths
   - Foreground or Occluder Subtraction
2. Pose Estimation
   - Accurate positioning of body joints
3. Depth from sensor measurements
   - Applications that require accurate depth estimation

Traditionally

- Using background subtraction

- Could be achieved from an individual image by 3D scene analysis

  https://www.youtube.com/watch?v=NyjyGuESkfM#t=1m33s

Event

with: David Hogg

Carried Object Detection

D Damen and D Hogg (2012). Detecting Carried Objects from Sequences of Walking Pedestrians. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
D Damen, D Hogg (2008). Detecting Carried Objects in Short Video Sequences. *European Conference on Computer Vision*

with: David Hogg

Carried Object Detection

D Damen and D Hogg (2012). Detecting Carried Objects from Sequences of Walking Pedestrians. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
D Damen, D Hogg (2008). Detecting Carried Objects in Short Video Sequences. *European Conference on Computer Vision*

with: Dan Lamb

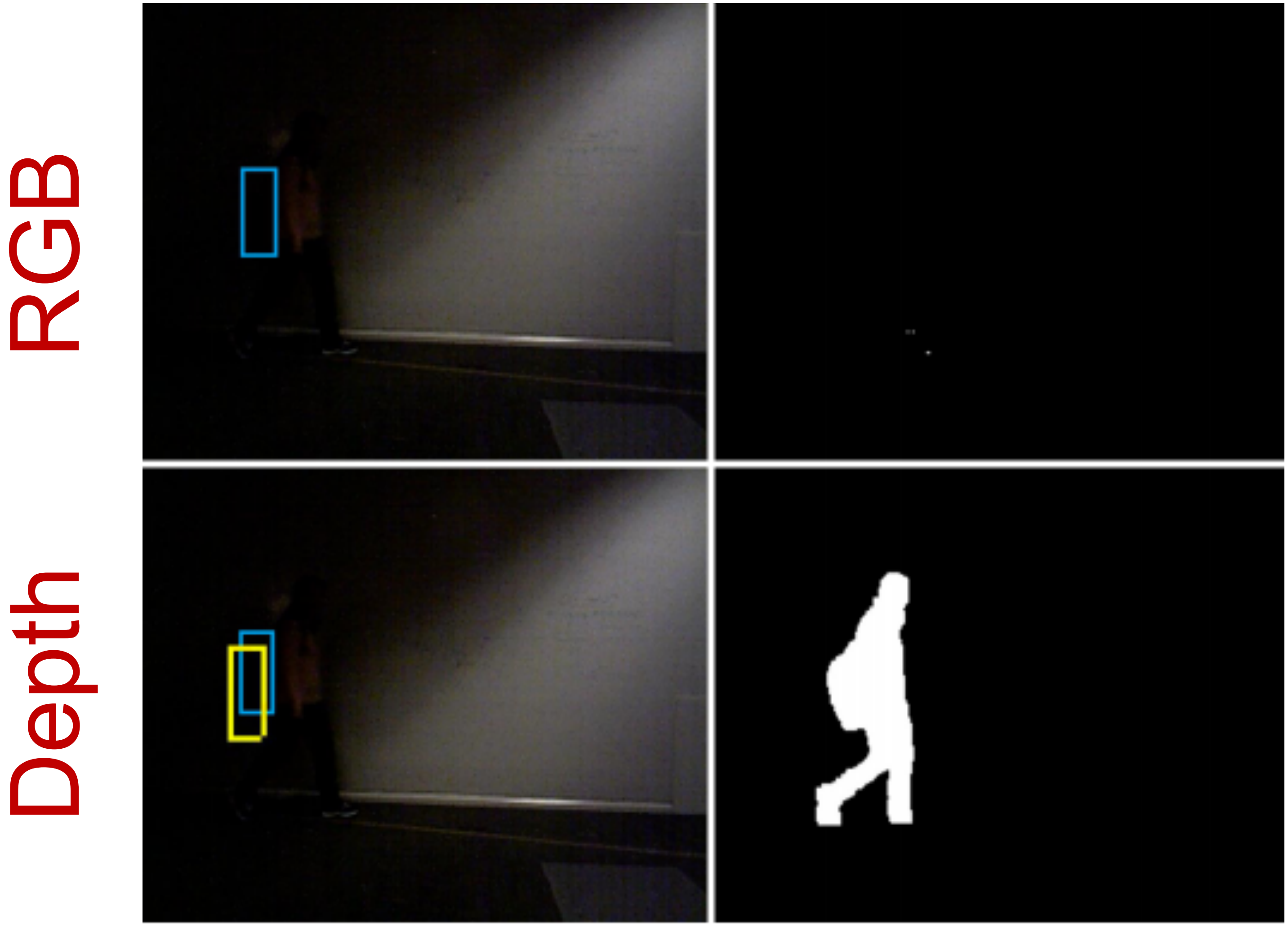
Carried Object Detection + RGBD

D Lamb (2014), MSc Thesis, University of Bristol

Event

with: Dan Lamb

Carried Object Detection ++

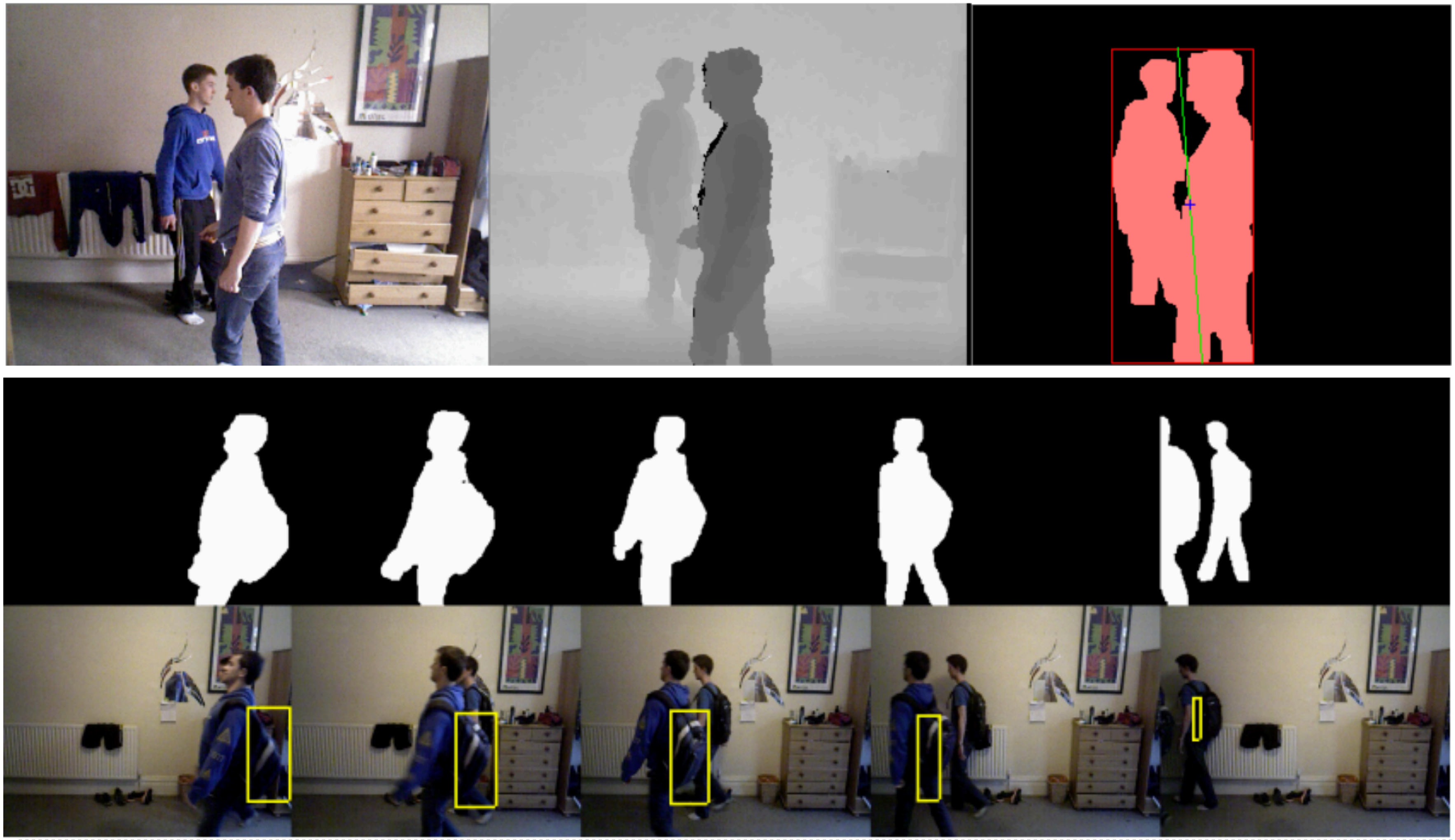


D Lamb (2014), MSc Thesis, University of Bristol

Event

with: Dan Lamb

Carried Object Detection ++



D Lamb (2014), MSc Thesis, University of Bristol

Usage of RGBD data for Action &Activity

Three main usages of RGBD sensors in action and activity recognition

1. Separation of Objects at various depths
    - Foreground or Occluder Subtraction
2. Pose Estimation
    - Accurate positioning of body joints
3. Depth from sensor measurements
    - Applications that require accurate depth estimation

with: Massimo Camplani
+ SPHERE Team

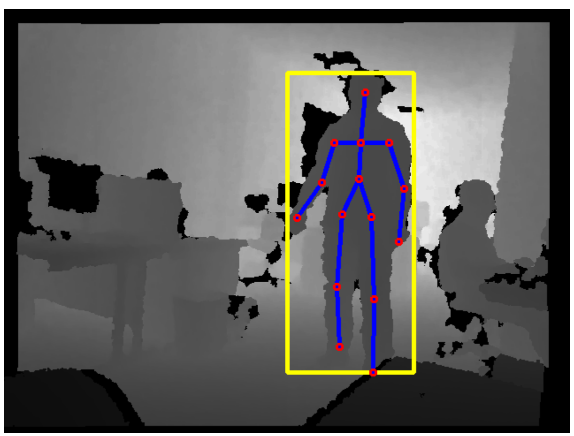
Skeleton Detection

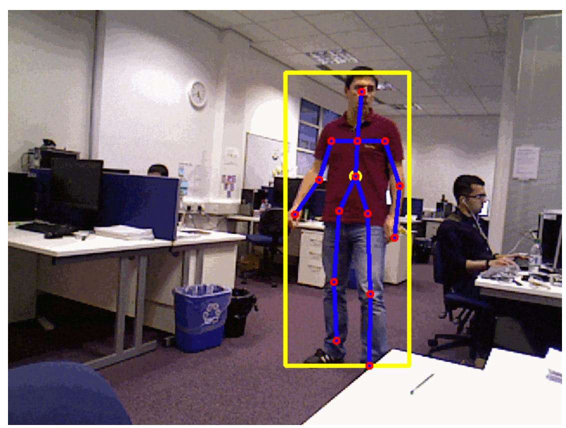
- OpenNI 2.0
  - nite::UserTracker::startSkeletonTracking()
  - nite::Skeleton::getJoint()
- Kinect SDK 2.0
  - skeletonData = new Skeleton[kinect.SkeletonStream.FrameSkeletonArrayLength];

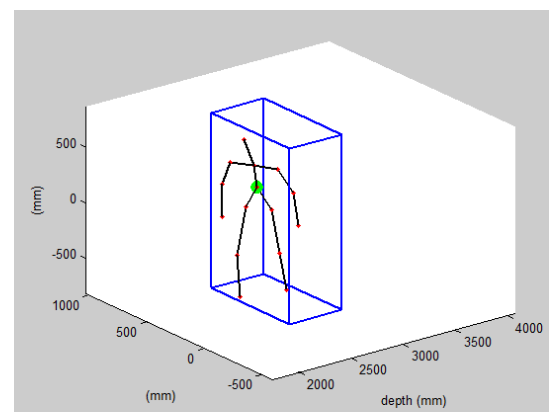
with: Massimo Camplani
+ SPHERE Team

Skeleton Detection

- OpenNI 2.0
  - nite::UserTracker::startSkeletonTracking()
  - nite::Skeleton::getJoint()
- Kinect SDK 2.0
  - skeletonData = new Skeleton[kinect.SkeletonStream.FrameSkeletonArrayLength];







Why skeleton detection?

- View-variant features
  - Hollywood 2 dataset, action class: sit_down

Why skeleton detection?

- Should be view-invariant… but!

Skeleton detection for action and activity recognition

- Depth-based features
- Joint-based features
- Hybrid features

with: Farnoosh Heidarivincheh
Majid Mirmehdi

Action Completion from RGB-D Data

F Heidarivincheh et al (2016). Beyond Action Recognition: Action Completion in RGB-D Data. *British Machine Vision Conference*

with: Farnoosh Heidarivincheh
Majid Mirmehdi

Action Completion from RGB-D Data

Pull*-vs-*pick

Action recognition

Having some predefined action classes, the aim is to recognize the class label of an action.

F Heidarivincheh et al (2016). Beyond Action Recognition: Action Completion in RGB-D Data. *British Machine Vision Conference*

Action

with: Farnoosh Heidarivincheh
Majid Mirmehdi

Action Completion from RGB-D Data

Action recognition

Having some predefined action classes, the aim is to recognize the class label of an action.

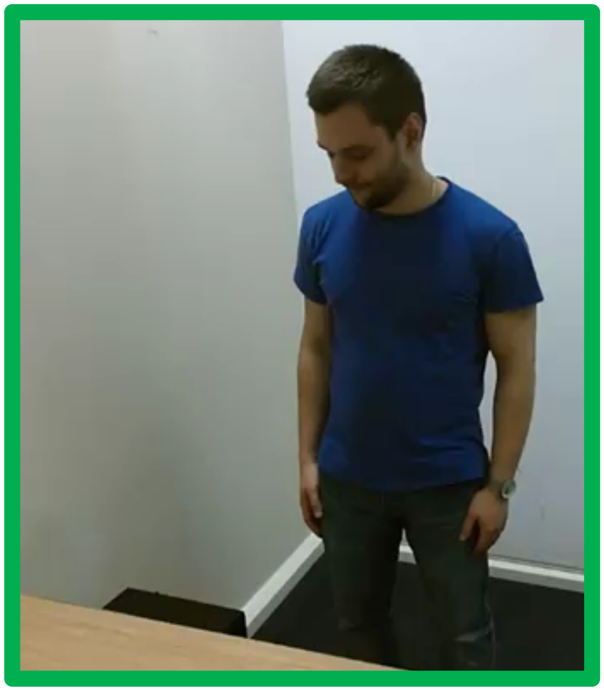
Pull*-vs-*pick

with: Farnoosh Heidarivincheh
Majid Mirmehdi

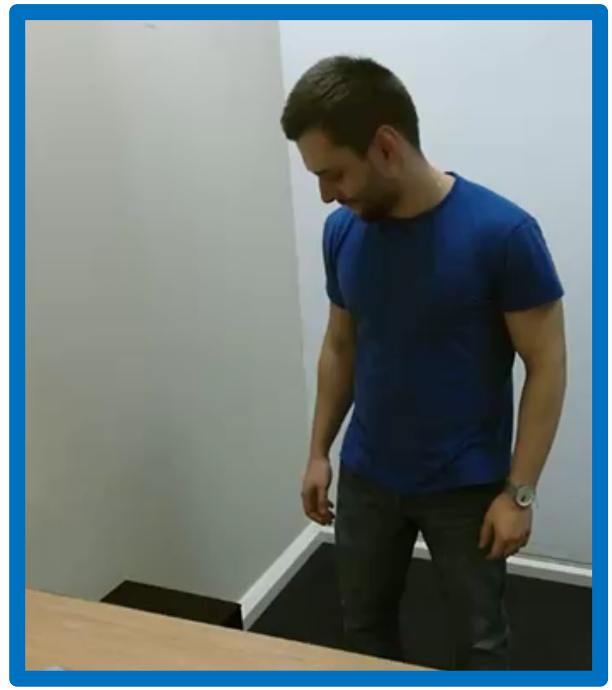
Action Completion from RGB-D Data

What if the observed action is not fully completed!?

Complete *pull*



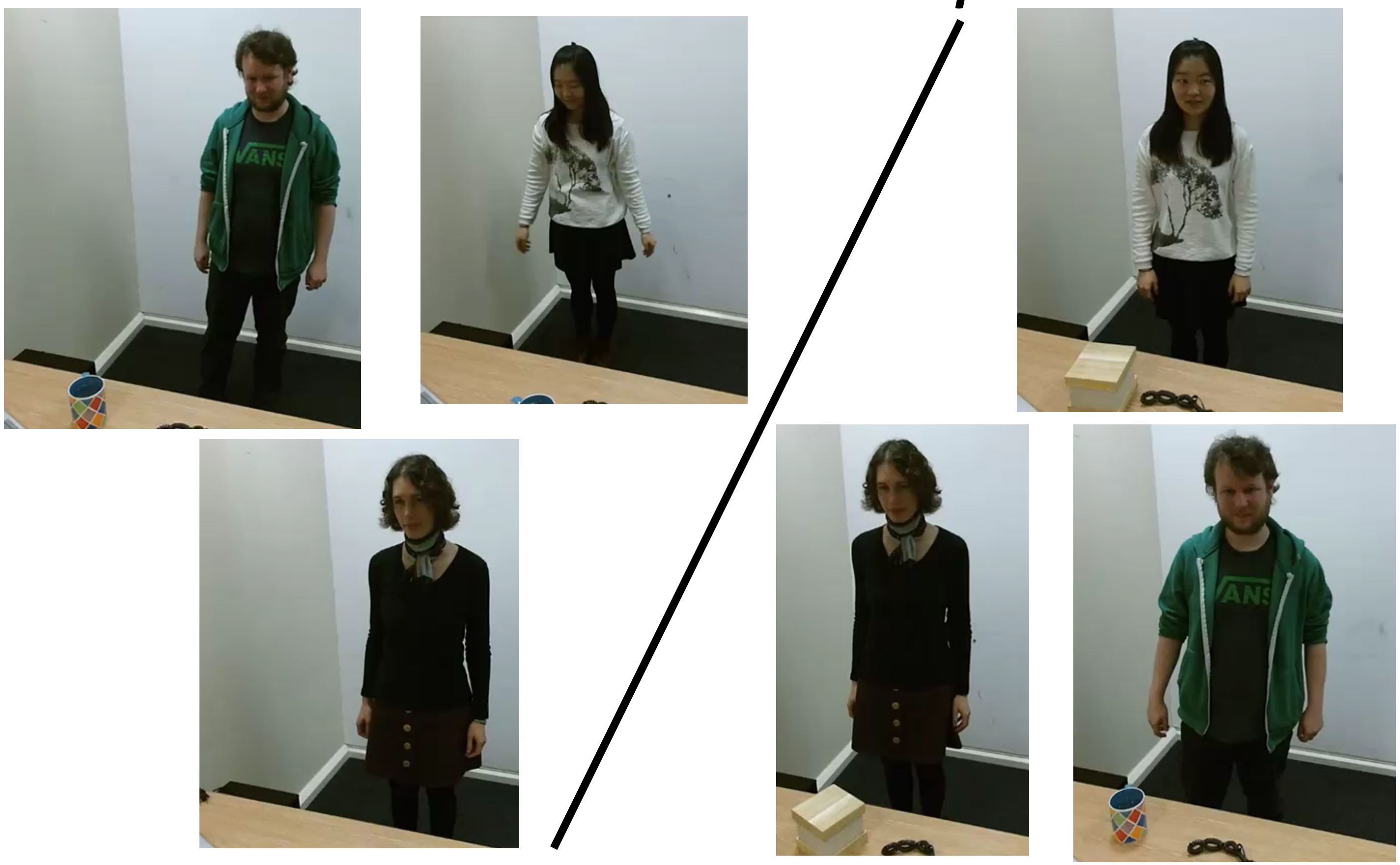
Incomplete *pull*



with: Farnoosh Heidarivincheh
Majid Mirmehdi

Action Completion from RGB-D Data

Pull-vs-***pick***

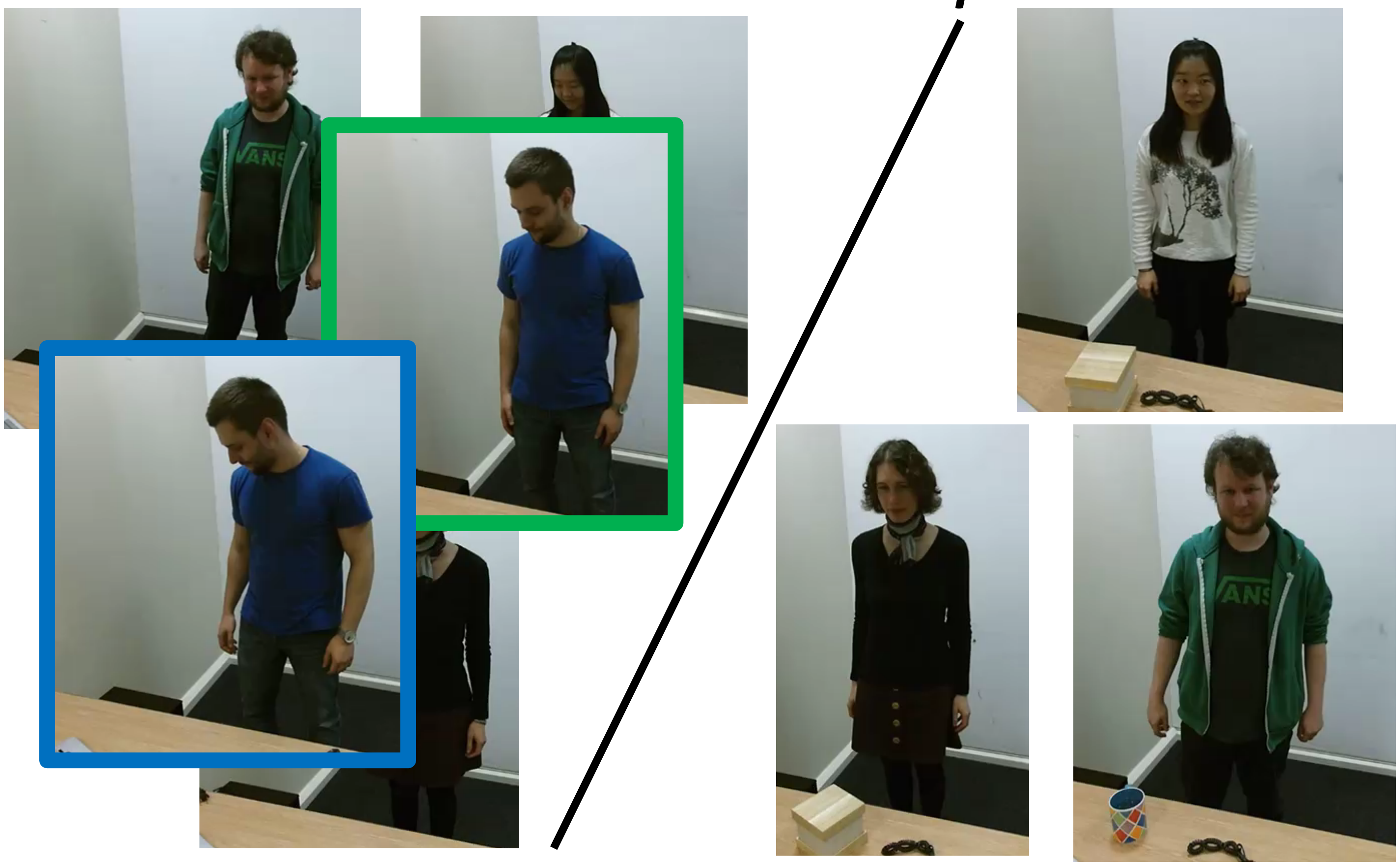


Complete *pull* and incomplete *pull* are introduced to *pull*-vs-*pick* classifier.

with: Farnoosh Heidarivincheh
Majid Mirmehdi

Action Completion from RGB-D Data

Pull*-vs-*pick



Both complete *pull* and incomplete *pull* are classified as *pull*.

with: Farnoosh Heidarivincheh
Majid Mirmehdi

Action Completion from RGB-D Data

Action Completion as a step beyond action recognition

- Action completion aims to recognise whether the action's goal has been successfully achieved.
- In many actions, an observer would be able to make the distinction between complete and incomplete by noticing **subtle differences in motion**.
- Incompletion could result from negligence or forgetfulness, difficulties in performing the action, or could be deliberate.
- We recognise incompletion when the action is **attempted but not completed**.

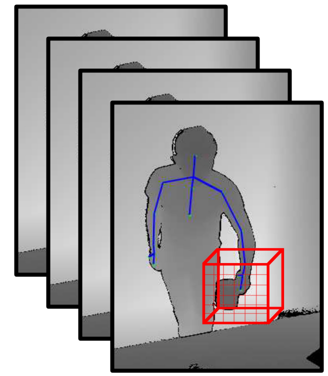
with: Farnoosh Heidarivincheh
Majid Mirmehdi

Action Completion from RGB-D Data

Features and temporal encoding

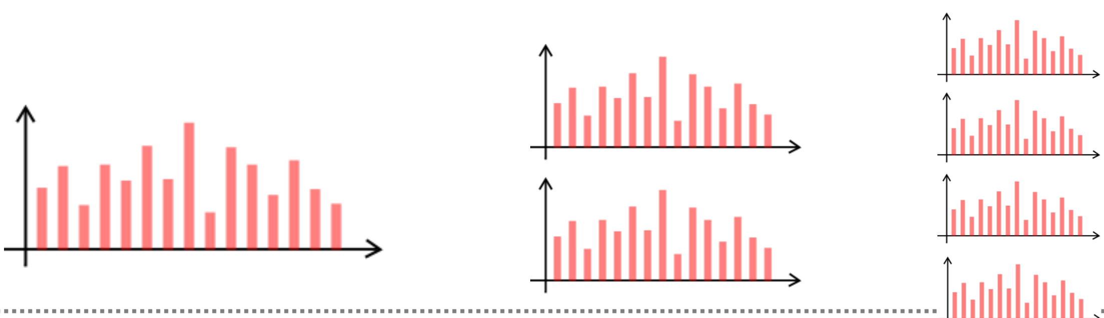
A pool of five depth features:

- Local Occupancy Pattern (LOP)[1]
- Joints Position (JP)[2]
- Joints Relative Position (JRP)[2]
- Joints Relative Angle (JRA)[2]
- Joints Velocity (JV)[2]



LOP: Depth information in the neighbourhood around each joint

Encoding temporal dynamics by Fourier temporal pyramid[1]



Different levels of Fourier temporal pyramid

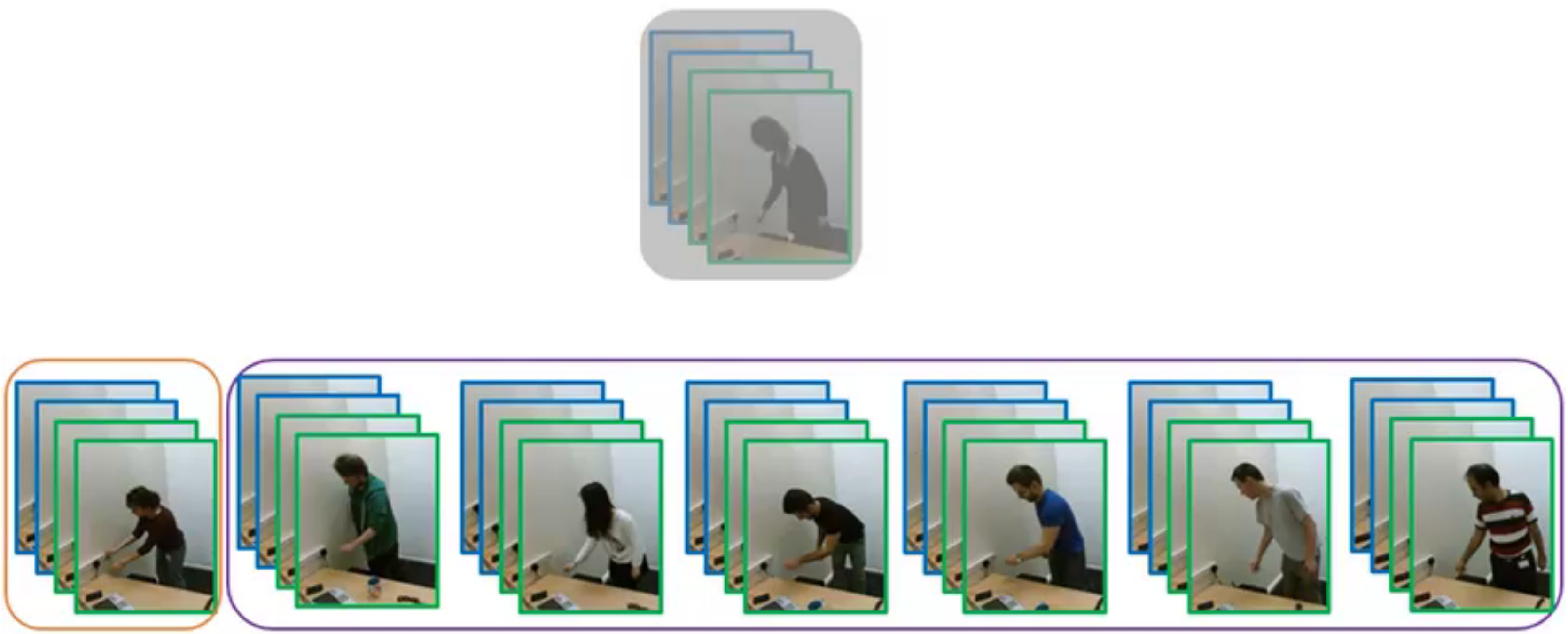
[1] Wang et al, "Mining actionlet ensemble for action recognition with depth cameras" in CVPR 2012.
[2] Tao et al, "A comparative study of pose representation and dynamics modelling for online motion quality assessment" in CVIU 2016

with: Farnoosh Heidarivincheh
Majid Mirmehdi

Action Completion from RGB-D Data

- Notion of completion differs per action → we need a pool of features.
- To choose the most discriminative feature per action:

  A general method: "Leave-one-person-out" cross validation on the training set

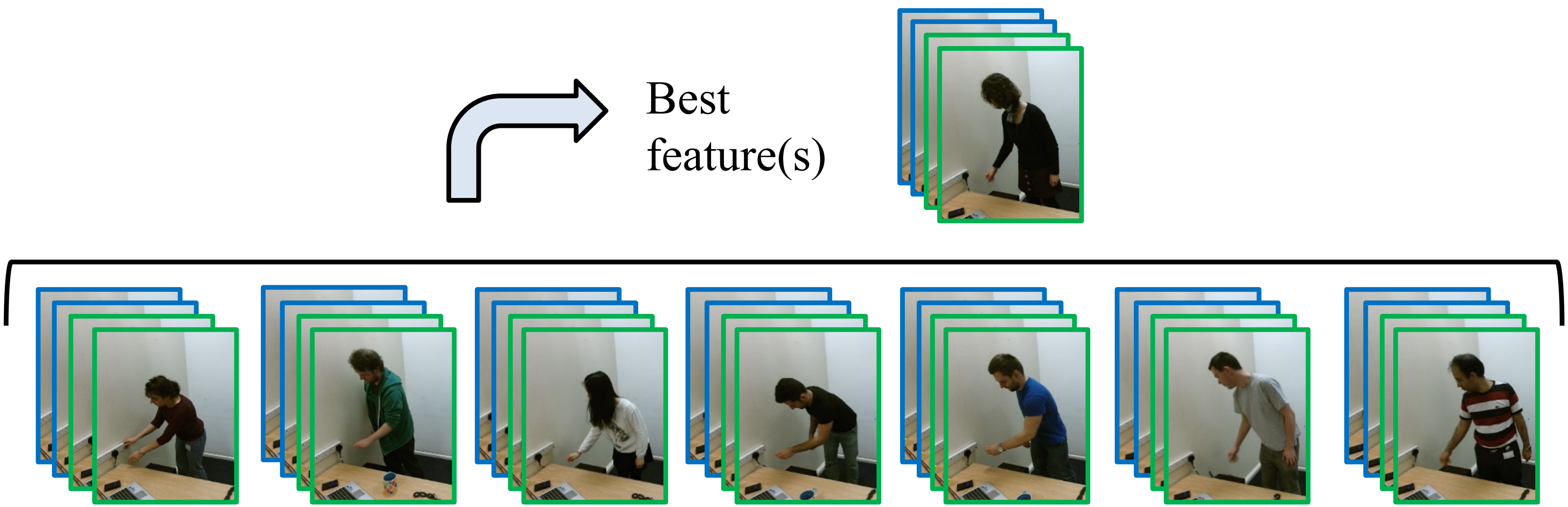


F Heidarivincheh et al (2016). Beyond Action Recognition: Action Completion in RGB-D Data. *British Machine Vision Conference*

with: Farnoosh Heidarivincheh
Majid Mirmehdi

Action Completion from RGB-D Data

- Evidence across folds is accumulated.
- Each feature in the pool of features is ranked by their accuracy.
- The feature(s) that performs the best is selected.



F Heidarivincheh et al (2016). Beyond Action Recognition: Action Completion in RGB-D Data. *British Machine Vision Conference*

with: Farnoosh Heidarivincheh
Majid Mirmehdi

Action Completion from RGB-D Data

Bristol Action Completion Dataset

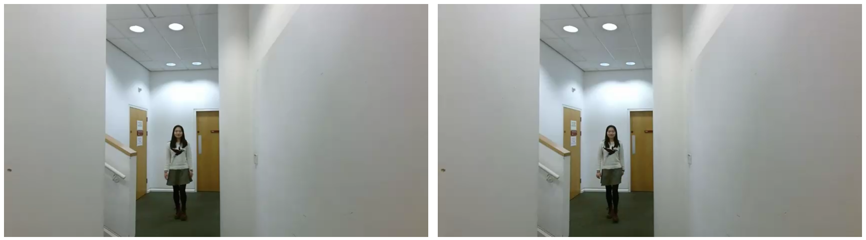
- Containing 414 sequences of complete and incomplete actions
- Comprising 6 actions: *switch, plug, open, pull, pick, drink*

	total #	# complete	# incomplete	$\mu(sec)$	$\sigma(sec)$
switch	67	35	32	3.87	0.72
plug	73	37	36	8.14	2.74
open	68	36	32	6.83	2.70
pull	71	34	37	6.43	1.70
pick	69	33	36	4.03	1.16
drink	66	34	32	8.83	2.09

with: Farnoosh Heidarivincheh
Majid Mirmehdi

Bristol Action Completion Dataset

complete | incomplete | complete | incomplete



switch

pull

plug

pick

open

drink

F Heidarivincheh et al (2016). Beyond Action Recognition: Action Completion in RGB-D Data. *British Machine Vision Conference*

with: Farnoosh Heidarivincheh
Majid Mirmehdi

Action Completion from RGB-D Data

Experiment A: Complete Action Recognition

- complete sequences were used in training and testing by a one-vs-all SVM.

	LOP	JP	JRP	JRA	JV
switch	**100**	99	99	**100**	**100**
plug	**99**	92.3	91.9	92.8	97.1
open	97.6	98.1	**100**	94.7	94.3
pull	**98.1**	91.4	91.4	94.7	92.3
pick	97.6	99.5	**100**	96.7	95.2
drink	99	97.1	98.1	99	**100**
Average	**98.6**	96.3	96.7	96.3	96.5

- Various features perform comparably with high % accuracy.

F Heidarivincheh et al (2016). Beyond Action Recognition: Action Completion in RGB-D Data. *British Machine Vision Conference*

with: Farnoosh Heidarivincheh
Majid Mirmehdi

Experiment B: Incomplete Action Recognition

- Complete samples were used for training.
- Incomplete test sequences were classified by finding their nearest neighbour.

LOP

	switch	*plug*	*open*	*pull*	*pick*	*drink*
∼switch	100	0	0	0	0	0
∼plug	2.7	91.9	0	0	0	5.4
∼open	0	0	75	11.1	8.3	5.6
∼pull	0	29.4	0	61.8	2.9	5.9
∼pick	0	0	15.2	0	27.3	57.6
∼drink	0	0	0	0	29.4	70.6

JP

	switch	*plug*	*open*	*pull*	*pick*	*drink*
∼switch	64.5	3.2	0	9.7	22.6	0
∼plug	0	83.8	0	10.8	5.4	0
∼open	0	5.6	86.1	5.6	2.8	0
∼pull	0	32.4	0	52.9	14.7	0
∼pick	0	33.3	15.2	9.1	42.4	0
∼drink	0	2.9	11.8	0	79.4	5.9

JRP

	switch	*plug*	*open*	*pull*	*pick*	*drink*
∼switch	61.3	12.9	0	6.5	19.4	0
∼plug	0	83.8	5.4	5.4	5.4	0
∼open	0	5.6	88.9	5.6	0	0
∼pull	0	32.4	11.8	38.2	14.7	2.9
∼pick	0	39.4	6.1	3	51.5	0
∼drink	0	2.9	11.8	0	85.3	0

JRA

	switch	*plug*	*open*	*pull*	*pick*	*drink*
∼switch	100	0	0	0	0	0
∼plug	2.7	86.5	0	10.8	0	0
∼open	0	5.6	88.9	5.6	0	0
∼pull	0	44.1	0	50	2.9	2.9
∼pick	0	12.1	12.1	0	69.7	6.1
∼drink	0	0	11.8	0	50	38.2

JV

	switch	*plug*	*open*	*pull*	*pick*	*drink*
∼switch	83.9	0	12.9	0	0	3.2
∼plug	2.7	54.1	2.7	2.7	0	37.8
∼open	0	2.8	0	5.6	0	91.7
∼pull	0	26.5	2.9	44.1	0	26.5
∼pick	0	33.3	3	36.4	27.3	0
∼drink	0	47.1	32.4	0	2.9	17.6

- Only some features distinguish the subtle changes between complete and incomplete.

with: Farnoosh Heidarivincheh
Majid Mirmehdi

Action Completion from RGB-D Data

Experiment C: Complete-vs-Incomplete Action Recognition

- Complete and incomplete samples of the same action were used in training and testing

	LOP	JP	JRP	JRA	JV
switch	**100**	85.1	85.1	**100**	**100**
plug	83.6	87.7	78.1	79.5	**94.5**
open	**97.1**	95.6	**97.1**	95.6	**97.1**
pull	87.3	71.8	77.5	88.7	**94.4**
pick	92.8	94.2	**98.6**	**98.6**	95.7
drink	97	97	97	97	**100**

- Again, the features have different success rates for the various actions.

F Heidarivincheh et al (2016). Beyond Action Recognition: Action Completion in RGB-D Data. *British Machine Vision Conference*

with: Farnoosh Heidarivincheh
Majid Mirmehdi

Action Completion from RGB-D Data

Experiment D: Selecting Features for Action Completion

- A general model using cross validation on training data

	Subjects								
	1	2	3	4	5	6	7	8	total
switch	100	100	100	100	100	100	100	100	100
	LOP,JRA,JV	LOP,JRA,JV	LOP,JV	LOP,JV	LOP,JV	LOP,JRA,JV	LOP,JV	LOP,JV	
plug	83.3	100	87.5	100	88.9	100	100	100	94.5
	JV	JV	JV	JV	JV	JV	JV	JV	
open	100	85.7	100	100	100	87.5	90	100	95.6
	JV	JV	JP,JRP	LOP,JRP,JV	JRP	JRA	JV	LOP,JRP,JRA,JV	
pull	88.9	100	100	100	100	87.5	80	100	94.4
	JV	JV	JV	JRA,JV	JV	JV	JV	JV	
pick	90	100	100	100	100	100	50	100	92.8
	JRA	JRA	JRA,JV	JP,JRA	JRA	JRP,JRA	LOP,JRA	JRA	
drink	77.8	100	100	100	100	100	100	100	97
	LOP,JP,JRP,JRA,JV	JV	JV	JV	JV	JV	JV	JV	
								total	95.7

- Results show high success rates compared to the best performance in complete-vs-incomplete action recognition

F Heidarivincheh et al (2016). Beyond Action Recognition: Action Completion in RGB-D Data. *British Machine Vision Conference*

with: Farnoosh Heidarivincheh
Majid Mirmehdi

Action Completion from RGB-D Data

Examples of success

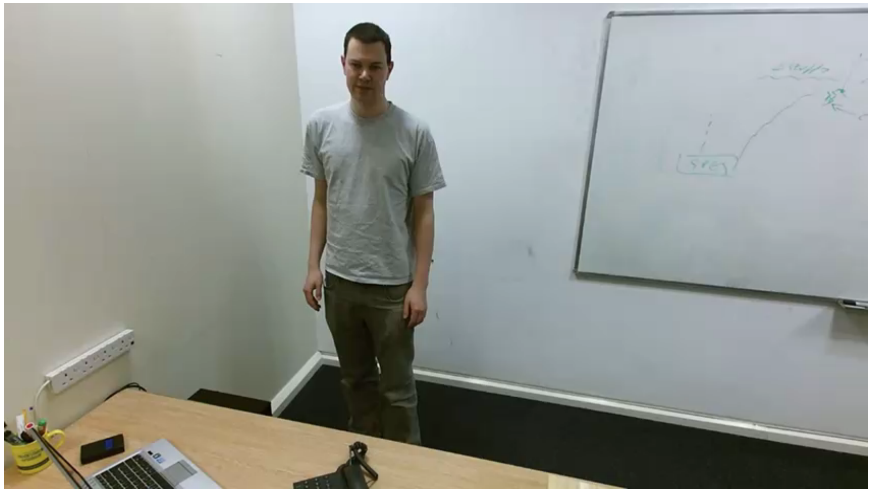
Complete *switch*
Classified as complete *switch* ✓

Incomplete *open*
Classified as incomplete *open* ✓

with: Farnoosh Heidarivincheh
Majid Mirmehdi

Action Completion from RGB-D Data

Examples of failure



Complete *drink*
Classified as incomplete *drink* ✗

Incomplete *pull*
Classified as complete *pull* ✗

Usage of RGBD data for Action &Activity

Three main usages of RGBD sensors in action and activity recognition

1. Separation of Objects at various depths
   - Foreground or Occluder Subtraction
2. Pose Estimation
   - Accurate positioning of body joints
3. Depth from sensor measurements
   - Applications that require accurate depth estimation

The need for (exact) depth measurements

1. Localisation and mapping
   - Wearable RGBD – Task monitoring
2. Tracking change in depth
   - Breathing monitoring and Remote Pulmonary Function Testing
3. Distance measurements (in metres)
   - Functional mobility testing
   - Routine analysis

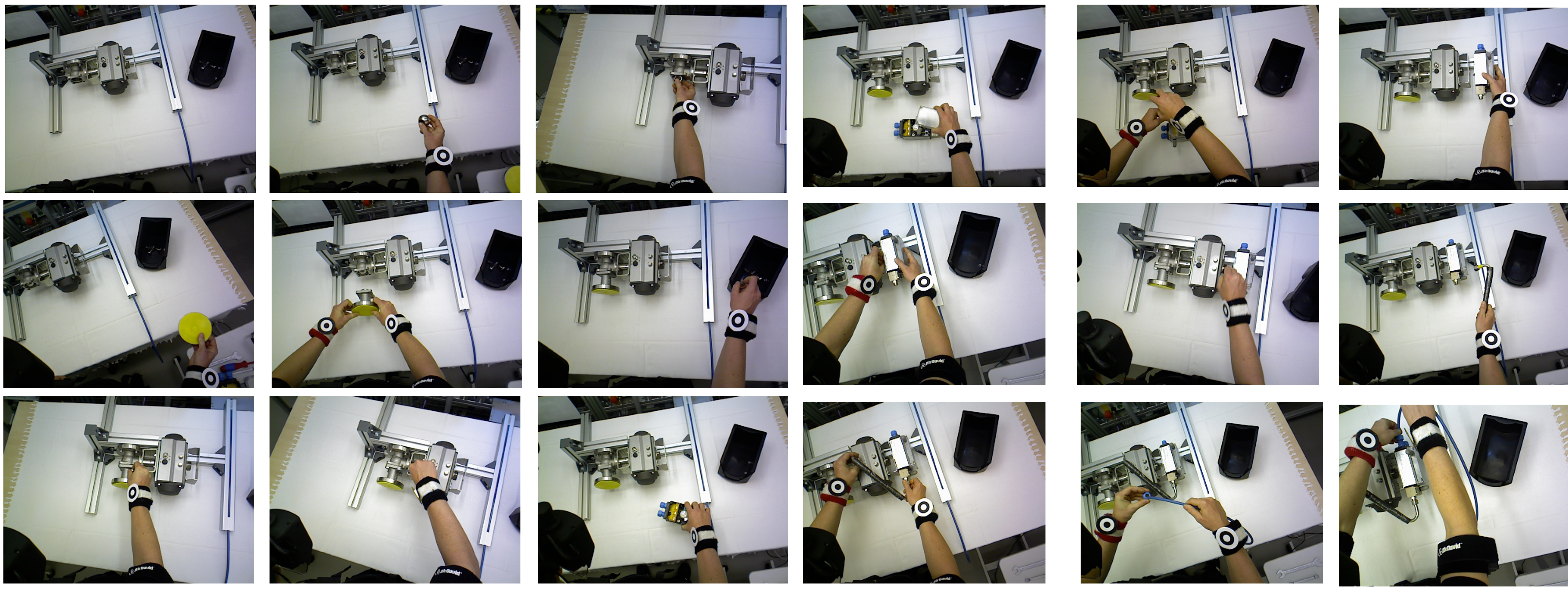
Task Monitoring

with: Andrew Gee
Andrew Calway
Walterio Mayol-Cuevas
+ collaborators

- EU FP7 (2010 – 2013)
- COGNITO: Cognitive Workflow Capturing and Rendering with On-Body Sensor Networks
- Fully-Wearable Sensors

Task Monitoring

with: Andrew Gee
Andrew Calway
Walterio Mayol-Cuevas
+ collaborators



University of BRISTOL

G Bleser et al (2015). Cognitive Learning, Monitoring and Assistance of Industrial Workflows Using Egocentric Sensor Networks. *PLOS ONE*

D Damen et al (2012). Real-time Learning and Detection of 3D Texture-less Objects: A Scalable Approach. *British Machine Vision Conference (BMVC)*

D Damen et al (2012). Egocentric Real-time Workspace Monitoring using an RGB-D Camera. *IEEE/RSJ International Conference on Intelligent Robots and Systems*

Task Monitoring

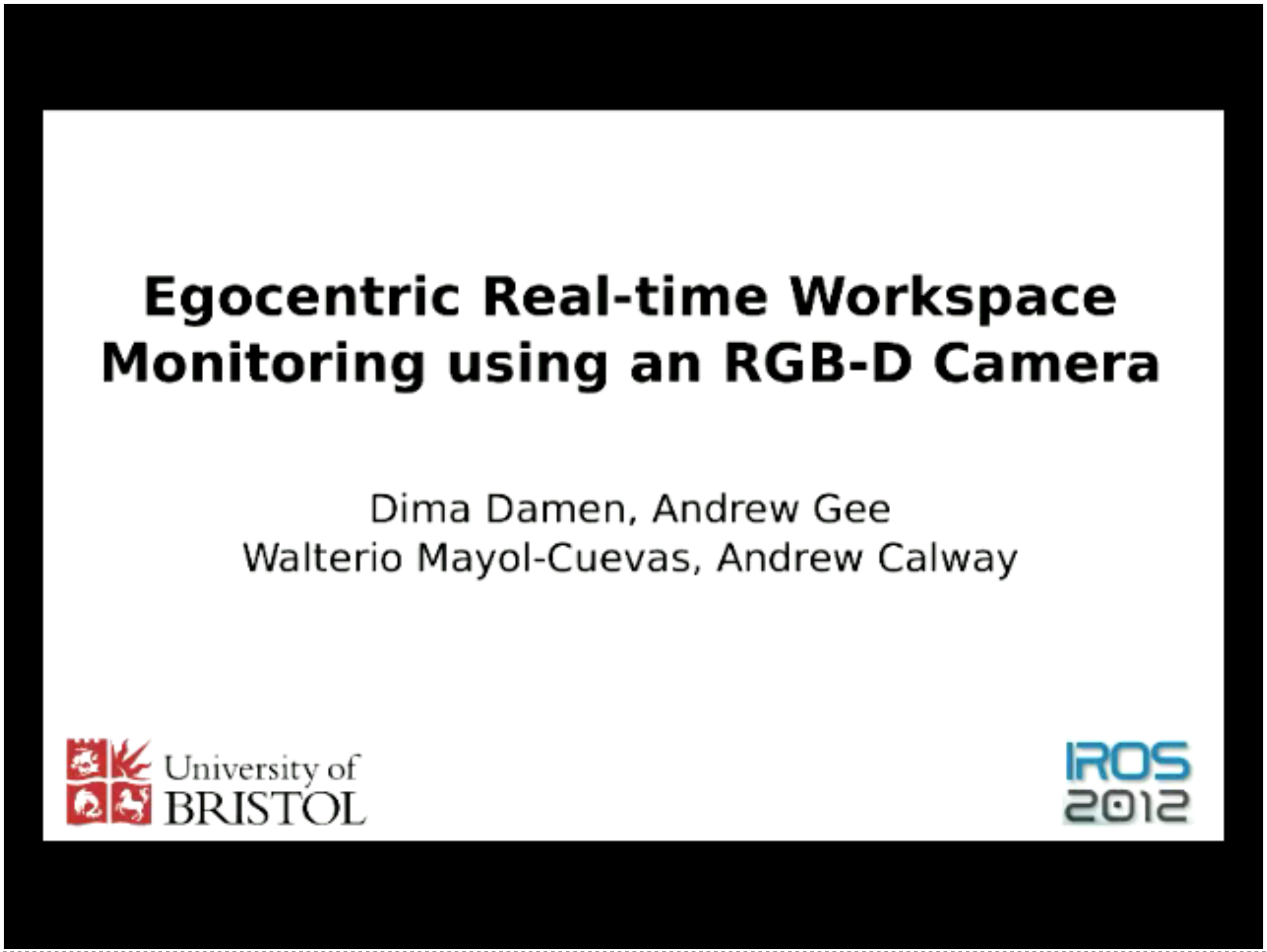
with: Andrew Gee
Andrew Calway
Walterio Mayol-Cuevas
+ collaborators

G Bleser et al (2015). Cognitive Learning, Monitoring and Assistance of Industrial Workflows Using Egocentric Sensor Networks. *PLOS ONE*
D Damen et al (2012). Real-time Learning and Detection of 3D Texture-less Objects: A Scalable Approach. *British Machine Vision Conference (BMVC)*
D Damen et al (2012). Egocentric Real-time Workspace Monitoring using an RGB-D Camera. *IEEE/RSJ International Conference on Intelligent Robots and Systems*

Activity

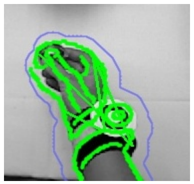
Task Monitoring

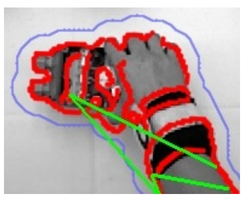
with: Andrew Gee
Andrew Calway
Walterio Mayol-Cuevas
+ collaborators

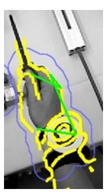


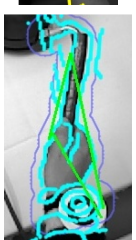
G Bleser et al (2015). Cognitive Learning, Monitoring and Assistance of Industrial Workflows Using Egocentric Sensor Networks. *PLOS ONE*
D Damen et al (2012). Real-time Learning and Detection of 3D Texture-less Objects: A Scalable Approach. *British Machine Vision Conference (BMVC)*
D Damen et al (2012). Egocentric Real-time Workspace Monitoring using an RGB-D Camera. *IEEE/RSJ International Conference on Intelligent Robots and Systems*

with: Andrew Gee
Andrew Calway
Walterio Mayol-Cuevas
+ collaborators

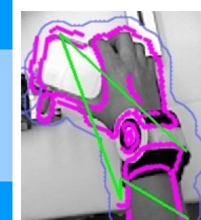


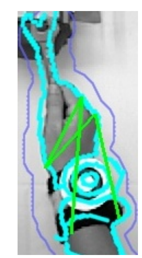






Obj	Recall	Precision
Ball	81%	12%
Bearing	23%	83%
Box	63%	92%
Box Cover	53%	48%
Screw Driver	15%	34%
Spanner	57%	29%
Rod	60%	49%

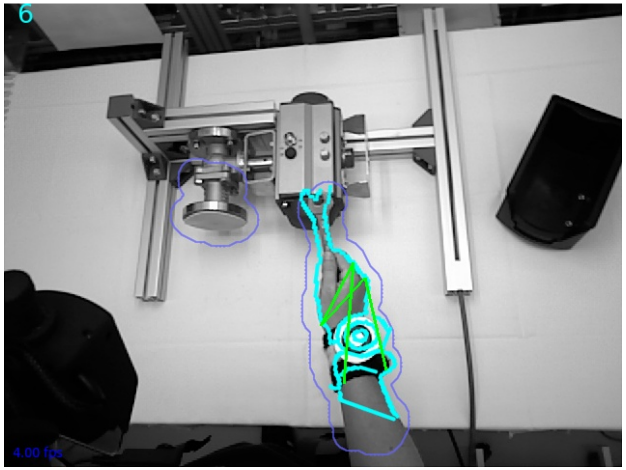


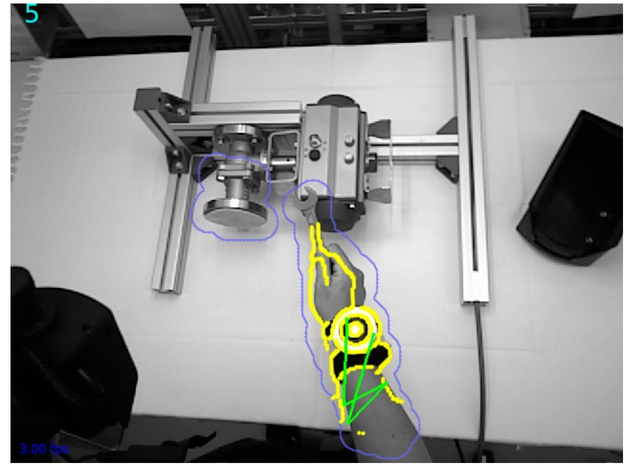


Nov 29th 2012

Task Monitoring

with: Andrew Gee
Andrew Calway
Walterio Mayol-Cuevas
+ collaborators





Obj	Recall	Precision
Screw Driver	15%	34%
Spanner	57%	29%

G Bleser et al (2015). Cognitive Learning, Monitoring and Assistance of Industrial Workflows Using Egocentric Sensor Networks. *PLOS ONE*
D Damen et al (2012). Real-time Learning and Detection of 3D Texture-less Objects: A Scalable Approach. *British Machine Vision Conference (BMVC)*
D Damen et al (2012). Egocentric Real-time Workspace Monitoring using an RGB-D Camera. *IEEE/RSJ International Conference on Intelligent Robots and Systems*

Activity

Task Monitoring

with: Andrew Gee
Andrew Calway
Walterio Mayol-Cuevas
+ collaborators

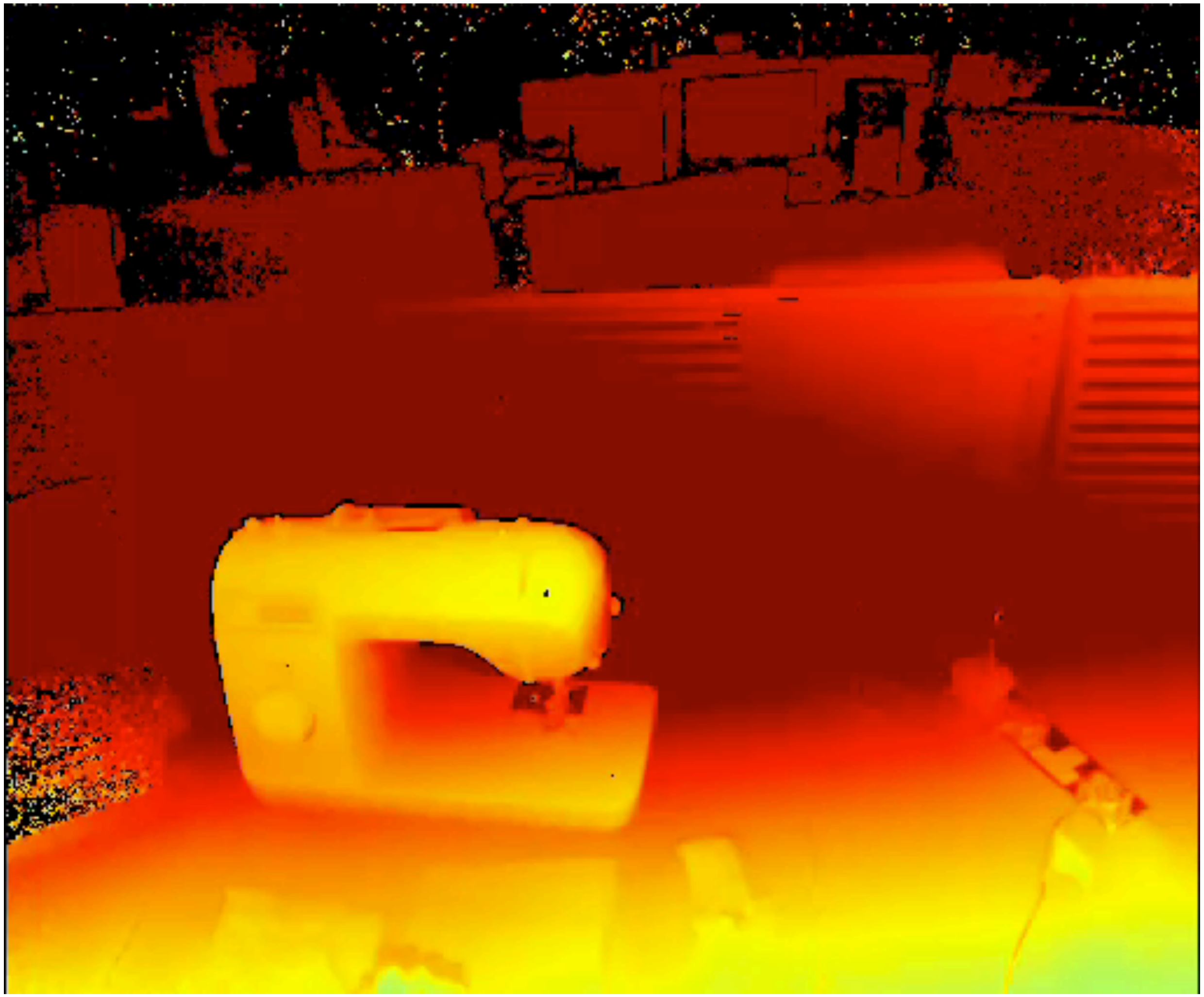
Cognito

G Bleser et al (2015). Cognitive Learning, Monitoring and Assistance of Industrial Workflows Using Egocentric Sensor Networks. *PLOS ONE*
D Damen et al (2012). Real-time Learning and Detection of 3D Texture-less Objects: A Scalable Approach. *British Machine Vision Conference (BMVC)*
D Damen et al (2012). Egocentric Real-time Workspace Monitoring using an RGB-D Camera. *IEEE/RSJ International Conference on Intelligent Robots and Systems*

Activity

Task Monitoring - 2017

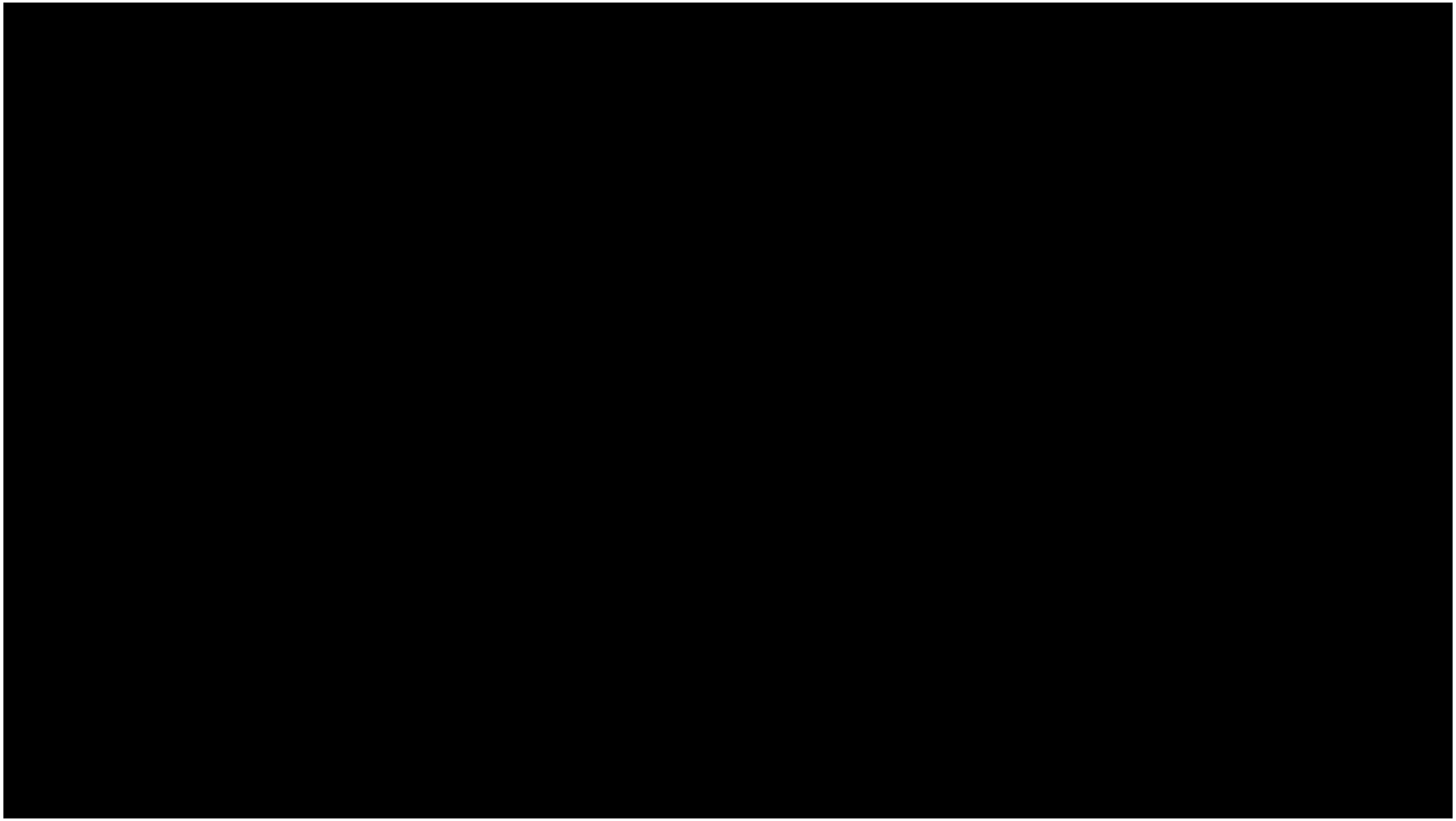
with: Longfei Chen
Kazuaki Kondo
Yuichi Nakamura
Walterio Mayol-Cuevas



L Chen et al (2017). Hotspots Detection for Machine Operation in Egocentric Vision. IAPR conference on *Machine Vision Applications*

Task Monitoring - 2017

with: Longfei Chen
Kazuaki Kondo
Yuichi Nakamura
Walterio Mayol-Cuevas

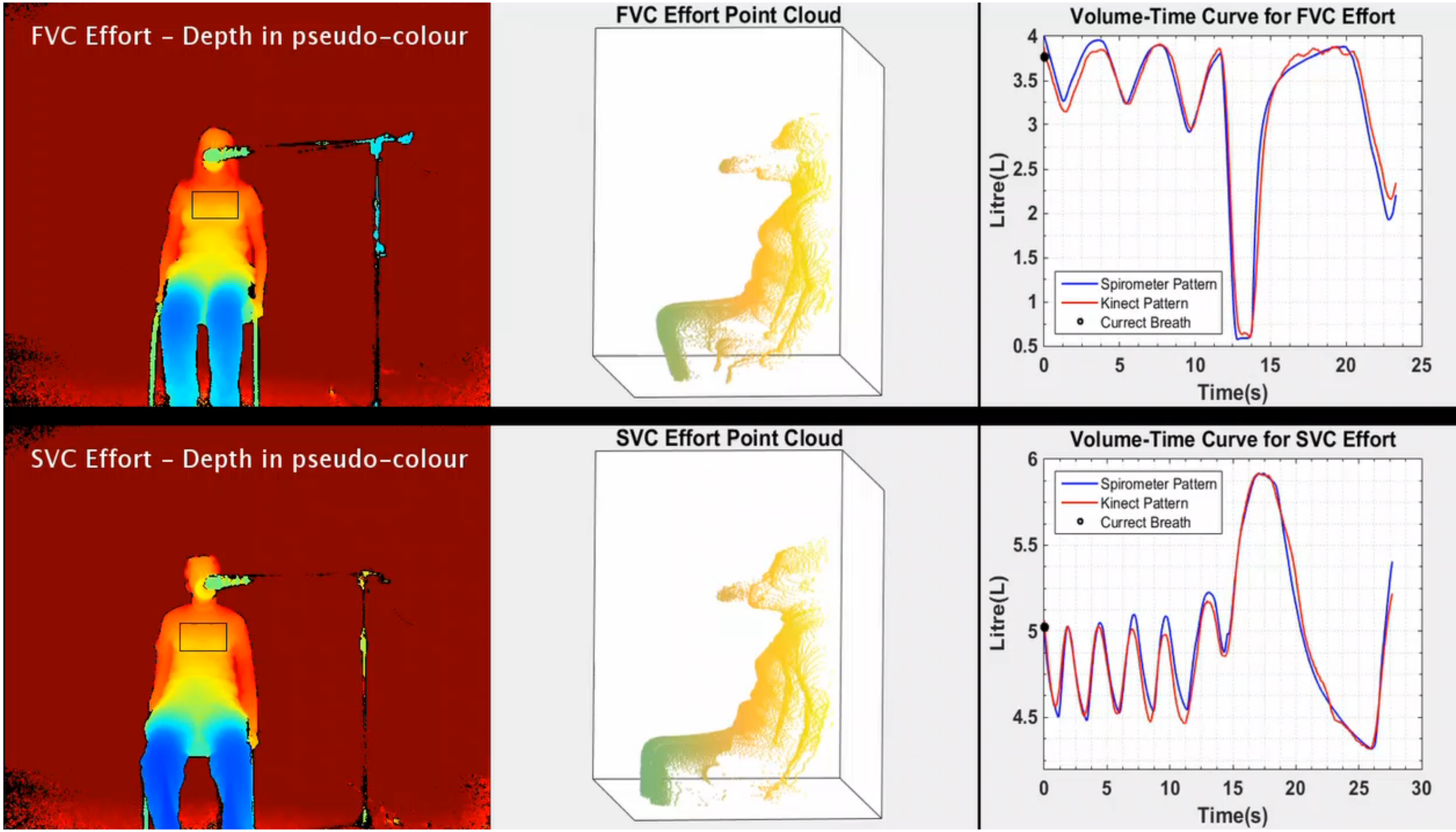


L Chen et al (2017). Hotspots Detection for Machine Operation in Egocentric Vision. IAPR conference on *Machine Vision Applications*

Action

with: Vahid Soleimani
Majid Mirmehdi
Sion Hannuna
Massimo Camplani

Remote Pulmonary Function Testing



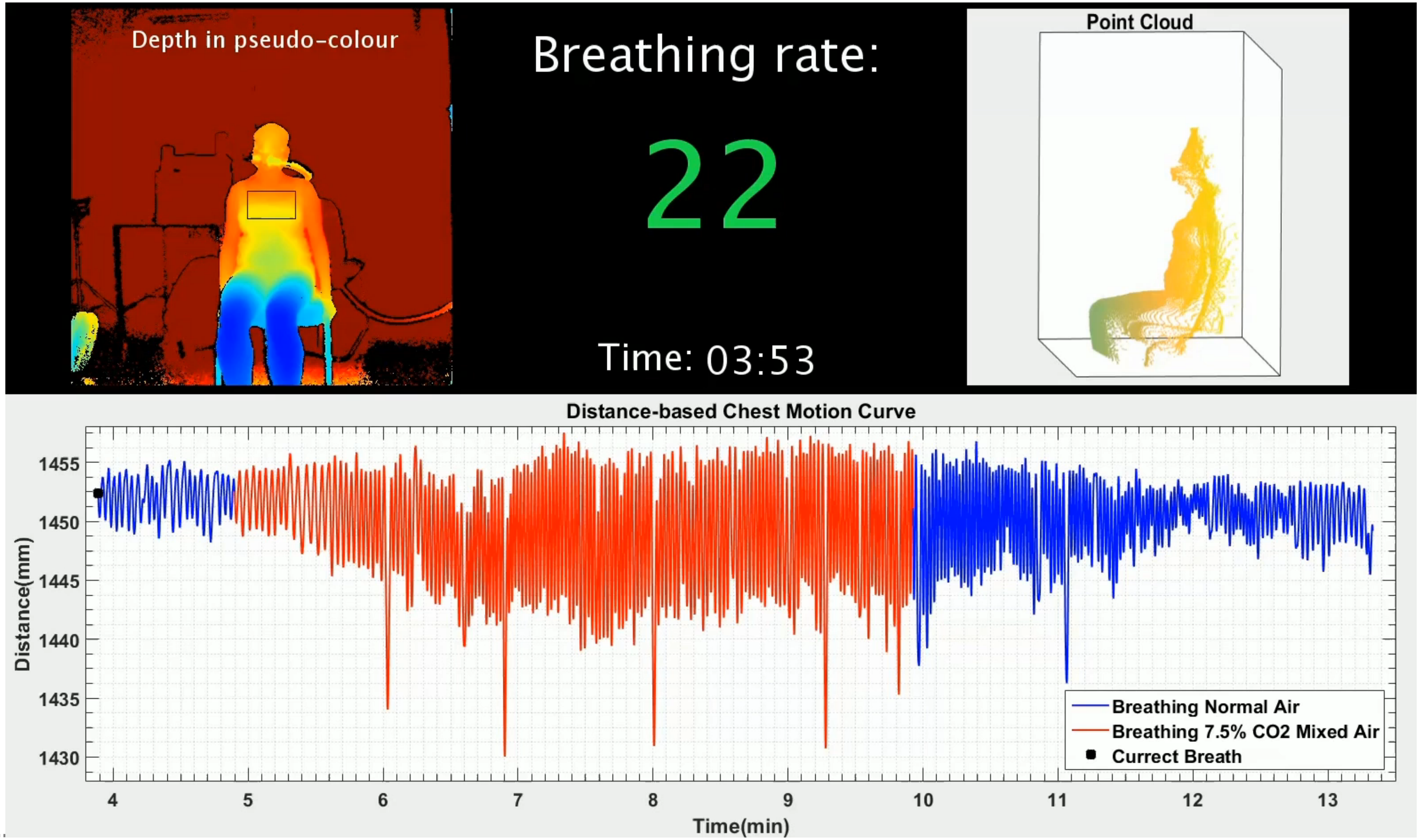
Soleimani et al (2015), Remote Pulmonary Function Testing using a Depth Sensor, *IEEE/CAS-EMB Biomedical Circuits and System Conferences*
Soleimani et al (2017), Remote, Depth-based Lung Function Assessment. *IEEE Transaction on Biomedical Engineering*

Anxiety Detection

with: Vahid Soleimani
Majid Mirmehdi
+ collaborators

Anxiety Detection

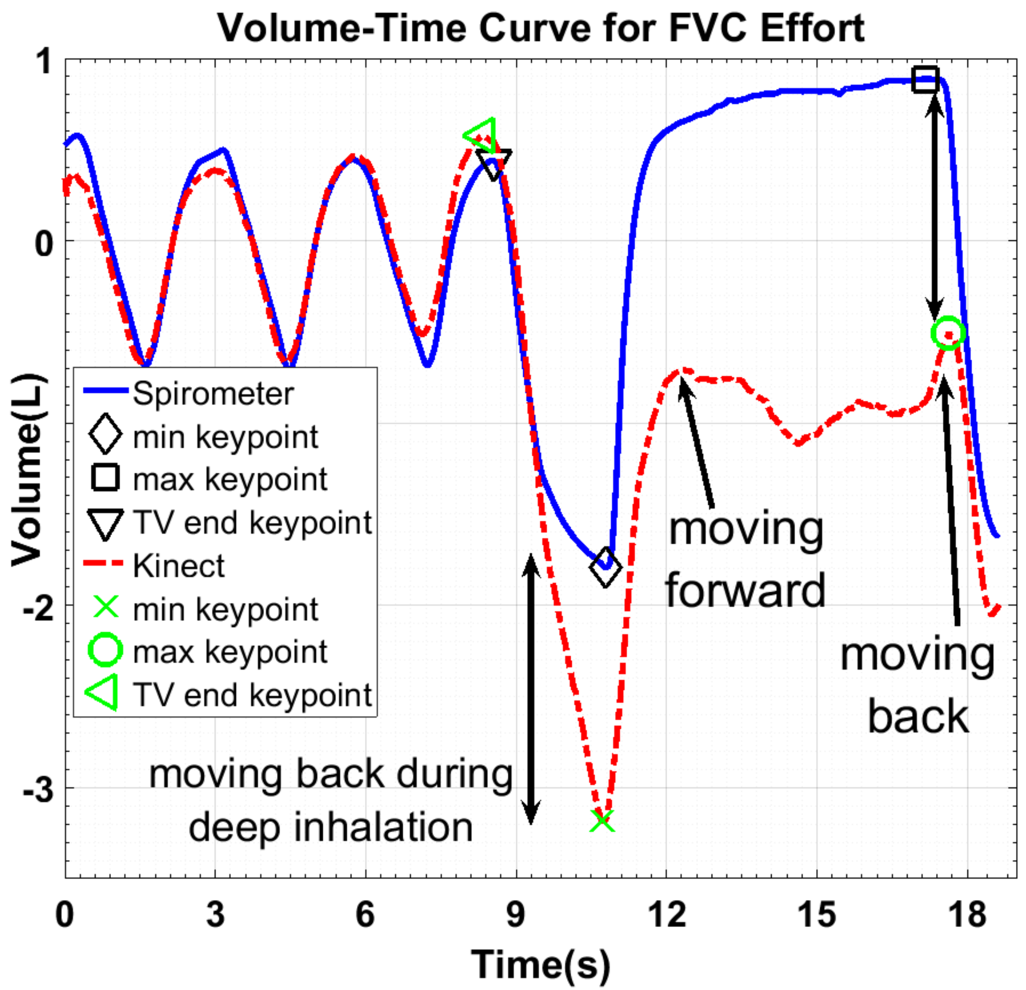
with: Vahid Soleimani
Majid Mirmehdi
+ collaborators



Action

Remote Pulmonary Function Testing

with: Vahid Soleimani
Majid Mirmehdi
Sion Hannuna
Massimo Camplani



Soleimani et al (2015), Remote Pulmonary Function Testing using a Depth Sensor, *IEEE/CAS-EMB Biomedical Circuits and System Conferences*
Soleimani et al (2017), Remote, Depth-based Lung Function Assessment. *IEEE Transaction on Biomedical Engineering*

Remote Pulmonary Function Testing

with: Vahid Soleimani
Majid Mirmehdi
Sion Hannuna
Massimo Camplani

- Two Kinects facing each other with ~3m distance.
- Subject sits in between on a backless chair.
- Since Kinects capture separate sides, there is no interference by this setup.
- Using 3 double sided chessboards to increase calibration accuracy.

Action

Two Facing Kinects

with: Vahid Soleimani
Majid Mirmehdi
Sion Hannuna
Massimo Camplani

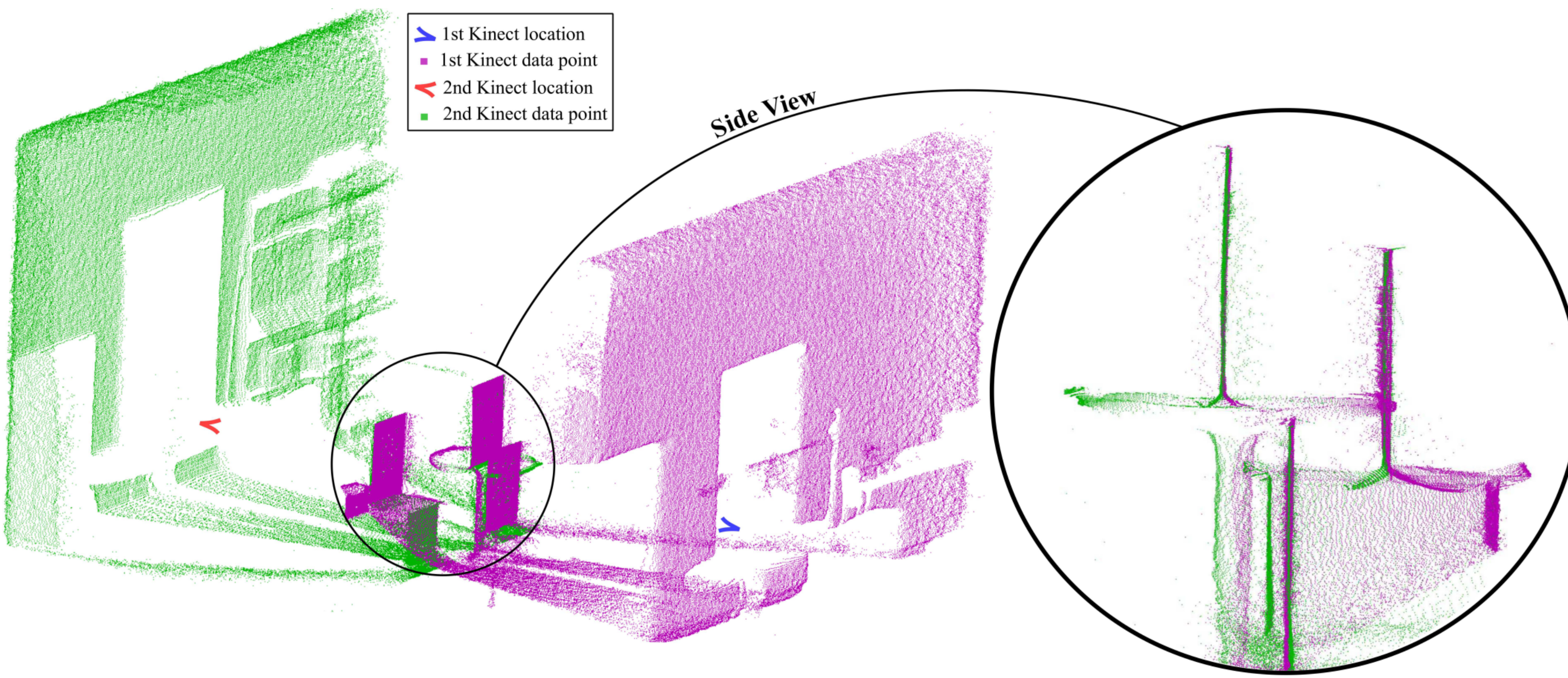
Double sided chessboards setup

Recording a subject performing breathing test

Action

Two Facing Kinects

with: Vahid Soleimani
Majid Mirmehdi
Sion Hannuna
Massimo Camplani



Point clouds are aligned and registered to a joint coordinate system.

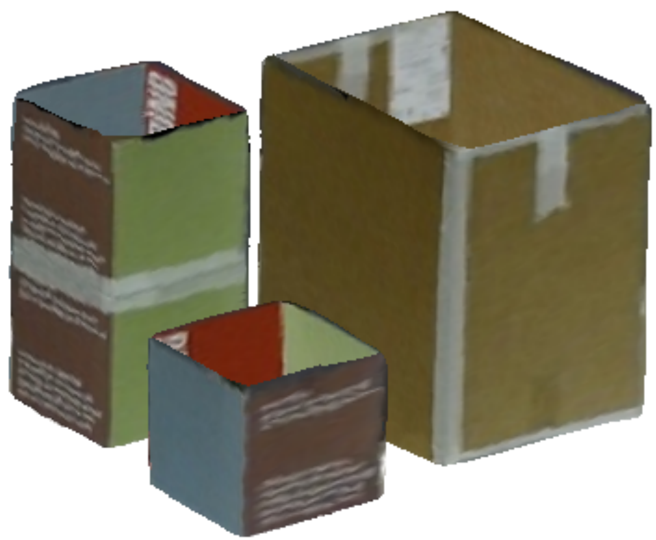
Soleimani et al (2016), 3D Data Acquisition and Registration using Two Opposing Kinects.

Action

Two Facing Kinects

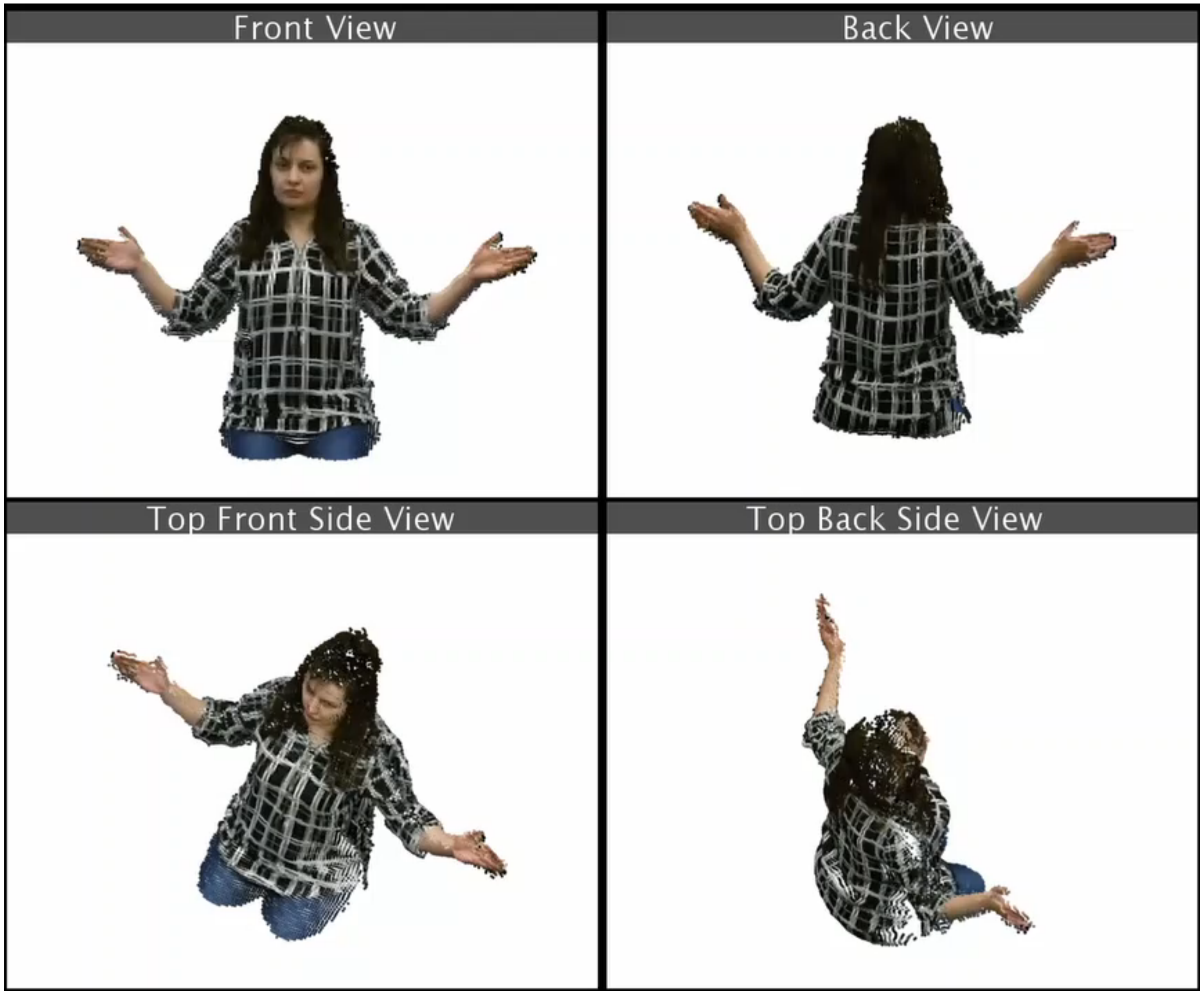
with: Vahid Soleimani
Majid Mirmehdi
Sion Hannuna
Massimo Camplani

- Quantitative assessment:
  - Using three differently sized boxes in three locations.
  - Performing *surface analysis* and automatically estimating *dimension*, *volume*, *surface planarity* and *angles*.



Two Facing Kinects

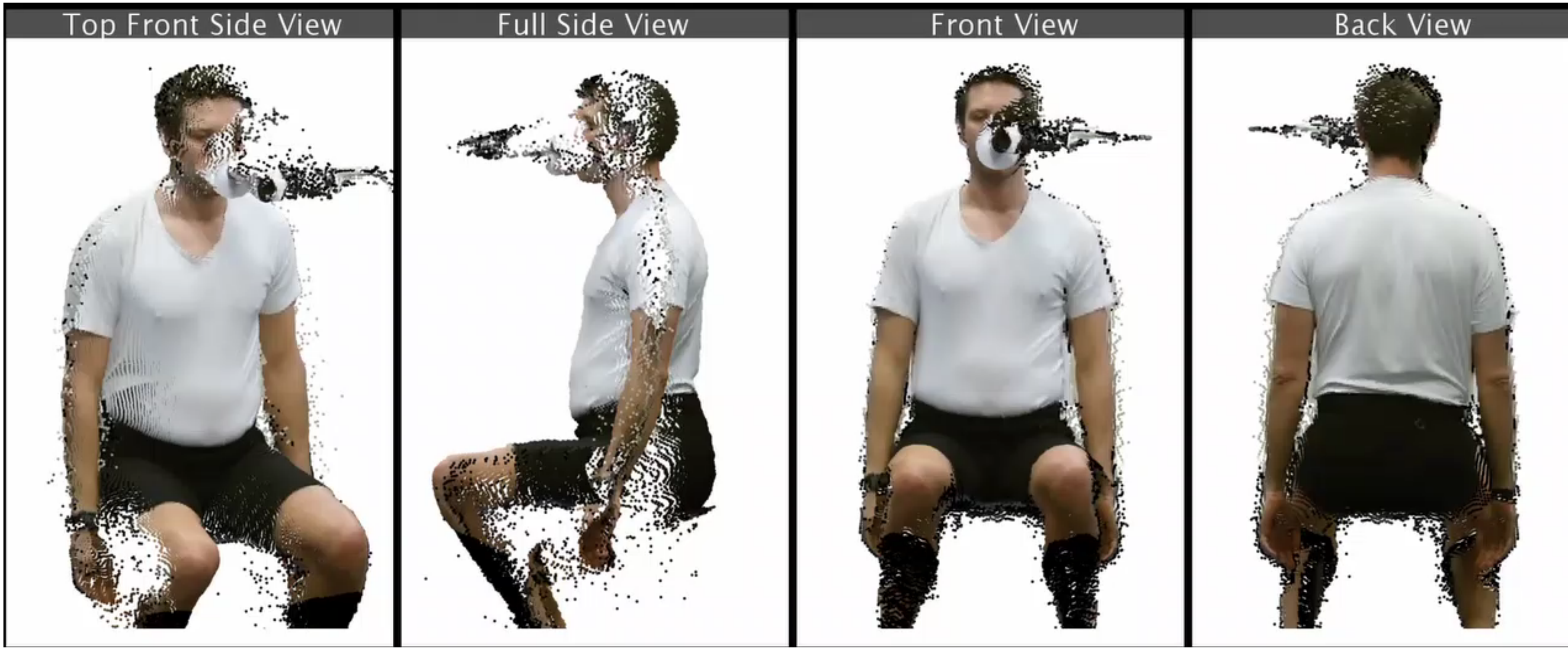
with: Vahid Soleimani
Majid Mirmehdi
Sion Hannuna
Massimo Camplani



Soleimani et al (2016), 3D Data Acquisition and Registration using Two Opposing Kinects.

Two Facing Kinects

with: Vahid Soleimani
Majid Mirmehdi
Sion Hannuna
Massimo Camplani

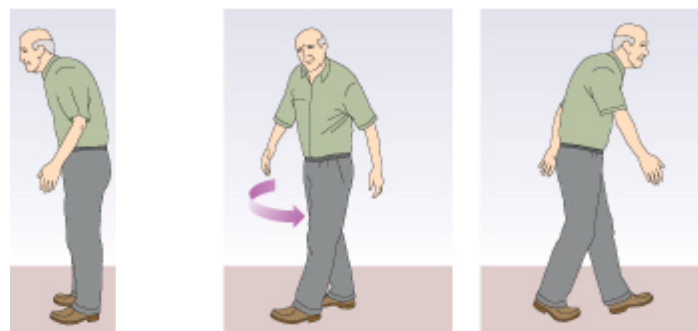


307 sequences of lung function assessment were recorded from 35 subjects using the proposed system.

Functional Mobility Testing

with: Hana Alghamdi
Majid Mirmehdi
+ SPHERE team
+ collaborators

Turn 180° Test : ask the patient to stand up, turn around until the patient facing the opposite direction and, walk towards a specified target.



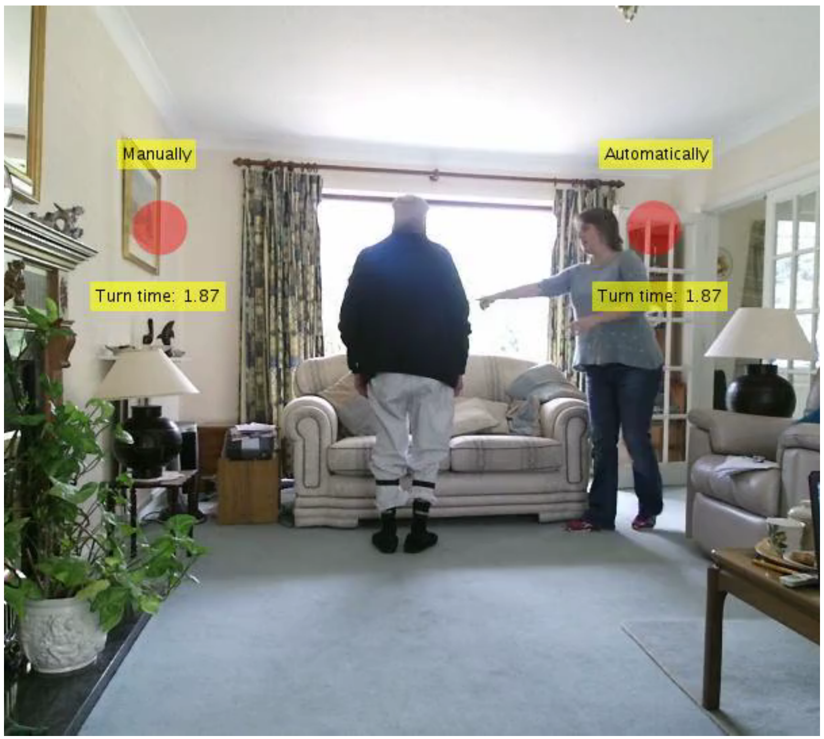
Different measures observed during turning:

- **Direction of turning**
- **Number of steps**
- **Turn Time (s)**
- **Turn Quality**
- **Turn Type**

with: Hana Alghamdi
Majid Mirmehdi
+ SPHERE team
+ collaborators

Functional Mobility Testing

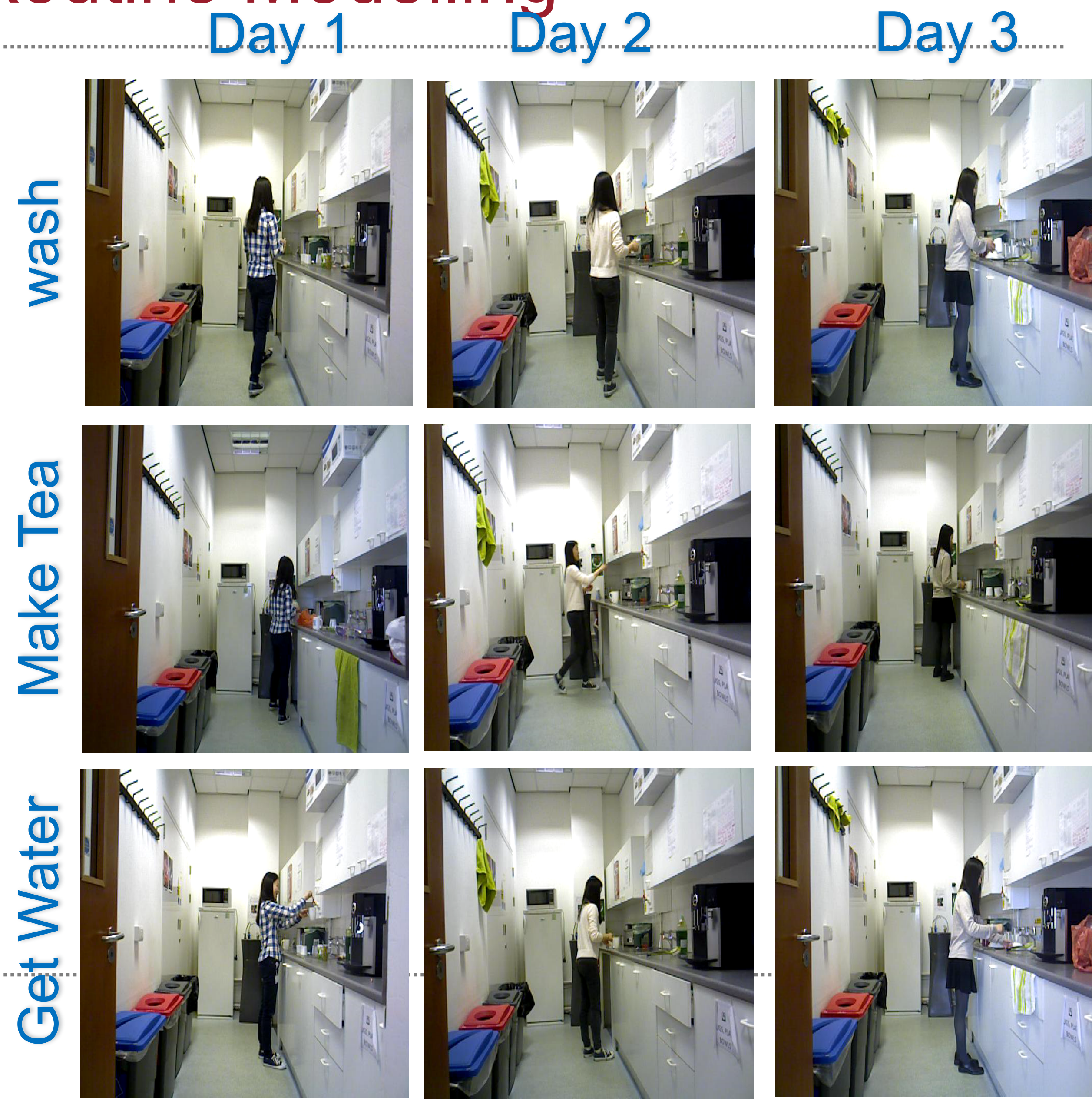
Best Case



Worst Case

with: Yangdi Xu
Dave Bull

Unsupervised Routine Modelling

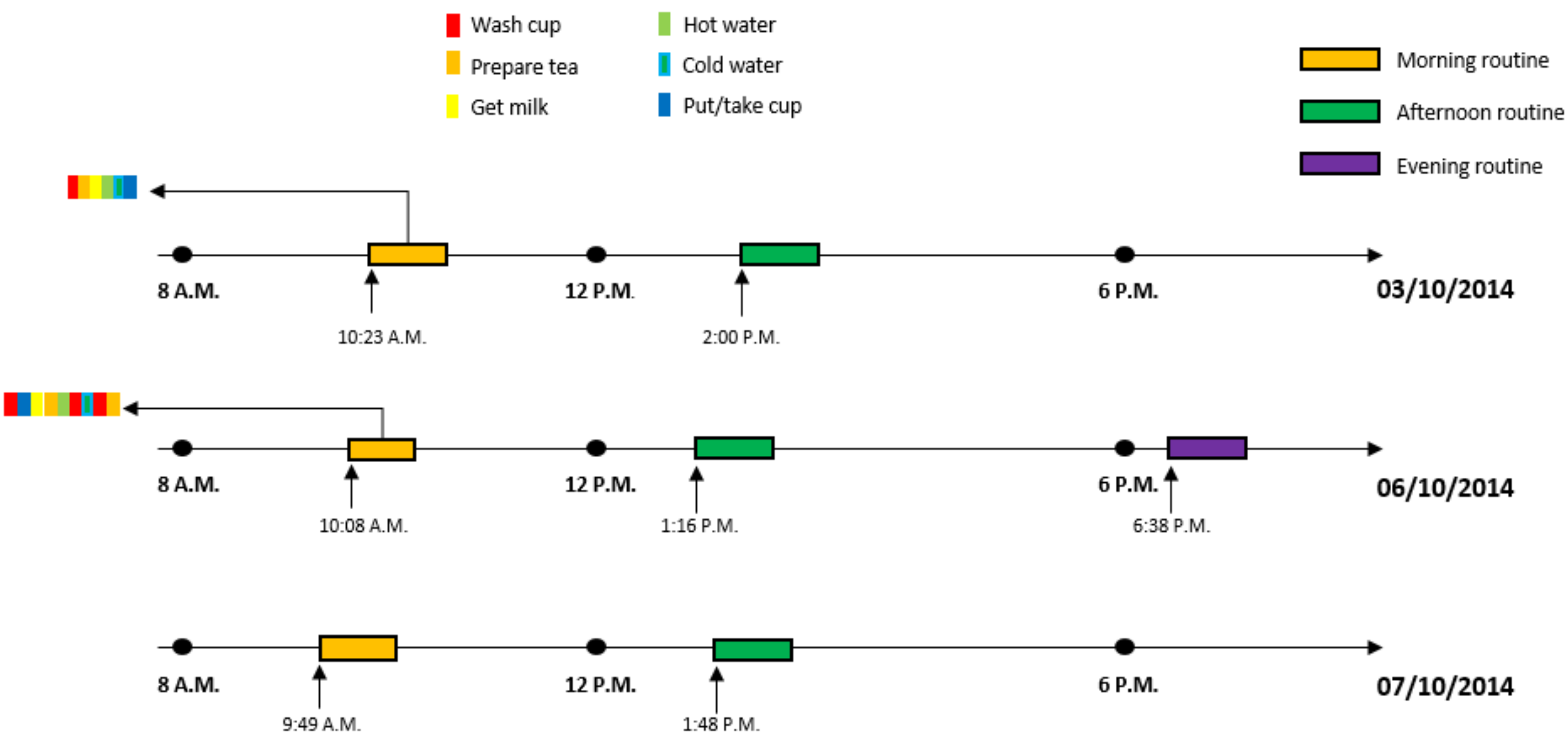


- A person's routine is the common or regular course of action, over a timescale (e.g. daily routine)

- Detecting routine changes or out-of-routine activities is essential for monitoring physical as well as mental wellbeing

with: Yangdi Xu
Dave Bull

Unsupervised Routine Modelling

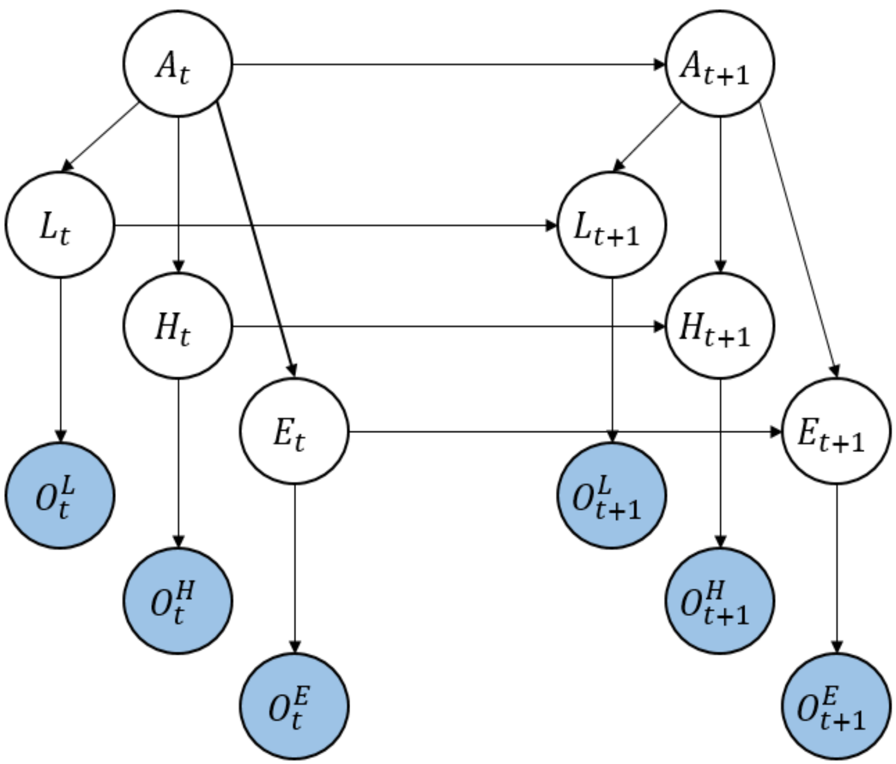


Xu et al (2015), Unsupervised Daily Routine Modeling from a Depth Sensor using Bottom-Up and Top-Down Hierarchies. *Asian Conference on Pattern Recognition*

with: Yangdi Xu
Dave Bull

Unsupervised Routine Modelling

Graphical Model

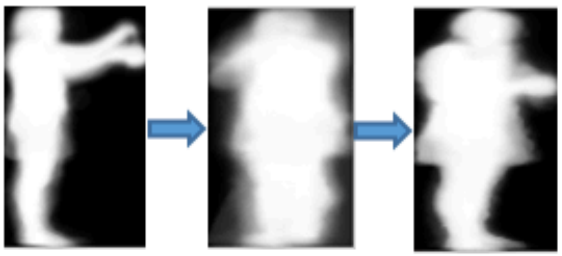
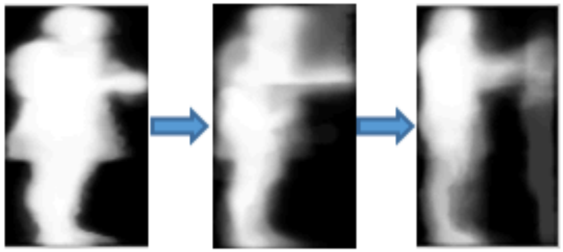


- A_t: Activity state
- L_t: Location state
- H_t: poses state
- E_t: Time envelope state

with: Yangdi Xu
Dave Bull

Unsupervised Routine Modelling

- Transitions in spatial and silhouette data are capable of discovering di from the data

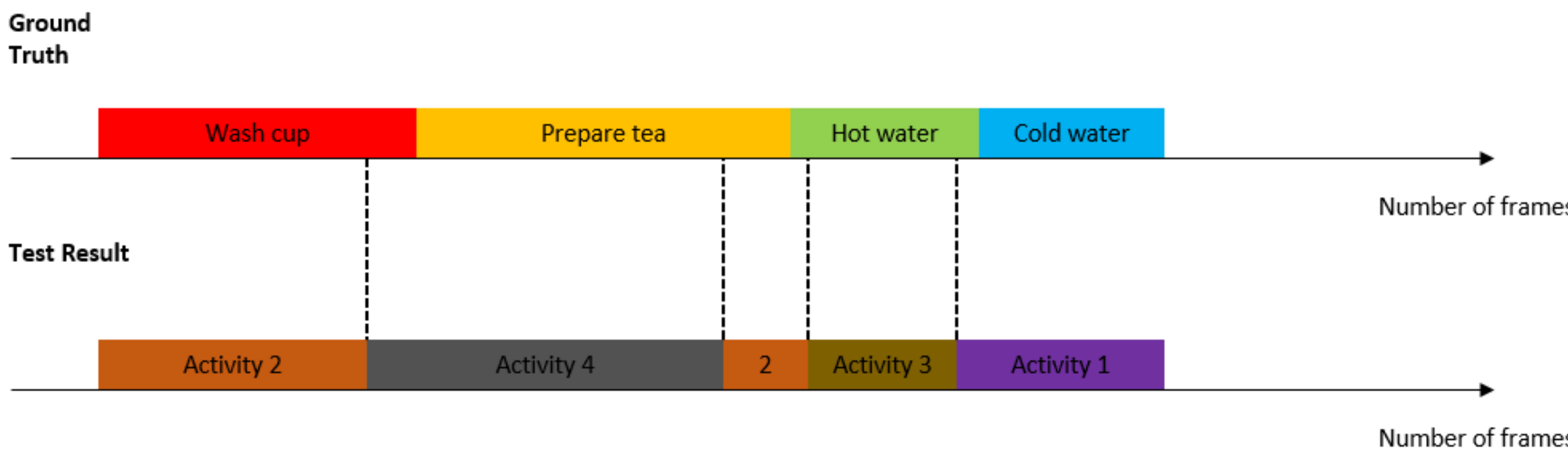
GT Label	Spatial	Silhouette
Get Water:	Boiler(r_2) -> water fountain (r_4) -> worktop (r_3)	
Make Tea:	No Frequent Transition	
Add Milk:	worktop (r_3) -> fridge (r_4) -> worktop (r_3)	No Frequent Transition

University of BRISTOL

Xu et al (2015), Unsupervised Daily Routine Modeling from a Depth Sensor using Bottom-Up and Top-Down Hierarchies. *Asian Conference on Pattern Recognition*

with: Yangdi Xu
Dave Bull

Unsupervised Routine Modelling



$$M(x,y) = \frac{\sum_{P_{gt}^{i}=x} \sum_{P_{gt}^{j}=y} C(P_{es}^{ij}, P_{gt}^{i}) + C(P_{es}^{ij}, P_{gt}^{j})}{|\{P_{gt}^{i}=x\}| \quad |\{P_{gt}^{j}=y\}|}$$

with: Yangdi Xu
Dave Bull

Unsupervised Routine Modelling

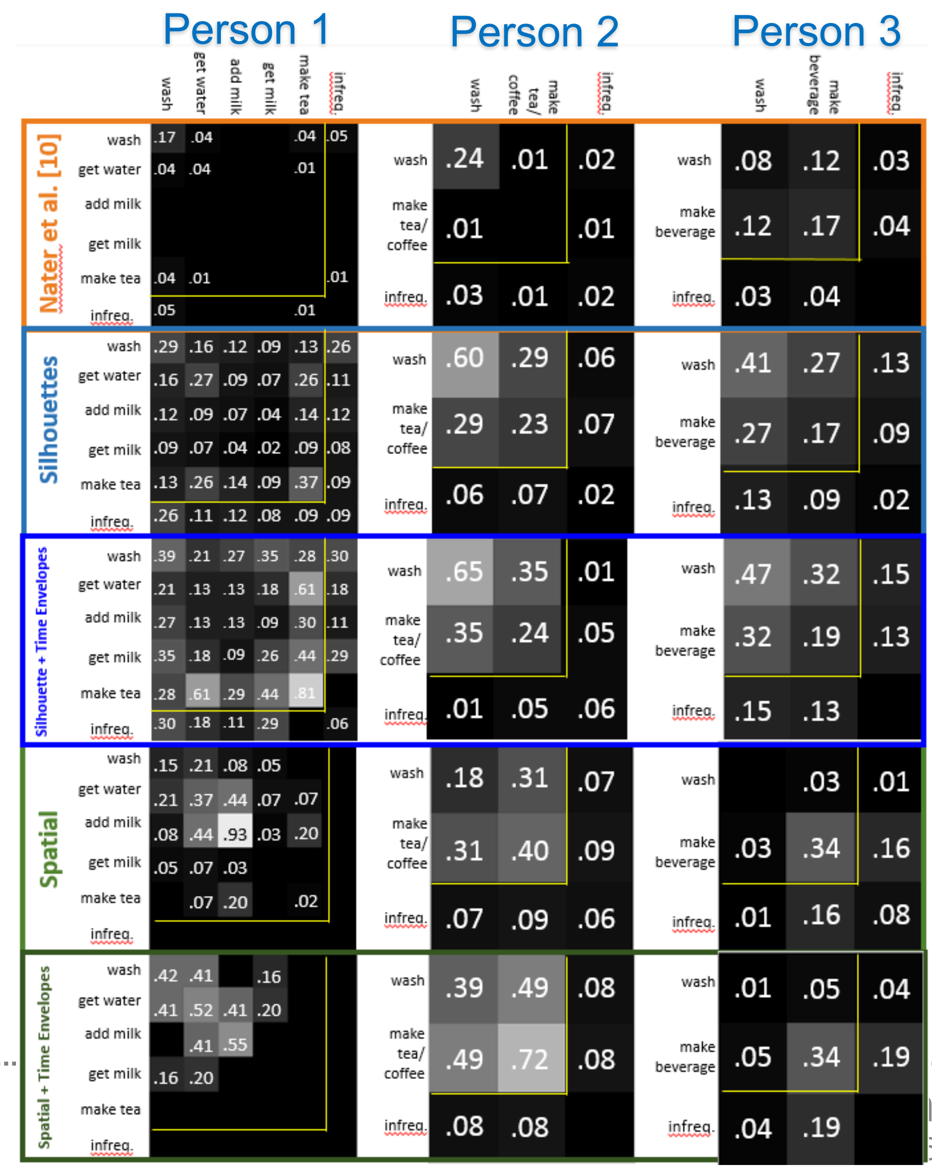
	wash	Prepare tea	Get milk	Get hot water	Get cold water	Put cup	Make porridge
wash	0.93	0.06	0.00	0.30	0.05	0.34	0.03
Prepare tea	0.06	0.70	0.09	0.13	0.16	0.15	0.77
Get milk	0.00	0.09	0.86	0.03	0.33	0.03	0.06
Get hot water	0.30	0.13	0.03	0.87	0.27	0.70	0.20
Get cold water	0.05	0.16	0.33	0.27	0.61	0.23	0.51
Put cup	0.34	0.15	0.03	0.70	0.23	0.50	0.24
Make porridge	0.03	0.77	0.06	0.20	0.51	0.24	0.00

Xu et al (2015), Unsupervised Daily Routine Modeling from a Depth Sensor using Bottom-Up and Top-Down Hierarchies. *Asian Conference on Pattern Recognition*

with: Yangdi Xu
Dave Bull

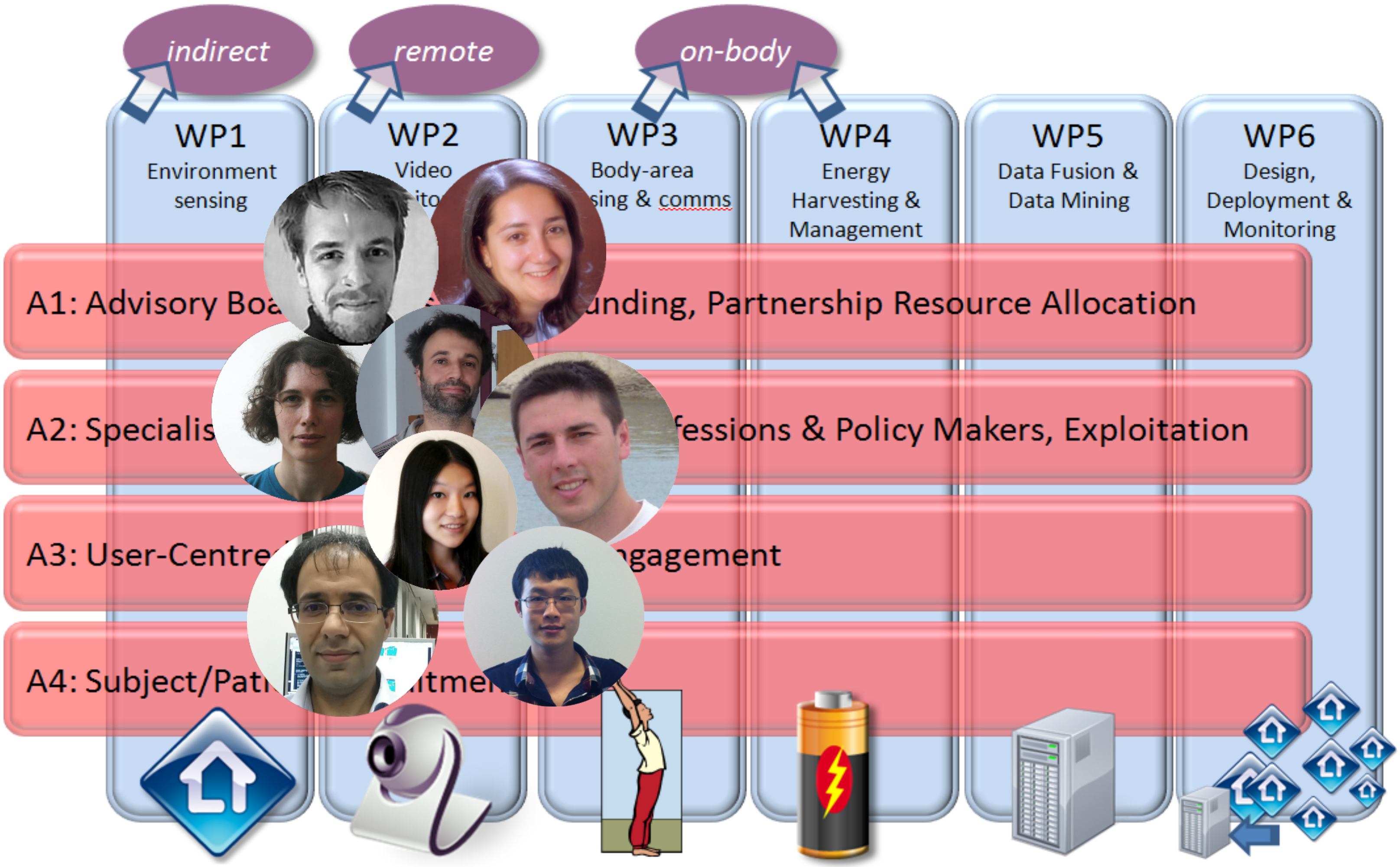
Unsupervised Routine Modelling

- Dataset of 3 people for 7 days
- Results show that using time envelope is helpful in discovering routine activities
  - More patterns are discovered
  - Better temporal overlap between discovered pattern and ground truth



Xu et al (2015), Unsupervised Daily Routine Modeling from a Depth Sensor using Bottom-Up and Top-Down Hierarchies. *Asian Conference on Pattern Recognition*

SPHERE


indirect
remote
on-body
WP1
Environment sensing
WP2
Video
WP3
Body-area
sing & comms
WP4
Energy Harvesting & Management
WP5
Data Fusion & Data Mining
WP6
Design, Deployment & Monitoring
A1: Advisory Boa
unding, Partnership Resource Allocation
A2: Specialis
fessions & Policy Makers, Exploitation
A3: User-Centre
gagement
A4: Subject/Pat

Hardware Platform (v2.0)

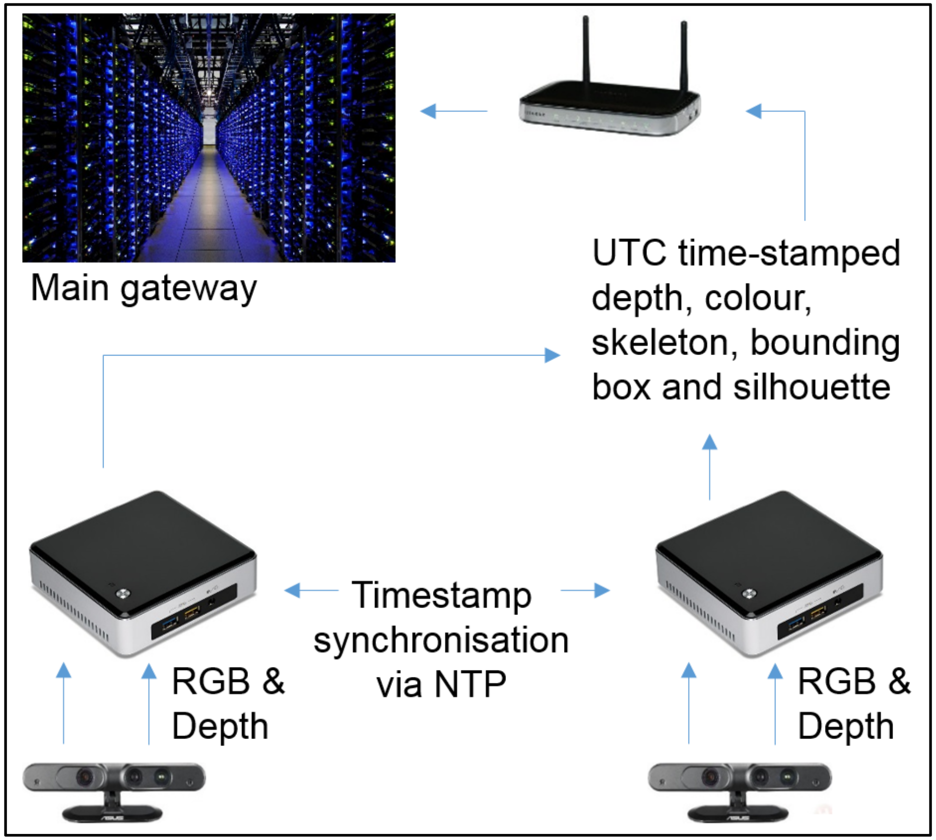
- RGB-D Asus Xtion
  - SOTA people detection and tracking with low computational burden
- The Intel Next Unit of Computing (NUC) with 8GB of RAM and an i5 processor
  - Small, attractive, powerful and able to support up to 4 Xtions at full resolution

with: Majid Mirmehdi
+ SPHERE team

SPHERE

Hardware Platform (v2.0)

- RGB-D Asus Xtion
  - SOTA people detection and tracking with low computational burden
- The Intel Next Unit of Computing (NUC) with 8GB of RAM and an i5 processor
  - Small, attractive, powerful and able to support up to 4 Xtions at full resolution



with: Majid Mirmehdi
+ SPHERE team

SPHERE

Looking forward: Recruitment of 100 homes?

Conclusion

- Current RGBD sensors are not ideal for action and activity recognition due to their per-frame calculation of depth information
- Three main usages of RGBD data in action and activity recognition
- Applications for action recognition where accurate depth estimation is required
- Storage requires for long-term usage is an obstacle for expanded usage of RGBD in action and activity recognition

Thank you…

Dima Damen

http://www.cs.bris.ac.uk/~damen

@dimadamen

http://www.linkedin.com/in/dimadamen

VI-Lab, University of Bristol

http://vilab.blogs.ilrt.org

SPHERE - a Sensor Platform for HEalthcare in a Residential Environment

http://www.irc-sphere.ac.uk/

@IRC_SPHERE

https://www.facebook.com/pages/Sphere