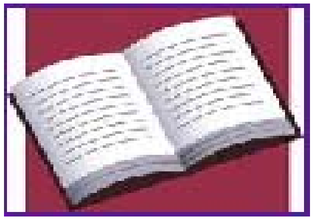
Mobile Computing and Context

Table of Contents



- Motivation
- Main ideas and results in analyzed papers
- Conclusions

Motivation

- activities recognition by automated systems lead to improvements in our life

- approaches build on intelligent infrastructures or use of computer vision

- current monitoring solutions are not feasible for a long-term implementation

Activity recognition using on-body sensing

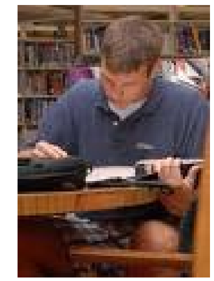
Common Ideas
Paper 1 and 2

1. Segmentation

Interesting	NULL	?

2. Classification

Classifier A: "He sleep" - 80% / "He learn" - 20%



Classifier B: "He sleep" - 75% / "He learn" - 25%

3. Fusion

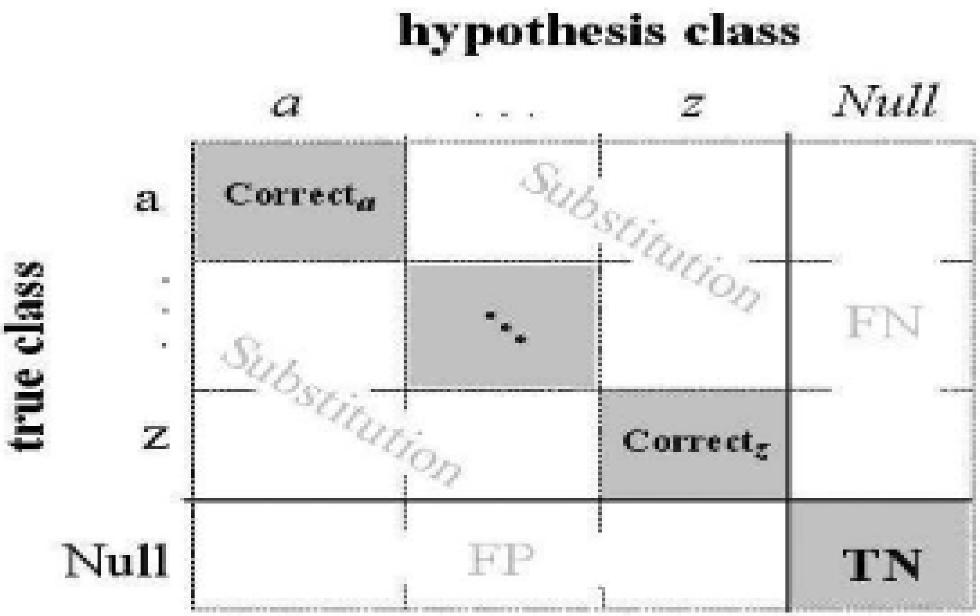
Classifier A + Classifier B:

? = He sleep

- on-body sensors are deployed strategically
- the selection of features and event detection thresholds play a key role
- prior training from data is required
- to analyze the recognition performance, Precision and Recall metrics were used

- the goal of each recognition approach is to find with higher accuracy true positive events

- high impact of false positive and false negative events



Multiclass Confusion Matrix

- Classification of NULL is a tough problem for any classifier

- Different fusion methods are used for accurate classification:

a) comparison of Top Choices (COMP)

b) methods based on class rankings

- Highest rank (HR)
- Borda Count
- Logistic Regression (LR)

c) agreement of the detectors (AGREE)

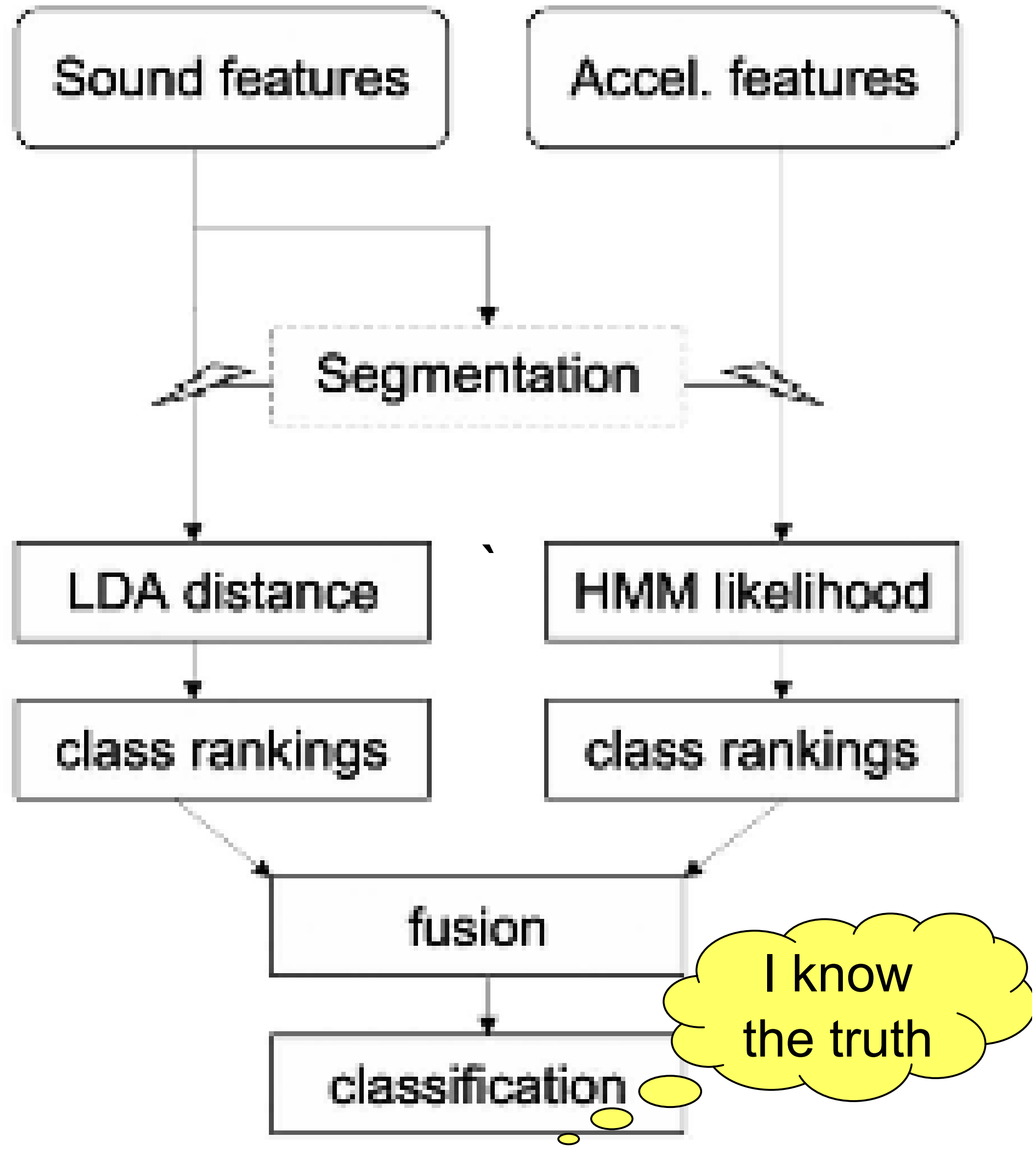
Activity Recognition of Assembly Tasks
Paper 1

- recognize the use of different tools involved in an assembly task in a wood workshop
- recognize of activities that are characterized by a hand motion and an accompanying sound
- microphones and accelerometers as on-body sensors

Broken up into segments

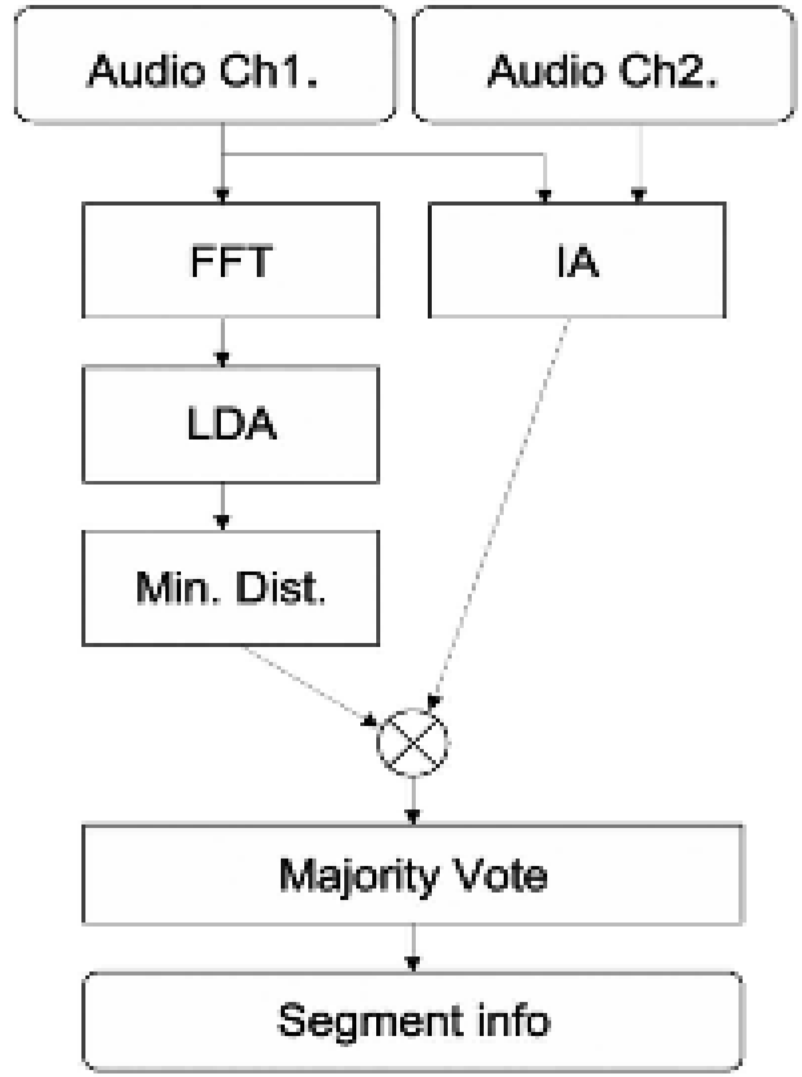
LDA distance and HMM likelihood, carried out over these segments

Covert into class ranking; combine using fusion methods

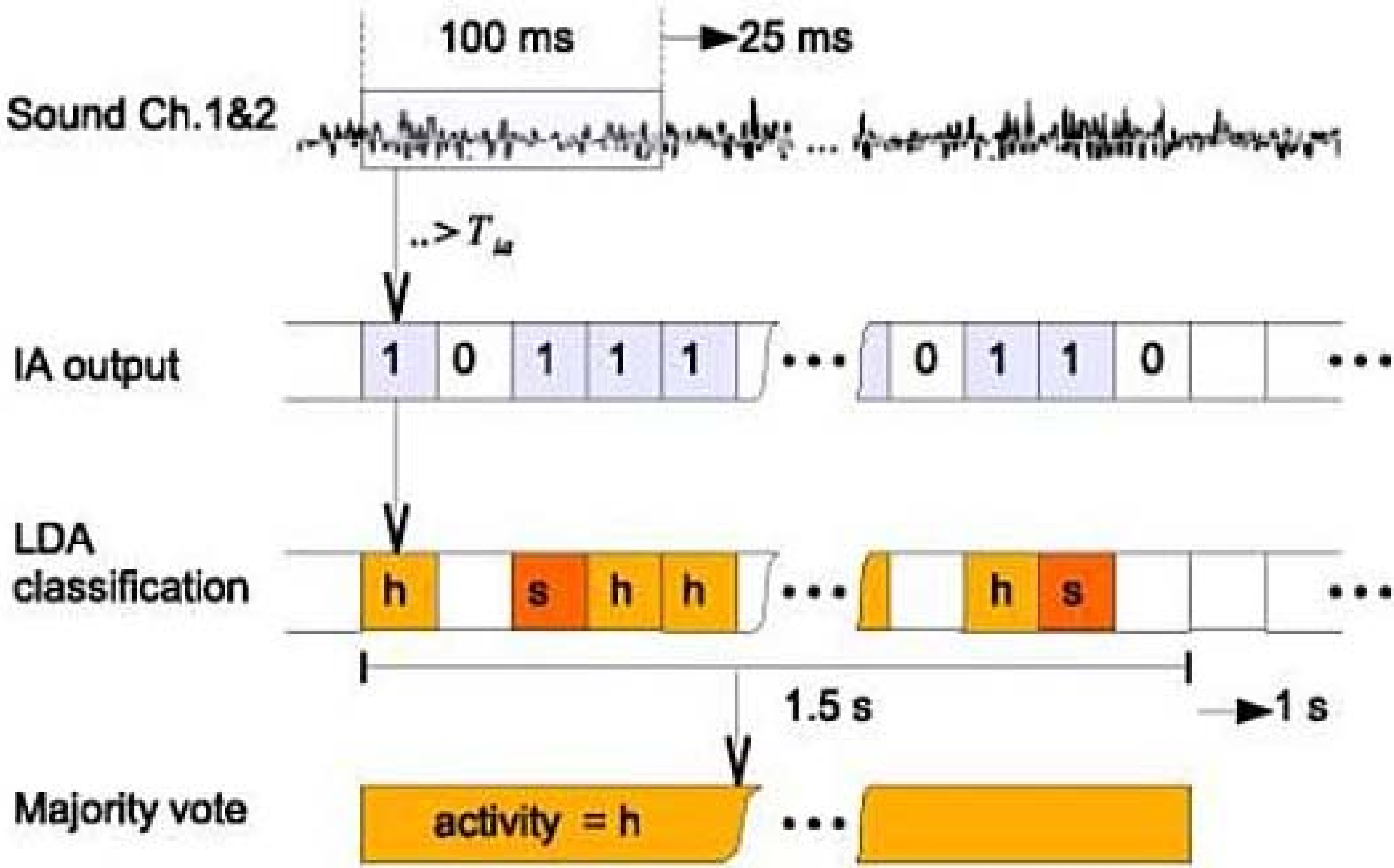


Overall recognition process

- Sound analysis used to identify relevant segments
- Using only IA produce fragmented results
- A different method of “smoothing” using majority vote was applied
- A relatively large window (1.5 s) was chosen to reflect the typical timescale of interest activities



Sound based segmentation



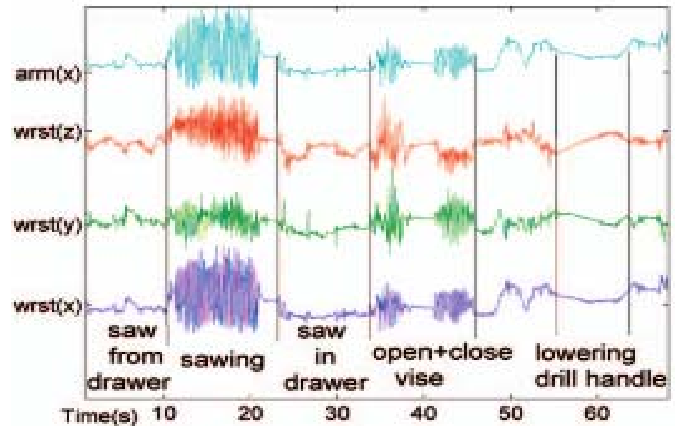
Jamie Ward, Diss. ETH 16520

sound classification

need when higher information about a segment is required

use the LDA distances; provides a list of class distance for each segment

acceleration classification

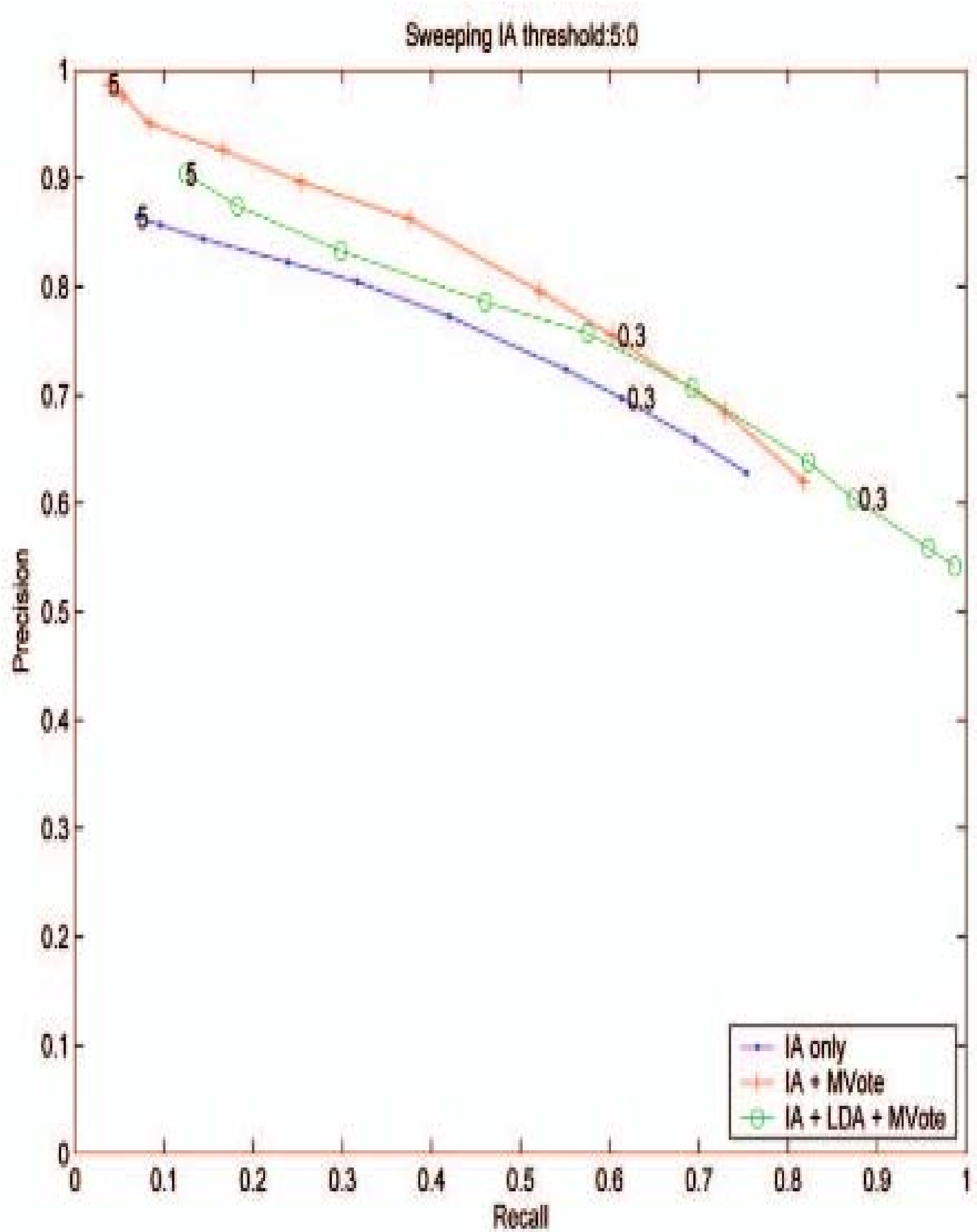


combination of features used to feed the HMM models

provides a list of HMM likelihoods for each segment

Fusion

Segmentation Results



$$Recall = \frac{true\ positive\ time}{total\ positive\ time} = \frac{TP}{TP+FN};$$

$$Precision = \frac{true\ positive\ time}{hypothesized\ positive\ time} = \frac{TP}{TP+FP};$$

	Sound		Accel.		COMP		LR	
Class (s)	%R	%P	%R	%P	%R	%P	%R	%P
hammer (196)	92	74	93	79	92	94	92	93
saw (306)	90	87	90	80	88	95	93	90
file (305)	77	80	80	82	65	94	82	90
drill (242)	95	54	99	41	95	64	96	59
sand (313)	82	67	87	92	77	93	83	94
grind (278)	83	69	63	66	62	80	75	73
screwd.(260)	52	20	53	87	51	86	53	81
vise (678)	65	55	74	49	61	69	73	53
drawer (659)	86	47	88	39	69	51	87	39
Pos.Average%	76	62	76	68	73	79	78	74
NULL(2778)	33	69	33	69	69	67	42	69

Continuous R and P for each Positive Class and the Average of These; User-Dependent Case

Continuous Time Results:

$$Recall = \frac{correct\ positive\ time}{total\ positive\ time} = \frac{correct}{TP+FN};$$

$$Precision = \frac{correct\ positive\ time}{hypothesized\ positive\ time} = \frac{correct}{TP+FP};$$

Three methods of evaluation:

- user-dependent
- user-independent (most severe)
- user-adapted

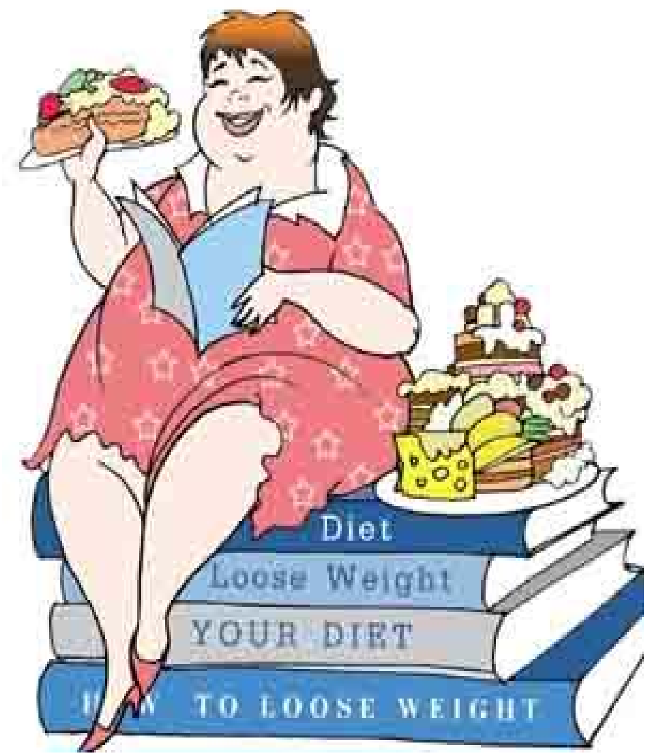
Lessons Learned

- using intensity differences works relatively well for detection of activities; however, short fragmented segments (apply smoothing)
- activities are better recognized using a fusion of classifiers
- less performance in user independent case; fused classifiers solve this problem.

- over one billion of overweight and 400 mil obese patients worldwide

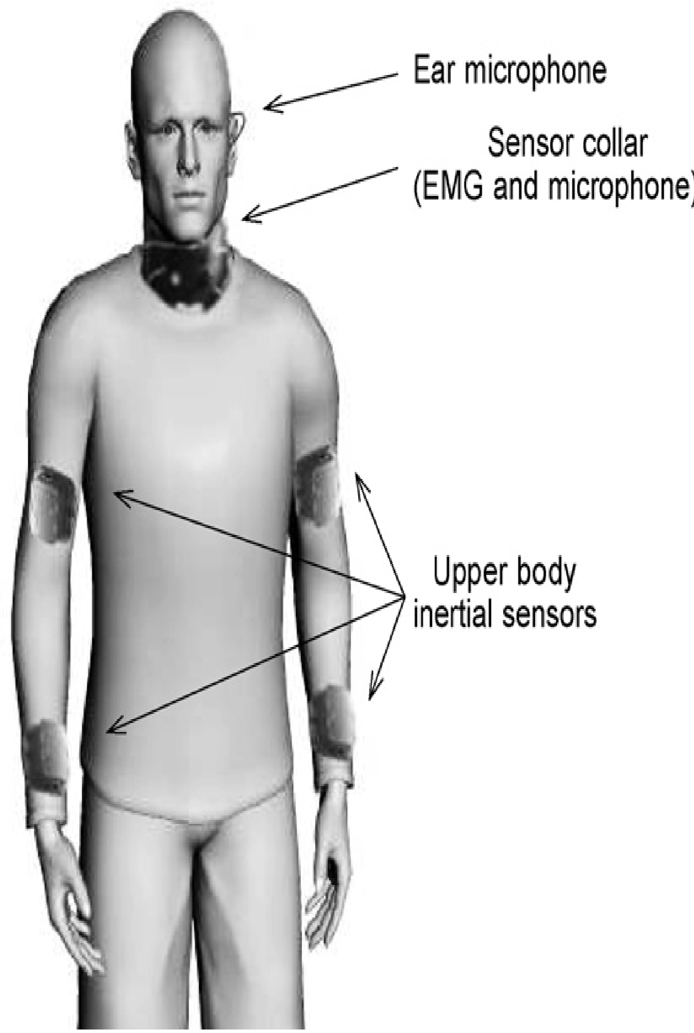
- several key risk factors have been identified, controlled by dieting behavior

- minimizing individual risk factors is a preventive approach to fight the origin of diet-related diseases



Three aspects of dietary activity

- characteristic arm and trunk movements associated with the intake of foods
- chewing of foods, recording the food breakdown sound
- swallowing activity

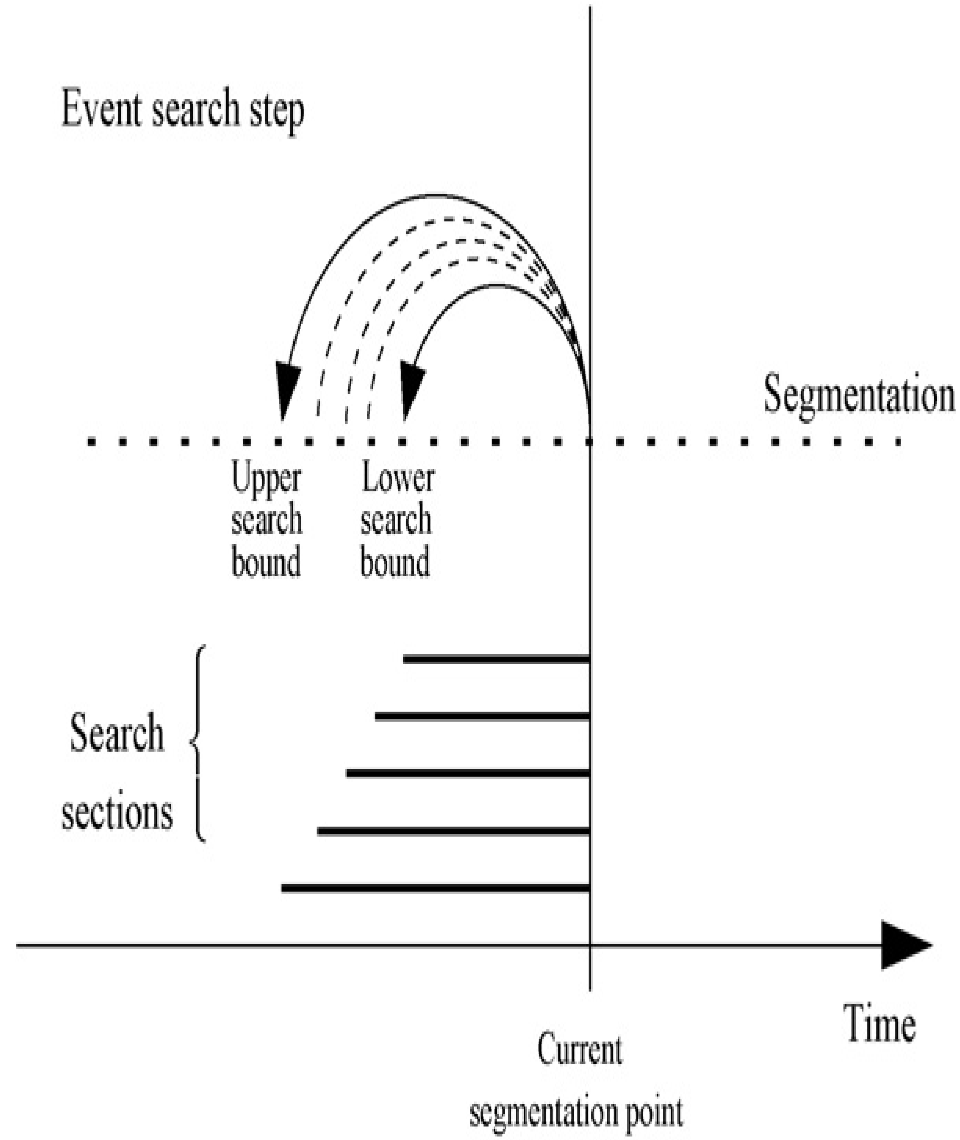


Sensor positioning at the body

- **Segmentation**
  using a fixed distance; manually annotation of events

- **Classification**
  similarity-based algorithm

- **Fusion**
  COMP, AGREE, LR
  use of confidence



Performance measurement

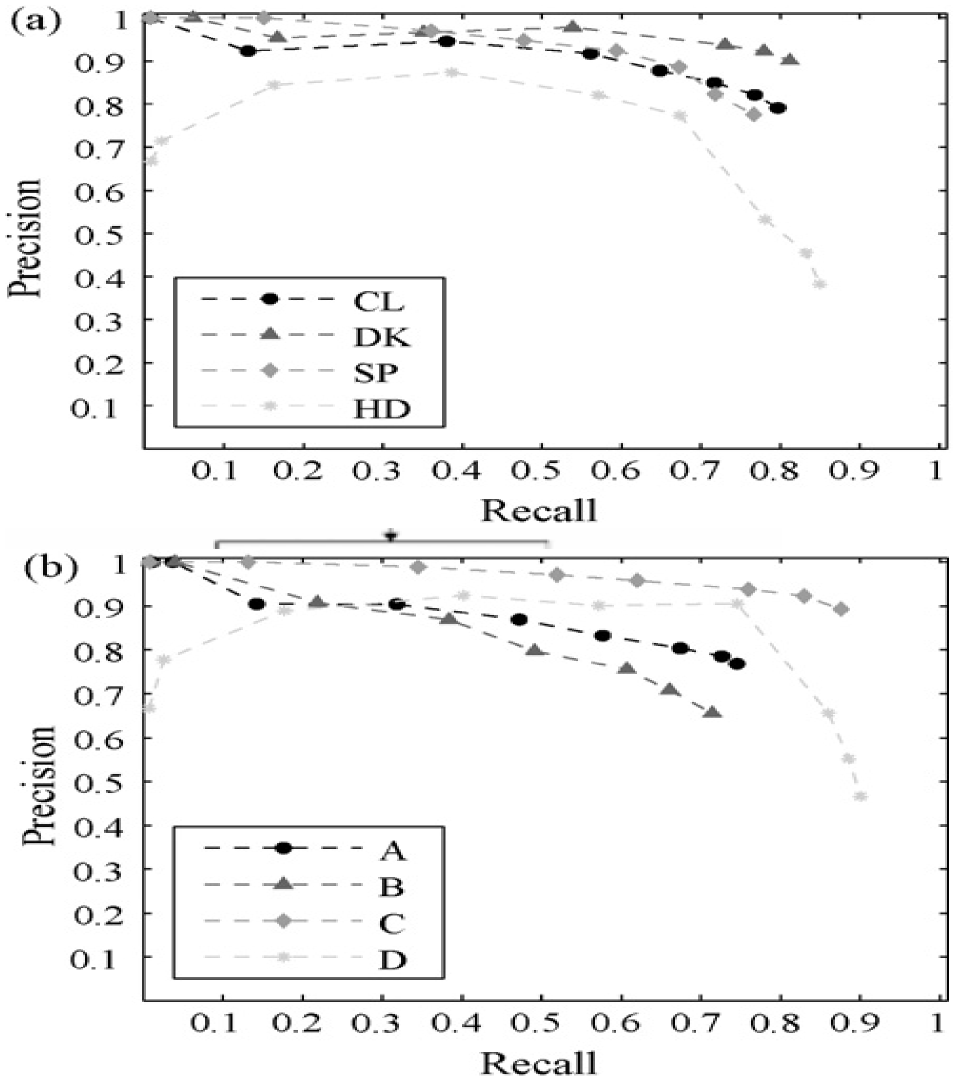
$$Recall = \frac{Recognised\ events}{Relevant\ events};$$

$$Precision = \frac{Recognised\ events}{Retrieved\ events};$$

R = 1 => perfect accuracy

P = 1 => 0 insertion errors

Movement Recognition



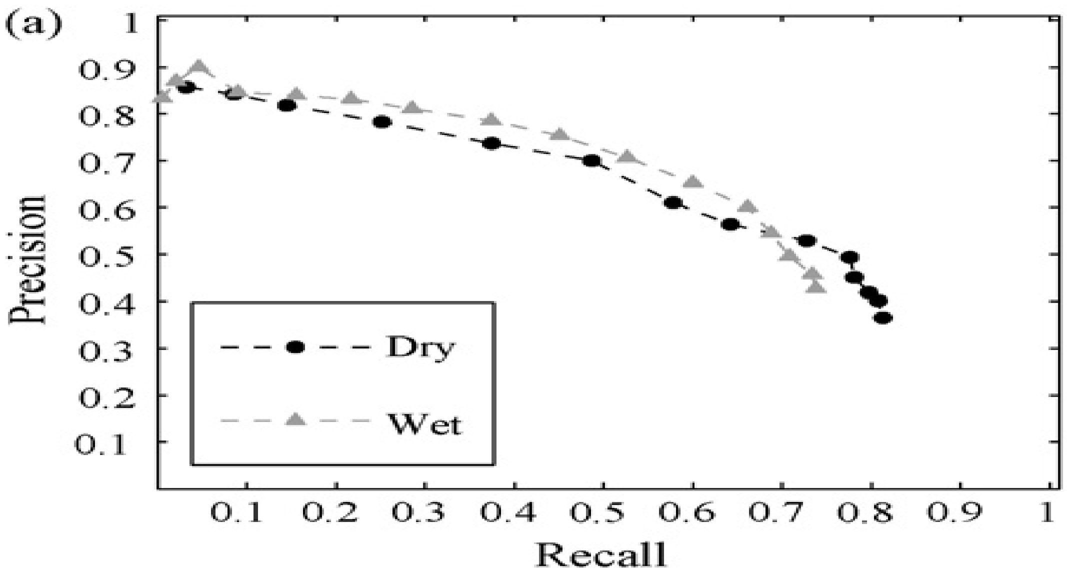
CL

DK

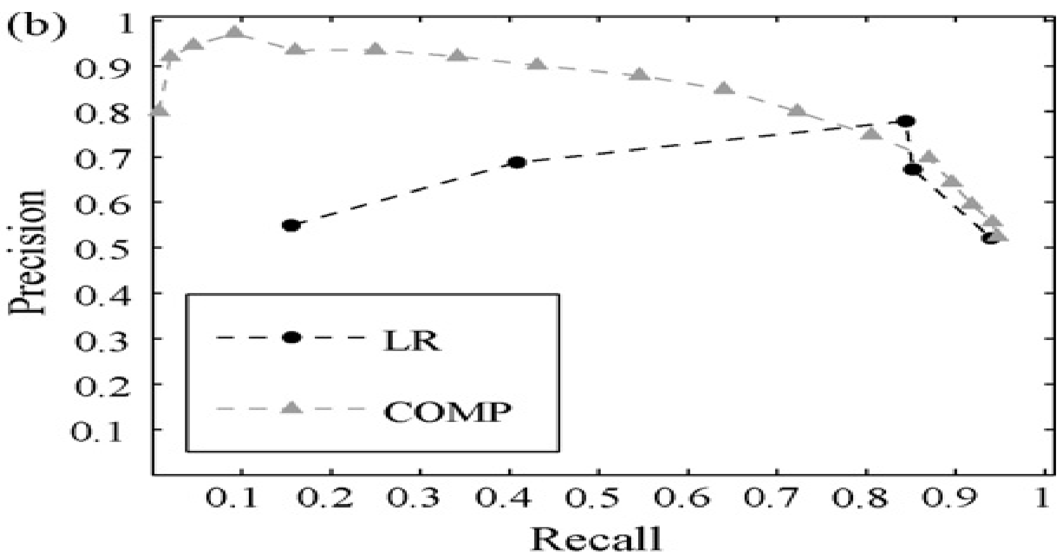
SP

HD

Chewing Recognition

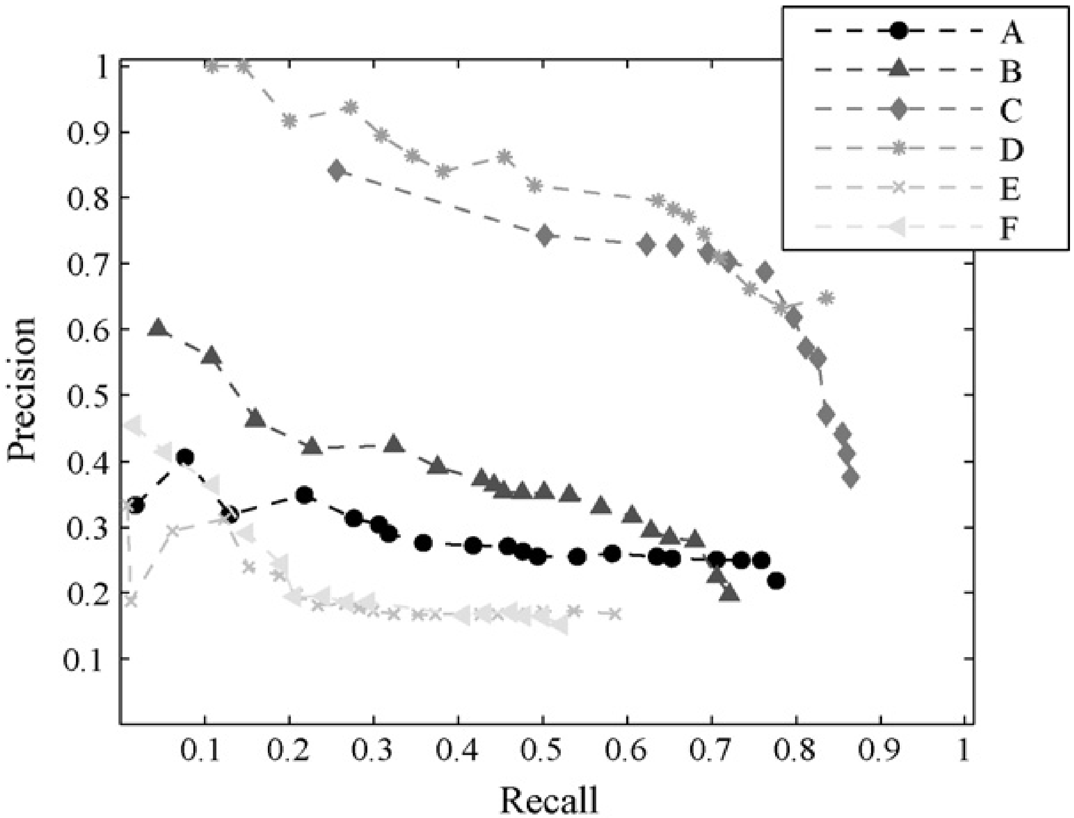


Dry



Wet

Swallowing recognition



Lesson learned

- food intake movements recognized with good accuracy
- chewing cycles were identified well; Still low detection performance with low amplitude chewing sounds
- it provides an indication for swallowing; Still incurs many insertion errors

Conclusion of Paper 1 and 2

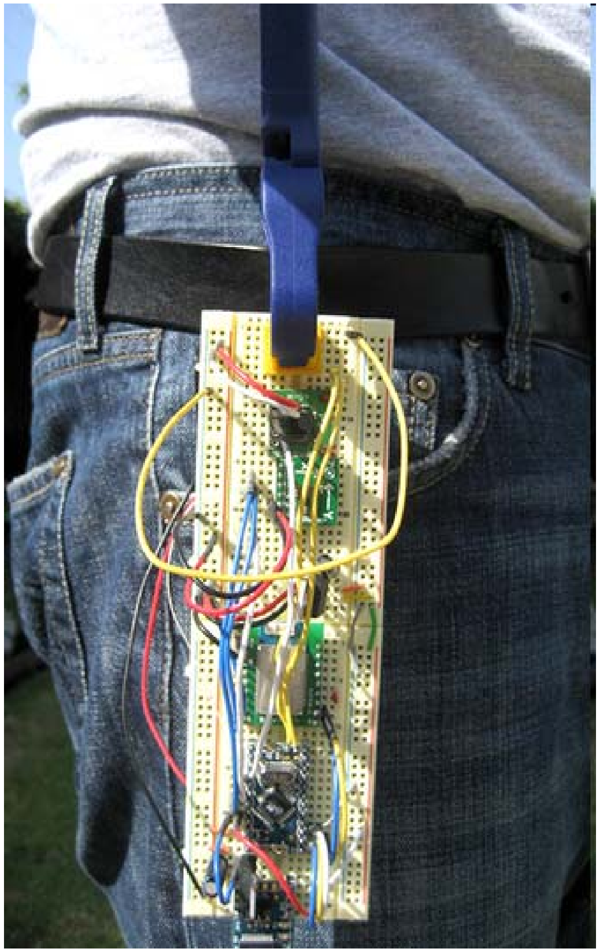
Pluses

- recognize different activities with good accuracy
- concepts used in "real-life" applications
- long term functionality

Conclusion of Paper 1 and 2

Minuses

- a lot of training
- sensitive to features & event threshold selection
- assumptions on NULL class
- uncomfortable systems for long-term use



However,

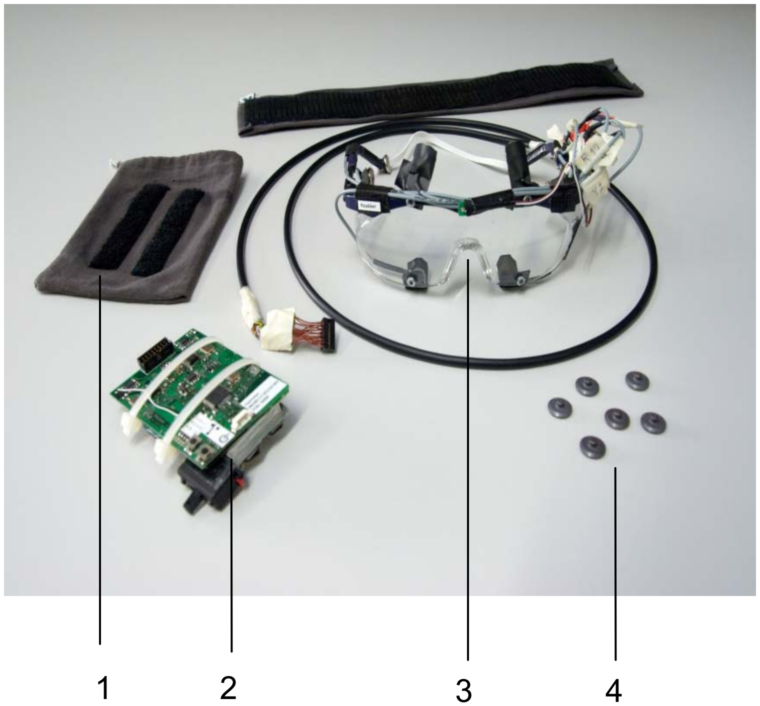
aspects like user attention and intentionality

cannot

be picked-up by usually sensors deployed

Recognition using EOG Goggles
Paper 3

- Identify eye gestures using EOG signals;
- Electrooculography (EOG) instead video cameras;
- Steady electric potential field from eyes;
- Alternate saccadic eye movement and fixations;
- Physical activities leads to artefacts;

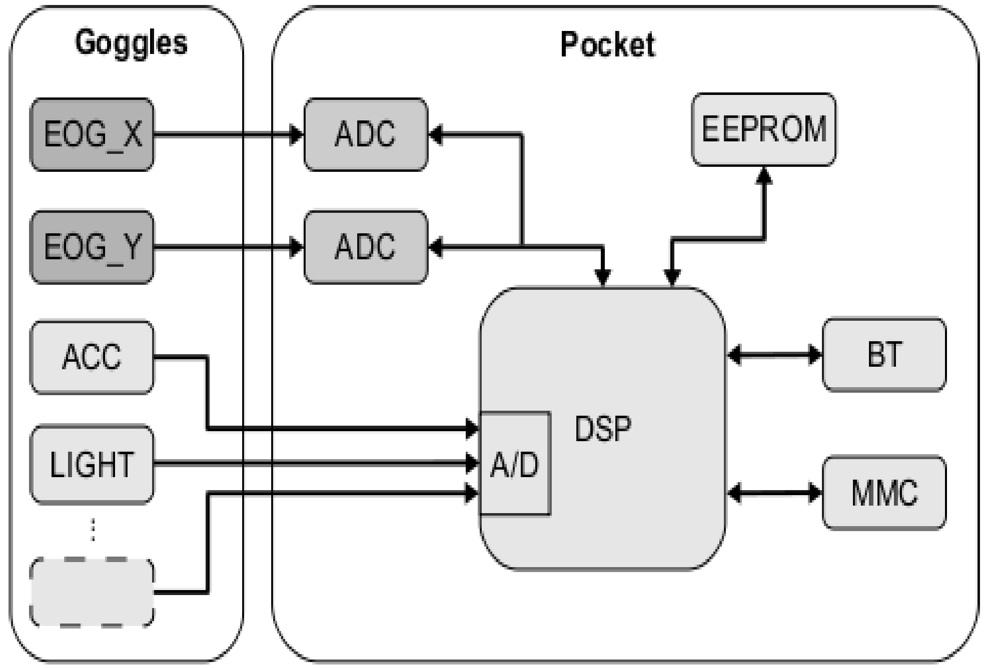


(1) armlet with cloth bag

(2) the Pocket

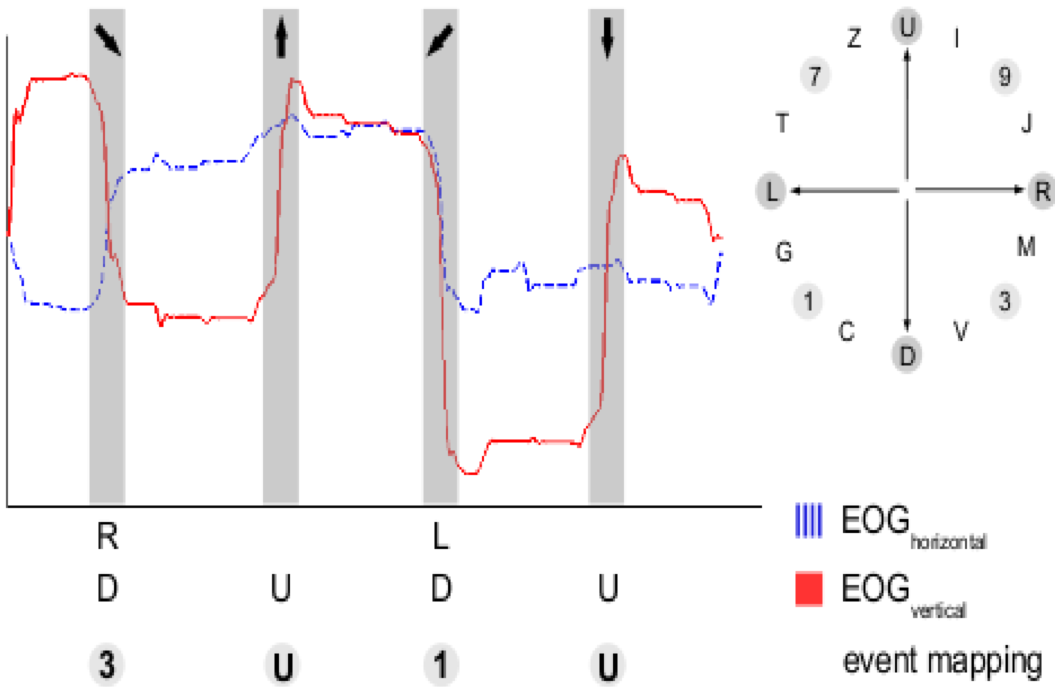
(3) the Goggles

(4) dry electrodes



Hardware architecture of the eye tracker

EOG gesture recognition



blink & saccade detection

⇩

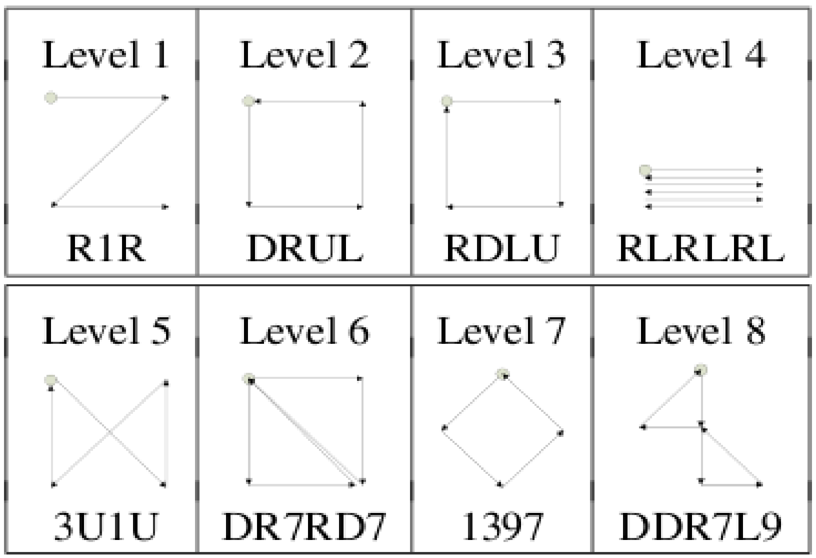
blink removal

⇩

stream of saccades events

median filter used to compensate artefacts

Eye gestures for stationary HCI



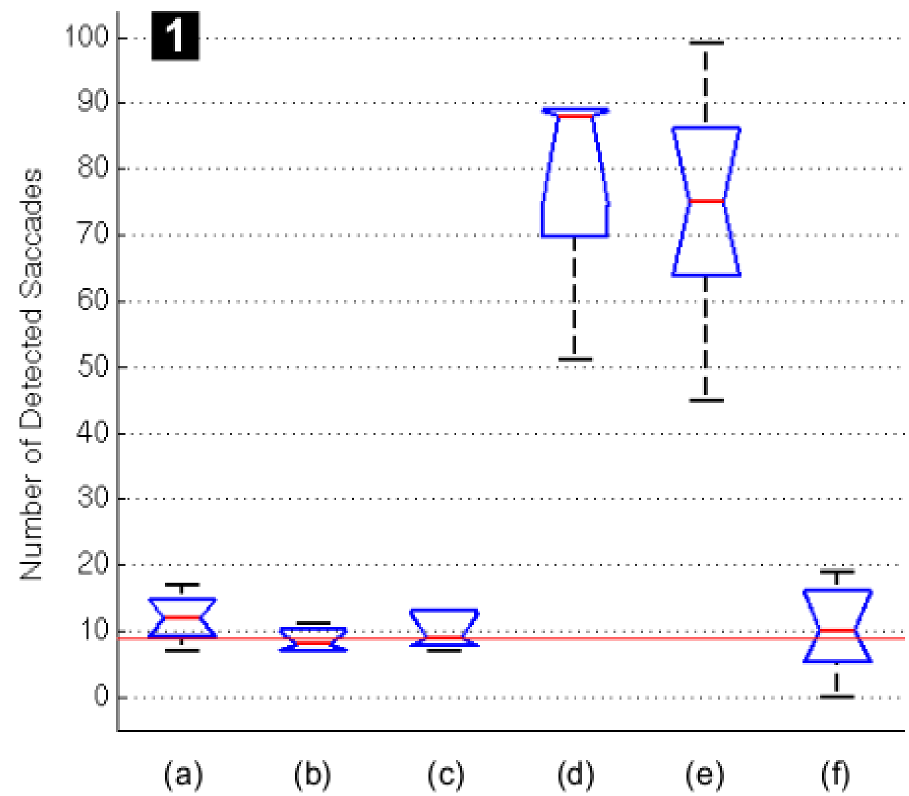
Eye gestures of increasing complexity

Gesture	$T_T[ms]$	$T_S[ms]$	T_S/T_T	$Acc[\%]$
R1R	3370	2890	0.858	85
DRUL	4130	3490	0.845	90
RDLU	3740	3600	0.963	93
RLRLRL	6680	5390	0.807	90
3U1U	4300	3880	0.902	89
DR7RD7	12960	5650	0.436	83
1379	6360	3720	0.585	84
DDR7L9	25400	5820	0.229	83

T_T: total time spent to complete the gesture
T_S: success time spent only on successful attempts
Acc: accuracy

Eye gestures for mobile HCI

- perform different eye movement on a head-up display
- investigate how artefacts can be detected and compensated
- an adapted filter performs well than a filter using a fixed window



(a) – (f) type of filter/medium used

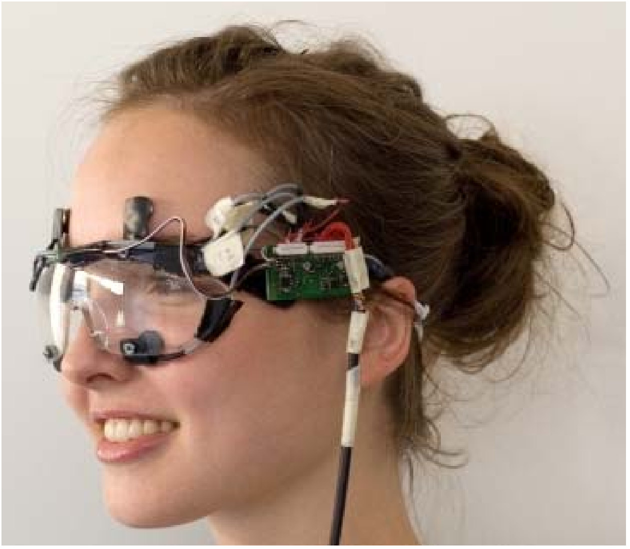
Lesson learned

- eye gesture recognition possible with EOG
- good accuracy of results in static scenarios
- artefacts may dominate the signal
- more complex algorithms for mobile scenarios

Conclusion of Paper 3

Pluses

- treat aspects which encompasses mere than physical activity
- much less computation power



Minuses

- uncomfortable for long-term use
- difficult for testing

Questions ?