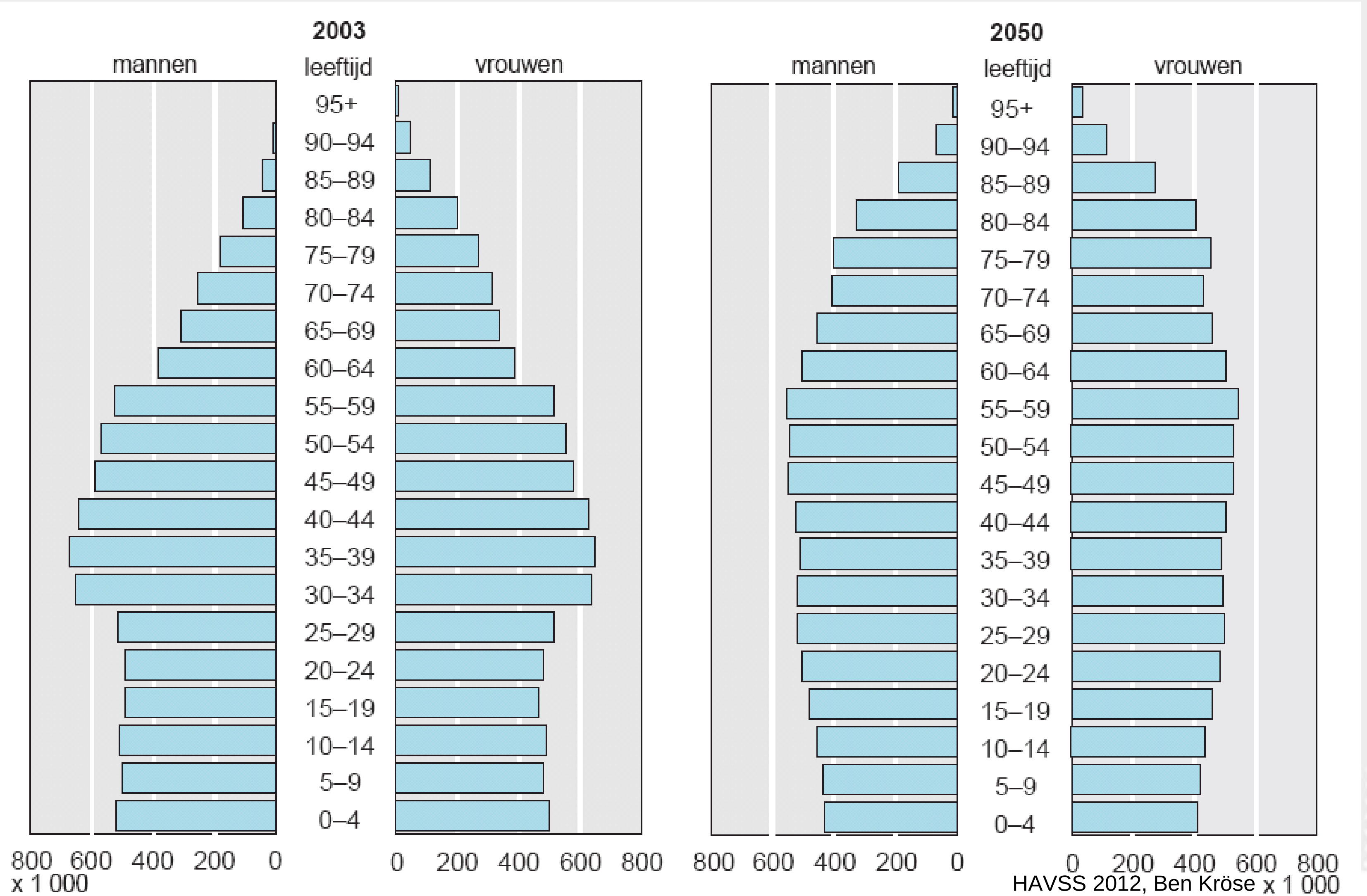
Activity recognition in ADL settings

Ben Kröse b.j.a.krose@uva.nl

Content

- Why sensor monitoring for health and wellbeing?
- Activity monitoring from simple sensors
- Cameras
- Co-design and privacy issues

Necessity for assistive technology



UNIVERSITY OF AMSTERDAM

Hogeschool van Amsterdam
University of Applied Sciences

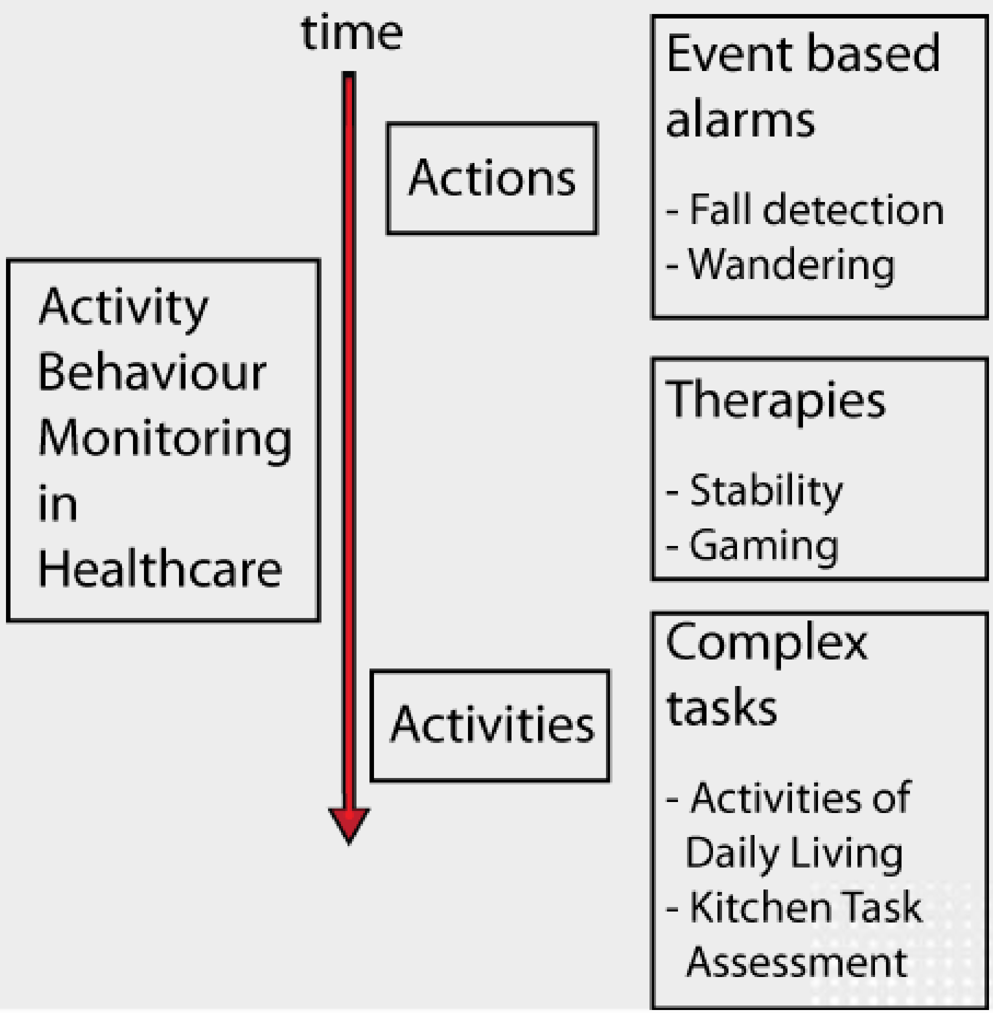
Why sensors for health and wellbeing?

- Assistive technology is needed…
  - Physical support
  - Cognitive support
  - Social support
- Assistive health systems need accurate assessments on the current state of the person;
  - Physical
  - Cognitive
  - Social

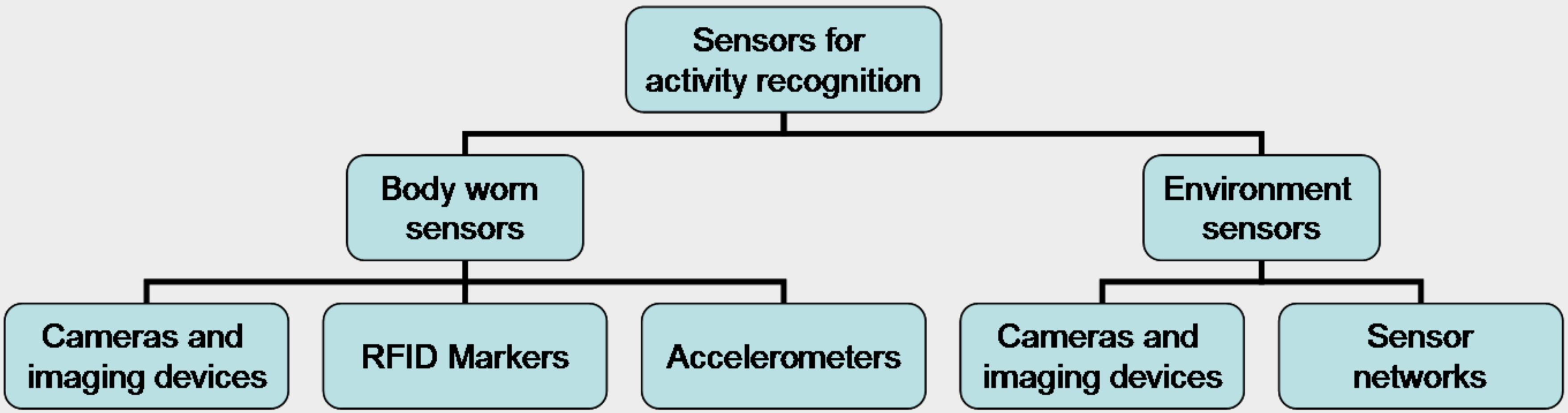
Why sensors for health and wellbeing?

- Sensing systems are needed that *monitor* the patient.
  - Monitoring systems vital signs **directly**
    - Heart rate
    - blood pressure
    - sugar level,
  - Monitoring the health state **indirectly**, by measuring the **activity behavior** of the patient

What sort of activities?



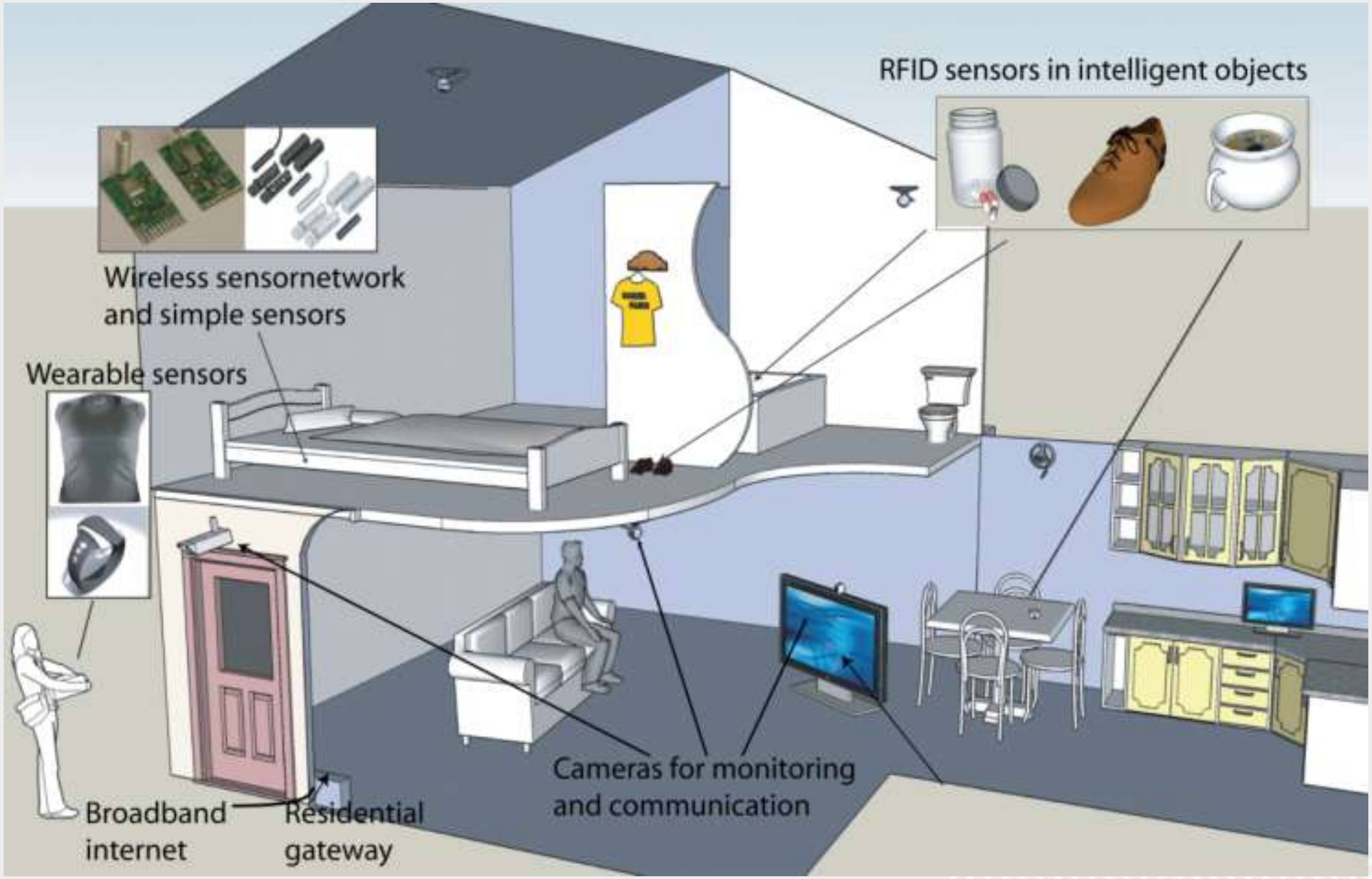
What sort of sensors?



UNIVERSITY OF AMSTERDAM

Hogeschool van Amsterdam
University of Applied Sciences

Smart Homes

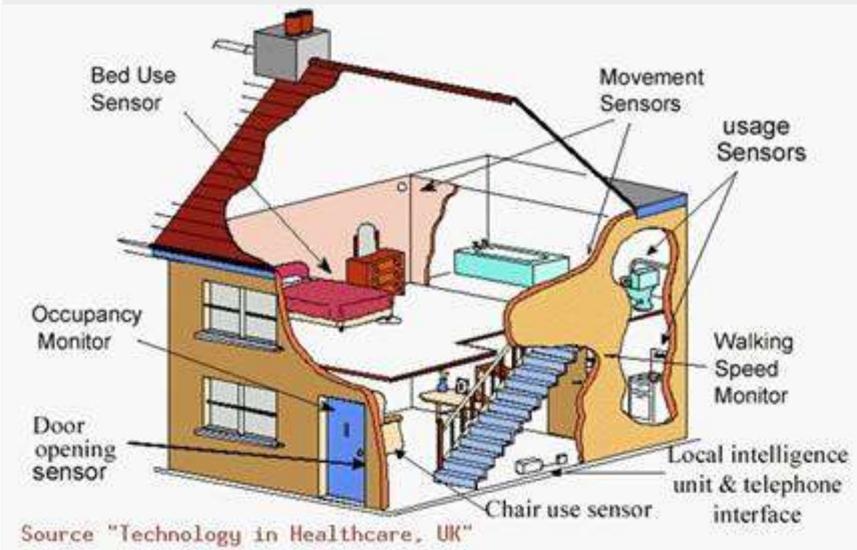


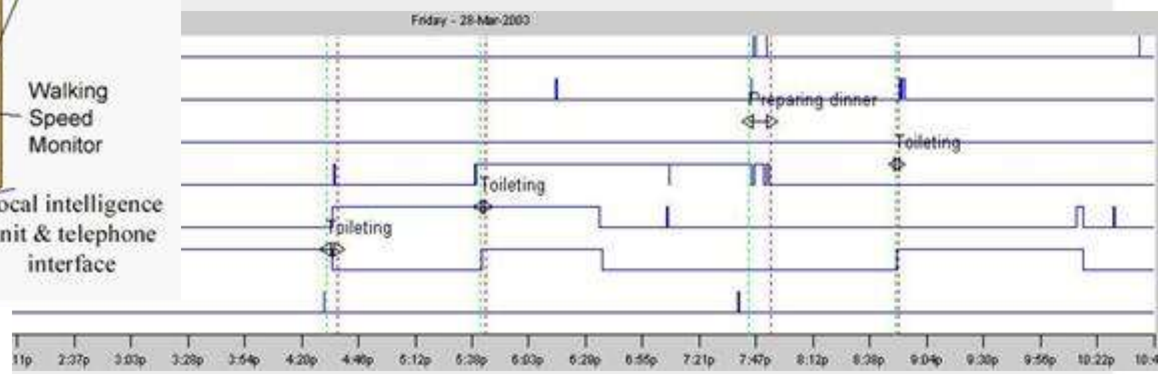
ACTIVITY RECOGNITION FROM SIMPLE SENSORS

UNIVERSITY OF AMSTERDAM

Hogeschool van Amsterdam
University of Applied Sciences

Monitoring of activities of elderly using simple sensors





- Psychogeriatric ward Naarderheem
- Assisted living appartments

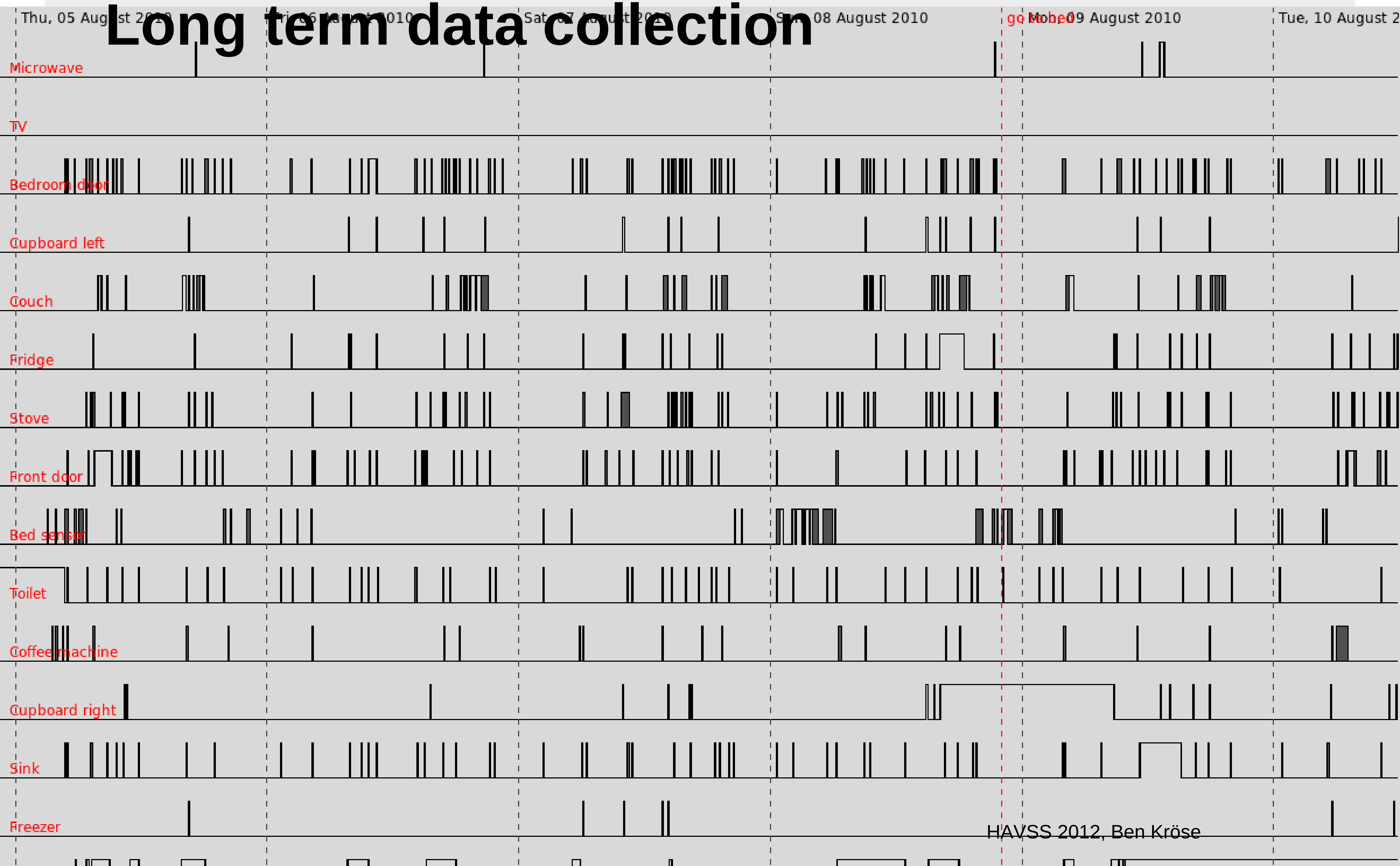
De Flank: Assisted living homes

What do we measure?

- Intern: Psychogeriatric ward Naarderheem: 8 rooms with sensors on bed, door, movement detectors.
- External Assisted living: 7 apartments with 15 sensors each:
  - Bed
  - Diverse kitchen cabinets
  - Doors
  - Electrical appliances
  - Couch
  - Motion sensors

Toilet flush sensor

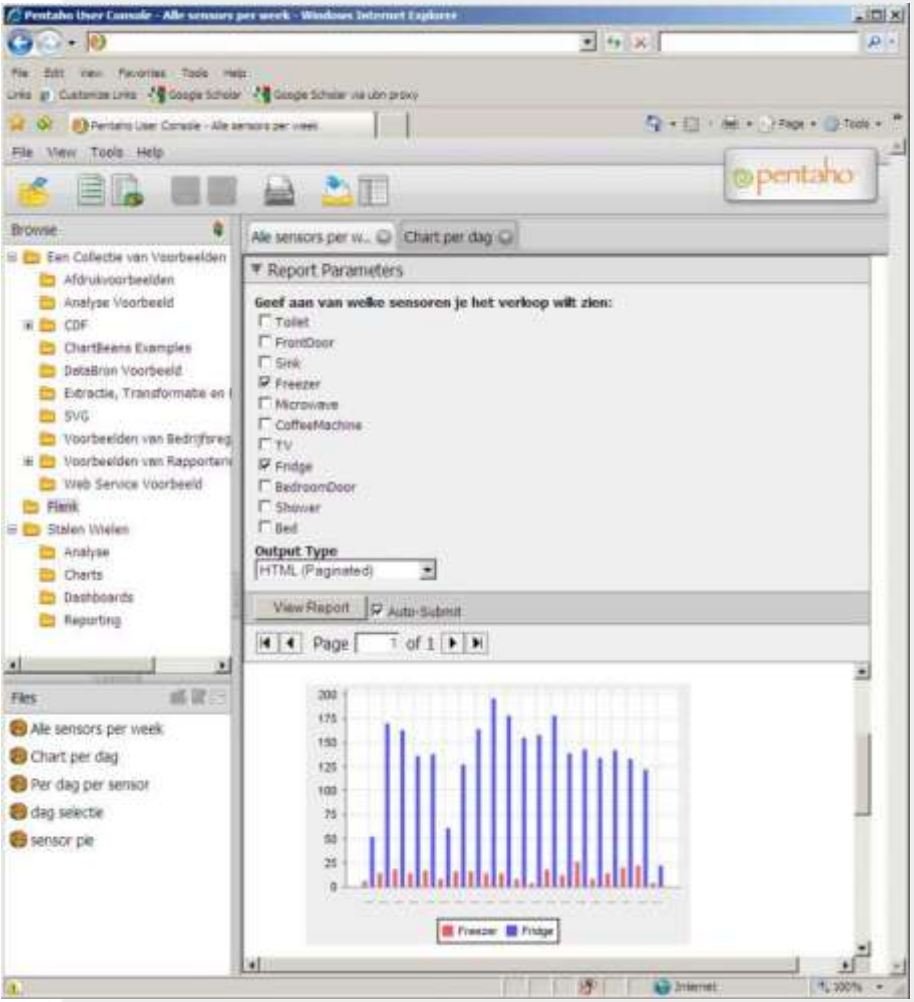
Long term data collection

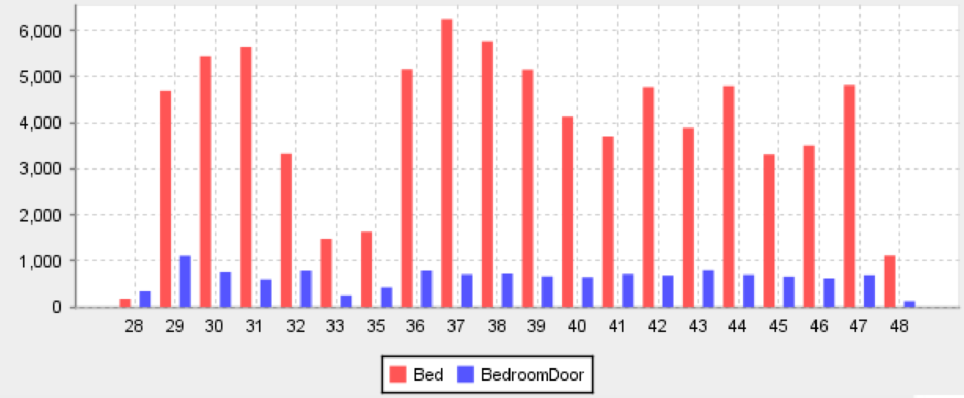


Feature selection

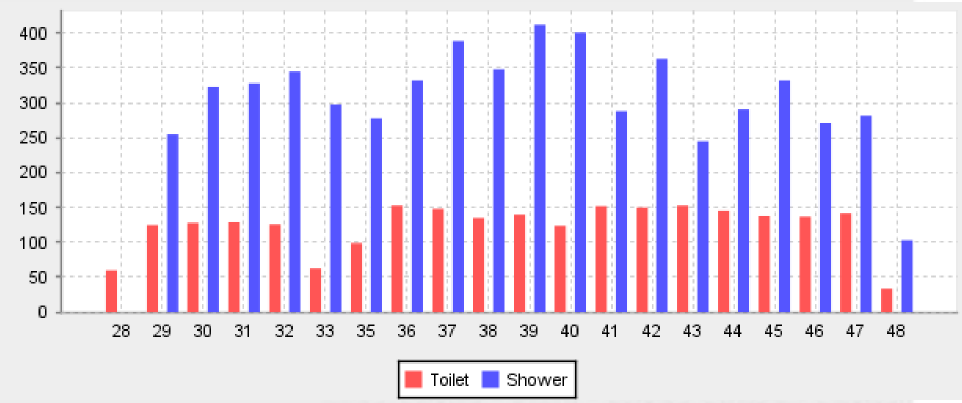
- Based on expert knowledge
  - Interview caregivers
  - Make features
  - Select best features
- Data driven
  - ICA, PCA
  - Clustering

Visualization sensordata





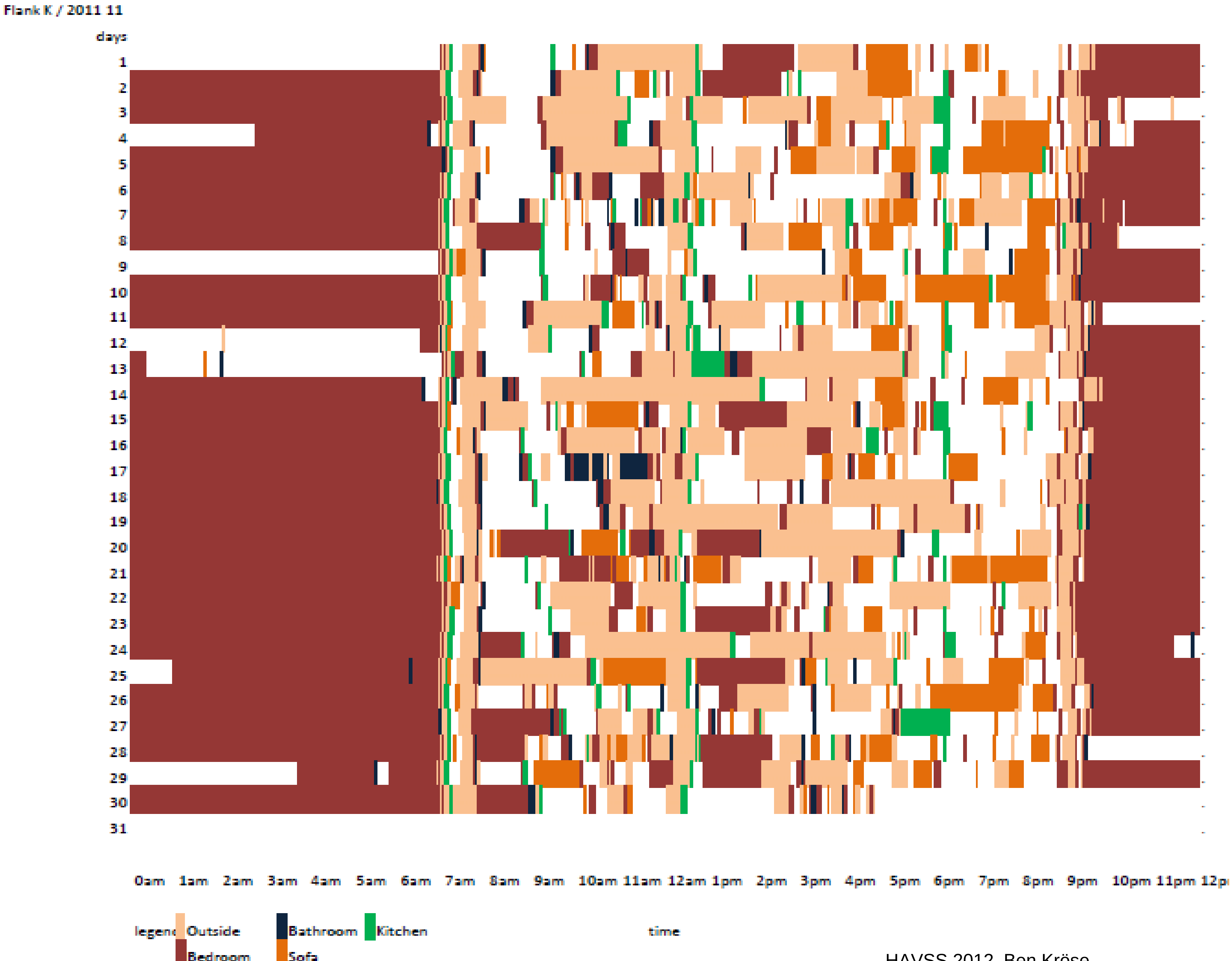
Badkamerpatroon:



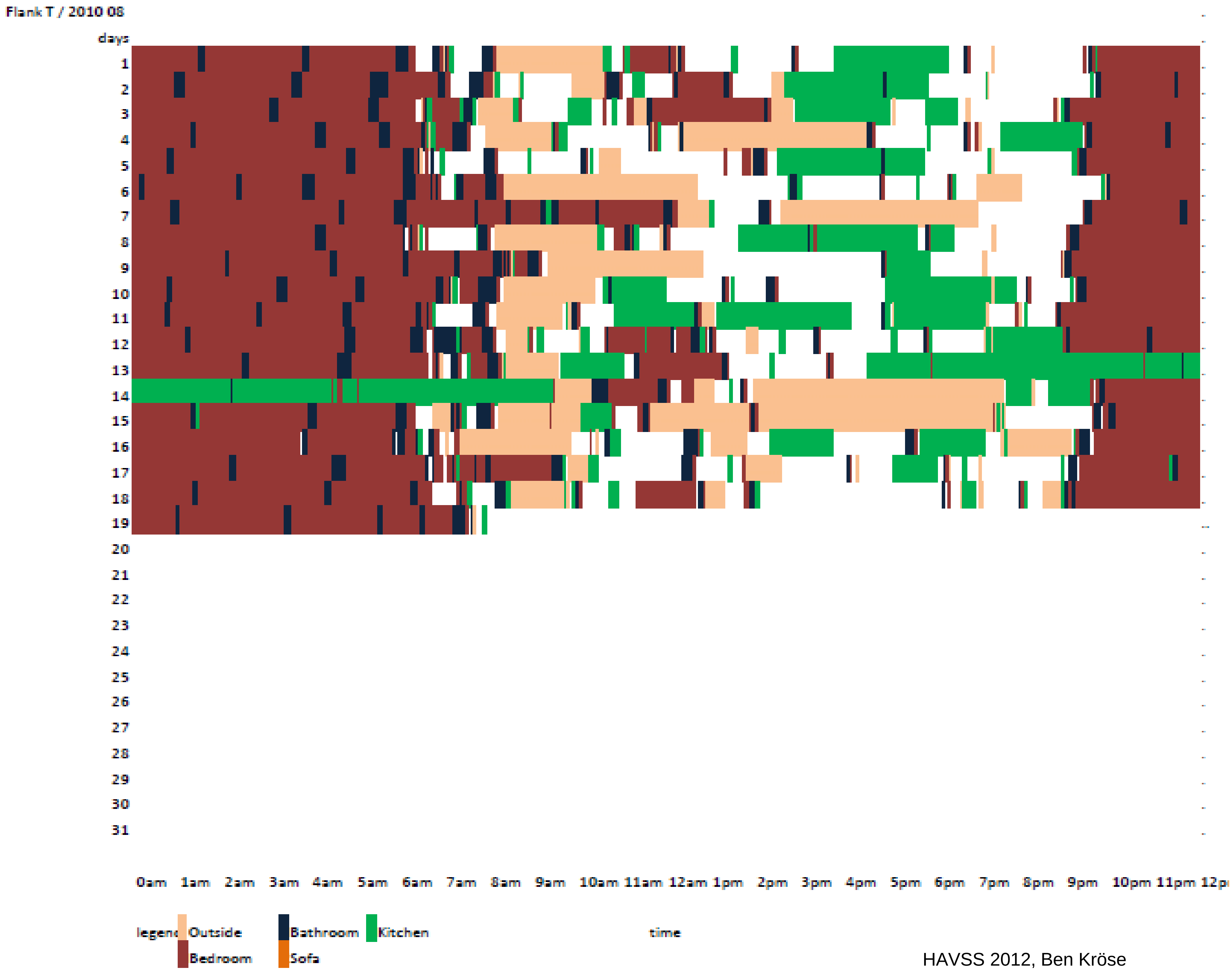
Comparison with measurements by caregivers

Occupational therapists use different measures to indicate ability of living independently:

- Subjective assessment
  - modified KATZ ADL index
    (self report on basic ADL, instrumental ADL)
- objective assessment
  - AMPS scale
    (Physical performance: gait- speedtest
    3-m measured walk, grip strenght-test
    Jamar Dynamometer)

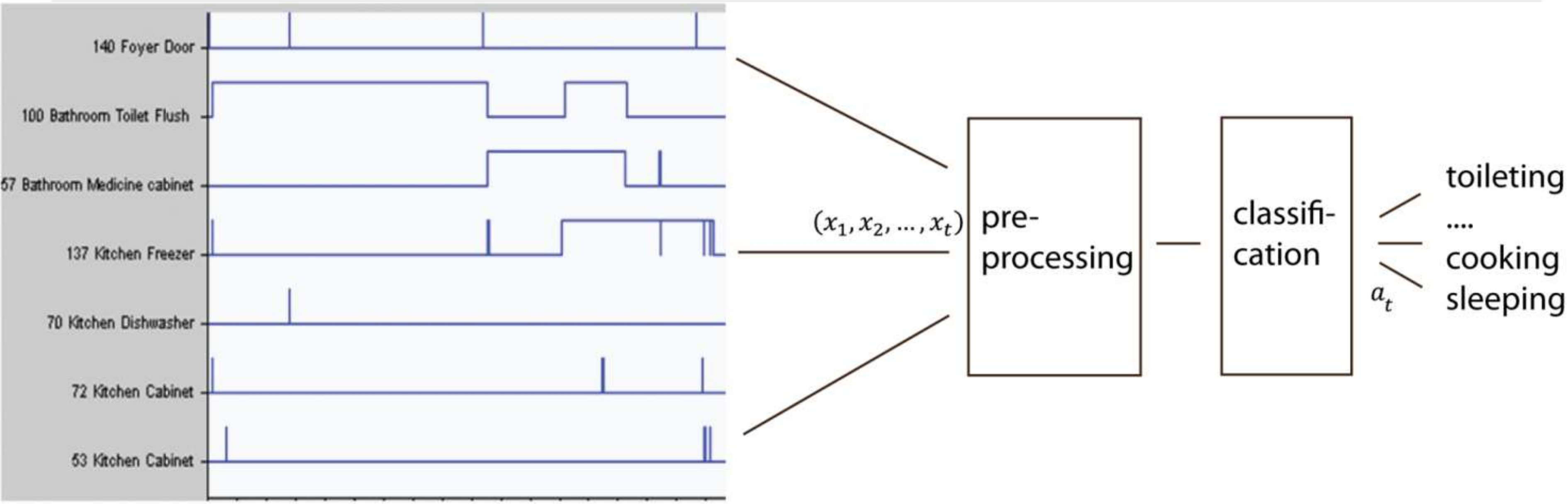


HAVSS 2012, Ben Kröse



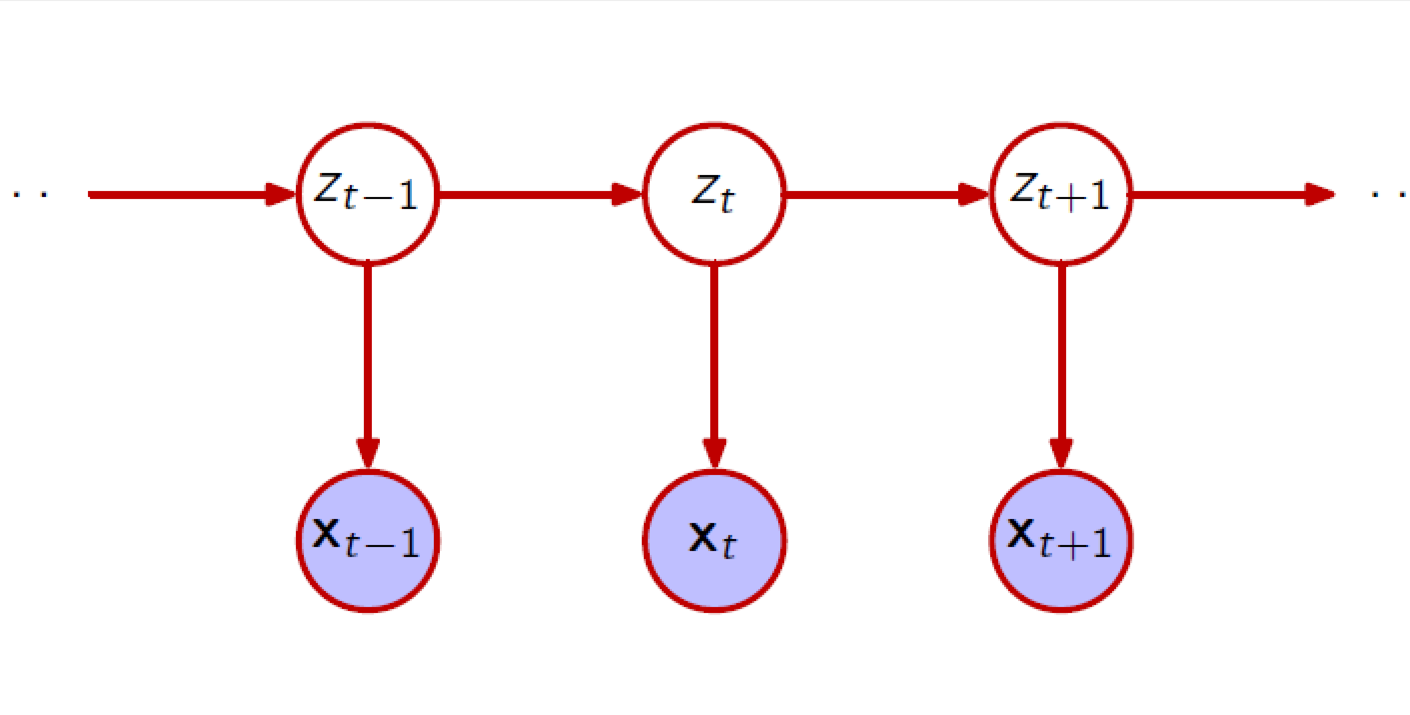
HAVSS 2012, Ben Kröse

Automatic recognition of activities



- Try to recognize ADL's from simple sensors

Hidden Markov Model

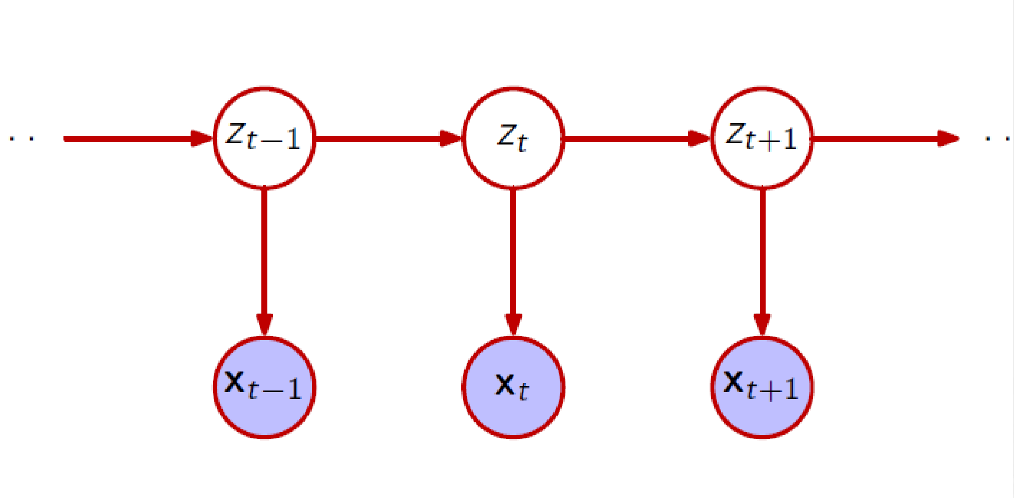


- Sensor pattern $\mathbf{x}_t$ depends on activity z_t
- activity z_t depends on activity z_{t-1}

Hidden Markov Model

- Activity at time *t*: z_t = {cooking, sleeping, ….}
- Sensor pattern $\mathbf{x}_{t}$: binary vector (0,0,1,1,…) indicating the sensor values at time = *t*

$$p(z_{1:T}, x_{1:T}) = p(z_1) \prod_{t=1}^{T} p(x_t|z_t) \prod_{t=1}^{T} p(z_{t+1}|z_t)$$



Hidden Markov Model

Parameters of the HMM:

- Transition probability: $\mathbf{A}=p(z_{t+1}|z_t)$
  - Modeled with a matrix A
- Observation model: $p(\mathbf{x}_t|z_t)$
  - Assume independence: $p(\mathbf{x}_t|z_t) = \prod_n p(x_{n,t}|z_t)$

    Each feature x_n is modeled by an independent Bernoulli distribution, where $\mu_{i,n}$ is the parameter of the n th feature

    $$p(x_{n,t}|z_t = i) = \mu_{i,n}{}^{x_n}(1 - \mu_{i,n}{}^{x_n})^{(1-x_n)}$$

UNIVERSITY OF AMSTERDAM

Hogeschool van Amsterdam
University of Applied Sciences

How to train the parameters?

- Need a training set consisting of examples:
  $\{z_1, \mathbf{x}_1 , z_2, \mathbf{x}_2 ,\ldots , z_N , \mathbf{x}_N \}$
- Estimate parameters with ML methods

Inference

- The inference problem for the HMM consists of finding the single best state sequence that maximizes $p(z_{1:T}, \boldsymbol{x}_{1:T})$
- Although the number of possible paths grows exponentially with the length of the sequence, the best state sequence can be found efficiently using the Viterbi algorithm.

Hidden Markov Model

Advantages:

- Fast and efficient
- Good precision in recognition
- Needs relatively little training data

Disadvantages:

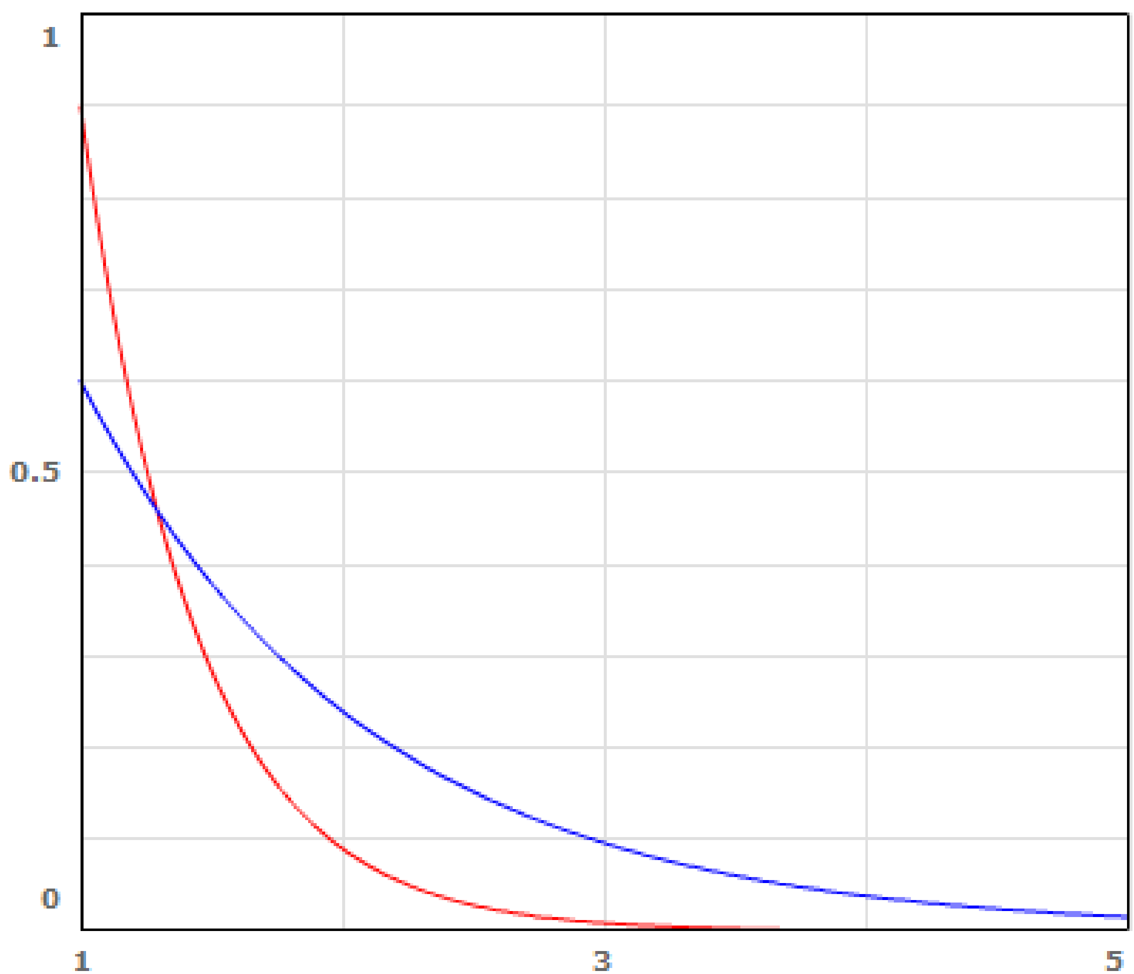
- Limited model of the duration of the activities
- No modeling of the structure or the duration of activities

Question

- What is the probability that the hidden state has a duration *d* given that the observation is constant?

Question

- What is the probability that the hidden state has a duration *d* given that the observation is constant
- $p(d) = (a_{ii})^{d-1}(1 - a_{ii})$


1
0.5
0
1
3
5

Conditional random field

Don't model the joint probability density

$$p(z_{1:T}, \boldsymbol{x}_{1:T})$$

but the conditional

$$p\left(z_{1:T} | \boldsymbol{x}_{1:T}\right) = \frac{1}{Z(\boldsymbol{x}_{1:T})} \prod_{c \in C} \varphi_C(z_C, \boldsymbol{x}_C)$$

as a product of 'clique potentials'

UNIVERSITY OF AMSTERDAM

Hogeschool van Amsterdam
University of Applied Sciences

Linear chain conditional random field

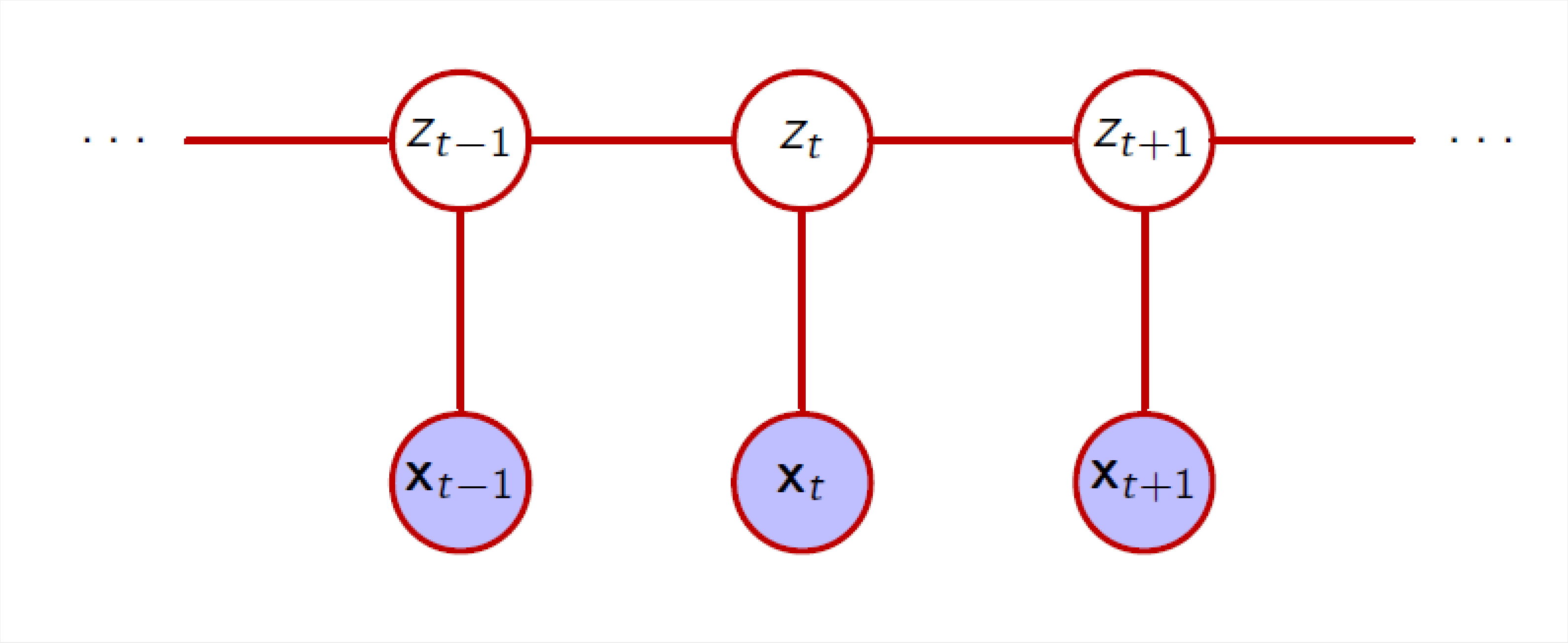
Clique potentials:

$$p\left(z_{1:T}|\boldsymbol{x}_{1:T}\right) = \frac{1}{Z(\boldsymbol{x}_{1:T})} \prod_{t=1}^{T} \varphi_t(z_t, z_{t-1}, \boldsymbol{x}_t)$$

With

$$\varphi_t(z_t, z_{t-1}, \boldsymbol{x}_t) = exp \sum_{k=1}^{K} \pi_k f_k(z_t, z_{t-1}, \boldsymbol{x}_t)$$

Linear chain conditional random field



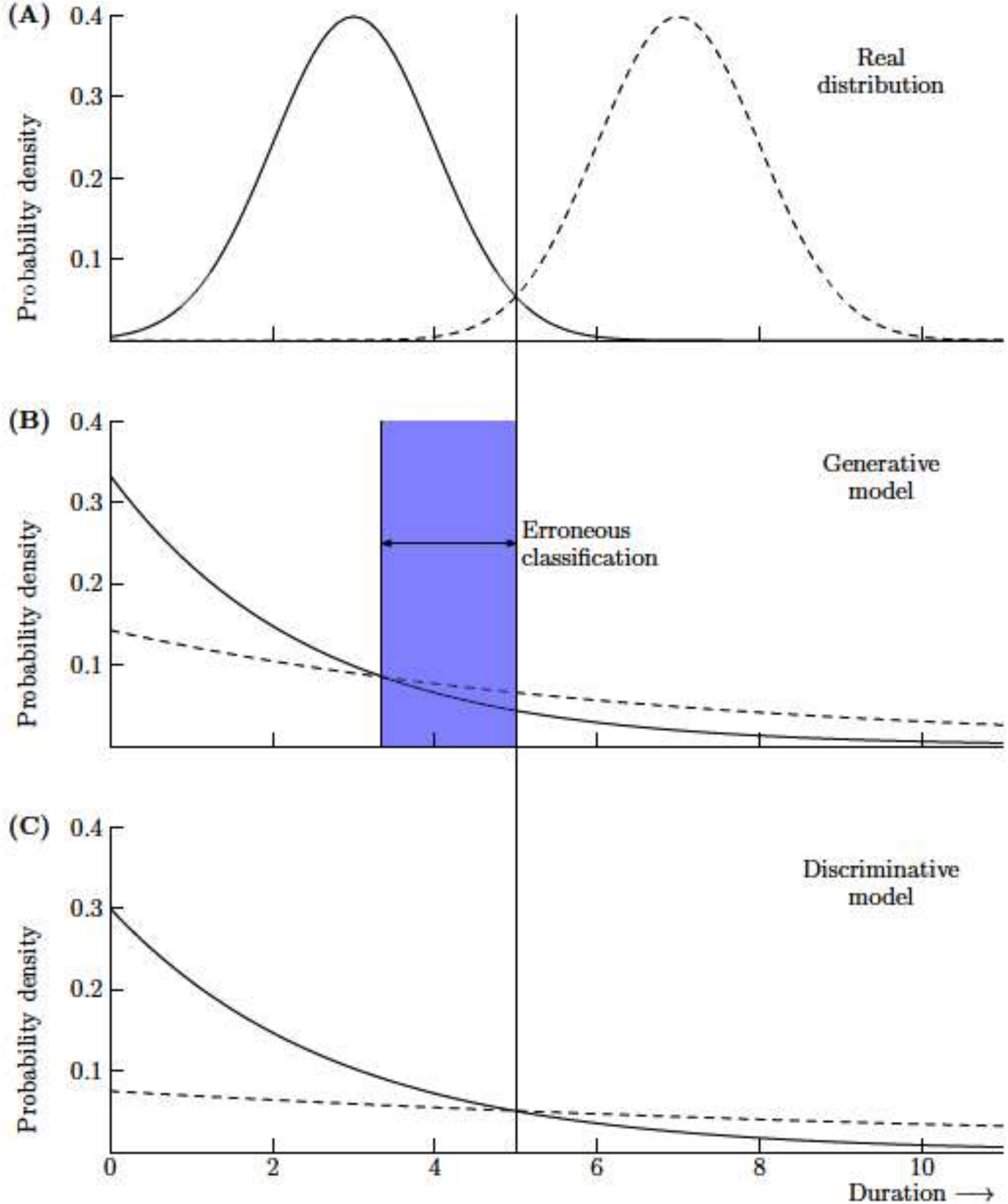
- Sensor pattern $\mathbf{x}_t$ and activity z_t are dependent
- activity z_t and activity z_{t-1} are dependent

Conditional Random Field

- Not a full probabilistic model
  - (more like a neural network)
- Also training is needed

Conditional Random Field

- Not a full probabilistic model
  - (more like a neural network)
- Also training is needed


(A)
Real
distribution
Probability density
(B)
Generative
model
Erroneous
classification
(C)
Discriminative
model
Duration ⟶

öse

Conditional random field

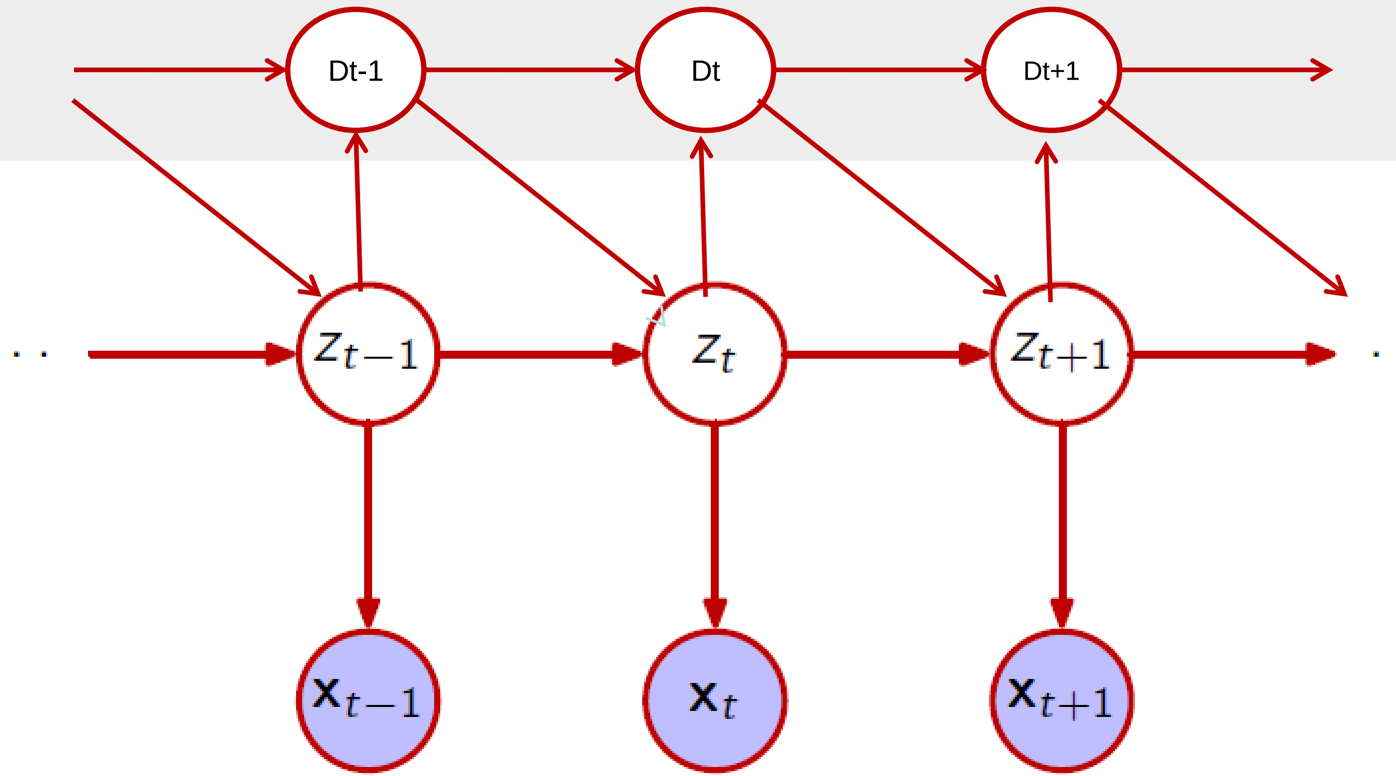
Advantages:

- Fast and efficient
- Good precision in recognition
- Better model of the duration of the activities

Disadvantages:

- No modeling of the structure or the duration of activities
- Needs many training samples
- Slow in training

Hidden Semi-Markov Model



d_t *is the remaining duration of a state*

- $$p(\mathbf{y}_{1:T}, \mathbf{x}_{1:T}, \mathbf{d}_{1:T}) = \prod_{t=1}^{T} p(\vec{x}_t \mid y_t) p(y_t \mid y_{t-1}, d_{t-1}) p(d_t \mid d_{t-1}, y_t)$$

Hidden Semi-Markov Model

Advantages:

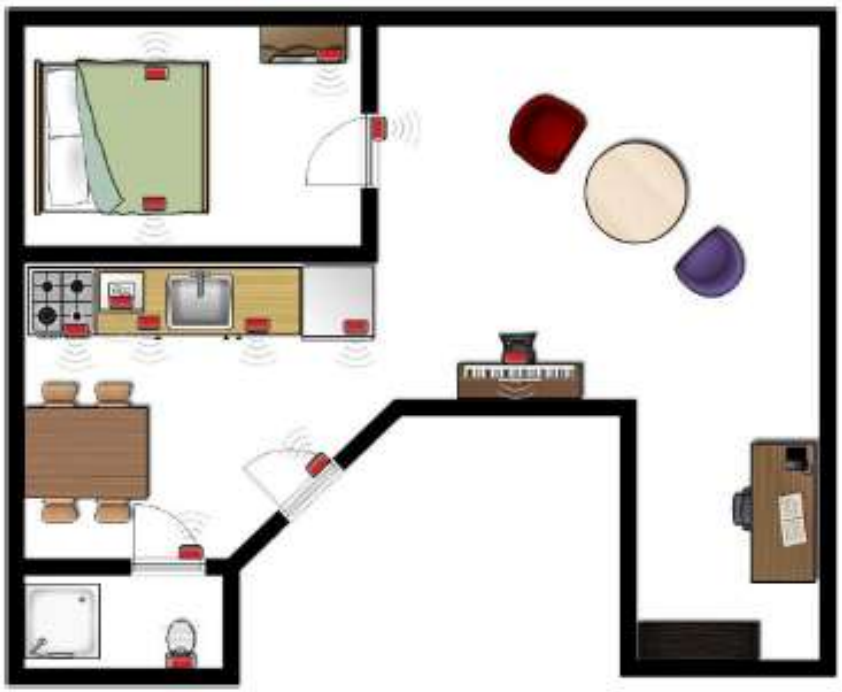
- Explicit modeling of the duration of an activity
- Good precision in recognition

Disadvantages:

- Needs many training samples
- Computational very expensive

Experiments: data

(a) House A



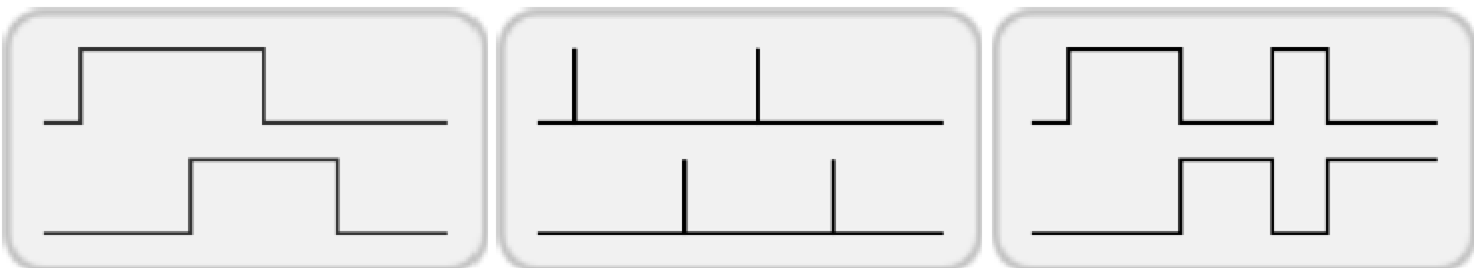
(b) House B

- Activiteiten

House A	House B	House C
Other	Other	Other
Leaving	Leaving	Leaving
Toileting	Toileting	Eating
Showering	Showering	Toileting
Brush teeth	Brush teeth	Showering
Sleeping	Sleeping	Brush teeth
Breakfast	Dressing	Shaving
Dinner	Prep. Breakfast	Sleeping
Snack	Prep. Dinner	Dressing
Drink	Drink	Medication
	Dishes	Breakfast
	Eat Dinner	Lunch
	Eat Breakfast	Dinner
	Play piano	Snack
		Drink
		Relax

Tab. 3.1: List of activities for each home.

Representation



(a) Raw (b) Changepoint (c) Last-fired

Fig. 3.3: Different feature representations.

Results

True \ Model	Other	Leaving	Toileting	Showering	Brush teeth	Sleeping	Breakfast	Dinner	Snack	Drink
Other	915	309	517	401	36	196	61	861	91	820
Leaving	30	19282	12	7	6	0	0	0	0	0
Toileting	46	4	259	13	15	19	0	2	1	6
Showering	7	1	13	229	0	0	0	1	0	0
Brush teeth	5	3	12	3	7	0	0	2	0	0
Sleeping	3	0	44	0	4	10778	0	0	0	0
Breakfast	11	0	3	0	0	1	31	22	9	10
Dinner	13	0	0	0	0	0	12	225	27	10
Snack	6	0	0	1	0	1	2	12	20	0
Drink	5	0	1	1	0	0	1	10	2	29

Tab. 3.7: Experiment 2, House A: Confusion matrix for the HMM using the last-fired

features.

True \ Model	Other	Leaving	Toileting	Showering	Brush teeth	Sleeping	Breakfast	Dinner	Snack	Drink
Other	3586	271	16	55	0	178	0	94	0	7
Leaving	8	19319	8	2	0	0	0	0	0	0
Toileting	59	10	220	9	0	63	0	4	0	0
Showering	182	6	6	57	0	0	0	0	0	0
Brush teeth	10	3	17	2	0	0	0	0	0	0
Sleeping	0	0	27	0	0	10802	0	0	0	0
Breakfast	23	0	0	0	0	3	53	0	3	5
Dinner	110	3	2	0	0	0	6	161	3	2
Snack	15	3	0	0	0	0	15	3	6	0
Drink	19	2	3	0	0	0	3	2	0	20

Tab. 3.8: Experiment 2, House A: Confusion Matrix for CRF using last-fired features.

		Precision	Recall	F-Measure	Accuracy
HMM	Raw	38 ± 20	46 ± 20	41 ± 20	59 ± 29
	Change	70 ± 16	74 ± 13	72 ± 14	92 ± 6
	Last	55 ± 17	70 ± 13	61 ± 15	90 ± 8
	Raw&Change&Last	64 ± 17	78 ± 11	70 ± 14	94 ± 4
	Raw&Change	47 ± 20	56 ± 20	51 ± 20	61 ± 29
	Raw&Last	63 ± 16	77 ± 12	69 ± 13	94 ± 4
	Change&Last	67 ± 18	79 ± 12	72 ± 15	94 ± 4
CRF	Raw	59 ± 19	56 ± 17	57 ± 17	90 ± 8
	Change	74 ± 17	68 ± 16	70 ± 16	91 ± 6
	Last	66 ± 16	66 ± 14	66 ± 15	96 ± 2
	Raw&Change&Last	72 ± 16	74 ± 13	73 ± 14	97 ± 3
	Raw&Change	75 ± 16	72 ± 13	73 ± 14	94 ± 5
	Raw&Last	67 ± 15	68 ± 14	67 ± 14	96 ± 3
	Change&Last	72 ± 15	74 ± 13	73 ± 14	97 ± 2

Tab. 3.4: Experiment 2, House A: Different feature representations for HMMs and CRFs.

Training time

		House A	House B	House C
Learning	HMM	1.3s	0.6s	1.0s
	CRF	1890.1s	1188.3s	3708.8s
Inference	HMM	3.9s	2.4s	3.2s
	CRF	4.7s	3.5s	5.2s

Tab. 3.9: Experiment 2: Computation times in seconds for learning and inference in HMMs and CRFs.

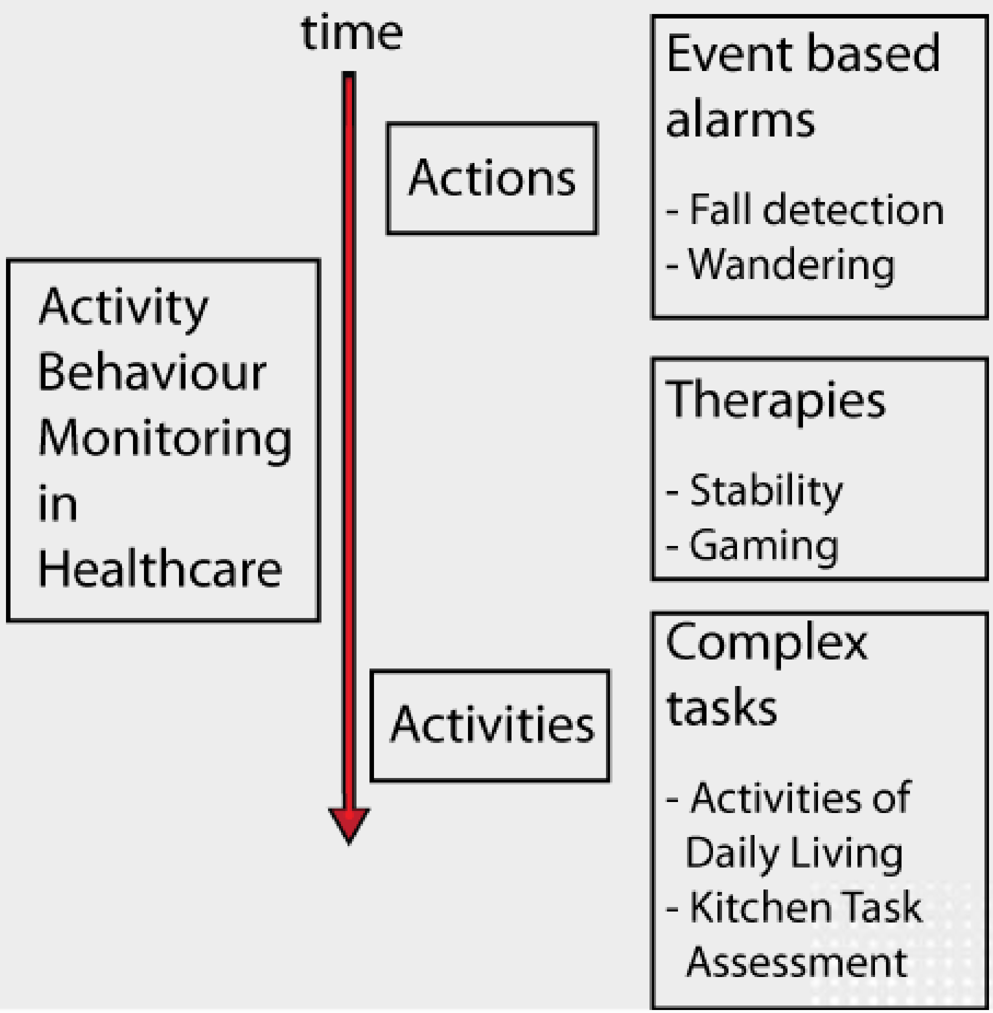
	Bathroom A			Kitchen A		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure
HMM	50 ± 13	67 ± 15	57 ± 12	56 ± 30	64 ± 33	59 ± 31
HSMM	70 ± 17	85 ± 14	76 ± 15	65 ± 27	75 ± 21	69 ± 24
CRF	73 ± 17	74 ± 14	73 ± 15	80 ± 21	79 ± 20	79 ± 20
SMCRF	75 ± 17	75 ± 14	75 ± 15	77 ± 23	77 ± 19	76 ± 21
	Bathroom B			Kitchen B		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure
HMM	64 ± 18	85 ± 14	72 ± 15	47 ± 23	56 ± 23	50 ± 22
HSMM	67 ± 20	91 ± 13	76 ± 15	47 ± 22	67 ± 20	54 ± 19
CRF	72 ± 16	75 ± 16	73 ± 15	42 ± 24	46 ± 22	44 ± 23
SMCRF	75 ± 17	77 ± 18	76 ± 17	52 ± 33	56 ± 29	54 ± 31
	Bathroom C			Kitchen C		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure
HMM	48 ± 32	57 ± 32	52 ± 32	46 ± 21	49 ± 22	46 ± 19
HSMM	60 ± 27	69 ± 27	64 ± 26	54 ± 23	61 ± 24	56 ± 22
CRF	53 ± 27	62 ± 22	57 ± 25	55 ± 28	57 ± 24	55 ± 25
SMCRF	60 ± 27	65 ± 24	62 ± 26	53 ± 26	57 ± 24	55 ± 25

Tab. 4.7: Experiment 2: Precision, recall and F-measure for hidden Markov model (HMM), hidden semi-Markov model (HSMM), conditional random field (CRF) and semi-Markov conditional random field (SMCRF). Experiments were performed on the kitchen and bathroom datasets. The changepoint and last sensor representation was used.

Cameras vs. other sensors in the home

- Body worn sensors:
  - Not always worn
  - Stigmatizing
- Simple ambient sensors
  - Nonintrusive
  - No detailed information
- Cameras
  - Privacy issues
  - Much information

What sort of activities?



SIMPLE ACTIONS

Simple actions: fall detection

- Most common cause of injury with persons 55+
- In the Netherlands annually 95.000 emergencies
- Of which 43.000 in and around the home
- 1.3% fatal
- Problem will increase with ageing population

Fall detection: existing solutions

- Wearable accelerometers

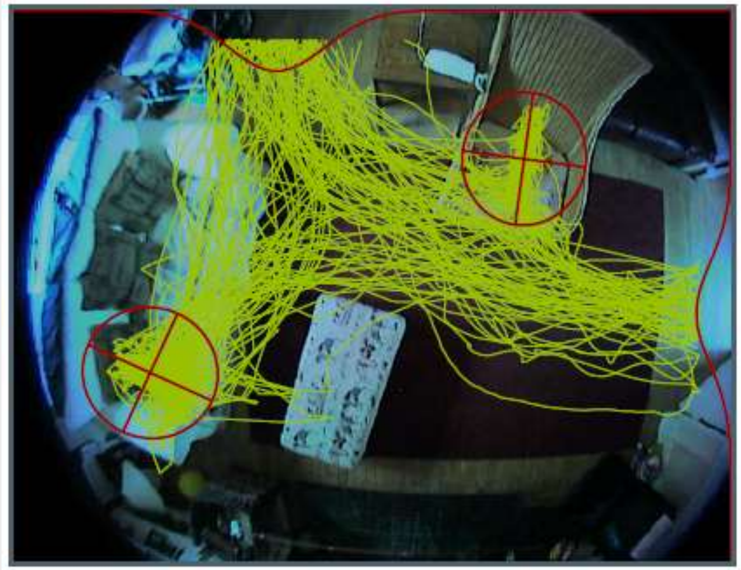
- Ambient detectors

Problems: not worn, restricted use

Fall detection with cameras

- Inactivity measurement
- Dynamics of the visual features

Cameras for fall detection



- Nait-Charif et al (2004): Inactivity based system:

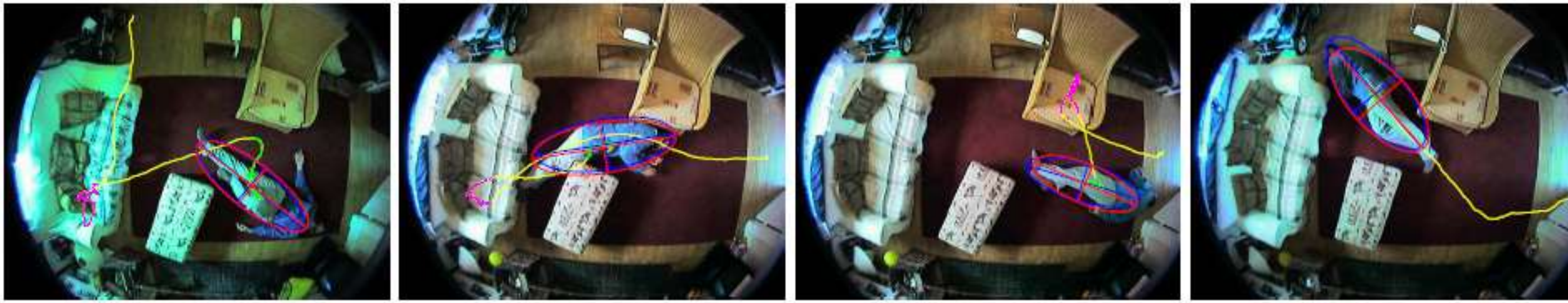


Figure 5. Segmented trajectories and detected unusual inactivity.

Activity summarisation and fall detection in a supportive home environment}},
author={Nait-Charif, H. and McKenna, S.J.},
booktitle={Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004},

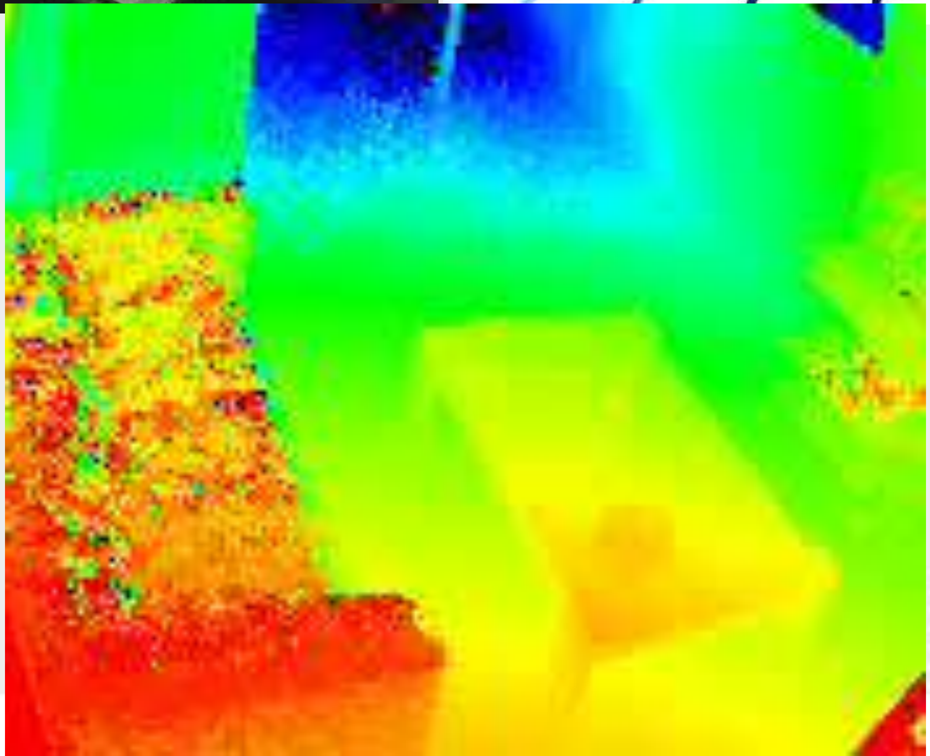
Cameras for fall detection

- 3D modeling of pose
  - Multiple cameras
  - Time of flight
  - Stereo

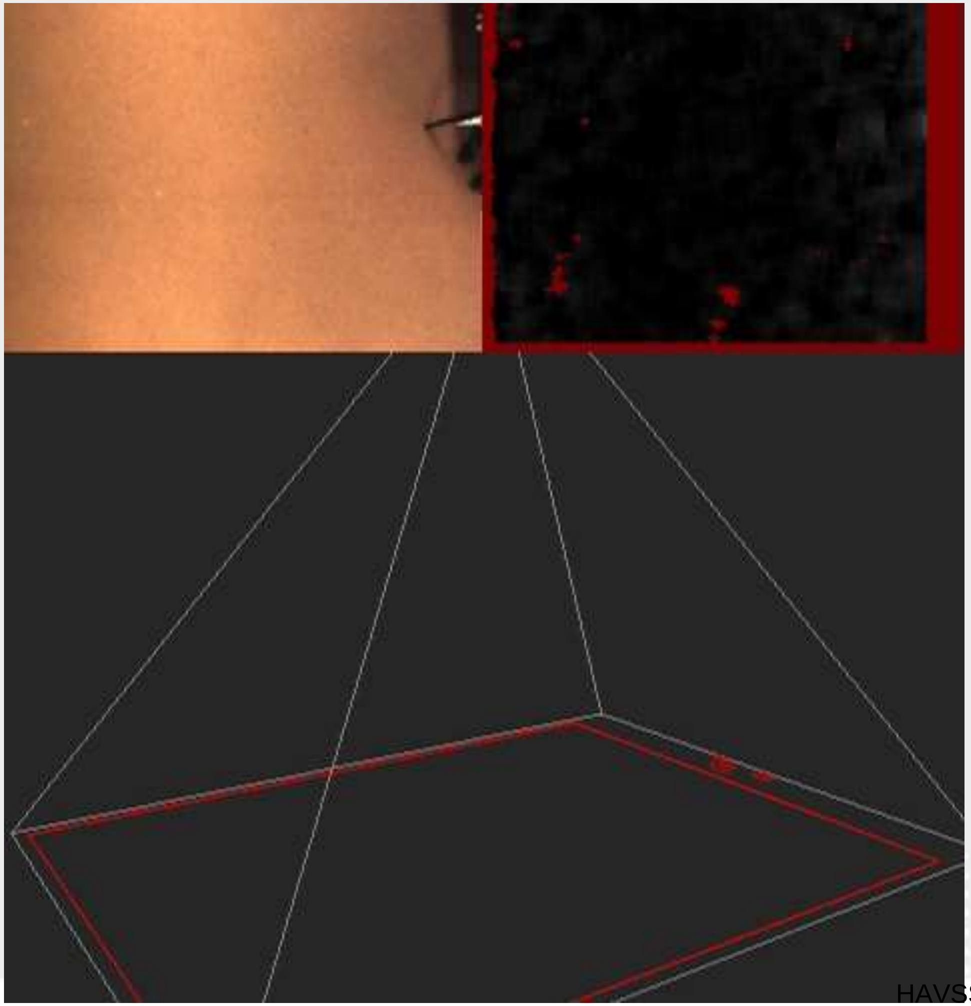
Fall lab

Hogeschool van Amsterdam
University of Applied Sciences

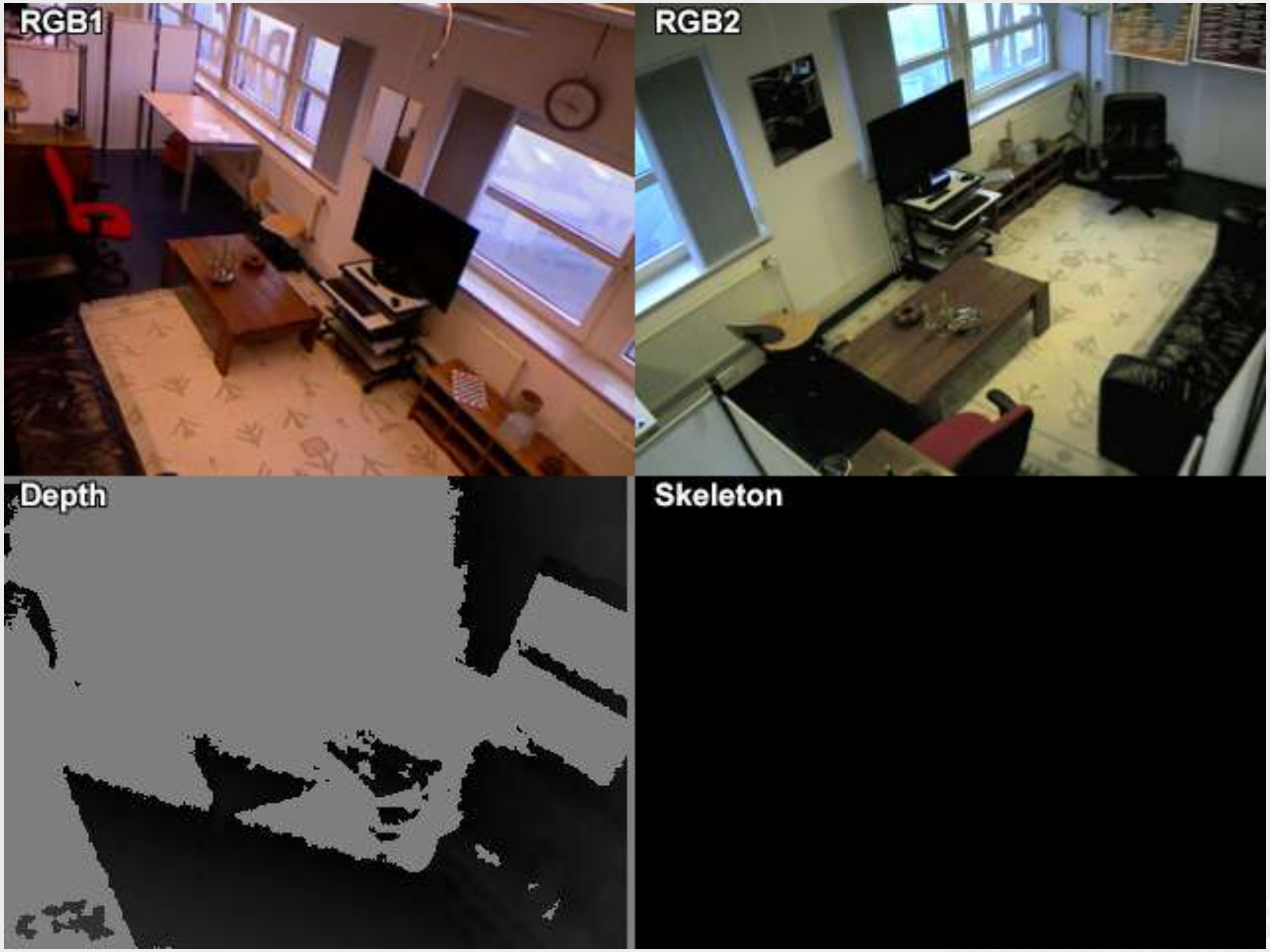




Oosterhout et al: 3D dynamics



Kinect



Compare overhead camera and Kinect

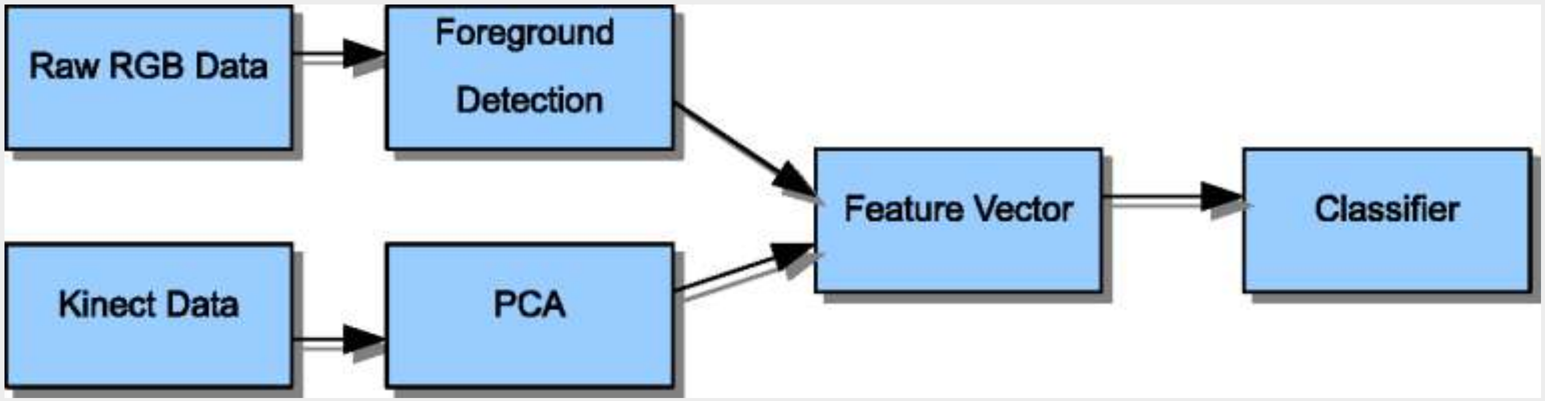
Kinect Data → PCA → Feature Vector → Classifier

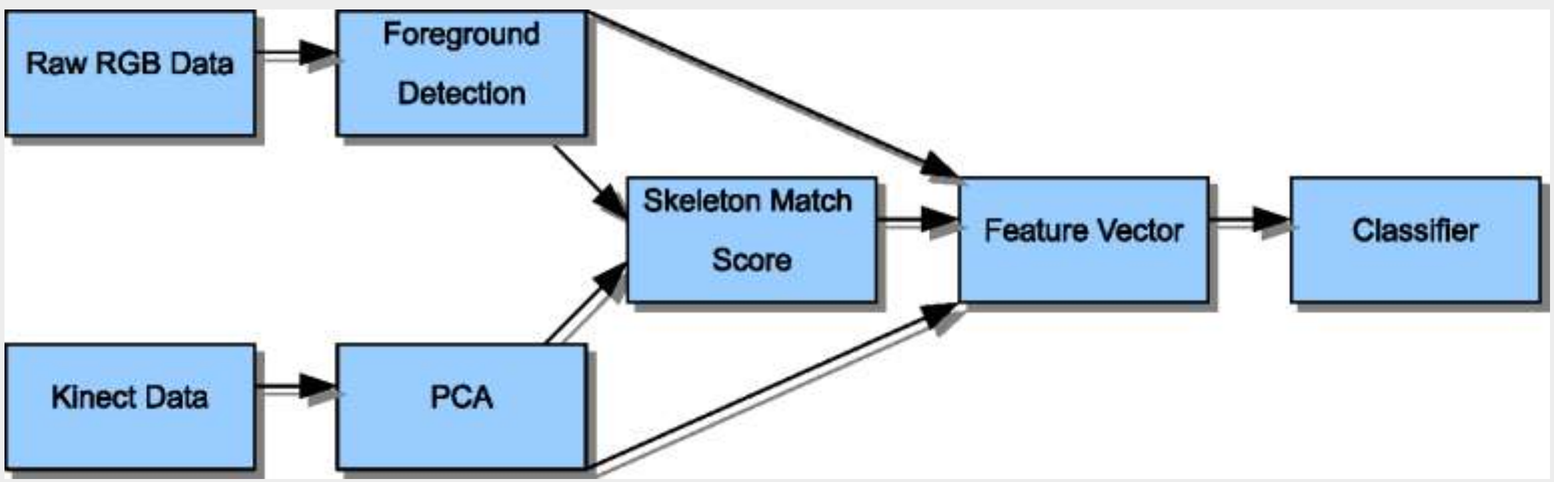
RGB Data → Foreground Detection → Feature Vector → Classifier



Results
(40 falls, 40 non-falls in total, 5-fold crossvalidation)

	TP	TN	FP	FN
Skeleton based	38.76/0.27	37.04/0.12	2.96/0.12	1.24/0.27
Bounding ellipse	36.28/0.38	40.00/0.00	0.00/0.00	3.72/0.38

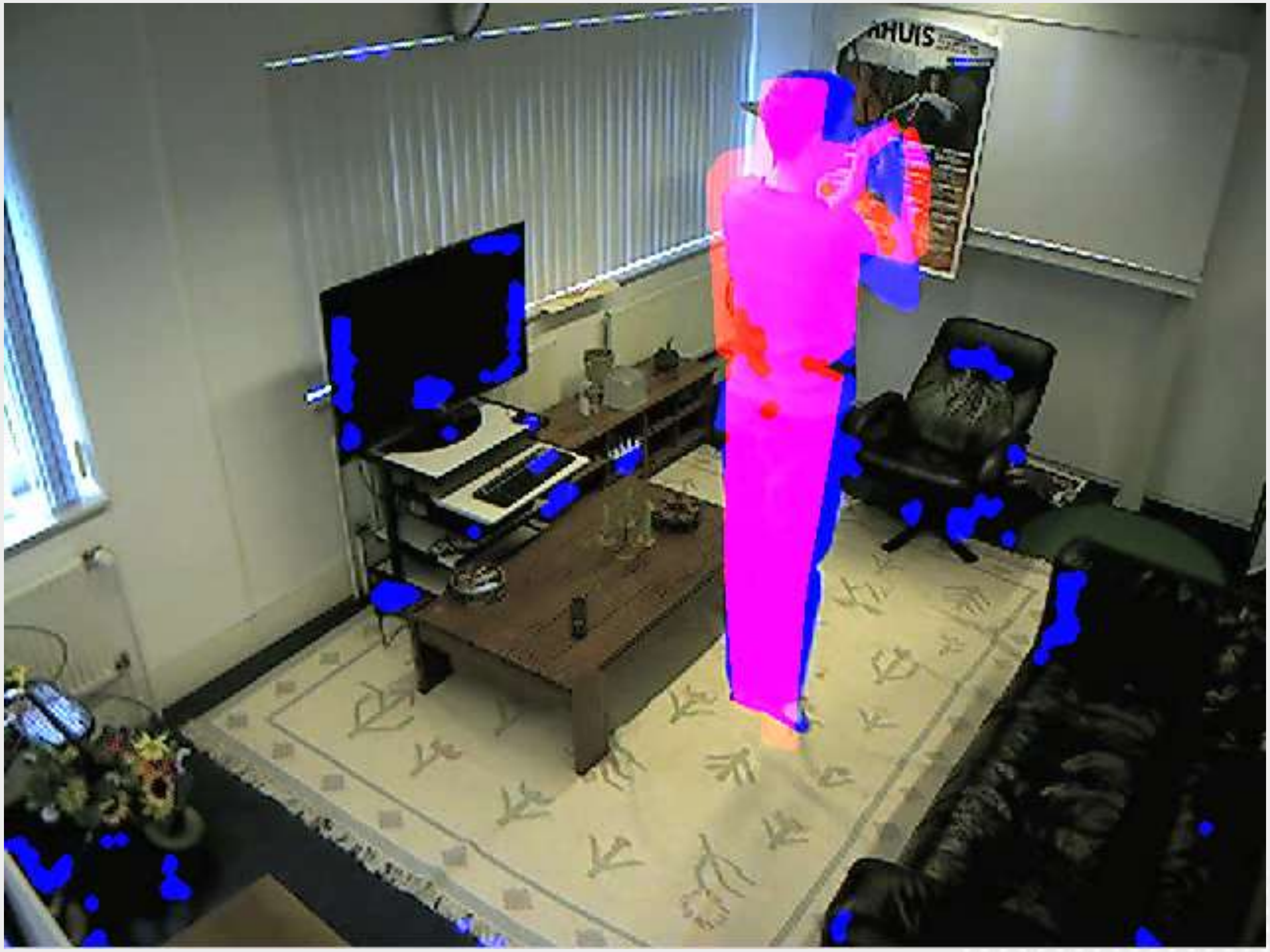

Raw RGB Data
Foreground Detection
Kinect Data
PCA
Feature Vector
Classifier


Raw RGB Data
Foreground Detection
Skeleton Match Score
Kinect Data
PCA
Feature Vector
Classifier

UNIVERSITY OF AMSTERDAM

Hogeschool van Amsterdam
University of Applied Sciences

Combining 2 camera's



Results (40 falls in total)

	TP	TN	FP	FN
Skeleton based	38.76	37.04	2.96	1.24
Bounding ellipse	36.28	40.00	0.00	3.72
Using all features	36.84	36.80		
Feature match	39.36	39.48		

Conclusions

- Research shows that methods from AI and computer vision are applicable to the health care domain
- There is a shortage of supervised data sets
- There is a shortage of realistic data sets
  - (elderly don’t fall like students)
- Privacy issues have to be taken serious and are a serious problem