Using Machine Learning for Real-time Activity Recognition and Estimation of Energy Expenditure

Emmanuel Munguia Tapia
PhD Thesis Defense

House_n
Massachusetts Institute of Technology

Do you know

- How many calories you expend each day?
- How many calories you need to to stay healthy?

Motivation

Obesity is a major health threat:

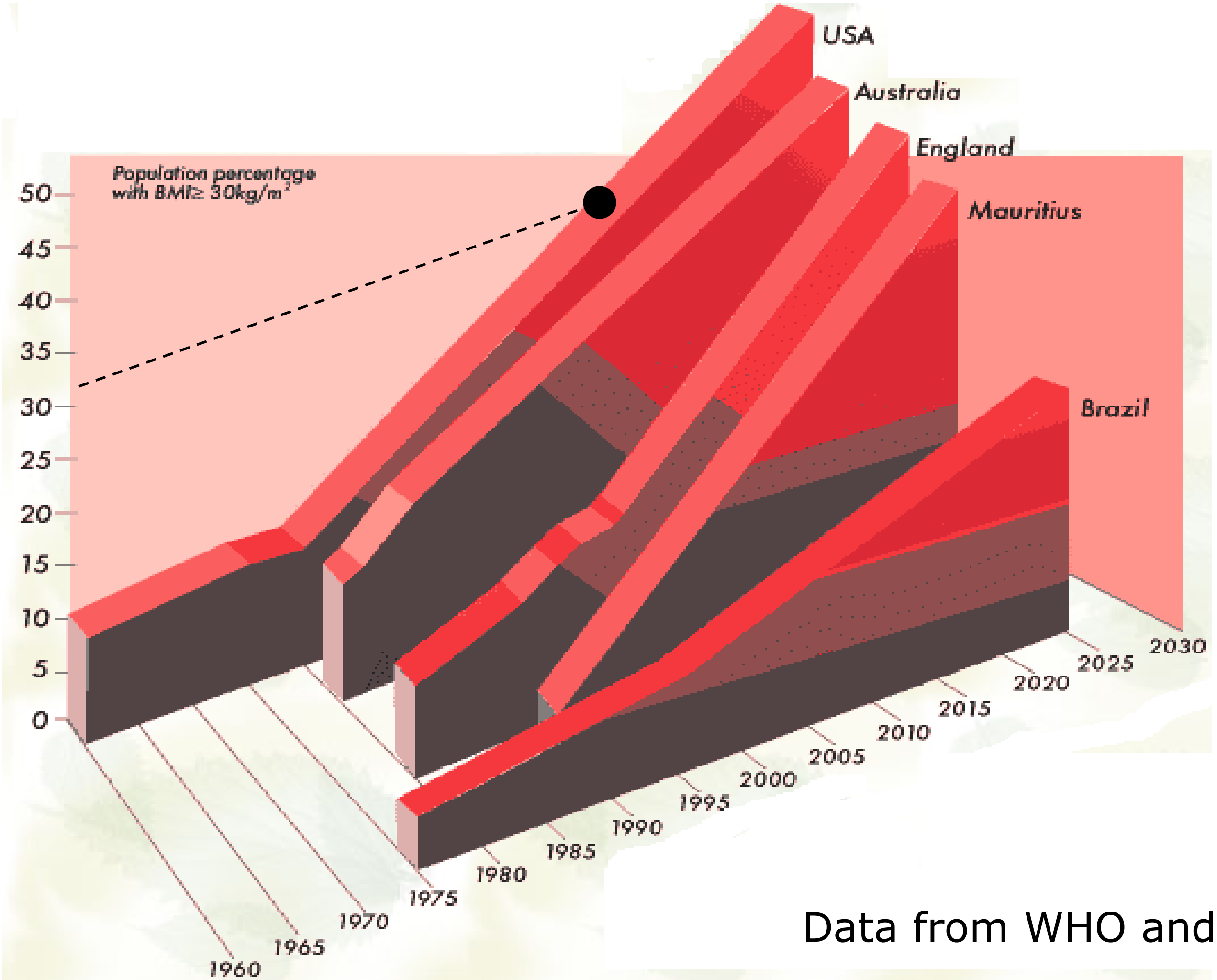
- 65% of U.S. adults are overweight
- 30% of U.S. adults are obese
- 16% of children are obese

National Center for Health statistics

Obesity is a risk factor for

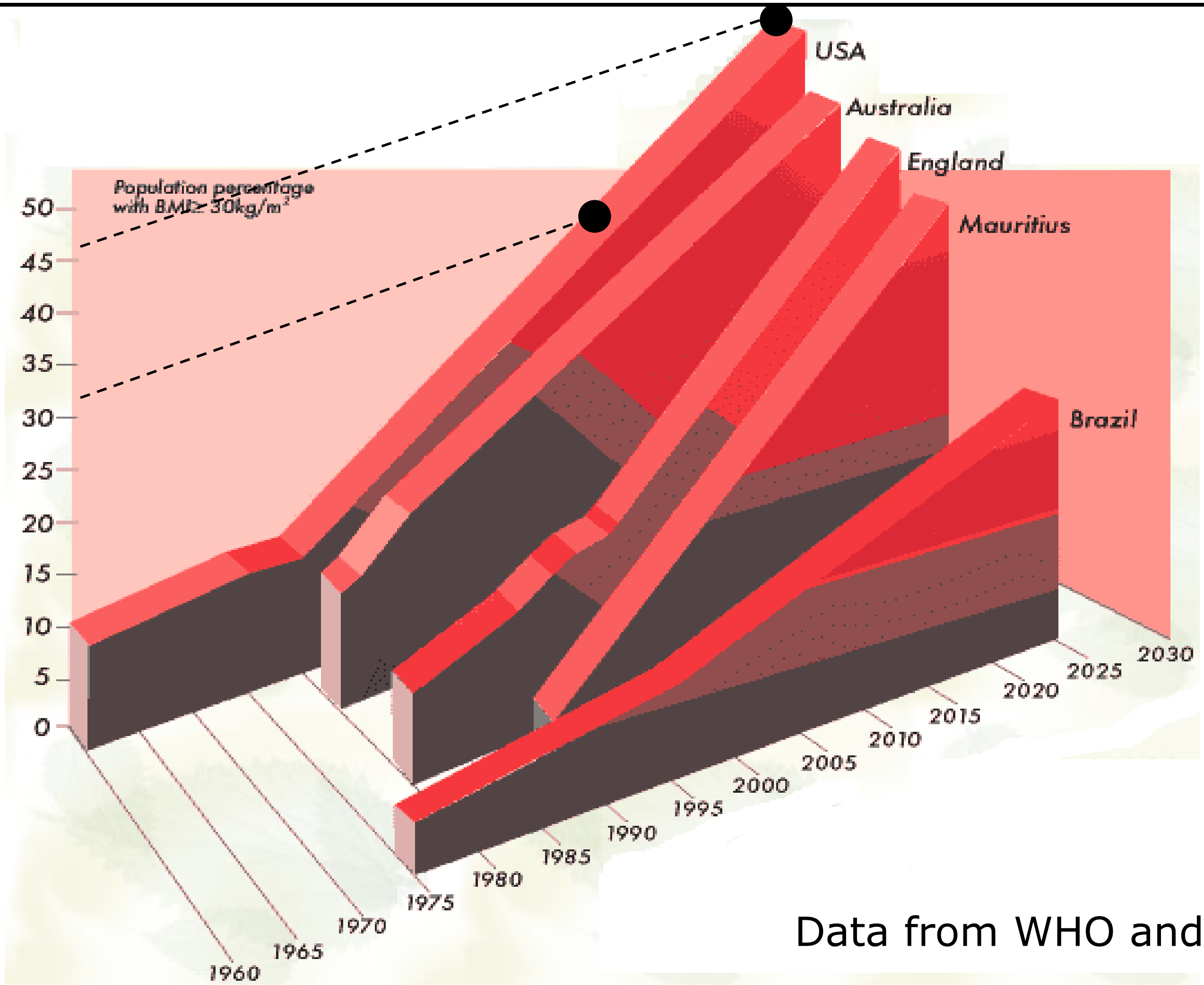
- Hypertension
- Type 2 diabetes
- Coronary heart disease
- Stroke
- Gallbladder disease
- Osteoarthritis
- Sleep apnea and respiratory problems
- Some cancers (endometrial, breast, and colon)

Projected prevalence of obesity



Data from WHO and IUNS

Projected prevalence of obesity



Data from WHO and IUNS

Energy (im)balance

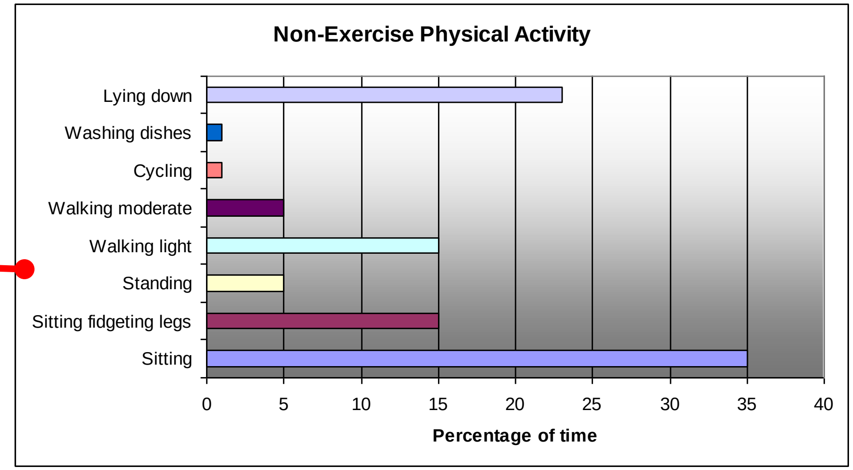
Body composition change ≈
Energy intake – Energy expenditure

Three ways to address the problem
- Magic pill
- Eat less or healthier
- Burn more calories

Two ways to help

- Knowing what people are doing
- Knowing how many calories are burned

If a mobile phone could…



1. Real-time feedback
2. Just-in-time interventions
3. Non-exercise activity thermogenesis

How is physical activity and energy expenditure presently measured?

In the lab…

Room indirect calorimetry

Portable indirect calorimeter

During free-living

Paper diaries

Electronic diaries

burdensome + time consuming

Electronic monitoring (during free-living...)

Actigraph (1axis)

Bodybugg (2axis)

Problems:

- Little or no contextual information
- Low performance on upper body and lower body activity

Goal of this work

Develop algorithms based on wireless wearable sensors that:

- Recognize activity type, intensity and duration
- Estimate energy expenditure
- Achieve reasonable performance
- Are amenable for real-time performance
- Work when sensors worn in convenient locations

This work explores the trade-offs that need to be made in order to achieve these goals

Activity Recognition Algorithms Experiments

Previous work

- Kern et al. 2003:
  8 activities, 18min, 1 researcher
- Blum et al. 2005
  8 activities, 24hrs, 1 researcher
- **Bao et al. 2005**
  20 activities, 30hrs, 20 subjects
- Olguin et al. 2006
  8 activities, 3 subjects
- Ravi et al. 2008
  8 activities, 2 subjects
- Huynh and Schiele 2005
  6 activities, 200min, 2 participants
- Lester et al. 2006
  10 activities, 3 subjects

Contributions

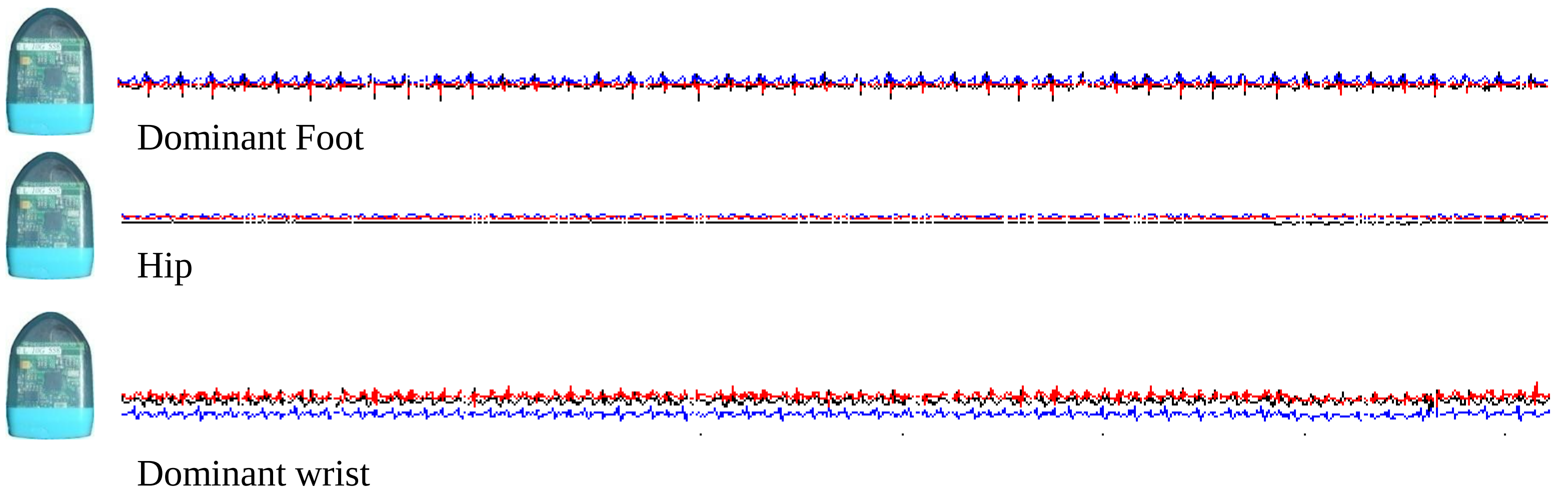
- 52 activities, 120hrs, 20 subjects
- Collected at a gym and residential home
- Recognize activity type *and* intensity
- Systematic experiments to determine
  - Algorithm parameters
  - Value of accelerometers versus heart rate
  - Location and number of the sensors
- Proof of viability of real-time system to recognize arbitrary activities

Demo: Activity recognition

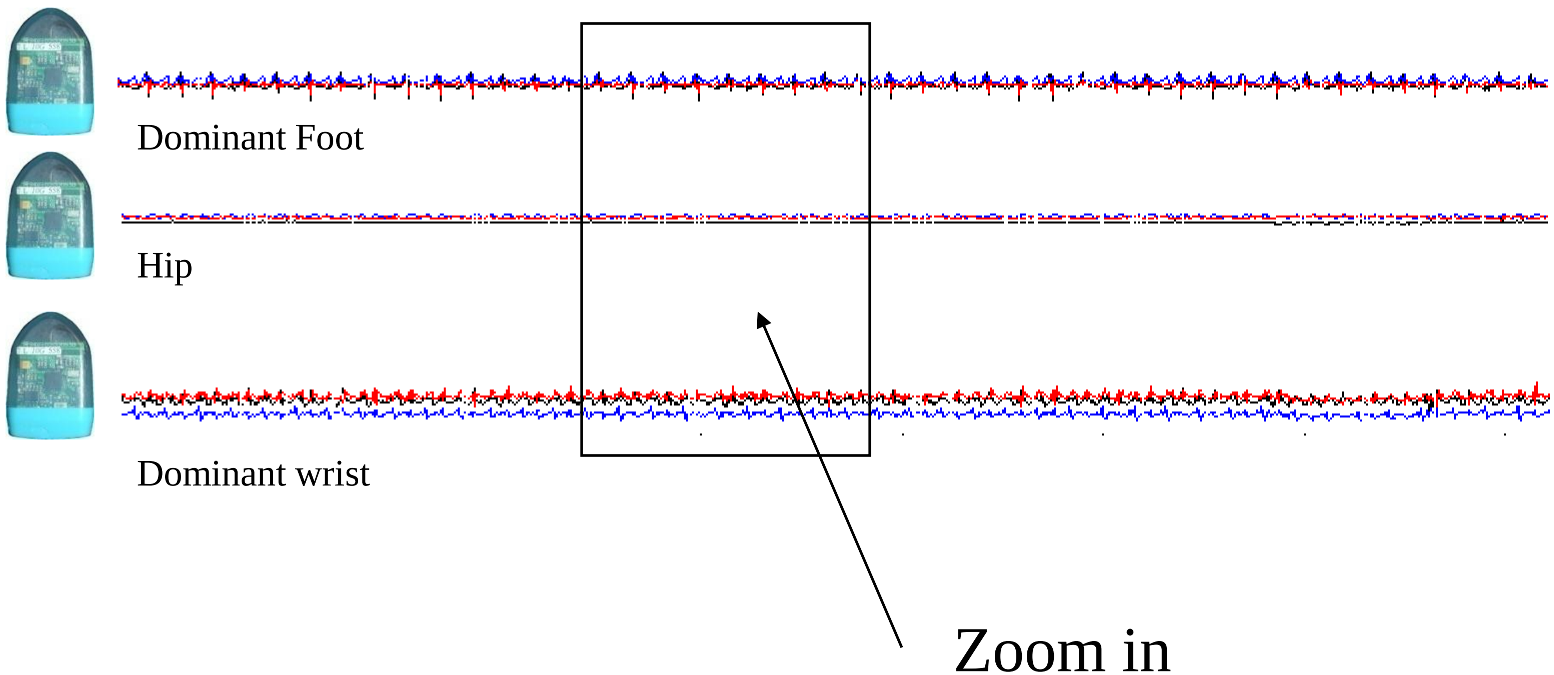
1. Wear three wireless accelerometers
2. Select 10 physical activities
3. Provide 2 minutes of data per activity

Activity Recognition Algorithm

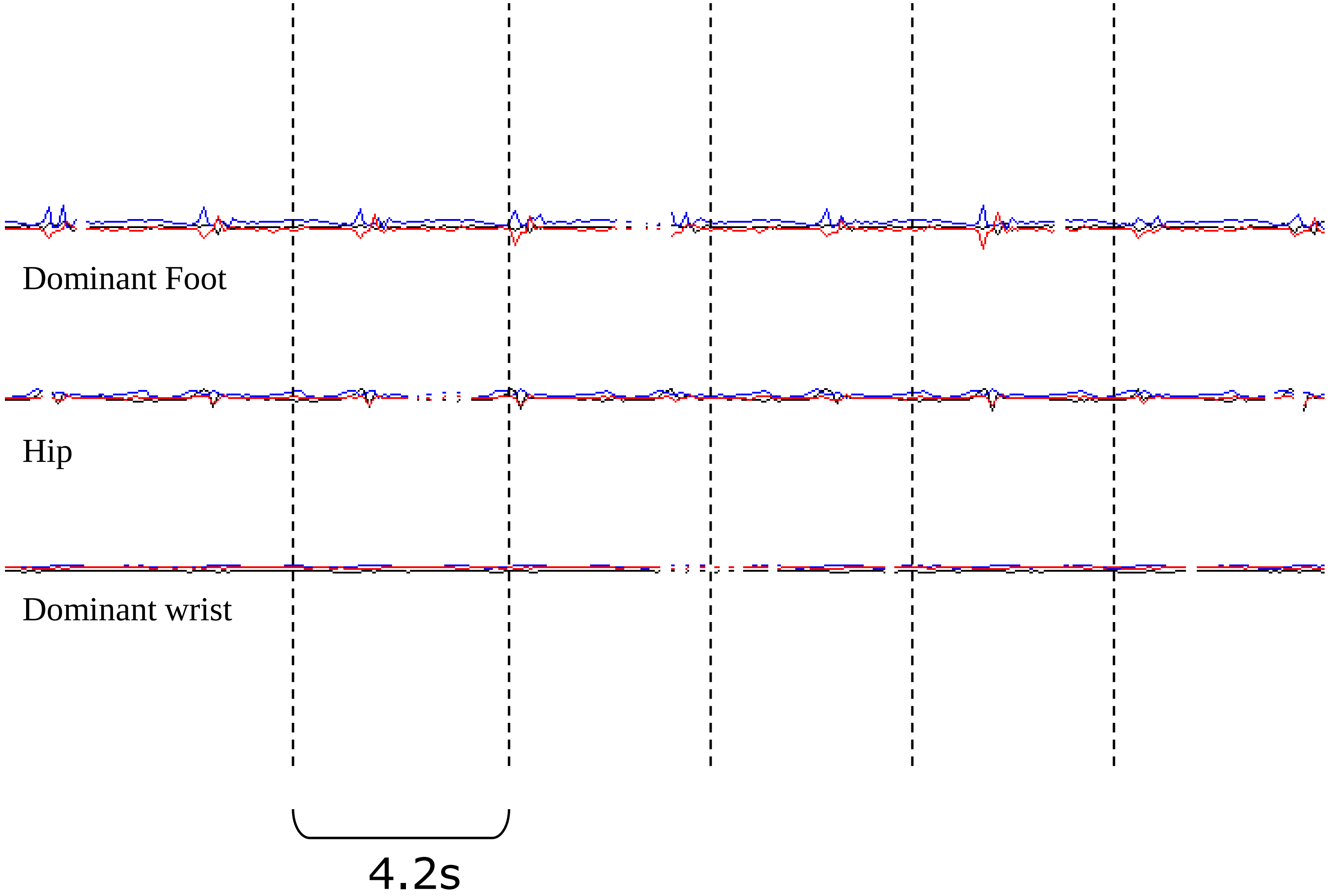
Walking treadmill 4mph 0%



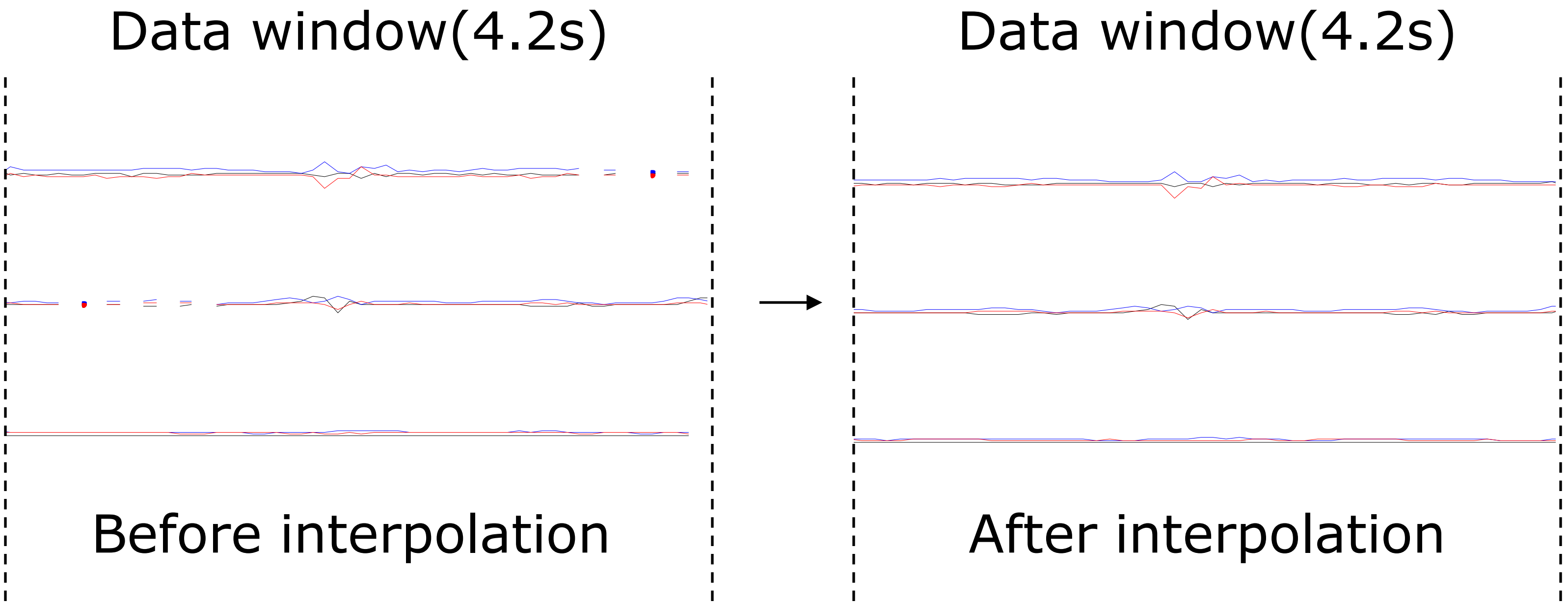
Walking treadmill 4mph 0%



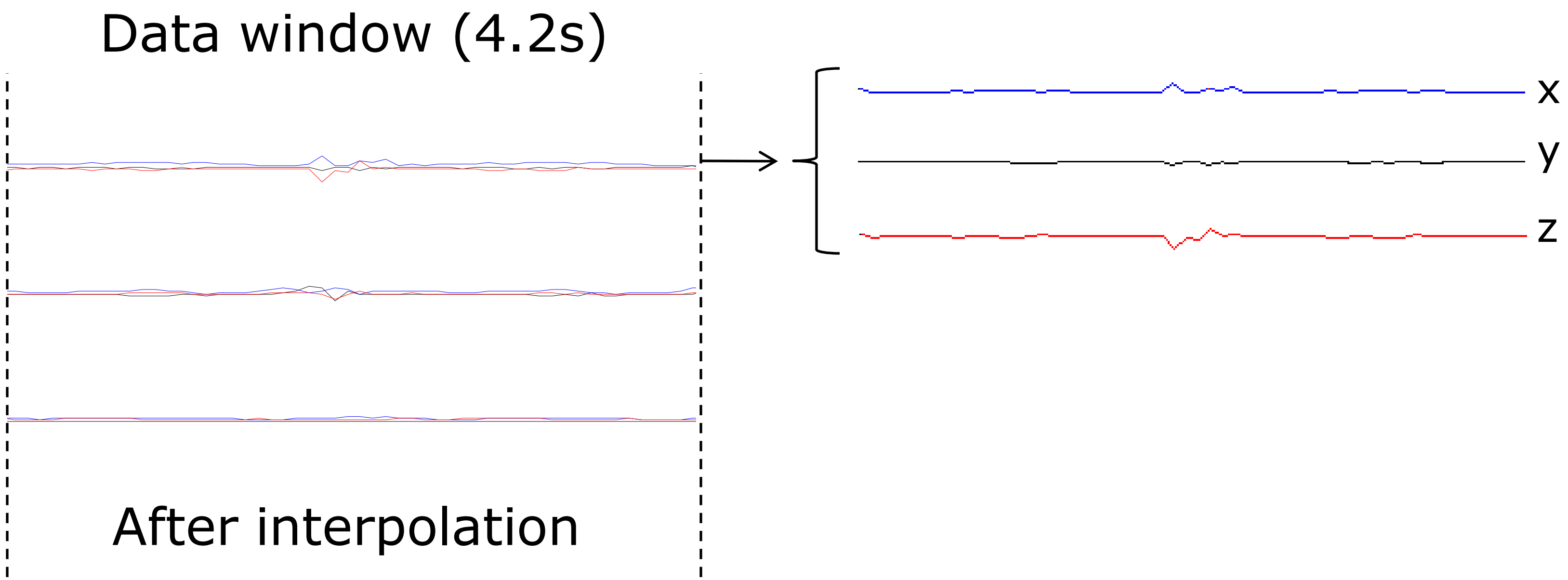
Segmentation: Sliding windows



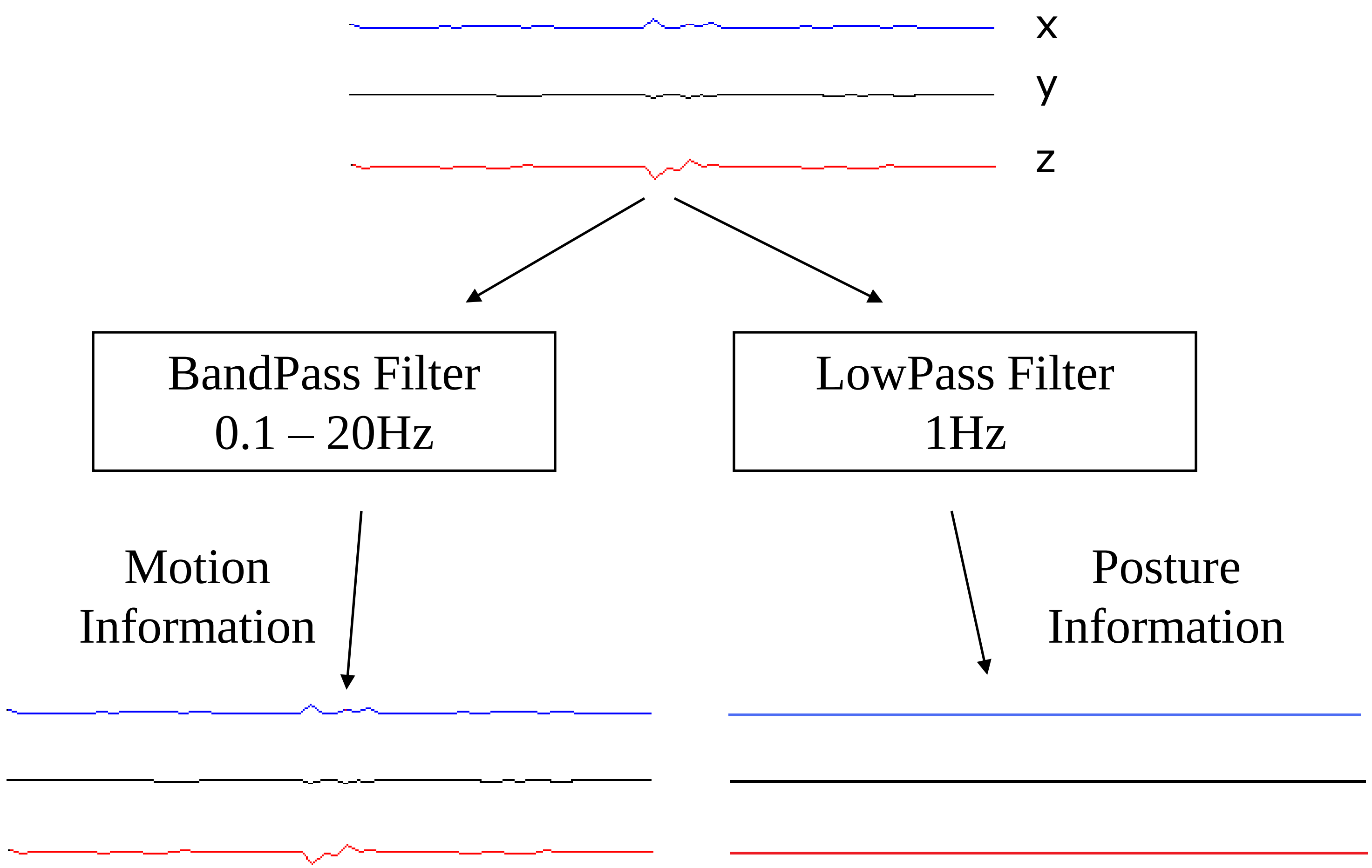
Interpolation: Cubic splines



Signal processing: Filtering



Signal processing: Filtering



Feature computation

For each of the 9 acceleration axis, compute the following features referred as *invariant reduced*

Signal variability

– Variance

Posture information

– Posture Distances

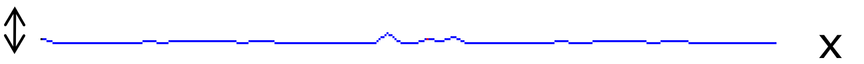
Activity intensity

– Energy between 0.3-3.5Hz

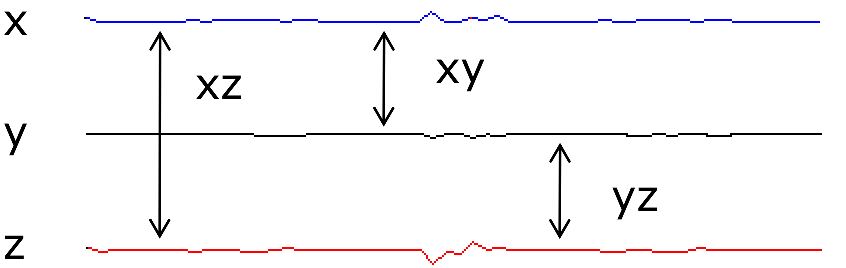
Frequency/periodicity of motion

– Top 5 peaks of the FFT

Time domain features

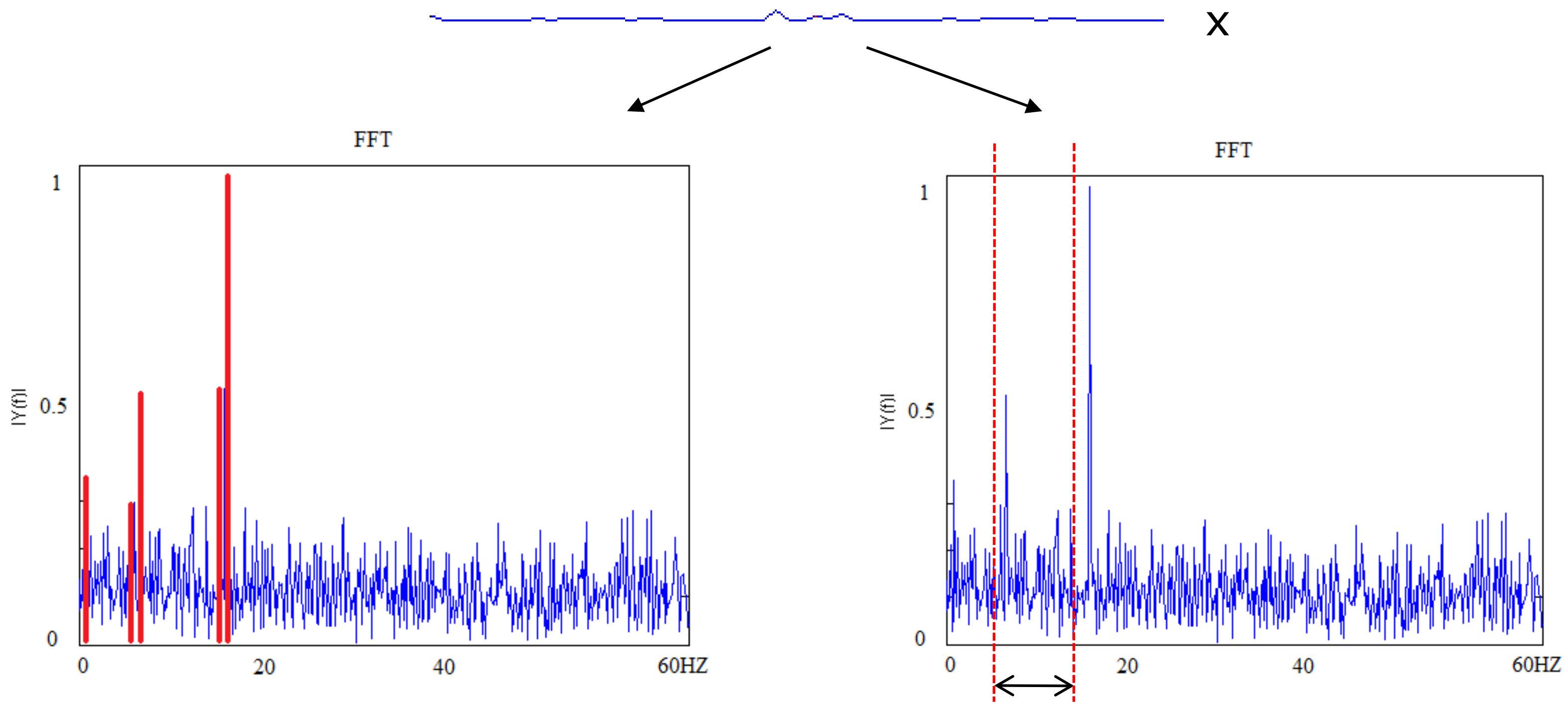


Variance



Posture distances

Frequency domain features



5 FFT peaks

Freq1,Mag1
Freq2,Mag2,
Freq3,Mag3,
Freq4,Mag4,
Freq5,Mag5

0.3 - 3.5Hz

Band energy

Training of classifier

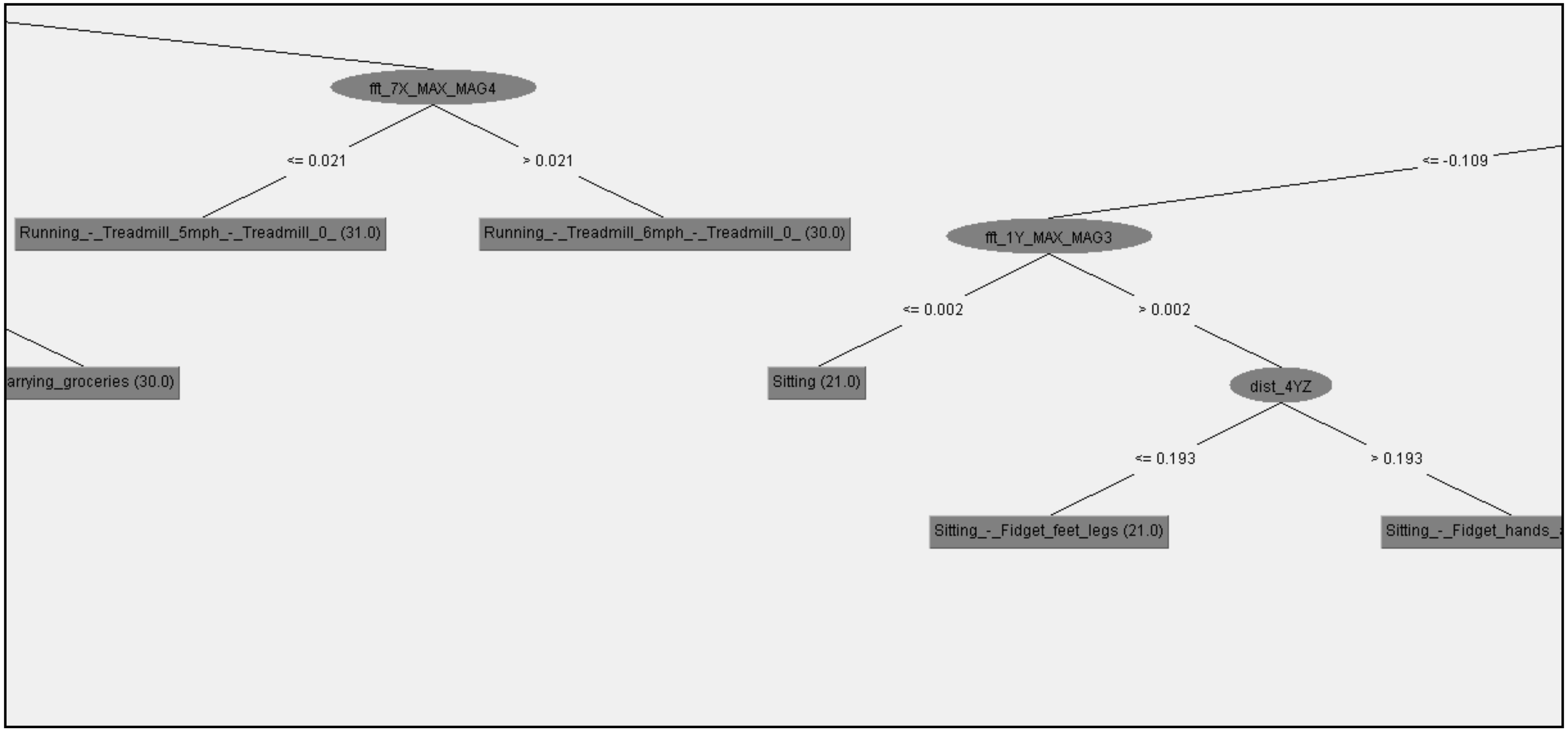
ACVar (9)
ACFFTPeaks (90)
ACBandEnergy (9)
DCPostureDist (9)

[val_1 val_2 … val_117]
Vector size: 117

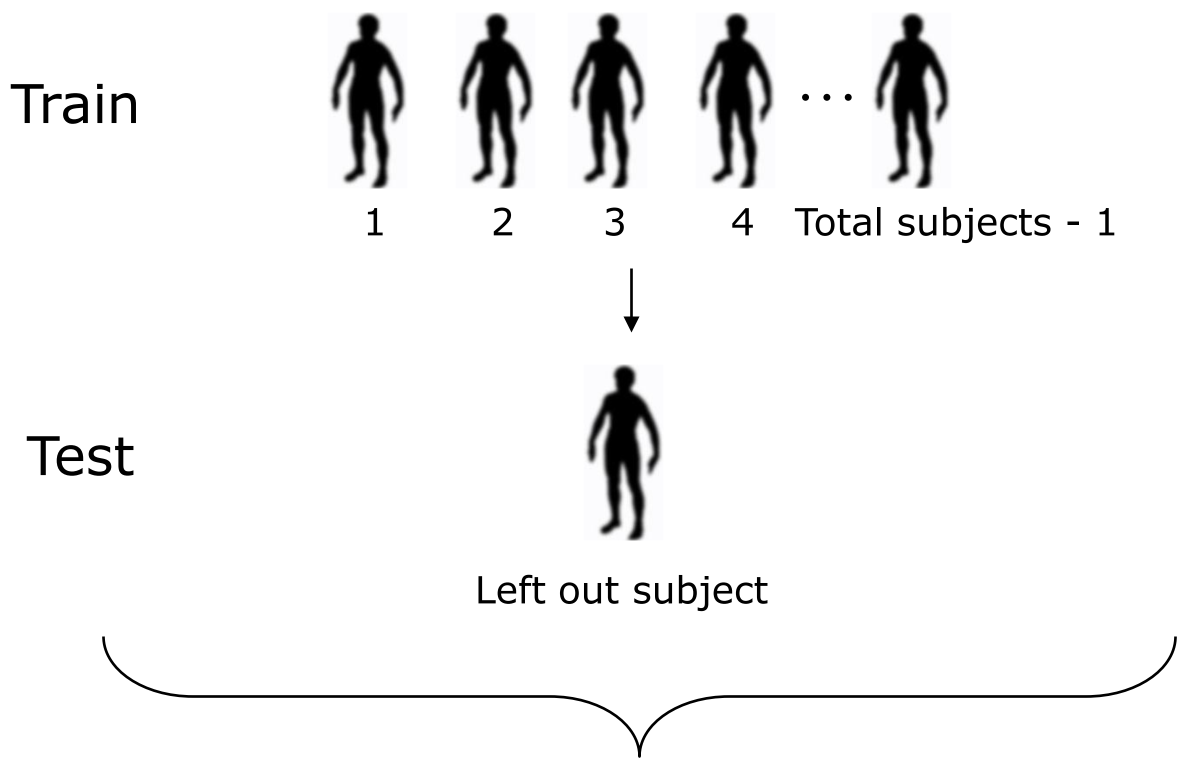
→

C4.5
Decision Tree
Classifier

C4.5 decision tree

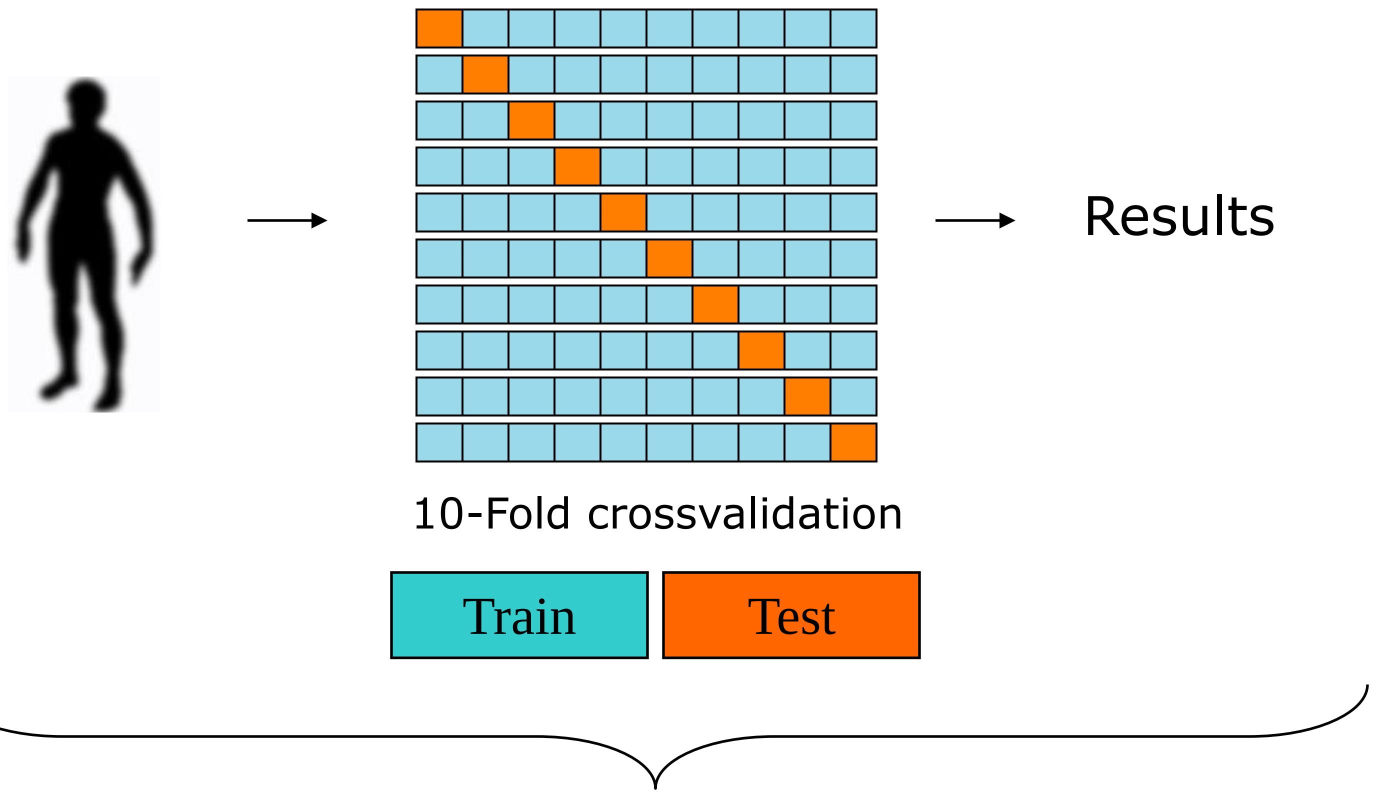


Subject independent evaluation



Repeat for as many subjects available
and average results

Subject dependent evaluation



Repeat for as many subjects we have
and average results

Target activities (52)

Type	Intensity
Lying down	Not applicable
Standing	Not applicable
Sitting	Not applicable
Sitting	Fidget feet legs
Sitting	Fidget hands arms
Kneeling	Not applicable
Walking	**Treadmill 2mph 0% grade**
Walking	**Treadmill 3mph 0% grade**
Walking	**Treadmill 3mph at 3% grade**
Walking	**Treadmill 3mph at 6% grade**
Walking	**Treadmill 3mph at 9% grade**
Running	Treadmill 4mph at 0% grade

Type	Intensity
Running	Treadmill 5mph at 0% grade
Running	Treadmill 6mph at 0% grade
Stairs Ascend stairs	Not applicable
Stairs Descend stairs	Not applicable
Cycling	80 rpm, light, moderate, hard
Cycling	60 rpm, light
Cycling	100 rpm, light
Rowing	30 spm, light, moderate, hard
Bicep curls	Light, moderate, hard
Bench weight lifting	Light, moderate, hard
Sit-ups	Not applicable
Crunches	Not applicable

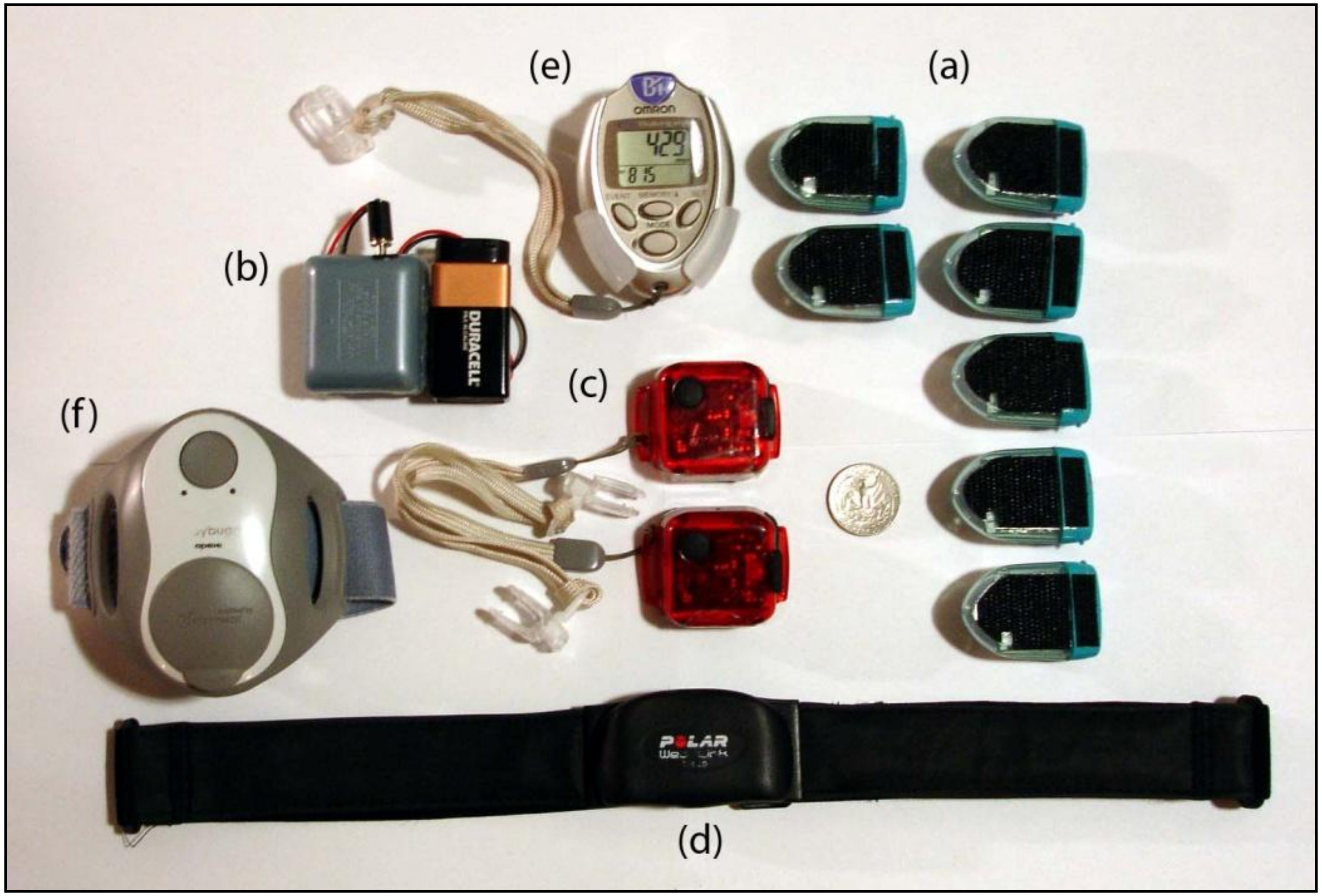
Gymnasium activity subset

Target activities (52)

Type	Type
Carrying groceries	Vacuuming
Doing dishes	Walking around block
Gardening	Washing windows
Ironing	Watching TV
Making the bed	Weeding
Mopping	Wiping/Dusting
Playing videogames	Writing
Scrubbing a surface	taking out trash
Stacking groceries	
Sweeping	
Typing	

Household activities subset

Sensing equipment



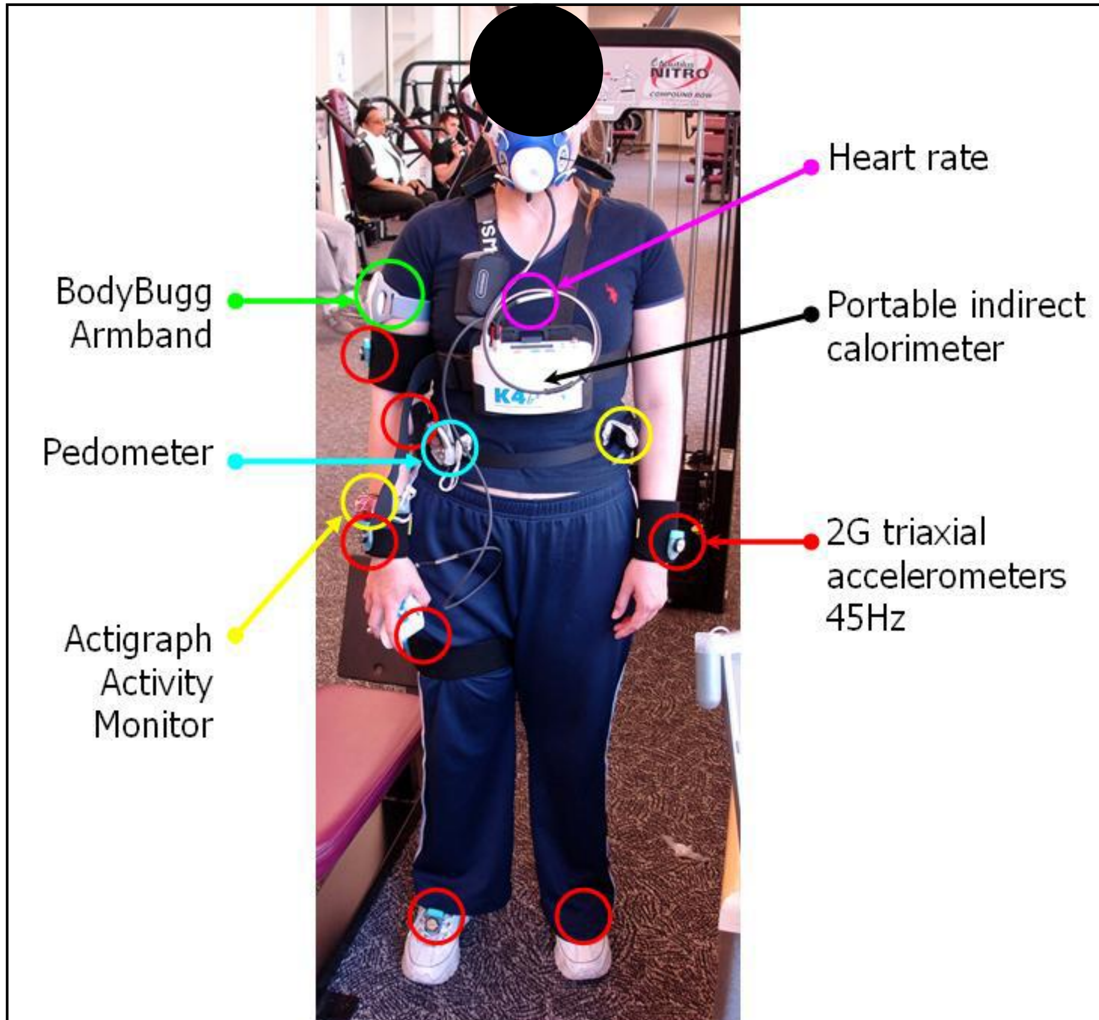
(a) Wireless accelerometers, (b) HR transceiver, (c) Actigraphs, (d) HR monitor, (e) pedometer, (f) Bodybugg armband

Sensor placement

20 Subjects
18 and 42 years old

Start/end times of activities annotated

3-4min per activity except physically demanding activities ~1min



Final system design

- Acceleration only
- Three sensors: hip, wrist, foot out of seven explored
- Computes features over each axis
- Feature set: minimizes dependency on sensor placement
- Classifier C4.5 classifier
- Sliding windows of 5.6s in length

Performance: 52 activities

Evaluation Method	Accuracy (%)	TP Range (%)	FP Range (%)
Subject dependent	87.9	80 - 93	0.1 – 0.2
Subject independent	50.6	34 – 77	0.5 – 1.3
Percent change	73%		

Random guess: 1.96%

Performance: 52 activities

Evaluation Method	Accuracy (%)	TP Range (%)	FP Range (%)
Subject dependent	87.9	80 - 93	0.1 – 0.2
Subject independent	50.6	34 – 77	0.5 – 1.3
Percent change	$\boxed{73\%}$		

Random guess: 1.96%

Performance: 52 activities

- Higher performance: Postures and exercises
- Lowest performance: household and resistance activities
- Confused:
  - Intensity levels
  - Household
  - Household with postures and ambulation
  - Activities involving upper body motion

How much training data?

Subject dependent performance

- Training: 75%, Testing: 25%
- Varied training from 75% to 7.5%
- 75% training data: Accuracy= 80.6%
- 60% training data: Accuracy= 76%

At 60% of data: 2min for most activities, 1min for physically demanding activities

Performance: Activity subsets

Activities to recognize	**Total**	**Activities Included**
All	51	All 51 activities
All with no intensities	31	No intensity levels for Bicep curls, bench weight lifting, walking, running, cycling, rowing, and sitting
Postures, ambulation and two MET intensity categories	11	Lying down, sitting, standing, kneeling, walking (2, 3mph), running (4,5, and 6 mph), moderate, vigorous
Postures and Ambulation with no intensity	8	Lying down, sitting, standing, kneeling, walking, running, ascending stairs, descending stairs
Postures	4	Lying down, sitting, standing, kneeling

Performance: Activity subsets

		Subject Dependent	Subject Independent
Activities to recognize	Random Guess (%)	Total Accuracy (%)	Total Accuracy (%)
All (51)	1.9%	87.9	50.6
All with no intensities (31)	3.2%	91.4	72.0
Postures, ambulation and two MET intensity categories (11)	9%	96.5	81.3
Postures and Ambulation with no intensity (8)	12.5%	98.4	92.9
Postures (4)	25%	99.3	98.0

Performance: Activity subsets

		Subject Dependent	Subject Independent
Activities to recognize	**Random Guess (%)**	**Total Accuracy (%)**	**Total Accuracy (%)**
All (51)	1.9%	87.9	50.6
All with no intensities (31)	3.2%	91.4	72.0
Postures, ambulation and two MET intensity categories (11)	9%	96.5	81.3
Postures and Ambulation with no intensity (8)	12.5%	98.4	92.9
Postures (4)	25%	99.3	98.0

If activity intensities are merged,
SI accuracy =72%

Performance: Activity subsets

		Subject Dependent	Subject Independent
Activities to recognize	**Random Guess (%)**	**Total Accuracy (%)**	**Total Accuracy (%)**
All (51)	1.9%	87.9	50.6
All with no intensities (31)	3.2%	91.4	72.0
Postures, ambulation and two MET intensity categories (11)	9%	96.5	81.3
Postures and Ambulation with no intensity (8)	12.5%	98.4	92.9
Postures (4)	25%	99.3	98.0

Performance: Activity subsets

		Subject Dependent	Subject Independent
Activities to recognize	**Random Guess (%)**	**Total Accuracy (%)**	**Total Accuracy (%)**
All (51)	1.9%	87.9	50.6
All with no intensities (31)	3.2%	91.4	72.0
Postures, ambulation and two MET intensity categories (11)	9%	96.5	81.3
Postures and Ambulation with no intensity (8)	12.5%	98.4	92.9
Postures (4)	25%	99.3	98.0

Performance: Activity subsets

		Subject Dependent	Subject Independent
Activities to recognize	**Random Guess (%)**	**Total Accuracy (%)**	**Total Accuracy (%)**
All (51)	1.9%	87.9	50.6
All with no intensities (31)	3.2%	91.4	72.0
Postures, ambulation and two MET intensity categories (11)	9%	96.5	81.3
Postures and Ambulation with no intensity (8)	12.5%	98.4	92.9
Postures (4)	25%	99.3	98.0

52 activities: Sensor subsets

Sensor Combination	Subject Dependent Accuracy
All sensors	87.9 ± 2.0
Hip + DWrist + DFoot	-1.8 %
Hip + DFoot	-3.5 %
Hip + DWrist	-4.9 %
DWrist + DThigh	-7.2 %
DWrist + DFoot	-7.7 %
Hip	-8.4 %
DFoot	-14.9 %
DThigh	-15.1 %
DUpperArm	-15.4 %
DWrist	-19.6 %

Sensor Combination	Subject Independent Accuracy
All sensors	50.6 ± 5.2
Hip + DWrist + DFoot	-7.9 %
DWrist + DThigh	-8.1 %
DWrist + DFoot	-13.0 %
Hip + DWrist	-15.6 %
Hip + DFoot	-18.9 %
DUpperArm	-26.5 %
DWrist	-27.7 %
Hip	-28.5 %
DFoot	-34.8 %
DThigh	-42.7 %

52 activities: Sensor subsets

Sensor Combination	Subject Dependent Accuracy
All sensors	87.9 ± 2.0
Hip + DWrist + DFoot	-1.8 %
Hip + DFoot	-3.5 %
Hip + DWrist	-4.9 %
DWrist + DThigh	-7.2 %
DWrist + DFoot	-7.7 %
Hip	-8.4 %
DFoot	-14.9 %
DThigh	-15.1 %
DUpperArm	-15.4 %
DWrist	-19.6 %

Sensor Combination	Subject Independent Accuracy
All sensors	50.6 ± 5.2
Hip + DWrist + DFoot	-7.9 %
DWrist + DThigh	-8.1 %
DWrist + DFoot	-13.0 %
Hip + DWrist	-15.6 %
Hip + DFoot	-18.9 %
DUpperArm	-26.5 %
DWrist	-27.7 %
Hip	-28.5 %
DFoot	-34.8 %
DThigh	-42.7 %

52 activities: Sensor subsets

Sensor Combination	Subject Dependent Accuracy
All sensors	87.9 ± 2.0
Hip + DWrist + DFoot	-1.8 %
Hip + DFoot	-3.5 %
Hip + DWrist	-4.9 %
DWrist + DThigh	-7.2 %
DWrist + DFoot	-7.7 %
Hip	-8.4 %
DFoot	-14.9 %
DThigh	-15.1 %
DUpperArm	-15.4 %
DWrist	-19.6 %

Sensor Combination	Subject Independent Accuracy
All sensors	50.6 ± 5.2
Hip + DWrist + DFoot	-7.9 %
DWrist + DThigh	-8.1 %
DWrist + DFoot	-13.0 %
Hip + DWrist	-15.6 %
Hip + DFoot	-18.9 %
DUpperArm	26.5 %
DWrist	-27.7 %
Hip	-28.5 %
DFoot	-34.8 %
DThigh	-42.7 %

52 activities: Sensor subsets

Sensor Combination	Subject Dependent Accuracy
All sensors	87.9 ± 2.0
Hip + DWrist + DFoot	-1.8 %
Hip + DFoot	-3.5 %
Hip + DWrist	-4.9 %
DWrist + DThigh	-7.2 %
DWrist + DFoot	-7.7 %
Hip	-8.4 %
DFoot	-14.9 %
DThigh	-15.1 %
DUpperArm	-15.4 %
DWrist	-19.6 %

Sensor Combination	Subject Independent Accuracy
All sensors	50.6 ± 5.2
Hip + DWrist + DFoot	-7.9 %
DWrist + DThigh	-8.1 %
DWrist + DFoot	-13.0 %
Hip + DWrist	-15.6 %
Hip + DFoot	-18.9 %
DUpperArm	26.5 %
DWrist	-27.7 %
Hip	-28.5 %
DFoot	-34.8 %
DThigh	-42.7 %

52 activities: Sensor subsets

Sensor Combination	**Subject Dependent Accuracy**
All sensors	87.9 ± 2.0
Hip + DWrist + DFoot	-1.8 %
Hip + DFoot	-3.5 %
Hip + DWrist	-4.9 %
DWrist + DThigh	-7.2 %
DWrist + DFoot	-7.7 %
Hip	-8.4 %
DFoot	-14.9 %
DThigh	-15.1 %
DUpperArm	-15.4 %
DWrist	-19.6 %

Sensor Combination	**Subject Independent Accuracy**
All sensors	50.6 ± 5.2
Hip + DWrist + DFoot	-7.9 %
DWrist + DThigh	-8.1 %
DWrist + DFoot	-13.0 %
Hip + DWrist	-15.6 %
Hip + DFoot	-18.9 %
DUpperArm	26.5 %
DWrist	-27.7 %
Hip	-28.5 %
DFoot	-34.8 %
DThigh	-42.7 %

Why 5.6s sliding windows?

Measured performance while varying window length from 1.4s to 91s

- Performance increases with longer windows
  - Improvement ~5% from 5.6s to 45s.
- Window length depends on activity type but this is computationally expensive
  - Long windows for household activities (e.g. 22-45s)
  - Short windows for postures (e.g. ≤ 5.6s)
- Long windows: low performance over short duration activities and long real-time delays.

Why not combine HR+ACC data?

	Subject Independent Evaluation		
Features subsets	**Accuracy** (%)	**TP Rate (%)**	**FP Rate (%)**
ScaledHR	13.8	4 – 16	1.6 – 2.3
Invariant Reduced	50	34 - 77	0.5 – 1.3
Invariant Reduced + ScaledHR	52	37 - 76	0.5 – 1.3

	Subject Dependent Evaluation		
Features subsets	**Accuracy** (%)	**TP Rate (%)**	**FP Rate (%)**
ScaledHR	38.4	24 – 39	1.4 – 1.6
Invariant Reduced	88	80 - 97	0.1 – 0.4
Invariant Reduced + ScaledHR	89.5	82 – 97	0.1 – 0.4

Why not combine HR+ACC data?

	Subject Independent Evaluation		
Features subsets	**Accuracy** (%)	**TP Rate (%)**	**FP Rate (%)**
ScaledHR	13.8	4 – 16	1.6 – 2.3
Invariant Reduced	50	34 - 77	0.5 – 1.3
Invariant Reduced + ScaledHR	52	37 - 76	0.5 – 1.3

	Subject Dependent Evaluation		
Features subsets	**Accuracy** (%)	**TP Rate (%)**	**FP Rate (%)**
ScaledHR	38.4	24 – 39	1.4 – 1.6
Invariant Reduced	88	80 - 97	0.1 – 0.4
Invariant Reduced + ScaledHR	89.5	82 – 97	0.1 – 0.4

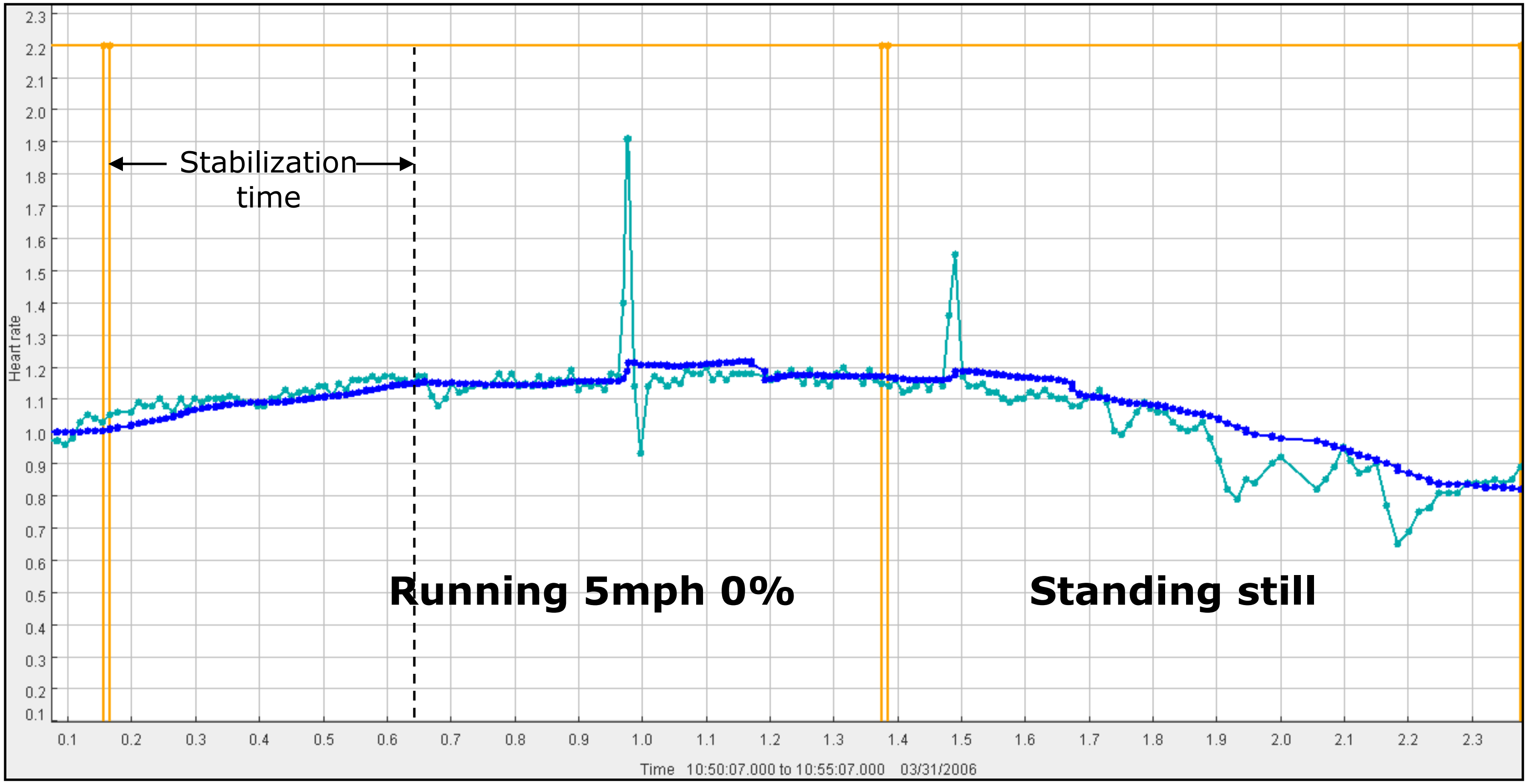
Why not combine HR+ACC data?

	Subject Independent Evaluation		
Features subsets	**Accuracy** (%)	**TP Rate (%)**	**FP Rate (%)**
ScaledHR	13.8	4 – 16	1.6 – 2.3
Invariant Reduced	50	34 - 77	0.5 – 1.3
Invariant Reduced + ScaledHR	52	37 - 76	0.5 – 1.3

	Subject Dependent Evaluation		
Features subsets	**Accuracy** (%)	**TP Rate (%)**	**FP Rate (%)**
ScaledHR	38.4	24 – 39	1.4 – 1.6
Invariant Reduced	88	80 - 97	0.1 – 0.4
Invariant Reduced + ScaledHR	89.5	82 – 97	0.1 – 0.4

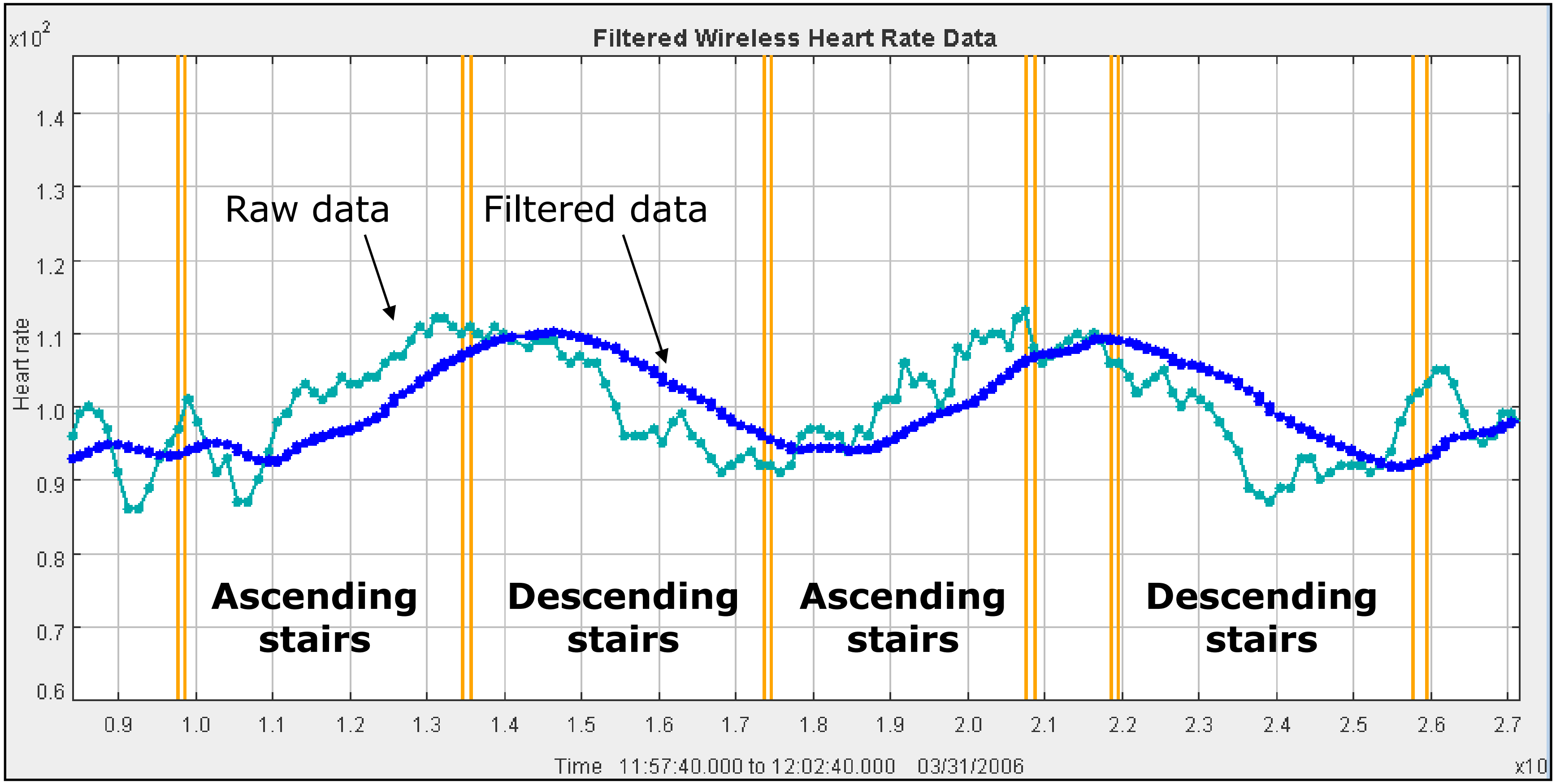
Percent change = 2-4%

Why such a low improvement?



Heart rate lags physical activity and remains altered once activity has ended. Thus, errors concentrated at start - end

Why such a low improvement?



Errors also occur for activities where heart rate constantly increases or decreases over time (e.g. physically demanding)

Real-time pilot study

Five participants were asked to:

- Wear 3 accelerometers and
- Type in 10 physical activities, exercises, postures, or activities of their choice
- Perform activities provided continuously for 2 minutes.

Real-time pilot study

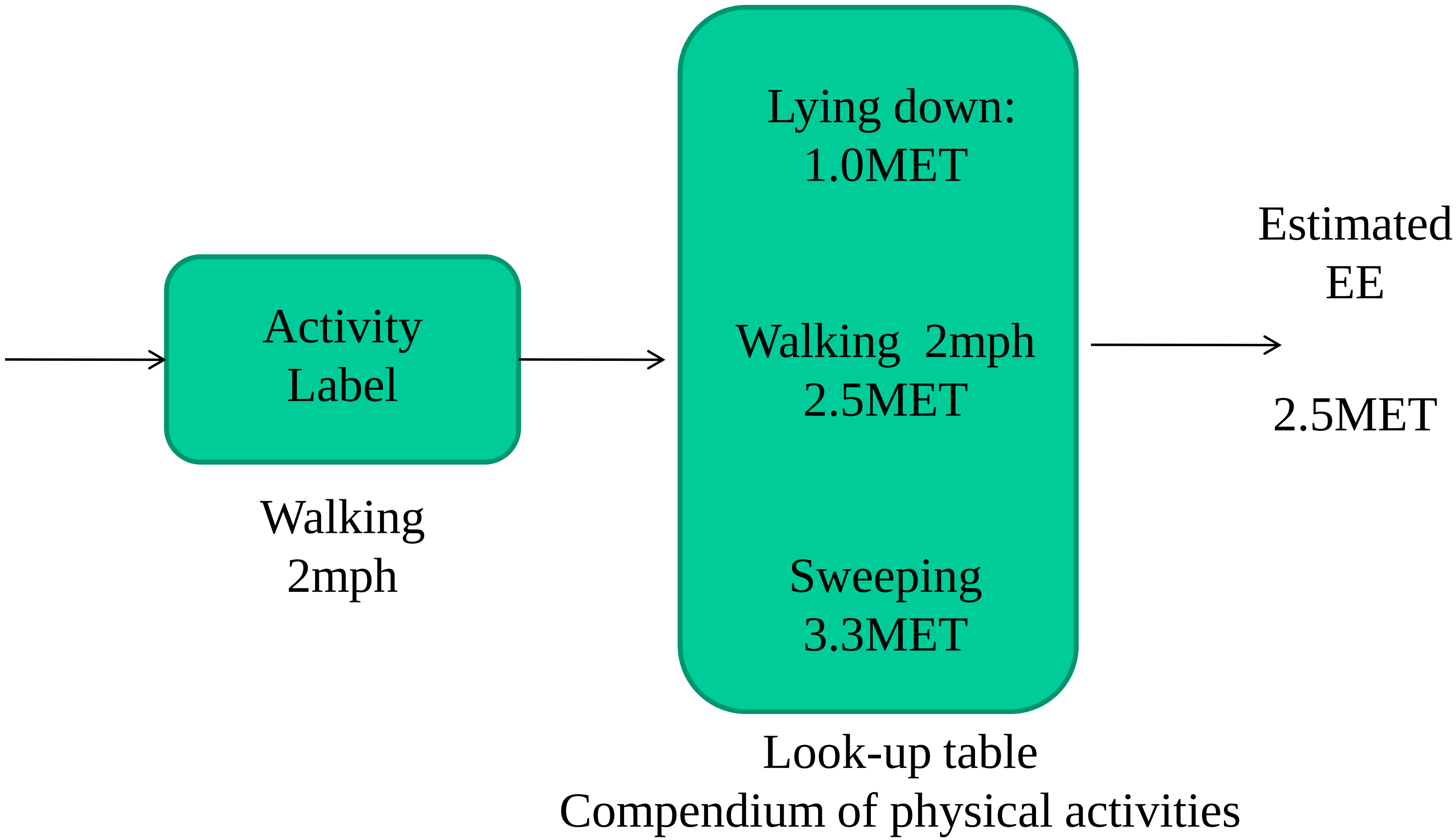
Subject	Activities performed		Total Accuracy (%)	True Positive Range (%)	False Positive Range (%)
1	Bouncing on a ball Waving hand Shaking my leg **Taekwondo Form #1** Side stretch	Jumping jacks Punching as I walk forward Lifting dumbbells Riding a bike Playing the drums	89.6	89.3 – 94.8	0.8 - 1.0
2	Walking Sitting still **Scratching head** Carrying box Washing dishes	Shaking hands Tossing ball in air Typing Talking on phone	91.7	84.5 – 98.2	0.4 – 0.17
3	Throwing **Bowling** Bouncing Typing Stepping	Stretching arm Walking **Tennis serve** Stretching legs Bending	78.9	70.7 – 93.2	1.3 – 3.8
4	Walk Type in computer Washing window Drawing in paper Wiping surface	Talking on the phone Sweeping Combing my hair **Hammering a nail** Eating	89.3	74.1 – 94.8	0.6 – 2.1
5	Walk Bicep curls Stretching **Applying cream** Brushing teeth	Wash dish **Knitting** Wash hands **Filing nails** Play piano	85.2	77.6 – 94.8	0.6 – 2.7

Energy Expenditure Algorithm Experiments

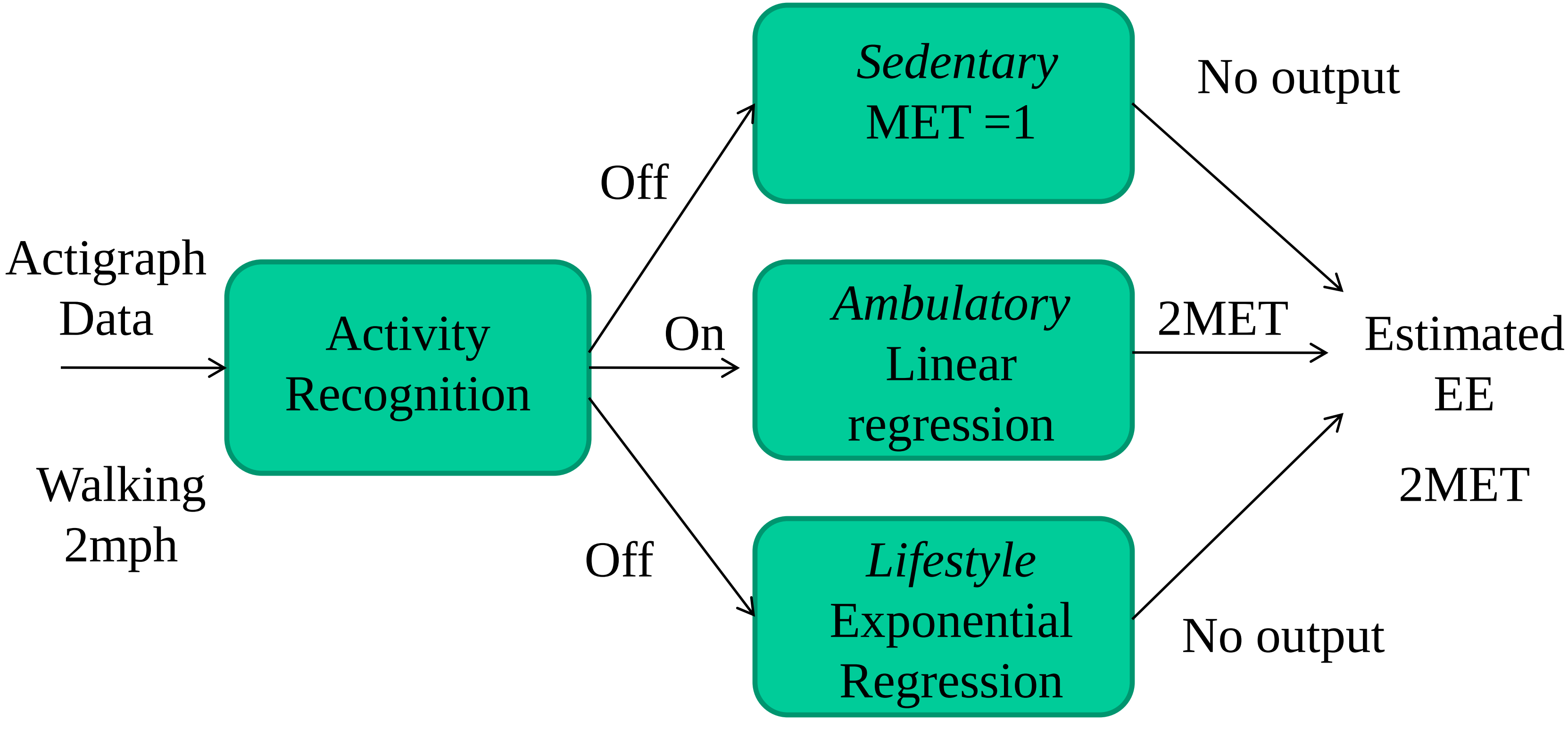
Accelerometer at the hip

- Hip accelerometer
- 1min windows
- Feature: overall motion
- Linear regression
- Predict EE in
  - METs
  - kcal/min
  - kJ/Min.

Compendium of physical activities



Crouter et al. 2007



17 activities, 20 subjects, 3hours/subject
r = 0.96, RMSE=0.73MET, MAED=0.75MET

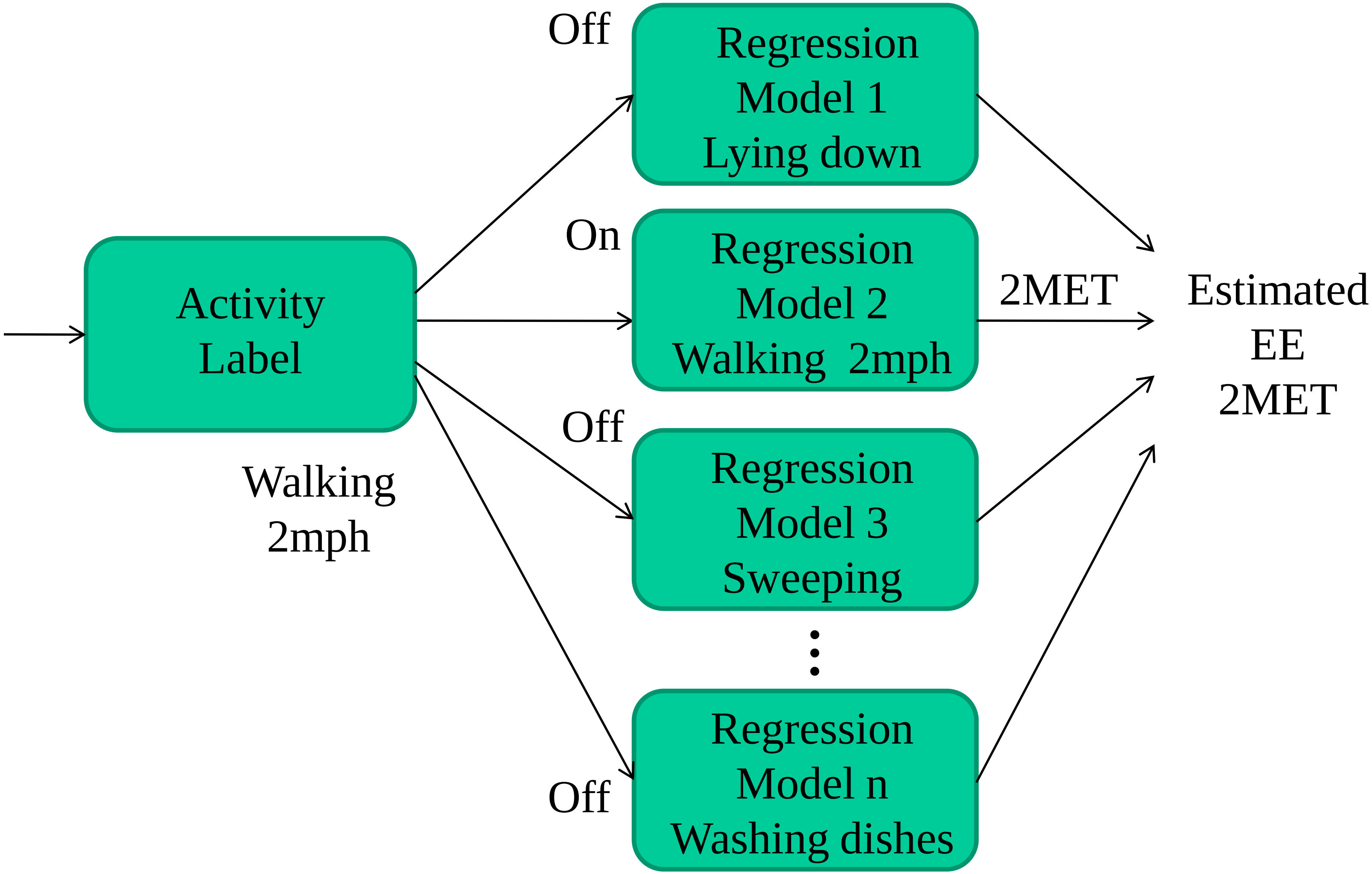
Crouter et al. 2007

EE in this work

This thesis extends the work of Crouter et al. by:

- Exploring the use of 51 activity dependent regression models
- The utilization of 7 accelerometers
- The exploration of 41 features
- The use of shorter window lengths
- The use of linear and non-linear regression models

Activity dependent regression



EE estimation assumptions

- Predicting EE in METs
  - 1MET = EE while lying down
  - METs normalize EE with respect to body mass
- Gross EE prediction
  - Gross=resting + motion energy expenditure
- Non-steady state EE is not eliminated
  - Might be difficult to reach during free-living
  - More realistic evaluation

MIT EE dataset

Reduced version of data collected

- Removed sessions containing any activity with low EE values (< 40%)
- Poor mask attachment
- 13 out of 40 sessions removed
- 15 gym and 12 household sessions

16 Participants

- men=7, woman=9
- 18-40 years old, Body mass 60-103kg

Evaluation measures

- **Correlation coefficient (r)**

$$r = [0,1]$$

- **Root mean squared error (RMSE)**

$$\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(p_i - a_i\right)^2}$$

Baseline EE experiments

Error Measures	Crouter et al. Actigraph	Compendium Comparable Activities (29 activities)	Compendium Closest Activities (52 activities)	Linear Regression	One Linear Regression model per activity	One non-linear regression model per activity
Total Correlation Coefficient	0.4	0.9 (125%)	0.8 (100%)	0.73 (82%)	0.87 (117%)	0.91 (127%)
Total root Mean Square Error	2.7	1.27 (-53%)	1.6 (-41%)	1.4 (-48%)	1.0 (-63%)	0.88 (-67.4%)
Maximum absolute Deviation	6.9	4.17	5.6	4.1	4.2	3.4

Performance over 52 activities using the *ACAbsArea feature* computed per sensor over one-minute sliding windows.

Baseline EE results

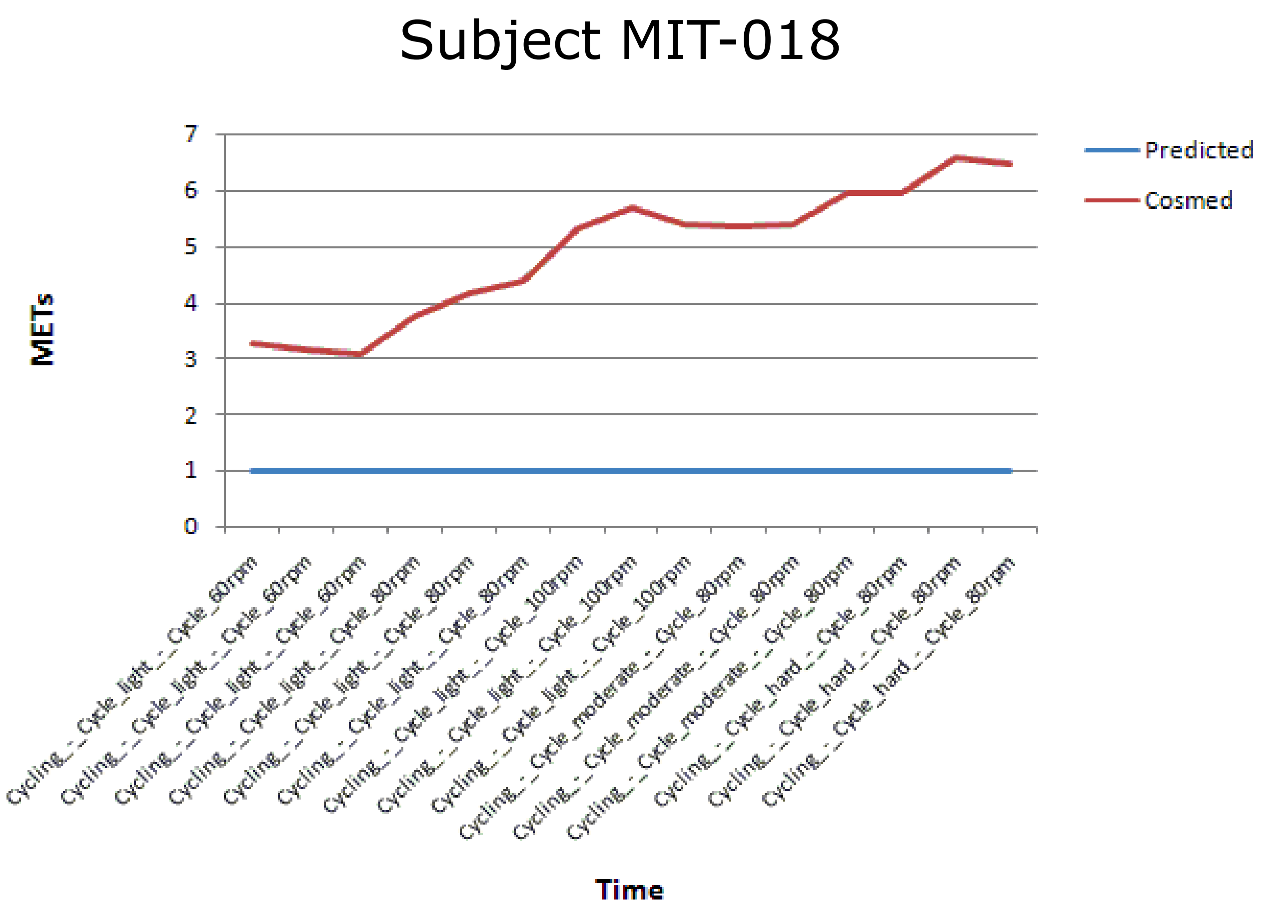
Error Measures	**Crouter et al. Actigraph**	Compendium Comparable Activities (29 activities)	Compendium Closest Activities (52 activities)	Linear Regression	One Linear Regression model per activity	One non-linear regression model per activity
Total Correlation Coefficient	0.4	0.9 (125%)	0.8 (100%)	0.73 (82%)	0.87 (117%)	0.91 (127%)
Total root Mean Square Error	2.7	1.27 (-53%)	1.6 (-41%)	1.4 (-48%)	1.0 (-63%)	0.88 (-67.4%)
Maximum absolute Deviation	6.9	4.17	5.6	4.1	4.2	3.4

- Performance lower than the obtained over 17 activities

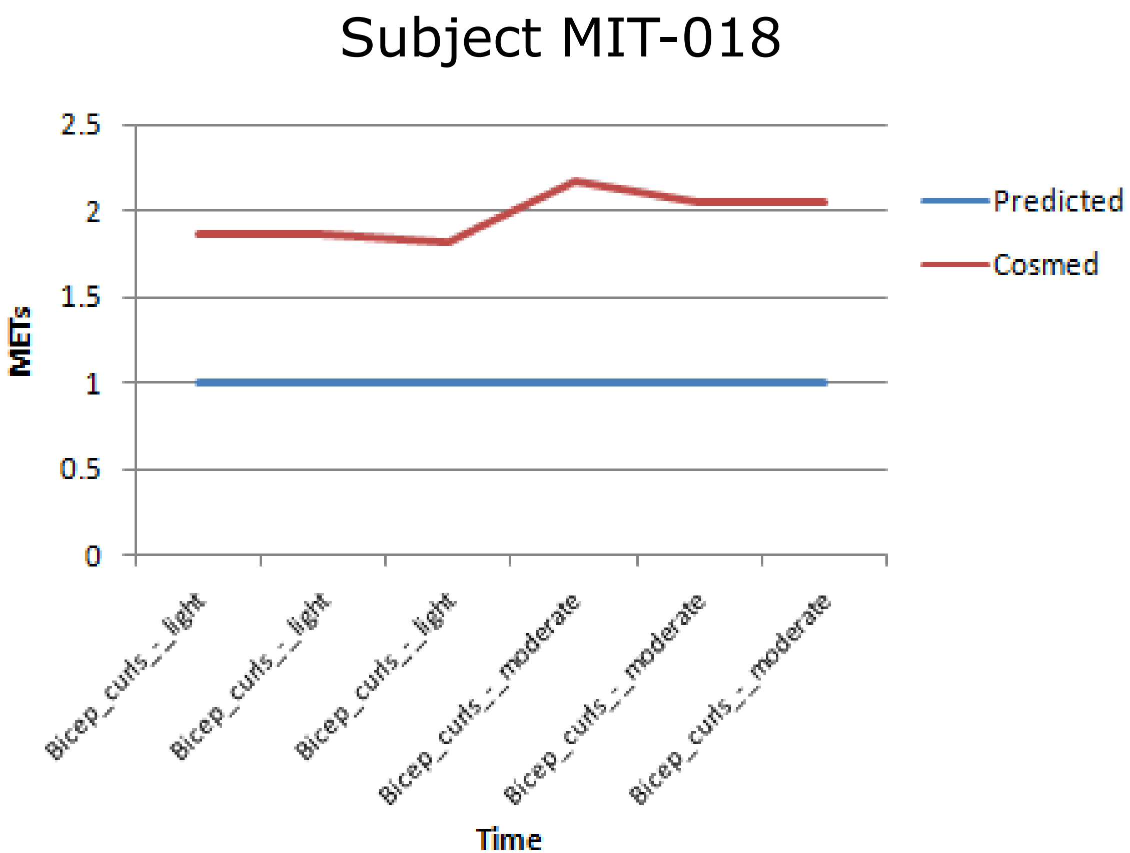
r = 0.92, RMSE=0.73MET

- High maximum error deviation!

Crouter: Lower body activity



Crouter: Upper body activity



Baseline EE results

Error Measures	**Crouter et al. Actigraph**	**Compendium Comparable Activities (29 activities)**	**Compendium Closest Activities (52 activities)**	**Linear Regression**	**One Linear Regression model per activity**	**One non-linear regression model per activity**
Total Correlation Coefficient	0.4	0.9 (125%)	0.8 (100%)	0.73 (82%)	0.87 (117%)	0.91 (127%)
Total root Mean Square Error	2.7	1.27 (-53%)	1.6 (-41%)	1.4 (-48%)	1.0 (-63%)	0.88 (-67.4%)
Maximum absolute Deviation	6.9	4.17	5.6	4.1	4.2	3.4

- This result indicates that knowledge of the activity being performed is important.
- Performance depends on mean EE listings in Compendium

Baseline EE results

Error Measures	Crouter et al. Actigraph	Compendium Comparable Activities (29 activities)	Compendium Closest Activities (52 activities)	Linear Regression	One Linear Regression model per activity	One non-linear regression model per activity
Total Correlation Coefficient	0.4	0.9 (125%)	0.8 (100%)	0.73 (82%)	0.87 (117%)	0.91 (127%)
Total root Mean Square Error	2.7	1.27 (-53%)	1.6 (-41%)	1.4 (-48%)	1.0 (-63%)	0.88 (-67.4%)
Maximum absolute Deviation	6.9	4.17	5.6	4.1	4.2	3.4

Performance improves over Crouter's mainly due to the use of six additional accelerometers.

Baseline EE results

Error Measures	Crouter et al. Actigraph	Compendium Comparable Activities (29 activities)	Compendium Closest Activities (52 activities)	Linear Regression	One Linear Regression model per activity	One non-linear regression model per activity
Total Correlation Coefficient	0.4	0.9 (125%)	0.8 (100%)	0.73 (82%)	0.87 (117%)	0.91 (127%)
Total root Mean Square Error	2.7	1.27 (-53%)	1.6 (-41%)	1.4 (-48%)	1.0 (-63%)	0.88 (-67.4%)
Maximum absolute Deviation	6.9	4.17	5.6	4.1	4.2	3.4

- Improvement over single linear regression model:

  R=19%, RMSE=-28.5%

- Activity dependent models help by allowing regression coefficients to be tuned for each activity

Baseline EE results

Error Measures	Crouter et al. Actigraph	Compendium Comparable Activities (29 activities)	Compendium Closest Activities (52 activities)	Linear Regression	One Linear Regression model per activity	One non-linear regression model per activity
Total Correlation Coefficient	0.4	0.9 (125%)	0.8 (100%)	0.73 (82%)	0.87 (117%)	0.91 (127%)
Total root Mean Square Error	2.7	1.27 (-53%)	1.6 (-41%)	1.4 (-48%)	1.0 (-63%)	0.88 (-67.4%)
Maximum absolute Deviation	6.9	4.17	5.6	4.1	4.2	3.4

- Improvement over activity-dependent linear regression:
  r=5%, RMSE=-12%
- Improvement over single linear model is:
  r=25%, RMSE=-37%

Performance per activity

- Weakest: activities with resistance/load
- Best: postures and household activities
- A single linear regression using 7sensors
  - Overestimates EE for postures
  - Predicts EE well for lower and upper body activities
- The Compendium of Physical activates
  - Overestimates EE for household activities and short duration short duration activities
  - Estimates EE better for activities that reached steady-state EE

Summary of results

EE estimation is improved by

- Accelerometers at upper and lower body
- Activity-dependent regression models

Questions to answer

- Fewer accelerometers?
- Performance of activity recognition algorithm?
- Heart rate data?

For full detail see thesis!

Final system design

The final EE estimation algorithm uses the following parameters:

- Only accelerometer data
- Three accelerometers: Hip, dominant wrist, and dominant foot.
- Feature: Top 5 FFT peaks per sensor.
- 5.6s sliding windows.

Sensor combinations

Sensor Combination	Correlation	RMSE
All sensors	0.71	1.28
Hip + DWrist + DFoot	-2.8%	+2.3%
DWrist + DFoot	-2.8%	+3.0%
Hip + DFoot	-4.2%	+3.0%
DWrist + DThigh	-12.7%	+5.2%
Hip + DWrist	-5.6%	+12.3%
DFoot	-8.5%	+5.2%
DThigh	-11.3%	+9.2%
DUpperArm	-15.5%	+11.1%
Hip	-33.8%	+13.5%
DWrist	-2.8%	+21.5%

Sensor combinations

Sensor Combination	Correlation	RMSE
All sensors	0.71	1.28
Hip + DWrist + DFoot	-2.8%	+2.3%
DWrist + DFoot	-2.8%	+3.0%
Hip + DFoot	-4.2%	+3.0%
DWrist + DThigh	-12.7%	+5.2%
Hip + DWrist	-5.6%	+12.3%
DFoot	-8.5%	+5.2%
DThigh	-11.3%	+9.2%
DUpperArm	-15.5%	+11.1%
Hip	-33.8%	+13.5%
DWrist	-2.8%	+21.5%

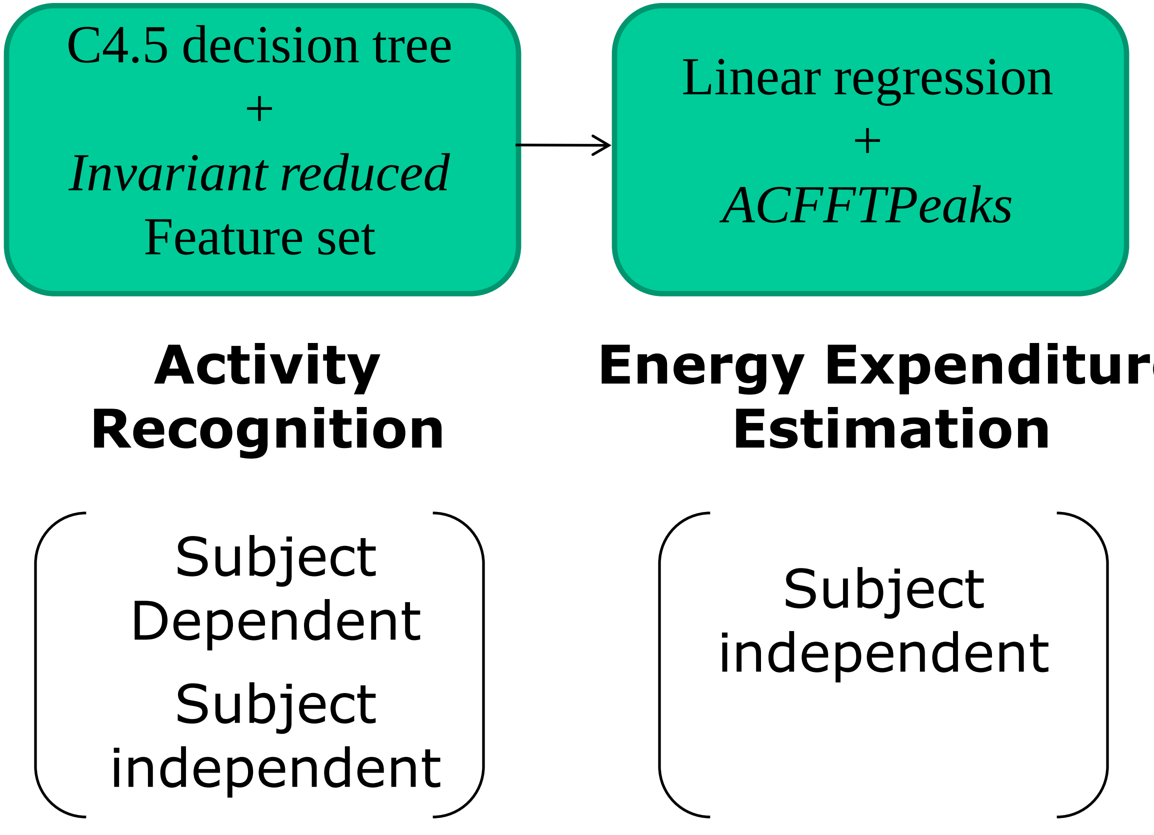
Sensor combinations

Sensor Combination	**Correlation**	**RMSE**
All sensors	0.71	1.28
Hip + DWrist + DFoot	-2.8%	+2.3%
DWrist + DFoot	-2.8%	+3.0%
Hip + DFoot	-4.2%	+3.0%
DWrist + DThigh	-12.7%	+5.2%
Hip + DWrist	-5.6%	+12.3%
DFoot	-8.5%	+5.2%
DThigh	-11.3%	+9.2%
DUpperArm	-15.5%	+11.1%
Hip	-33.8%	+13.5%
DWrist	-2.8%	+21.5%

Sensor combinations

Sensor Combination	Correlation	RMSE
All sensors	0.71	1.28
Hip + DWrist + DFoot	-2.8%	+2.3%
DWrist + DFoot	-2.8%	+3.0%
Hip + DFoot	-4.2%	+3.0%
DWrist + DThigh	-12.7%	+5.2%
Hip + DWrist	-5.6%	+12.3%
DFoot	-8.5%	+5.2%
DThigh	-11.3%	+9.2%
DUpperArm	-15.5%	+11.1%
Hip	-33.8%	+13.5%
DWrist	-2.8%	+21.5%

Activity dependent regression



51 activity dependent models

Method	Activity Feature set	Energy feature set	Correlation	RMSE
LR	-	ScaledHR	0.84	1.01
LR	-	ACFFTPeaks	0.72	1.28
51 activities ARSI LR	Invariant reduced	ACFFTPeaks	0.77	1.31
51 activities ARSD LR	Invariant reduced	ACFFTPeaks	0.88	0.99

Estimation of energy expenditure over 51 activities using three sensors at the hip, dominant wrist and dominant foot.

51 activity dependent models

Method	Activity Feature set	Energy feature set	Correlation	RMSE
LR	-	ScaledHR	0.84	1.01
LR	-	ACFFTPeaks	0.72	1.28
51 activities ARSI LR	Invariant reduced	ACFFTPeaks	0.77	1.31
51 activities ARSD LR	Invariant reduced	ACFFTPeaks	0.88	0.99

Results during subject dependent evaluation are very close to the ones obtained when activity is assumed to be known (<2%).

51 activity dependent models

Method	Activity Feature set	Energy feature set	Correlation	RMSE
LR	-	ScaledHR	0.84	1.01
LR	-	ACFFTPeaks	0.72	1.28
51 activities ARSI LR	Invariant reduced	ACFFTPeaks	0.77	1.31
51 activities ARSD LR	Invariant reduced	ACFFTPeaks	0.88	0.99

- Heart rate data outperforms best accelerometer-based feature.
- Activity-dependent models using subject dependent activity recognition achieve a performance close HR data.

Interesting findings

- Features other than overall amount of motion improve performance
  - Use of 5 FFT peaks + energy + mean crossing rate features instead of overall motion feature at hip sensor improves

    r=+13%, RMSE=-21%

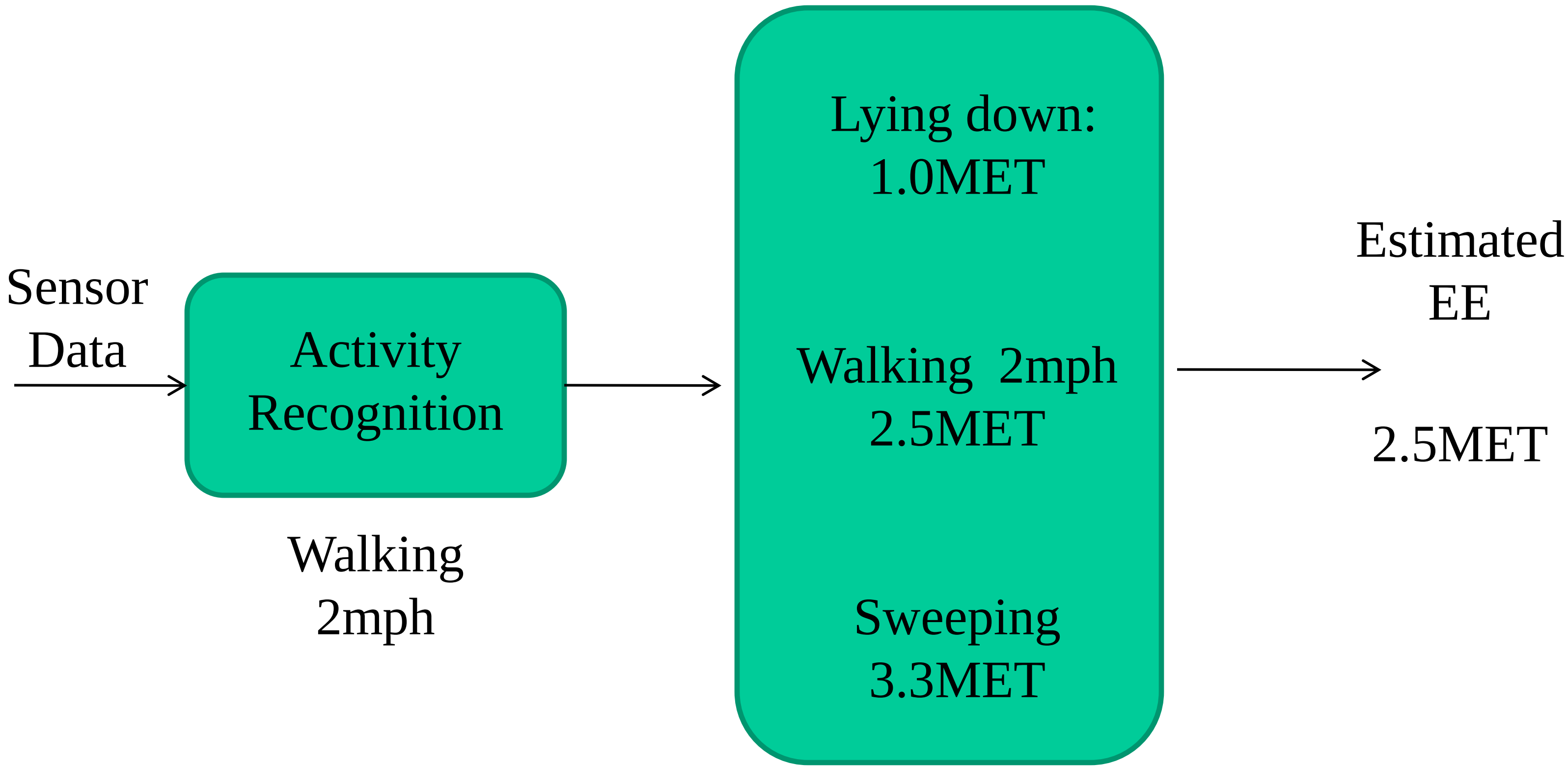
Interesting findings

- Features other than overall amount of motion improve performance
  - Use of 5 FFT peaks + energy + mean crossing rate features instead of overall motion feature at hip sensor improves

    r=+13%, RMSE=-21%

- Addition of heart rate data to best accelerometer feature (5 FFT peaks) improves performance
  - SI: r=+22%, RMSE=-31%

Activity dependent mean values



Look-up table
Mean EE values per activity

Activity dependent mean values

Method	Feature set	Correlation	RMSE
51 activities ARSI Mean	Invariant reduced	0.80 ± 0.08	1.15 ± 0.31
51 activities ARSD Mean	**Invariant reduced**	$\mathbf{0.90 \pm 0.04}$	$\mathbf{0.84 \pm 0.23}$

- This appears to be the best EE estimation strategy at least on the dataset explored
- Improvement with respect to activity-dependent linear regression models

Subject dependent: r=4%, RMSE=-12%
Subject independent: r=2%, RMSE=-15%

Problems with AR-based EE

- Mean EE estimation would overestimate EE for short duration activities (physically intense).
- Misclassifications could affect EE estimates.
- Spurious misclassifications need to be filtered.

Alternatives

- Train on large set of mutually exclusive activities
- Recognize the ‘unknown’ activity and
  - use generic EE model for this activity
  - prompt user at the end of day for *unknown* periods

Contributions: Activity recognition

- Recognition of **52 activities** and subsets on 20 non-researchers
- Recognition of activity **intensity**
- 2min **Subject dependent training** is a promising strategy
- **Three sensors** at hip, wrist, foot
- Acceptable performance without **HR**
- **Real-time system** than can be trained to recognize arbitrary activities

Contributions: EE estimation

- **Activity-dependent** models improve performance
  - Accelerometer and heart rate
  - Performance is close to ACC+HR
- Estimation of **mean EE values** outperforms linear regression models when using activity-dependent models
- EE estimation using **HR** outperforms EE estimation using accelerometer data
- Exploration of impact of parameters

Future work directions

- Create the user interfaces necessary to allow interactive training.
- Allow users to fix the recognition. algorithm (more data/modify models).
- Experiments when data is collected over several days for same subjects (n>40).
- Include activity duration EE estimation.
- Use activity transition information to improve EE estimates.

Committee members

- **Dr. William L. Haskell**, Professor of Medicine at Stanford University
- **Prof. Alex (Sandy) Pentland**, Toshiba Professor in Media Arts and Sciences
- **Dr. Joseph A. Paradiso**, Associate Professor of Media Arts and Sciences
- **Dr. Stephen S. Intille**, Research Scientist
- **Kent Larson**, Principal Research Scientist MIT Department of Architecture

Thank you!

Any Questions?

Contact:

Emmanuel Munguia Tapia

emunguia@mit.edu