Chapter 1

Introduction

1.1 Problem statement

*** Give context problems, means to reach (e.g. monitoring of patients, getting conclusions over health and activity).

1.2 Literature review

This section will provide a literature review per coherent subject of this thesis. In subsection 1.2.1 the research in segmentation and classification of human activities will be discussed.

1.2.1 Segmentation and classification

In recent years much research is done in the field of segmentation and classification of signals into human activities. Due to the nature of many applied algorithms and the field of research (computer vision, data mining, robotics), often there is no clear distinction between the segmentation and classification phase. As a result, these two subjects will be discussed together although, where possible, the different techniques for the two will be mentioned. All the reviewed papers have a global setup in common; a single- or multiple- sensory device is worn by subjects. From the raw signals features are extracted which are used to train some models. During the testing phase the same extracted features are compared with the trained models to classify the activity. Some methods rely on clear segmentations of the signals, whilst other ignore this possible feature and classify on sliding windows.

During the last ten years much research is performed, as a result of smaller and cheaper measuring devices. *** cite needed ***. With the introduction of smartphones with sensory hardware the research has shifted to these devices, as they create a more natural and ubiquitous setting for the test subject. *** cite needed ***.

A distinction can be made in the sensor setup, like wearing multiple sensors, a single sensor and the location of the sensors. Research have been performed with multiple sensors, refered to as Body Sensor Networks, as done in [1–4]. Locations used include the upper arm, waist, thigh, wrist and ankle. Single sensors have be worn at different locations, including the upper arm [5], the front leg pocket [6–9], the waist [10–12], the chest [13, 14], the shoulder [15] and the dominant wrist [16, 17]. Other approaches have not used body-worn sensors at all, but instead relied on other meaures, such as motion capture data [18, 19] and visual [20] and acoustic data in office surroundings [21]

In [3] the effectiveness of different body locations for sensors is discussed. It is shown that using two locations (the hip and wrist or thigh and wrist) instead of five does not significantly reduce the activity recognition. The (dominant) wrist is the preferred location when only a single sensor is used.

Some experiments have only been performed in laboratory settings [16], thus probably resulting in less robust trained models. In [8] the test subjects have performed the testing phase outside, and thus the method is tested in a more natural environment. *** Elaborate on lab/outside ***.

A more objective comparison can be made based on the segmentation and classification technique used. In [3] and [10] extensive comparisons are made between much used algorithms which are widely implemented, such as in the WEKA toolkit [22]. These include the base-classifiers Decision Trees (also used in [6, 7]), Decision Tables, K-nearest neighbors (also used [7, 8]), Support Vector Machines (also used in [9, 23, 24]) and Naive Bayes (also used in [17]). The research of [3] also implemented a number of meta-level classifiers, of which Plurality Voting resulted in the highest accuracy. A very well known and applied algorithm from the field of data mining is k-means clustering. A drawback of this method is that the number of expected clusters needs to be predetermined. To overcome this shortcoming an interative implementation can be used, which selects the best clustering. A variant is known as Fuzzy c-means clustering, in which a data point can belong to one or more clusters, each with a degree of membership. The method of k-means clustering (or a modified version) is used in [5, 12, 19].

Very much used in machine learning problems are neural networks. *** cite needed *** Although these kind of networks typically perform well on spatial-typed problems *** cite needed ***, adjustments have been made to apply them for temporal classification.

The methods of [6, 16] create a multi-level neural network to generate complex discriminators as activity classifiers. A variant of neural networks known as Kohonen Self Organizing Maps is used in [5] to create a vector codebook, which is further processed by a k-means clustering.

Inspired on the problems for continuous speech recognition have Hidden Markov Models been applied to activity recognition tasks. A layered approach proved to be robust for small changes in [20]. In [25] five activities have been modeled as HMMs. The method of [26] introduces HMMs with Markov Jump Linear Systems to segment a signal, with a fixed number of models. In [11] a layer of HMMs of on top of a layer of static classifiers to estimate the current activity. This has a commonality with [15] in which HMMs are used to capture regularities and smooth activity transitions. The application of [2] constructs separate HHMs for each activity and then joins them in a single model to estimate the likelihood of an activity.

A common property of the above mentioned classification algorithms is that there is no explicit segmentation phase. Data points can be segmented a posteriori, by joining adjacent data points which have the same label. An approach in which the data points are processed per segment include [27], in which known motifs in the sensor signals are discovered. In [1] the signal is segmented (without classification) by applying adaptive threshold values on the energy of the signal. The estimates of the direct density-ratio for probability distributions are compared to create change-points in [28]. By making minimal model assumptions a non-parametric segmentation is obtained, without classification. A further application of probability distributions is discussed in [13], which uses an non-parametric Bayesian method to create an Infinite Gaussian Mixture Model to represent activities. Very interesting is the ability to recognize past and new activities without training. The method of [18] considers the intrinsic dimensionality of poses by using Probabilistic Principal Component Analysis and creates cut-points. This method also is independent of trained models and thus also is a non-parametric approach. The method of [29] creates a piecewise linear representation of the signal as thus approximates it. This does not result in a segmentation in the sense of some homogeneity over the data points, but it can support one. In contrast, the method of [14] splits and merges segments the signal in a constant number k segments using a cost function to minimize the heterogeity of a segment.

*** TABLE ***
*** Create a table of all the papers mentioned? Columns would be:
Refs — single/multiple — body locations — activities — lab/outside — training/testing method — result ***

*** Elaborate on measures, e.g. precision, recall, confusion matrix etc. *** *** Precision is defined as the ratio of reported correct cuts versus the total number of reported cuts. Recall is defined as the ratio of reported correct cuts versus the total number of correct cuts. The closer precision and recall are to 1, the more accurate the algorithm is. ***

*****

1.2.2 List of papers

List of all papers, shortly categorized

- "Discovering characteristic actions from on-body sensor data" [27], cited: 63. Motifs, HHM, DTW, 87% accuracy. Known motifs. Worn on wrist, single.
- "Recognition of human activities using layered hidden Markov models" [20], cited: 3. HMM, layerd (primitive and abstract actions). Uses vision for workplace-activities. Direct classification. No sensors
- "Accelerometer-Based Gait Analysis, A survey", [23], cited: 4. Compares methods to distinguish walking of normal, fast, slow. Focus on gait, but compares classification methods such as SVM, PCA, KSOM.
- "An Automatic Segmentation Technique in Body Sensor Networks based on Signal Energy", [1], cited: 13. Automatic segmenting, adaptive threshold. Pure segmenting, no classification. Nice and clean method? Weakness: single action segmented as multiple. Used multiple sensors.
- "Towards HMM based Human Motion Recognition using MEMS Inertial Sensors", [25], cited: 13. Uses HHM, Fourier transform for features. 5 fixed activities, trained HMM. Correct rates from 90% to 100%. Classification, not just segmentation. Single sensor.
- "Change-Point Detection in Time-Series Data by Direct Density-Ratio Estimation", [28], cited: 41. Non-parametric approach, online method, no strong model assumptions. Uses multiple datasets, origin referenced. Much follow-up work done. Artificial datasets. (no sensors)
- "Unsupervised, Dynamic Identification of Physiological and Activity Context in Wearable Computing", [5], cited: 106(!). Combination of KSOM, iterative K-means clustering, transient states removal (using markov model). On-line algorithm. Single sensor on upper arm.

- "Activity Recognition using Cell Phone Accelerometers", [6], cited: 122 (!), 2011. Uses WEKA: decision trees, logistic regression and multilayer neural networks. Results up to 98% (jogging). Much difficulty with stairs up and down. Single accelerometer, front leg pocket. Offline method. Many recent references.

- "Activity recognition from accelerometer data", [10], cited: 434 (!), 2005. Compares 18 (offline) classifiers, base level (WEKA) and meta-level. Single device, pelvic region.

- "Activity recognition from user-annotated acceleration data", [3], cited: 1008 (!!), 2004. Five bi-axial sensors. Decision tables, instance-based learning, C4.5 (trees, highest accuracy 89.30%) and naive bayes (weka), twenty activities.

- "Using Hierarchical Clustering Methods to Classify Motor Activities of COPD Patients from Wearable Sensor Data", [4], cited: none, 2005. Uses Linear Discriminant Analysis for cluster classification (using Cluster Quality Index to determine k) and simple rule-based separation for high level ambulatory. Hierarchical Dendogram to merge clusters when similarity to high. Sensors on forearms and legs.

- "Non-Parametric Bayesian Human Motion Recognition Using a Single MEMS Tri-Axial Accelerometer", [13], cited: none, 2012. Recognizes the number of human activities, single sensor on the chest. No training data. Infinite Gaussian Mixture Model, collapsed Gibbs sampler. Compares with parametric Fuzzy C-mean (data point belongs to multiple clusters), unsupervised K-means, non-parametric mean-shift. Outperforms all significantly, high hit rate and low false alarms.

- "An Online Algorithm for Segmenting Time Series", [29], cited: 487, 2001. Reviews algorithms to get a piecewise linear representation, proposes a novel sliding-window and bottom up approach, on-line. Mere segmentation. No classification; only approximation of signal. (Dataset available).

- "Segmenting Motion Capture Data into Distinct Behaviors", [18], cited: 242, 2004. On-line methods: cut-points on increased intrinsic dimensionality, distribution of poses is observed to change. Batch: cut when consecutive frames belong to different Gaussian mixture models. Fixed fourteen motions, compared with manual segmentation. Uses PCA and Probabilistic PCA (models the non-pc subspaces as noise). PPCA is the best. (paper anne)

- "Aligned Cluster Analysis for Temporal Segmentation of Human Motion", [19], cited: 54, 2008. Uses extension on kernel k-means clustering and Dynamic Time

Alignment Kernel (kernel of DTW) for temporal invariance, robust temporal matching metric. Coordinate descent algorithm solves ACA. Fixed number of clusters, trapped in local minima. (paper anne)

- "Bayesian Nonparametric Methods for Learning Markov Switching Processes", [26], cited: 10, 2010. Uses HMM for state-space model for segmentation, Markov Jump Linear systems. Unbounded number of Markov modes (parameters). Number of models is fixed, though? (paper anne). No sensors.
- "Layered Representations for Human Activity Recognition", [21], cited: 215, 2002. Layers of Hidden Markov Models HMM, multiple levels of granularity (Based on intuition) and context. Less retraining, only lower models. Classified also. No sensors, visual and akoustic data in office surroundings.
- "Time series segmentation for context recognition in mobile devices", [14], cited: 151, 2001. Splits and merges segments, while keeping k constant using cost function on segments for internal heterogeneity. Multiple instances determine number of k. Microphone and accelerometer data, in front of chest.
- "A practical approach to recognizing physical activities", [11], cited: 259, 2006. Uses method of [15], activity classification algorithm. Selects most useful features and then recognizes walking, sitting, etc. First layer static classifier on features or data (energy, mean, variance, correlation, etc), then a layer of HMM to estimate activity. All offline. Shows trained model is robust to location of wearing on body.
- "A hybrid discriminative/generative approach for modeling human activities", [15], cited: 252, 2005. Used in method above. Combines boosting to select and reduce useful features and learn static classifiers with HMM to capture regularities and smooth activities (quick switching is unlikely). Single sensor, multiple measures (accelerometer, audio, etc), worn on the shoulder.
- "Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers", [16], cited: 67, 2008. Separates dynamic from static activities. Uses multilayer feedforward neural networks to generate complex discriminating surfaces as activity classifiers. Feature subset selection approach is developed. Neural pre-classifier with constant threshold criterion. Uses Common Principal Component Analysis. Recognizes eight activities with 95% overall accuracy. Laboratory environment. Worn on dominant wrist.
- "Single-accelerometer-based daily physical activity classification", [17], cited: 46, 2009. Uses Naive Bayes classifier, sensor worn on wrist. Five activities. PCA to reduce dimensions and create independence for NB. Comparable results as Decision Trees, but with flexibility to add activities.

- "Distributed Continuous Action Recognition Using a Hidden Markov Model in Body Sensor Networks", [2] cited: 17, 2009. Single HMM, sensor network cluster movements, HMM constructs continuous actions using postures and actions, merges left-to-right HMM for each action to a single. Transcript generation per sensor uses Gaussian Mixed Model, multiple models with m mixtures are trained with Expectation-Maximization (EM), best is chosen.

- "Offline and online activity recognition on mobile devices using accelerometer data", [7], cited: none, 2012. Compares offline and online methods, cellphone in pocket. Reduces feature extraction (mean and standard deviation (for resources of cellphone). Training is done by database and labeling, classification with K-nearest neighbors, decision trees (C4.5) and decision rules. Online and offline KNN-3 is best, 99% and 97%. Six daily activities. Trousers pocket.

- "Recognizing human activities user-independently on smartphones based on accelerometer data", [8], cited: none, 2012. Classifiers K-nearest neighbors and QDA (quadratic discriminant analysis). Online regocnition implemented on phone. five daily activities. Model training by decision tree (active/inactive). Results for QDA, online: 95%. Activities performed outside. Trousers pocket.

- "Activity recognition from acceleration data based on discrete consine transform and SVM", [9], cited: 25, 2009. Uses Discrete Cosine Transform (DCT) to get features, PCA to reduce them an multiple SVMs to classify windows. Four daily activities, 97% precision. Trousers pocket.

- "Physical Activity Recognition Using a Single Tri-Axis Accelerometer", [12], cited: 16, 2009. Five daily activities, FFT, fuzzy c means classification, offline. 99% accuracy. No smartphone, other device.

1.3 System architecture

*** [GRAPHICS diagram: sensors –¿ signal pre-processing –¿ feature extraction –¿ analysis [* segmentation, * labeling] –¿ conclusions]