Improved representation of movements for human motion analysis using mobile inertial sensors: Including expert knowledge to achieve characterization.

PhD student: Anita Sant'Anna
Supervisors: Nicholas Wickström, Thorsteinn Rögnvaldsson
Halmstad University - Sweden

November 2, 2009

Technical report: IDE0955

Abstract

Despite the many advancements in human motion analysis, this is still a challenging research subject, specially on what concerns the analysis of inertial sensor data. This document presents a review of current motion analysis methods, identifies their shortcomings, and proposes new paths for improving the analysis of inertial sensor data. The proposed solution is based on a linguistic approach and aims at: 1) finding an appropriate representation technique for movements; 2) including expert knowledge as a substitute for supervised training; and 3) characterizing movements based on the different levels of abstraction involved. This report introduces the proposed approach, discusses the possible paths to its development and its contributions to the research community.

1 Introduction

From a technological point of view, human motion analysis opens way for a number of applications in fields such as tele-monitoring and human-machine interaction. Some of these applications could be: monitoring the eating habits of a patient at home (Amft *et al.*, 2005); supervising rehabilitation exercises at home (Jovanov *et al.*, 2005); or creating accessible computer interfaces for disabled patients (Starner *et al.*, 2000).

Motion analysis systems may be mainly characterized as mobile or stationary, according to the type of sensors employed. Stationary systems are mostly based on vision sensors, which can be very accurate but are limited to a specific room or area equipped with sensors. Mobile systems are composed of inertial sensors, which can be placed in garments or embedded in portable objects. They have potential for a much larger reach than stationary systems, but a lot of development is still needed in order to achieve their full potential.

This work focuses on improving the analysis of motion data from inertial sensors in order to contribute to the development of wearable sensor systems. More specifically, this document aims at identifying the shortcomings of available systems and hypothesizing about possible paths for improvement.

The rest of this documents is organized as follows. Related work is presented in Section 2; the identified problem is stated in Section 3; the proposed approach to solve this problem is explained in Section 4; preliminary work which led to this proposal is explained in Section 5; future work is presented in Section 6; and finally, Section 7 concludes this document.

2 Related Work

This section will not present an endless list of related published papers, such a list would never be complete and representative of the whole field of movement analysis. Instead, it will analyze previously published works and structure them in a generalized framework.

2.1 Framework

The framework presented here is limited to motion analysis techniques, that is, those techniques concerned with analyzing human movements *per se*, disregarding the environment. One example of what will be considered here is the classification of different activities from inertial sensor data, e.g.(Bao & Intille, 2004). On the other hand, tracking a specific person in a video sequence, e.g.(Hampapur *et al.*, 2005), is out of the scope.

Previously published works on motion analysis are spread out over many different research areas. The many available options of sensors and analysis methods also help contribute to a highly diversified body of work. The main common thread through so many different applications and research areas is the motion data itself. Therefore, the framework presented here is organized in terms of data abstraction levels.

Abstraction levels are discriminated according to the type of information they contain. Low abstraction levels contain data in a format which is difficult for a human to interpret, whereas high abstraction levels contain information which is easily mapped to human concepts. The lowest abstraction level considered here is *raw data*, followed by *features*. A slightly higher abstraction level is *atomic patterns*, which may be combined into *composite patterns* creating an even higher abstraction level. At the highest level are *actions*, in terms of action verbs closely coupled to our human concepts. Figure 1 illustrates the considered levels of abstraction in ellipses, and commonly used general data analysis approaches in boxes.

2.2 Data Abstraction Levels

2.2.1 Raw Data

Raw data refers to the data acquired directly from the sensors. As previously mentioned, there are two main types of motion analysis systems: fixed and mobile. Fixed systems are typically composed by vision sensors. The data acquired from such sensors could be video sequences e.g. (Weinland *et al.*, 2006); or 3-dimensional trajectories of certain markers placed on the subject's body, i.e. motion capture systems (Gehrig & Schultz, 2008). Fixed

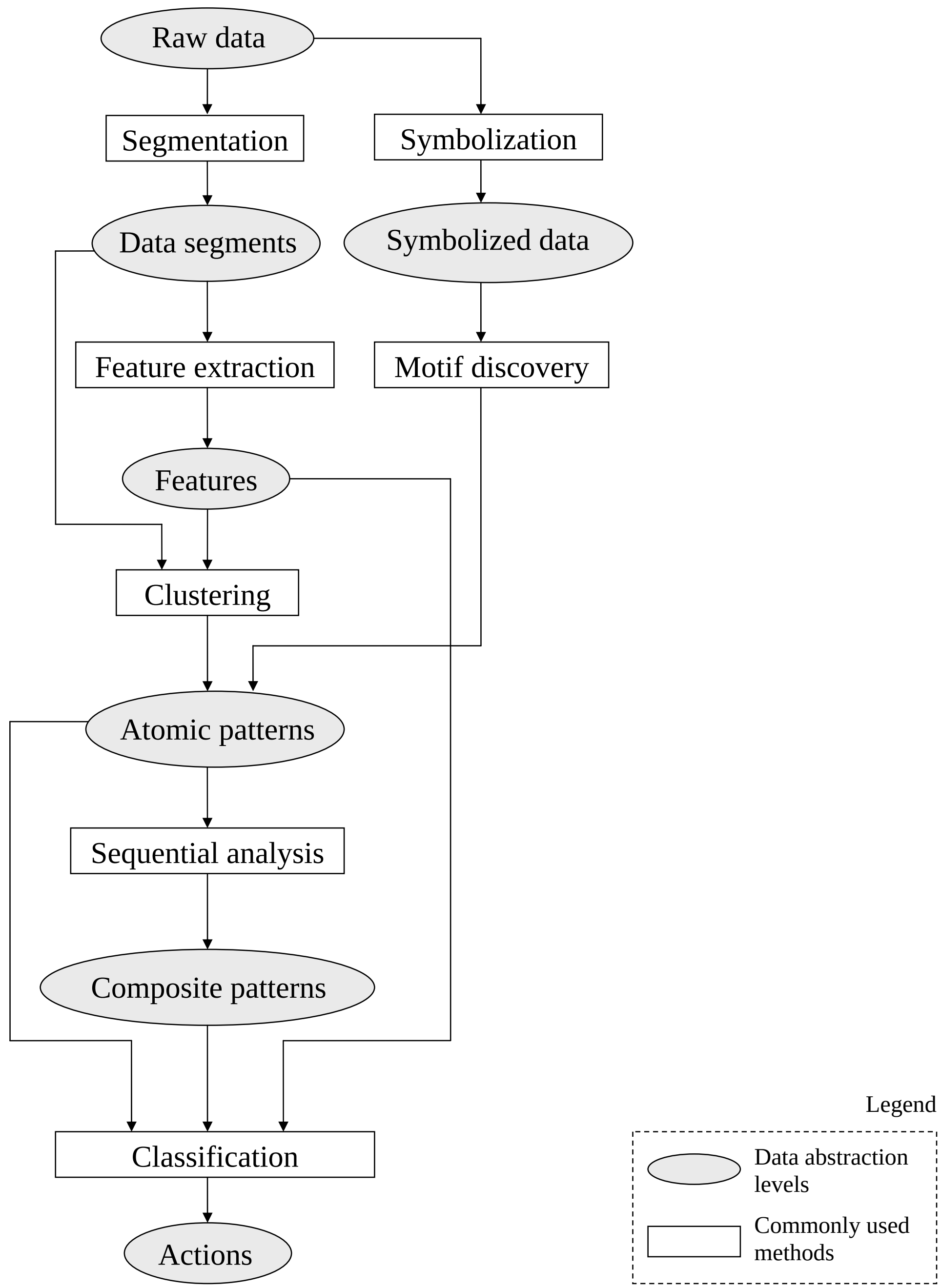

Figure 1: Visual representation of the framework. The diagram illustrates the different levels of motion data abstraction in ellipses and commonly used data analysis approaches in squares. The arrows represent how one abstraction level can be transformed into another through certain data analysis and processing methods.

systems can also be composed of pressure sensitive mats, and infrared sensors. Pressure sensitive mats are particularly useful for clinical gait analysis (McDonough *et al.*, 2001), and biometric identification through gait (Jung *et al.*, 2003). Mobile systems, on the other hand, are typically composed of inertial sensors such as accelerometer, e.g.(Mathie *et al.*, 2004); gyroscopes, e.g.(Zhu *et al.*, 2008); goniometers, e.g.(Gibbs & Asada, 2005); and pressure sensors, e.g. (Zhang *et al.*, 2005).

Independent of the type of data considered, raw data must be processed in order to convey relevant information. As illustrated in Figure 1, the next level of data abstraction is *features*, which can be achieved through segmentation followed by feature extraction. Raw data may also be transformed through symbolization and motif discovery into *atomic patterns*.

2.2.2 Segmented and Symbolized Data

There are two intermediate abstraction level represented in Figure 1, data segments and symbolized data. These are considered the outputs of segmentation, e.g. (Zhou *et al.*, 2008), and symbolization, e.g. (Daw *et al.*, 2003), respectively and they will not receive too much attention here. However it is worth noting that data segments may serve as input to clustering.

2.2.3 Features

Features are here defined as specific measurable heuristic properties of the different movements under study. That is, variables which in some way are able to convey information about different movements. Features are typically obtained from the temporal and/or frequency domains.

Features are typically used to represent segments of the original signal. The main reason for using specific features to characterize segments is so similarities can be found between the different segments, under the assumption that similar segments correspond to similar movements. Another reason for using features is to reduce the amount of data to be considered, i.e. a segment of length M can be represented by N features, with $N \ll M$. The grouping of these similar features (and therefore segments) is normally achieved through clustering.

The use of features to represent motion data is specially relevant to clinical applications, where unhealthy subjects are to be differentiated from healthy ones. Many examples are found within gait analysis studies. Aziz and Arif (Aziz & Arif, 2006) have used a normalized corrected Shannon Entropy of symbolized gait data (joint angular displacement) as a measure of complexity of gait pattern. This feature was shown to be significantly different between healthy and neurodegenerative disease patients (Aziz & Arif, 2006). More recently, Schuartz and Rozumalski (Schuartz & Rozumalski, 2008) proposed a Gait Deviation Index by comparing a set of features from patient data and average values from control patients.

2.2.4 Atomic Patterns

Atomic (or primitive) patterns are the shortest subsequences of the original data which represent specific movements and/or can be combined to represent complex movements. The form of these atomic patterns is highly dependent on the type of data being used and the overall goal of the system. Some examples of atomic patterns are short sequences of symbolized data, e.g.(Minnin *et al.*, 2006), or certain sets of features, e.g.(Zhu *et al.*, 2008). When dealing with video images, patterns may also be certain "key" frames, e.g.(Li *et al.*, 2008).

The main idea behind defining atomic patterns is that human movements are complex but can be decomposed into a number of basic forming patterns, much like in a language, words are compositions of letters. This link between language and movement is supported by the mirror neuron theory (Rizzolatti & Craighero, 2004) which state that the same brain mechanisms are used when observing and when performing an action. This means that there is a tight link between our abstract concepts of actions and the motor mechanisms which perform them. Lately, motion analysis research have been trying to explore language-based approaches order to find an effective and general representation for motion data (Guerra-Filho & Aloimonos, 2007).

2.2.5 Composite Patterns

Once the atomic patterns have been identified, they can be either directly used to classify actions or combined to form composite patterns. The organization of composite patterns has lately been investigated through state-space analysis, e.g.(Li *et al.*, 2008), and/or hierarchical clustering techniques, e.g.(Ogale *et al.*, 2007). Once again, drawing a parallel with natural languages, if the atomic patterns are letters, then composite patterns are words, and the methods used to create composite from atomic patterns are the grammar (Guerra-Filho & Aloimonos, 2007).

2.2.6 Actions

Once atomic and composite patterns have been identified and properly represented, they can be matched to a data base of labeled templates. Action labels are normally verbs used to describe actions. Such a data base would normally contain one template for each action we would like to identify.

We would like to emphasize at this point the difference between characterizing a movement and classifying a movement. When we characterize a patterns we describe it terms of certain known features. When we classify a pattern, we are finding which of a pre-determined set of classes is the most similar to the pattern we are analyzing. For example, one may decide someone is walking slowly by analyzing the velocity of the walk (characterization), or one may train a classifier with 2 classes, slow walk and fast walk. However, when the classifier indicates that a pattern corresponds to slow walk, there is no way of knowing what makes it slow, only that this patten is somehow more similar to the patterns in the slow walk category.

2.3 Commonly used methods

2.3.1 Symbolization

Symbolization deals with representing the original signal as a sequence of more general symbols, typically by using less quantization levels. The choice of quantization boundaries may depend on the different characteristics to be imposed on the final symbols. The most frequent choices are equidistant or equiprobable symbols.

Symbolization improves motion analysis by removing noise and outliers from the data while keeping its most striking dynamics. For a more detailed review on symbolization of time series, the reader is directed to (Daw *et al.*, 2003).

Recently, a symbolization method called Symbolic Aggregate Approximation (SAX) has been shown to possess several interesting features for symbolizing time series. The most important characteristics of this method are the ability to reduce the dimensionality/numerosity of the original data, and the ability to provide a distance measure which is a lower bound for the corresponding distance measure defined for the original signal (Lin *et al.*, 2007).

2.3.2 Segmentation

Segmentation divides the continuous signal in the temporal dimension into manageable sized chunks. It is worth noting that the segmentation and symbolization differ on the dimension where the the signal is divided. Segmentation occurs on the temporal dimension whereas symbolization occurs on the quantization values of the dependent variable.

Many automatic solutions to the segmentation problem have been proposed. One approach is piece-wise linear segmentation, e.g. (Keogh, 2004), which tries to fit line segments over the original data. Yet another approach is to use signal characteristics such as zero crossings, e.g.(Guerra-Filho & Aloimonos, 2007), or variance, e.g.(Zinnen *et al.*, 2007).

When dealing with video images, one common segmentation technique is based on key frames, where the amount of movement from one frame to the next is minimum, e.g.(Weinland *et al.*, 2006). Another technique based on clustering was introduced in (Zhou *et al.*, 2008). There are also manual approaches to segmentation, where the designer/user decides by hand where a certain movement begins and where it ends, e.g.(Bao & Intille, 2004).

2.3.3 Feature Extraction

The choice of features is very important to ensure a good representation of the movements in question. Typically, features are chosen by trial and error. One common practice is to use Principal Component Analysis (PCA) to find the most representative directions in the feature space and reduce the dimensionality of the data, e.g.(Ben Abdelkader *et al.*, 2001).

2.3.4 Finding Motifs

The success of the system highly depends on the choice of atomic patterns. As mentioned previously, this choice depends on the type of data being used and the overall goal of the system. One common approach in the literature is to automatically find repetitive atomic patterns in the signal, normally after symbolization. Several techniques for knowledge discovery in time series may be used, e.g.(Minnin *et al.*, 2006; Lin *et al.*, 2002).

2.3.5 Clustering

Clustering is commonly used to find groups of similar features or segments. The main issue in clustering is finding an appropriate distance measure. The most common is the Euclidean distance but this measure is not the most adequate for all cases. For example, the Euclidean distance has been shown inappropriate for clustering time series segments (Chiu *et al.*, 2003).

This task is even more complicated if the segments are of different lengths. To cope with different segment lengths, some works have used Dynamic Time Warping. Vlachos *et al.* (Vlachos *et al.*, 2002) presented a measure based on Longest Common Subsequence and compared it against Euclidean distance combined with Dynamic Time Wrapping.

The most commonly used clustering technique in the literature is K-means clustering. However, more elaborate clustering techniques have been used, e.g.(Zhou *et al.*, 2008; Barbič *et al.*, 2004)

2.3.6 Sequential Analysis

We grouped under the heading Sequential Analysis, all methods which have as purpose the organization of atomic patterns into composite patterns in some way. Directed graphs and hierarchical clustering are popular choices of methods. Recently published works have used combinations of Bayesian Networks, Hidden Markov Model (HMM) and other graphical representations to model the interaction between several atomic patterns. For example, Li *et al.* (Li *et al.*, 2008) encoded actions in a weighted directed graph, where nodes were key frames (salient postures) shared by all actions. A complete movement was then encoded as one or multiple paths in the graph. Zhu *et al.* (Zhu *et al.*, 2008) used HMM to model several atomic hand gestures based on inertial sensor data, and then used Hierarchical HMM to model possible combinations of the atomic models.

Hierarchical clustering has also been used to organize the relationships between atomic patterns. Ogale *et al.* (Ogale *et al.*, 2007) worked on automatically parsing sequences of key frames in order to create a context-free grammar for human action. These are only a few examples of how atomic patterns maybe combined in to composite patterns.

2.3.7 Classification

Once we have defined a set of patterns (atomic or composite) to represent actions, the problem of action recognition is a traditional classification problem. Several classification approaches have been used in the literature, they may be divided into 4 main groups: closest distance classification, closest model classification, trained classifiers and probabilistic approaches.

Closest distance classification calculates the distance between the pattern in question and sets of previously defined classes, and assigns the pattern to the class that lies the closest. This distance measure can be calculated in different feature spaces or between data string subsequences. Some examples of closest distance classifiers are KNN, e.g.(Bharatula *et al.*, 2005), and pattern matching, e.g.(Stiefmeier *et al.*, 2007). The distance measure most commonly used is the Euclidean distance but other similarity (dissimilarity) measures such as Chi-square have been used, e.g.(Lucena *et al.*, 2007).

Closest model classification is commonly used in the works that model patterns with HMM and variants. In such cases, each possible class is modeled with a HMM and the pattern in question is mapped to the class whose model is the most likely to produce the observed sequence, e.g.(Dong *et al.*, 2007; Li *et al.*, 2008; Zhu *et al.*, 2008; Wang *et al.*, 2001; Wilson & Bobick, 1999).

Trained classifiers refer to those classifiers which must initially be trained with labeled data and can then be used to classify unknown data into one of the previously defined classes. A typical example is ANN, e.g.(Zhang *et al.*, 2005). Also assigned to this category are approaches such as Linear discriminant analysis, e.g.(Gehrig & Schultz, 2008), and C4.5 decision trees, e.g.(Bharatula *et al.*, 2005).

The last group of classifiers are the ones based on Bayes' Theorem, namely naive Bayes classifiers, e.g.(Ogale *et al.*, 2007), and Bayesian Networks, e.g.(Bharatula *et al.*, 2005).

2.4 Discussion

The framework presented here has so far considered fixed and mobile systems equally. However, the differences between such systems also deserves attention.

The type of raw data gathered from vision sensors is able to convey information about many things at once. They often record whole body movements and the data is, in a way, already labeled since a human looking at the video can easily identify which action is being performed. The analysis of this type of data is fairly intuitive since it is so closely linked to our main observatory sense. However, data may only be acquired while the subject is in the same location as the sensors. The obvious lack of mobility can only be compensated by deployment of a large number of sensors.

Mobile systems can acquire data independent of location, and are suitable for continuous monitoring applications. Inertial sensors can be miniaturized and placed in garments or portable objects. The developments in textile sensors also contribute to the spreading of wearable sensors. However, the efficiency of the system highly depends on hardware design choices concerning types and number of sensors, power consumption, security of wireless communications, compliance, among others. In addition, inertial sensors can only provide information about the body section to which they are attached. The analysis of such sensor data is, therefore, a lot less intuitive.

Mobile and fixed systems are complementary, each being more appropriate for certain applications. This work, however, will focus on the development of mobile, wearable systems.

3 Problem

The development of mobile systems depends mainly on two things: hardware design and appropriate processing of the data. A lot of work has been done in designing small, low power sensor nodes for wearable mobile systems, e.g.(Bharatula *et al.*, 2005). The most limiting factor today being, perhaps, the size/capacity ratio of batteries.

The processing of the data acquired is, in turn, strongly dependent on the chosen representation of movement. Ideally, this representation should describe (model) any observed movement, even if there is no link to a known labeled template in the database. The ability to model unlabeled movements enables continuous and incremental expansion of the system, since new observed movements can be labeled and added to the data base one by one.

The methods developed so far for motion analysis from inertial sensors are limited in their representation of movement, i.e. only a certain number of pre-defined activities can be recognized, e.g.(Ravi *et al.*, 2005; Bao & Intille, 2004; Logan & Healey, Aug. 2006). These methods strongly rely on feature extraction and supervised learning techniques. The resulting model is a "black box" which takes the sensor signal as input and outputs classes of activities. Using such models, one cannot understand the reasoning" behind classification. In addition, these works have focused on classification, where different performances of the same activity could only be identified if they were considered different classes of activity, e.g.(Zhang *et al.*, 2005).

Although classification is the step towards the highest level of data abstraction, i.e. actions, many applications could profit from the characterization of movements in terms of lower abstraction levels. As mentioned in Section 2.2, characterization at feature level enables differentiation between healthy and unhealthy patients. Characterization at composite pattern level can provide enough information to differentiate between performances, e.g. from one subject to another, or before and after surgical intervention. The ability to describe different performances is key to the effectiveness of, for example, long term monitoring systems.

In short, the issues not previously addressed in the literature can be listed as:

- Development of a representation which allows:
  - general movement representation, independent of labeled data;
  - description/comparison of different performances;
- movement characterization, not only classification;
- traceable classification, instead of "black box" approach.

4 Proposed Approach

Given the unexplored issues stated in Section 3, we propose to develop a method which fulfills the following criteria:

1. fully automatic, no supervised training needed;
2. capable of general movement representation, independent of labeled data;
3. capable of describing/comparing different performances of the same activity;
4. incorporates previous knowledge about the system, linking classification to human reasoning instead of the "black box" approach.

We believe this may be achieved through the path highlighted in Figure 2. Raw data is segmented and features are extracted from each segment. The features are clustered in order to find atomic patterns, which in turn are organized into composite patterns. An expert system is in charge of characterizing unknown patterns and linking known patterns to actions.

Although the design of expert systems is long and highly dependent on the expert who defines the rules, we expect the use of an expert system will contribute to the method in two main ways:

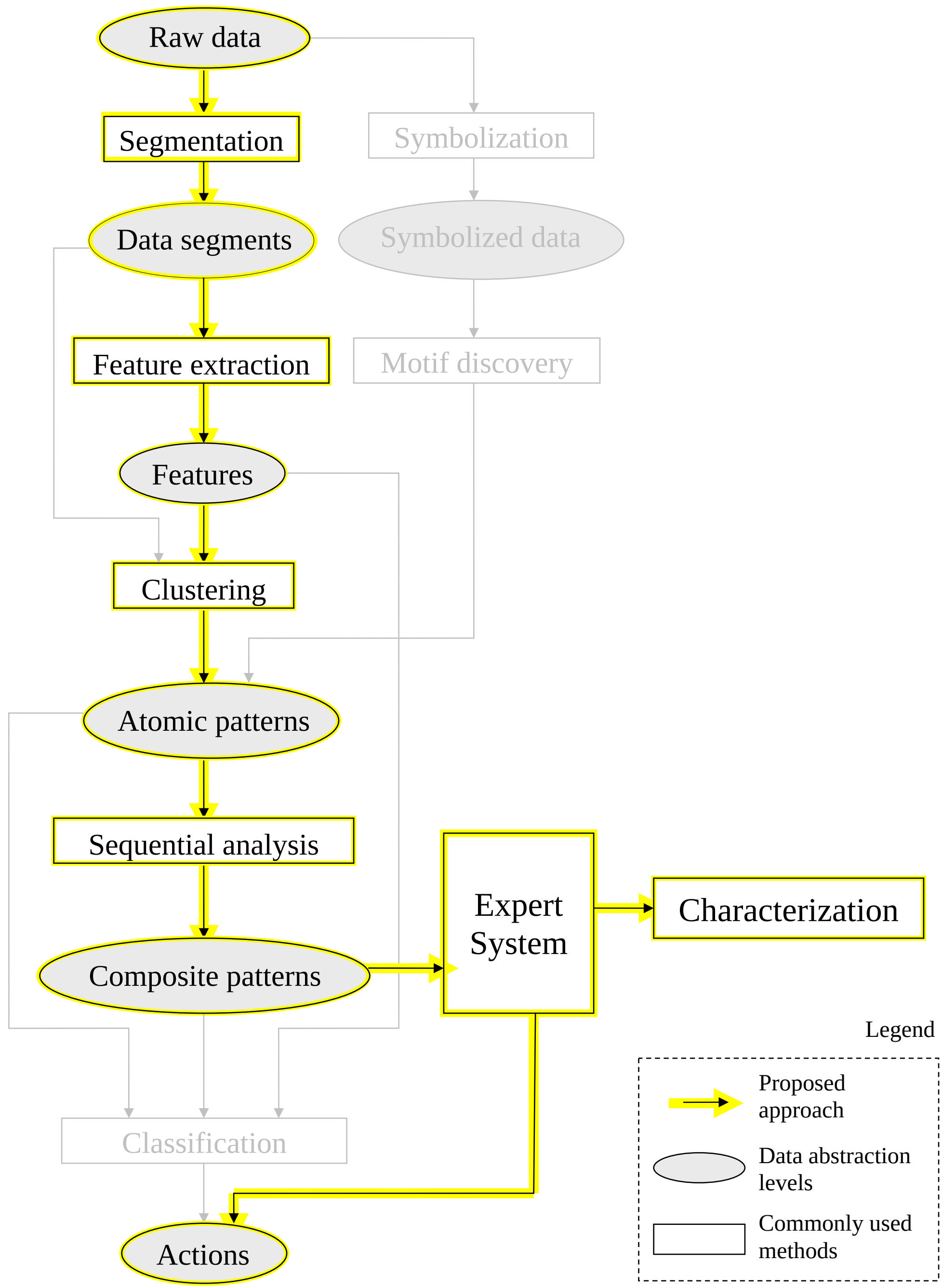

Figure 2: Position of our proposal withing the framework of motion analysis. Ellipses represent the different data abstraction levels and rectangles represent commonly used methods. The arrows represent how one abstraction level can be transformed into another through certain data analysis and processing methods. The highlighted path corresponds the proposed approach.

Results are traceable: Unlike the classification techniques discussed in Section 2.3, an expert system can "explain" its output. The importance of such characteristic is evident, for example, in the clinical setting where the human expert may want confirmation or a better understanding of the result.

Simple alterations do not require retraining of the system: If, for example an extra sensor is added to the system, no retraining is needed. The expert system can handle the extra data with the addition of extra rules. The accuracy of the system may also be adjusted incrementally by addition or modification of rules.

5 Preliminary Work

As a proof of concept, we have looked into one particular motion: gait. Gait analysis is important when assessing e.g. risk of falling (Harris *et al.*, 2005), lower limb prosthetics or orthoses adjustments (Barth *et al.*, 1992), and recovery after knee or hip surgery (Lindemann *et al.*, 2006). Gait analysis has recently also been associated with assessing the risk of developing cognitive impairments in old age (Van Iersel *et al.*, 2004).

5.1 Experimental Setup

Gait acceleration data was collected with two SHIMMER sensor nodes (shimmer-research.com) each equipped with a tri-axial accelerometer, which is sampled at 50Hz and the data is streamed continuously via Bluetooth to a nearby computer. Six subjects participated in the experiments. The subjects had the SHIMMER nodes attached to both shins, close to the ankles. When the subject was standing still, the x axis of the accelerometer corresponds to the vertical axis, the y axis of the accelerometer corresponds to the horizontal axis tangential to the subject's coronal plane, and the z axis corresponds to the horizontal axis tangential to the subject's sagittal plane.

The subjects were asked to walk a straight line on the pressure mat according to three different instructions: 1) to walk at a comfortable self-paced speed, *normal walk data set*; 2) to walk at a very slow speed taking shorter steps, *slow walk data set*; and 3) to walk as comfortably as possible while having the right knee immobilized with a brace in order to simulate limping, *limp walk data set*. The data obtained from the pressure mat was used as ground truth for heel-strike (HS) and toe-off (TO) instances.

5.2 Method

The method described here aims to extract not only temporal information about heel-strike (HS) and toe-off (TO) instances, but also dynamic information about the walking pattern from acceleration signals. This is achieved through symbolization of the signal, and analysis of the context and distribution of each symbol. The symbolized data represents a higher abstraction level, which is more easily coupled to expert knowledge. The proposed method is, therefore, automatic and does not depend on labeled training data.

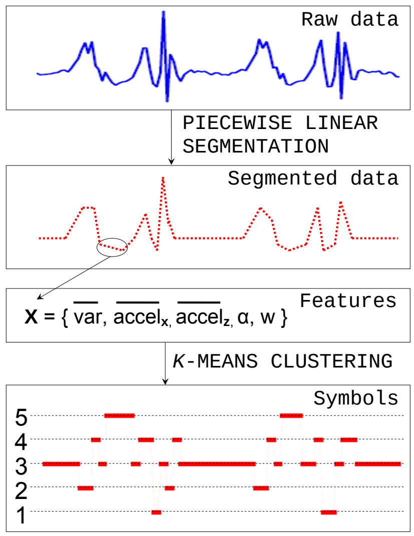

Figure 3: **Symbolization**. Graphical representation of the steps taken towards symbolization of the signal.

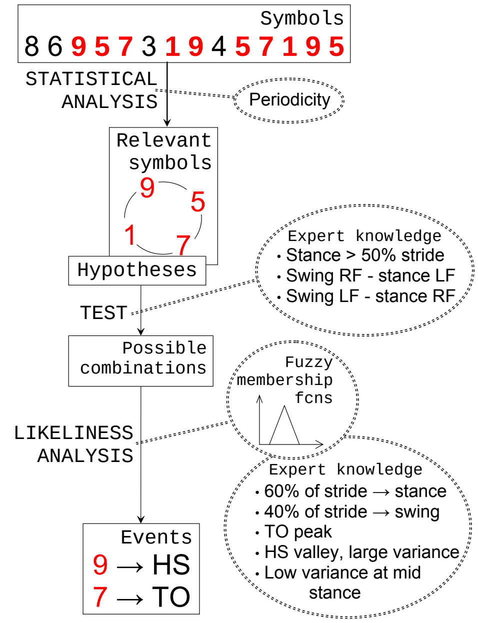

Figure 4: **Context analysis**. Graphical representation of the steps taken towards context analysis of the signal.

5.2.1 Segmentation

The symbolization of the signal is illustrated in Figure 3. After filtering the acceleration signal with a low-pass filter at 20 Hz, piecewise linear segmentation (implemented as described in (Keogh, 2004)) was performed on the resultant acceleration signal: $A_{res} = \sqrt{A_x^2 + A_z^2}$, where A_x and A_z are accelerations in the accelerometer's local coordinate system. The third axis was not considered because it introduced extra variability in the data due to the mirrored nature of the feet's movements, and it was judged dispensable when the subjects walk a straight line.

5.2.2 Feature Extraction

The features extracted from each segment were: mean segment variance of the resultant acceleration, $v\bar{a}r$; mean segment acceleration on both axes, $acc\bar{e}l_x$ and $acc\bar{e}l_z$; the tangent of the angle between the approximated line segment and the horizontal axis, α; and the number of samples in the segment, w.

5.2.3 Clustering

The segment features were standardized (zero mean and unit standard deviation) and divided into different groups using k-means clustering with randomized initial conditions. Clustering was performed considering from 2 to 10 clusters. The optimum number of clusters was chosen based on the minimum Davies-Bouldin index (Petrovic, 2006). The limit of 10 clusters was chosen because most of the initial trials resulted in less than 10 clusters. A unique symbol was assigned to the segments within each cluster. For simplicity, the symbols are integers contained in the interval between 1 and the number of clusters.

At this stage, it is not known how or if the symbols are semantically significant. However, it is expected that the acceleration peaks observed at HS and TO (Selles *et al.*, 2005) are characterized by at least one symbol. Considering that K-means clustering is sensitive to initial conditions, it could happen that the symbols do not reflect HS and TO. In that case, the clustering phase is repeated.

5.2.4 Sequential Analysis

The physical characteristics of the system are reflected on the symbolic data as certain symbol sequences or symbol distributions. Some of the gait characteristics expected to be found in the signal are described in Table 1. This knowledge was incorporated into the model in order to find the symbols corresponding to HS and TO. The algorithm is divided into three main steps: finding relevant symbols, hypotheses testing, and estimating "likeliness", as illustrated in Figure 4.

1	During normal walk, approximately 60% of the total stride time corresponds to stance
2	TO is reflected on the resultant acceleration signal as a peak
3	HS is reflected on the resultant acceleration signal as a valley and large variance
4	The foot moves the least at mid-stance (very small variance of the resultant acceleration)

Table 1: **Gait Characteristics**. Expert knowledge about some physical characteristics of gait.

Finding relevant symbols:

Taking advantage of the cyclic nature of gait, the most common period out of all symbols and symbol transitions is considered the average stride period, $StrP$. Relevant symbols are symbols which are likely to express striking characteristics of the original signal, such as HS and TO, and are expected to appear approximately once every cycle. Symbols (or transitions) with period similar to or half of the estimated stride period are considered relevant symbols. From this point on, relevant transitions are represented as extra relevant symbols and analyzed similarly.

Once the relevant symbols have been identified, the original acceleration signal can be represented as a cyclic sequence composed of the relevant symbols (see Figure 4). It is important to determine where the cycle begins, e.g. which relevant symbol corresponds to HS. In order to determine which symbols should be associated with HS (and TO), all possible symbols are considered and evaluated according to certain assumptions, as follows. For each foot, the N relevant symbols $\boldsymbol{S} = \{S_1, S_2, ..., S_N\}$ are organized into all possible pairwise combinations. These combinations are hypotheses of which symbols could correspond to HS and which could correspond to TO. The combinations for the right foot (Eq. 1) and left foot (Eq. 2), are then recombined to express all possible permutations, considering both feet in parallel (Eq. 3). The single-foot and parallel hypotheses are exemplified in Figure 5.

$$C^{RF} = \{(S_1, S_2), (S_1, S_3), ...(S_N, S_{(N-1)})\} = \{C^{RF}_{1,2}, C^{RF}_{1,3}, ...C^{RF}_{N,(N-1)}\} \tag{1}$$

$$C^{LF} = \{(S_1, S_2), (S_1, S_3), ...(S_N, S_{(N-1)})\} = \{C^{LF}_{(1,2)}, C^{LF}_{(1,3)}, ...C^{LF}_{(N,(N-1))}\} \tag{2}$$

$$C_{parallel} = \{(C^{RF}_{1,2}, C^{LF}_{1,2}), (C^{RF}_{1,2}, C^{LF}_{(1,3)}), ...(C^{RF}_{1,2}, C^{LF}_{(N,(N-1))}), ...(C^{RF}_{N,(N-1)}, C^{LF}_{(N,(N-1))})\} \tag{3}$$

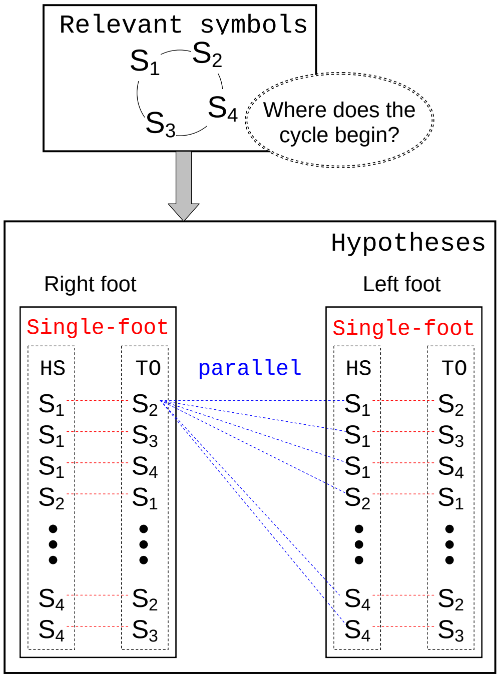

Figure 5: **Creating hypotheses**. The single step hypotheses are all pairwise combinations of the relevant symbols, and parallel hypotheses are all permutations of the right foot and left foot single step hypotheses.

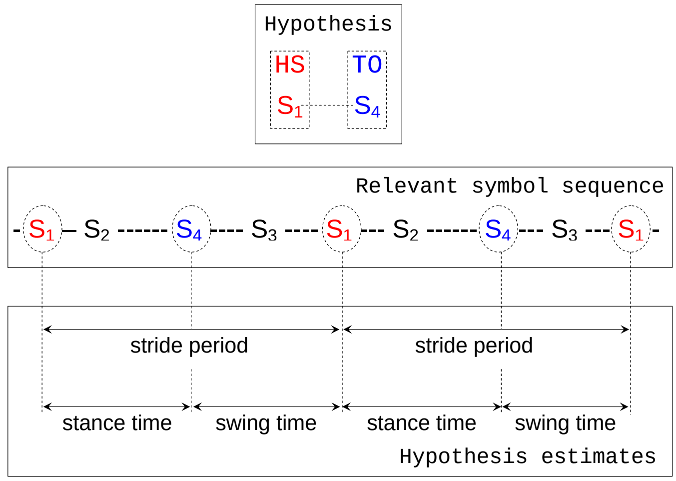

Figure 6: **Hypothesis estimates**. For each hypothesis, the symbols corresponding to HS and TO are used to calculate the stride period, stance and swing. These estimates are used for testing hypotheses and as inputs to the partial likeliness membership functions.

Hypotheses Testing:

The single-foot hypotheses created in the previous step are tested according to the assumption that stance lasts over 50% of the stride time. The average stride period, considering the correspond HS symbol, is calculated for each hypothesis and compared to the corresponding average stance time, as illustrated in Figure 6. If the average stance time is shorter than half the average stride time, the hypothesis is discarded. Then, the remaining parallel hypotheses are tested based on the assumption that swing in one foot can only take place during stance in the other foot. The parallel hypotheses that fail this test are discarded.

Estimating likeliness:

The remaining hypotheses are considered possible and a measure of "likeliness" tries to estimate which hypothesis is more likely to be true. The likeliness of each hypothesis k is estimated according to the four observations stated in Table 1. An extra assumption is added to ensure that the most likely true hypothesis is able to detect an adequate number of strides, given the size of the data. Each observation is represented by a fuzzy membership function which maps measures such as average symbol acceleration and variance to a likeliness value. The four membership functions used are illustrated in Figure 7, and the calculus of the partial likeliness value associated with each assumption is shown in Table 2. For each hypothesis k, the final likeliness value, $L(k)$, is obtained from the product of the corresponding six partial likeliness values for each foot, $L(k) = \prod_{i=1}^{6} L_i^{RF}(k). \prod_{i=1}^{6} L_i^{LF}(k)$. The hypothesis with the largest final likeliness value is chosen as true and the corresponding symbols are used to identify HS and TO instances in the data.

Observations	Input variable for each hypothesis k	Partial likeliness value for each hypothesis k
Approximately 60% of the stride time corresponds to stance	relative stance time, $StnT_{relative}(k) = \frac{StnT(k)}{StrP(k)}$, where $StrP$ is the stride period and $StnT$ is the stance time	$L_1(k) = F_C(StnT_{relative}(k))$
TO is reflected on the resultant acceleration as a peak	average resultant acceleration at TO, $TO_{accel}(k)$	$L_2(k) = F_B(TO_{accel}(k))$
HS is reflected on the resultant acceleration signal as a valley and large variance	average resultant acceleration at HS, $HS_{accel}(k)$, and the maximum acceleration variance between HS and HS + 10 samples, $HS_{var}(k)$	$L_3(k) = F_A(HS_{accel}(k))$ $L_4(k) = F_A(1\text{-}HS_{var}(k))$
The foot moves the least at mid-stance	average variance of resultant acceleration at mid-stance, $MidStance_{var}$	$L_5(k) = F_A(MidStance_{var}(k))$
Detected number of strides must be approximately 40% of the maximun number of strides which could fit within the recorded data	relative number of detected strides, $N_{relative}(k) = \frac{N(k)StrP(k)}{M}$, where, N is the number of detected strides and M is the length of the recorded data set	$L_6(k) = F_D(N_{relative}(k))$

Table 2: **Partial likeliness values** calculated from hypothesis estimates and membership functions. Each likeliness value corresponds to a piece of expert knowledge.

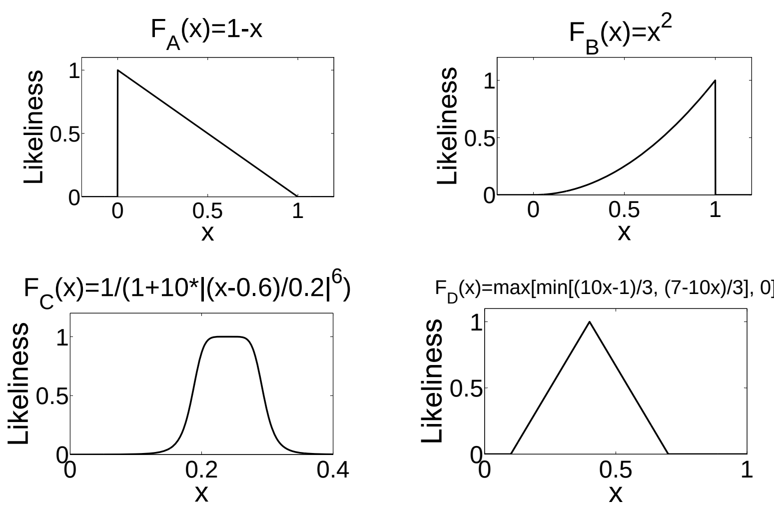

Figure 7: **Membership functions** used to estimate the likeliness of each hypothesis. The inputs to each of these functions is explained in Table 2.

5.3 Results

The detected HS and TO instances were compared to the ground truth data from the pressure sensitive mat. The resulting mean error and standard deviation for each data set, in seconds, are presented in Table 3. Considering the sampling frequency, 50Hz, the mean errors correspond to approximately 2 samples for the normal walk data set, and approximately 5 samples for the slow and limp walk data sets.

Data set	Event	mean (standard deviation)	95% confidence interval
Normal Walk	HS	0.05 (0.05)	[-0.06, 0.15]
	TO	0.03 (0.02)	[-0.01, 0.07]
Slow Walk	HS	0.10 (0.11)	[-0.12, 0.32]
	TO	0.09 (0.15)	[-0.20, 0.37]
Limp Walk	HS	0.10 (0.07)	[-0.03, 0.22]
	TO	0.09 (0.08)	[-0.07, 0.24]

Table 3: Mean error (in seconds) and standard deviation for HS and TO instances for all data sets.

The method presented here is totally automatic and does not depend on labeled training data as do most previous works based on supervised learning techniques. The use of expert rules for context analysis allows the system to be gradually improved, through alteration or addition of new rules, without needing to retrain the whole system. Although this expert system depends on careful tuning of the rules, these rules may be coded to reflect the reasoning made by clinicians when assessing gait, rendering the system easy to interpret. In addition, the symbolic representation of the signal allows for direct comparison of symbol sequences between subjects, or the evolution of symbol sequences for the same subject over time.

6 Future Work

There are two overall goals to this project:

Clinical application: Development and deployment of a clinically relevant and novel application-oriented system, which will demonstrate the relevance of the work in practice. The development of the system will follow the same criteria mentioned in Section 4 but limited to a specific set of movements.

The system developed here will be able to acquire, store and process motion information with wearable, mobile sensors. The system will enable the acquisition of large quantities of data, which may contribute to the inference analysis of risk factors and clinical indicators previously unknown. In addition, the development of a cheap motion data acquisition and analysis system can provide better clinical assistance to unprivileged areas.

Generalization of motion representation: Extension of the representation used in the previous step to include more general movements. This will probably rely on the appropriate choice of methods used in the processing and organization of the motion data, e.g. segmentation and clustering techniques, and sequential analysis. The techniques developed here have roots in and therefore will contribute to research fields such as knowledge discovery in time series.

These overall goals will be achieved through a series of experiments and studies. Some of the experiments planned for the near future are described below.

6.1 Assessment of Motor Dysfunction

The idea here is to collaborate with an orthopedic laboratory or clinic in order to develop a system to automatically extract relevant gait information. This work would result in a reliable assessment system, which is more consistent and precise than current visual techniques and a lot cheaper, financially and with regards to other resources, than full gait analysis using motion capture (mocap) systems.

Some potential partners for this study are the Scandinavian Orthopedic Laboratory, and the Orthopedic Clinic at the municipal hospital in Halmstad. The data obtained for this study could be recordings of inertial sensor data along with mocap data of patients' walking patterns.

If we are able to provide a good measure of quality of gait, our method may have the opportunity to become a standard index in gait analysis and contribute to the expansion of the Swedish National Hip/Knee Arthroplasty Register. This work could also result in an innovative product, such as an intelligent prosthesis, or a clinical gait analysis mobile kit. However, poor collaboration ties with the clinic and failure to acquire good quality data are potential threats to this study.

6.2 Further Methodological Development

This work will be guided by data from healthy subjects performing different activities. The goal is to generalize the motion representation technique to cope with different activities. Technical development could focus on one or more of the previously mentioned main tasks: segmentation, feature extraction, clustering, grammar structure, expert system, or characterization.

One potential partner is the Robotics Institute at Carnegie Mellon University. They have been developing techniques for analyzing human movements with mocap systems, and recently have started looking at inertial sensor data as well.

This study will help increase robustness and generalization of our approach. However, once more, we may not have much control over the data acquisition experiments. This study will give us the opportunity to visit Carnegie Mellon for a few months and get familiar with the work they are doing there. However, our ties with Carnegie Mellon are fairly weak at the moment, and there is a small possibility that the collaboration is terminated even before it starts.

6.3 Investigating Other Types of Data

When detached from the motion description applications, the motion language methodology can be seen as a compression technique, in the sense that we are looking for the underlying structures (patterns) in the data so we can represent the original signal as a series of smaller symbols.

One way to further develop the motion language as a general compression technique is to study its performance on other types of data sets. Possible data could be EEG signals, ECG signals, seismographic data, epilepsy accelerometry data.

If successful this study could generalize our approach to other research areas. However, this will not be directly related to our first topic of interest, motion analysis.

There is a chance that analyzing other types of signal will be inspiring, and lead us towards currently unexpected directions. On the other hand, there is a chance that this non-motion data will prove irrelevant to the development of the motion analysis approach.

7 Conclusion

Despite the numerous and varied works in human motion analysis previously published, there are a few gaps in knowledge still to be filled. Those gaps have been identified as:

1. the development of a motion language representation which allows representation of movement data independent of labeled data, and description/comparison of different performances;

2. movement characterization in addition to classification;

3. traceable classification, as opposed to supervised trained classifiers.

The proposed solution can be described as:

1. finding an appropriate representation technique for movements based on the motion language approach;

2. including expert knowledge to the system as a substitute for supervised training;

3. trying to characterize movements based on the different levels of abstraction involved.

The development of this solution will gradually take place through a few application-oriented and method-oriented small studies. At the end of this long term project, the following should have been achieved:

1. deployment of a system for clinical application;

2. generalization of the motion language to include characterization of unknown movements;

The project hopes to contribute to the research community mainly by:

1. enabling the acquisition of large quantities of data, which may contribute to the inference analysis of risk factors and clinical indicators previously unknown;

2. contributing to research fields such as knowledge discovery in time series;

The expected date of completion for the project is the Autumn term of 2011.

References

Amft, O., Junker, H., & Troster, G. 2005. Detection of eating and drinking arm gestures using inertial body-worn sensors. *Pages 160– 163 of: Wearable Computers, 2005. Proceedings. Ninth IEEE International Symposium on.*

Aziz, W., & Arif, M. 2006. Complexity analysis of stride interval time series by threshold dependent symbolic entropy. *European Journal of Applied Physiology*, **98**, 30–40.

Bao, Ling, & Intille, Stephen S. 2004. Activity Recognition from User-Annotated Acceleration Data. *Pervasive 2004*, April, 1–17.

Barbič, J., Safonova, A., Pan, J., Faloutsos, C., Hodgins, J., & Pollard, N. 2004. Segmenting motion capture data into distinct behaviors. *In: ACM International Conference Proceeding Series*, vol. 62.

Barth, D. G., Schumacher, L., & Thomas, S. S. 1992. Gait Analysis and Energy Cost of Below-Knee Amputees Wearing Six Different Prosthetic Feet. *Journal of Prosthetics and Orthotics Online*, **4**(2), 63.

Ben Abdelkader, C., Cutler, R., Nanda, H., & Davis, L. 2001. *Audio- and Video-Based Biometric Person Authentication.* Springer Berlin / Heidelberg. Chap. EigenGait: Motion-Based Recognition of People Using Image Self-Similarity, pages 284–294.

Bharatula, N. B., Stäger, M., Lukowicz, P., & Tröster, G. 2005. Empirical Study of Design Choices in Multi-sensor Recognition Systems. *In: 2nd International Forum on Applied Wearable Compiting.*

Chiu, B., Keogh, E., & Lonardi, S. 2003. Probabilistic Discovery of Time Series Motifs. *Pages 493–498 of: 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.*

Daw, C. S., Finney, C. E. A., & Tracy, E. R. 2003. A review of symbolic analysis of experimental data. *Review of Scientific Instruments*, **74**(2).

Dong, L., Wu, J., & Chen, X. 2007. A body activity tracking system using wearable accelerometers. *Pages 1011–1014 of: 2007 IEEE International Conference on Multimedia and Expo.*

Gehrig, D, & Schultz, T. 2008. Selecting Relevant Features for Human Motion Recognition. *Pages 1–4 of: 19th International Conference on Pattern Recognition, 2008. ICPR 2008.*

Gibbs, P. T., & Asada, H. H. 2005. Wearable Conductive Fiber Sensors for Multi-Axis Human Joint Angle Measurements. *Journal of NeuroEngineering and Rehabilitation*, **2**(7).

Guerra-Filho, Gutemberg, & Aloimonos, Yiannis. 2007. A Language for Human Action. *Computer*, **40**(5), 42–51.

Hampapur, A., Brown, L., Connell, J., Ekin, A., Haas, N., Lu, M., Merkl, H., & Pankanti, S. 2005. Smart video surveillance: exploring the concept of multiscale spatiotemporal tracking. *IEEE Signal Processing Magazine*, **22**, 38– 51.

Harris, J.E., Eng, J.J., Marigold, D.S., Tokuno, C.D., & Louis, C.L. 2005. Relationship of Balance and Mobility to Fall Incidence in People With Chronic Stroke. *Physical Therapy*, **85**(2), 150.

Jovanov, E., Milenkovic, A., Otto, C., & Groen, P. C. 2005. A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation. *Journal of Neuroengineering and Rehabilitation*, **2**(6).

Jung, J.-W., Bien, Z., Lee, S.-W., & Sato, T. 2003. Dynamic-footprint based person identification using mat-type pressure sensor. *Pages 2937 – 2940 of: Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 3.

Keogh, E. 2004. Data Mining in Time Series Databases Tutorial. *In: Proceedings of the IEEE Int. Conference on Data Mining.*

Li, W., Zhang, Z., & Liu, Z. 2008. Graphical Modeling and Decoding of Human Actions. *Pages 175–180 of: 2008 IEEE 10th Workshop on Multimedia Signal Processing.*

Lin, J, Keogh, E., Lonardi, S, & Patel, P. 2002. Finding Motifs in Time Series. *Pages 53–68 of: Proceedings 2nd International Workshop on Temporal Data Mining.*

Lin, J., Keogh, E., Wei, L., & Lonardi, S. 2007. Experiencing SAX: a Novel Symbolic Representation of Time Series. *Data Mining and Knowledge Discovery*, **15**(2), 107–144.

Lindemann, U., Becker, C., Muche, R., Aminian, K., Dejnabadi, H., Nikolaus, T., Puhl, W., Huch, K., & Dreinhofer, KE. 2006. Gait analysis and WOMAC are complementary in assessing functional outcome in total hip replacement. *Clinical Rehabilitation*, **20**(5), 413.

Logan, B., & Healey, J. Aug. 2006. Sensors to Detect the Activity of Daily Living. *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE.*

Lucena, M. J., Fuertes, J. M., & Blanca, N. P. 2007. Human Motion Characterization Using Spatio-Temporal Features. *Pages 72–79 of: Iberian Conference on Pattern Recognition and Image Analysis.*

Mathie, Merryn J., Coster, Adelle C. F., Lovell, Nigel H., & Celler, Branko G. 2004. Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement. *Physiological Measurement*, **25**(2), R1–R20.

McDonough, A.L., Batavia, M., Chen, F.C., Kwon, S., & Ziai, J. 2001. The validity and reliability of the GaitRite system's measurements: A preliminary evaluation. *Archives of Physical Medicine and Rehabilitation*, **82**(3), 419–425.

Minnin, D., Starner, T., Essa, I., & Isbell, C. 2006. Discovering Characteristic Actions from On-Body Sensor Data. *10th IEEE International Symposium on Wearable Computers*, 11–18.

Ogale, A. S., Karapurkar, A., & Y., Aloimonos. 2007. *Dynamical Vision*. Vol. 4358. Springer Berlin / Heidelberg. Chap. View-invariant modeling and recognition of human actions using grammars, pages 115–126.

Petrovic, S. 2006. A Comparison Between the Silhouette Index and the Davies-Bouldin Index in Labelling IDS Clusters. *Pages 53–64 of: Proceedings of the 11th Nordic Workshop of Secure IT Systems*.

Ravi, Nishkam, Dandekar, Nikhil, Mysore, Prreetham, & Littman, Michael L. 2005. Activity Recognition from Accelerometer Data. *American Association for Artificial Intelligence*.

Rizzolatti, G., & Craighero, L. 2004. The Mirror-Neuron System. *Annual Review of Neuroscience*, **27**(1), 169–192.

Schuartz, M. H., & Rozumalski, A. 2008. The gait deviation index: A new comprehensive index of gait pathology. *Gait and Posture*, **28**, 351–357.

Selles, Ruud W., Formanoy, Margriet A. G., Bussmann, Johannes B. J., Janssens, Peter J., & Stam, Henk J. 2005. Automated estimation of initial and terminal contact timing using accelerometers; development and validation in transtibial amputees and controls. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **13**(1), 81 – 88.

Starner, T., Auxier, J., Ashbrook, D., & Gandy, M. 2000. The gesture pendant: a self-illuminating, wearable, infrared computer vision system for home automation control and medical monitoring. *Pages 87–94 of: Wearable Computers, 2000. The Fourth International Symposium on*.

Stiefmeier, T., Roggen, D., & Tröster, G. 2007. Fusion of String-Matched Templates for Continuous Activity Recognition. *Pages 41–44 of: 11th IEEE International Symposium on Wearable Computers*.

Van Iersel, M. B., Hoefsloot, W., Munneke, M., Bloem, B. R., & Olde Rikkert, M. G. M. 2004. Systematic review of quantitative clinical gait analysis in patients with dementia. *Z Gerontol Geriat*, **37**, 27–32.

Vlachos, M., Kollios, G., & Gunopulos, D. 2002. Discovering similar multidimensional trajectories. *Pages 673–684 of: 18th International Conference on Data Engineering, 2002. Proceedings*.

Wang, Tian-Shu, Shum, Heung-Yeung, Xu, Ying-Qing, & Zheng, Nan-Ning. 2001. Unsupervised Analysis of Human Gestures. *Pages 174–181 of: PCM '01: Proceedings of the Second IEEE Pacific Rim Conference on Multimedia*. London, UK: Springer-Verlag.

Weinland, D., Ronfard, R., & Boyer, E. 2006. Automatic Discovery of Action Taxonomies from Multiple Views. *Pages 1639– 1645 of: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2.

Wilson, A.D., & Bobick, A.F. 1999. Parametric hidden Markov models for gesture recognition. *Pages 884–900 of: IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21.

Zhang, K., Sun, M., Lester, D. K., Pi-Sunyer, F. X., Boozer, C. N., & Longman, R. W. 2005. Assessment of human locomotion by using an insole measurement system and artificial neural networks. *Journal of Biomechanics*, **38**, 2276–2287.

Zhou, F., De la Torre Frade, F., & Hodgins, J.K. 2008. Aligned Cluster Analysis for Temporal Segmentation of Human Motion. *In: IEEE Conference on Automatic Face and Gestures Recognition*.

Zhu, C., Cheng, Q., & Sheng, W. 2008. Human Intention Recognition in Smart Assisted Living Systems Using a Hierarchical Hidden Markov Model. *Pages 253–258 of: 4th IEEE Conference on Automation Science and Engineering*.

Zinnen, Andreas, Laerhoven, Kristof Van, & Schiele, Bernt. 2007. Toward Recognition of Short and Non-repetitive Activities from Wearable Sensors. *Pages 142–158 of: AmI-07: European Conference on Ambient Intelligence*.