# Using Kinect for Assesing the state of Multiple Sclerosis patients

Dimitris Kastaniotis, George Economou and Spiros Fotopoulos

Physics Department
University of Patras
Patras, Greece
dkastaniotis@upatras.gr, spiros@physics.upatras.gr , economou@physics.upatras.gr

Gerasimos Kartsakalis and Panagiotis Papathanasopoulos

Department of Neurology, Neuropsychology Unit, University of Patras Medical School
Patras, Greece
kartsakalis@gmail.com, papat@med.upatras.gr

*Abstract*— **In this work a prototype video-based system for assessing the state of patients with Multiple sclerosis is proposed. In particular we introduce an automated system for capturing and analyzing gait sequences from patients performing the well known 2-minute walking test. The contribution of this work is twofold. First we provide a computerized approach for performing the 2-minute walk test and showing that there is a great correlation between the estimated by the system walking distance and the distance measured by the physicians (Pearson's Rho= 0.7292, p<0.001). Second we present some preliminary results indicating that extracted gait style information is able to differentiate between healthy controls (HC) and Multiple Sclerosis patients even when the Extended Disability Status Scale (EDSS) is low. In order to exploit style information we first incorporated a view invariant representation of the skeleton. Then we represented a selected set of limbs using the Euler Angles and we mapped the sequences of features into the dissimilarity space achieving fixed length representation. Classification performed using Linear Discriminant Analysis resulted into an 88.2% correct classification rate. For our experiments we used a total of 9 MS and 8 HC matched in gender and age.**

*Keywords-component; Kinect; Multiple Sclerosis; gait analysis;*

## I. INTRODUCTION

Multiple Sclerosis (MS) is a common chronic inflammatory and neurodegenerative disease in which the insulating nerve cells in the brain and spinal cord are damaged [1]. Of major importance is the effect of the disease on gait and balance, and thus several clinical tests are focusing in the quantification of patient's status. Among others [2], the 25 feet test and the 2-minute walking test are the most common objective methods for quantifying gait disability in clinical MS trials. While these methods quantify average measures, they ignore the variability of several parameters that appear in gait. Indeed, variability of some gait attributes has proven to be very useful when analyzing gait [3].

Following the trends in Computer Aided Diagnosis (CAD), motion capture systems (MOCAPs) have been utilized in the context of human motion analysis for automated health care monitoring. The main goal of these systems is to automate the process of gait observation and reduce subjectivity and error appearing in measurements. These systems though are very expensive and invasive. From a computer vision perspective the problem of video based pose estimation has been one of the most challenging problems. Recently with the launch of low cost depth sensors the landscape of pose estimation has changed. Sensors like the Microsoft's Kinect provide depth as well as optical information in real time (30fps on PC and 200fps on a gaming console). Using advanced machine learning techniques one is able to take an estimation of a pose given the depth image in real time [4]. The implementation of the algorithm proposed in [4] one is able to track up to two skeletons in real time. An implementation of this algorithm is available to users via Microsoft's Kinect SDK.

In this context several approaches utilizing gait information have been proposed. Here we categorize the applications into invasive and non-invasive. An overview of these methods is presented in [5]. Intrusive methods usually require the placement of a number of sensors on human body while non-intrusive only sense the environment and estimate the pose via a data processing procedure. Initially, intrusive techniques were the only solution against the problem of pose estimation [5]. With the advent of accurate and low cost sensors the balance between intrusive and non intrusive started to change. Indeed, there is a demanding need to perform accurate measurements of gait sequences with low cost. Authors in [6] shown that one can reliably perform gait analysis using such a sensor. Following this finding several applications were developed focusing on human action, and gender recognition [7],[8] and promoting further the applicability of the sensor in a wide range of applications.

In the context of medical diagnosis though only a few works have been published. Authors in [6] first posed the problem of non-invasive gait analysis using a Kinect sensor. Others have tried to assess the state of patients by using pose, depth and color information. For example authors in [9] utilized a Kinect sensor in order to detect the maximum speed and proposed a test named Short Maximum Speed Test which they compare with the 25-Foot walk test. Other utilized depth and color information [10] while the Kinect Aided Diagnosis system was presented as a computerized approach for capturing and analyzing data during the two minute walking test [11].

While all aforementioned methods perform very well in terms of accuracy, none of them studies gait fluctuations. These methods are not utilizing features that contain information related to gait variability. Indeed there are several attributes in the way we walk that can be used in order to exploit information about gait fluctuation. Authors in [5] have shown that stride, joint angles etc are some of the features that can be utilized for medical diagnosis purposes. Furthermore, instead of using average measures (total distance, maximum speed etc) one could benefit from studying the dynamical characteristics of gait. Authors in [12] have noticed the increasing interest into studying the variability of gait parameters. Based on evidence that gait variability is connected with fall risk, some applications target the problem of elderly people fall prediction by detecting balance disorders [13]

In this work we present for the first time a prototype system that is designed [11] for capturing, analyzing and classifying gait sequences from subjects diagnosed with Multiple sclerosis (MS), performing the well known 2-minute walking test. The contribution of this work is twofold. First we study whether or not perceptive computing-based analysis of walking style is feasible and reliable during the 2-minute walking test clinical trials. Results indicate that the proposed architecture can reliably estimate walked distance. Second, based on recent findings that gait variability encompasses clinical attributes, here we show that exploiting information from gait style, via selected joint angles, can be efficiently used to discriminate between HC and MS patients, even for small EDSS values. Indeed gait encompasses a plethora o biological cues that humans are able to effortlessly analyze in short time. Preliminary experimental results performed utilizing 18 subjects (9 MS Patients and 8 Healthy controls) indicate that our system can achieve significant recognition rates 88.2% using information extracted from walking style. Our system named KAD (Kinect Assessment Dissorder) [11] addresses two modern issues that appear in gait analysis. The non- intrusive and low cost pose tracking of human skeleton as well as the analysis of gait variability in order to exploit higher order information that is related to the clinical status of patients.

## II. PROPOSED METHOD

### A. System Architecture

. In our experimental setting, during the 2-minute walking test patients walk in a corridor and the physician measures the walked distance. Thus in order to estimate the total distance, our system needs to cover as much area as possible in order to continuously track the individual and estimate the average speed. In favor of achieving maximum coverage of the corridor we placed by three Microsoft Kinect for windows sensors in different positions along the corridor. Using the Microsoft Kinect SDK we designed specialized software that allows us to manage the tracked skeleton from all three sensors. Additionally to the skeleton captured we also record and store the depth information. It should be noticed also that our system does not using at all color information. This allows us to retain confidentiality as visual inspection of the gait

sequences via depth and or skeleton does not provide information about the identity of the person.

Retest reliability analysis using intra-class correlation coefficient (ICC) utilized in order to study the consistency between all three sensors. Using the estimated average speeds by all sensors the computed correlation coefficient ICC > 0.99 shown excellent results validating the hypothesis that all sensors behave in a consistent and stable manner. In the following paragraphs we provide the details of our data processing pipeline.

### B. Feature Extraction and Representation

Be compliant with view invariance is of great importance as the sensors position and/or individual route could vary a lot. Without view invariance the feature representation would not be representative and could potentially lead to high bias of the classification model. In this work, we propose the use of a coordinate system that is attached to the human body as shown in the following figure. Following such a representation gives more flexibility to the system as the sensors do not need to be calibrated before every trial and the subject can follow a different route without affecting the feature extraction procedure. The following figure represents the two coordinate systems (Kinect sensor (green) and the Torso based (orange)).
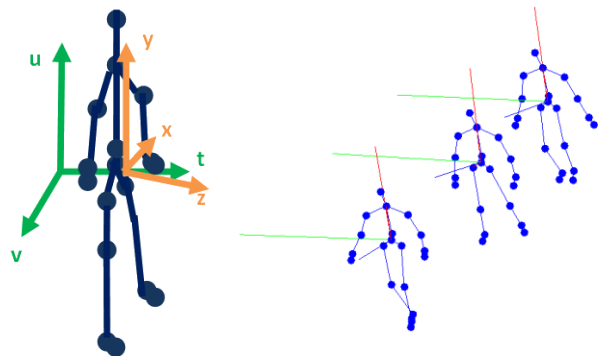


Figure 1. Left: The view invariant coordinate system (xyz) with respect to the sensors system (uvt). Right: Three frames selected from a sequence depicting the attached to the torso (xyz) coordinate system.

Our feature extraction pipeline is composed of three steps. In the first step we transform our skeleton data into a new coordinate system attached to the human body. This allows us to achieve a rotation invariant representation. In the second step we represent 8 skeletal primitives using the Euler Angle representation. In the third step we map the Euler Angle distributions into the dissimilarity space in order to get a fixed length feature representation.

The procedure for estimating the view invariant "torso" coordinate system is summarized in the following steps:

  a. Use the joints connecting the center of the hip with the center of shoulders as the first axis v1 (red vector).

b. Take the orthogonal to the v1 component of the vector connecting the two shoulders as the second one v2 (green vector).
c. Compute the third vector –v3- as the dot product of the v1 and v2 vectors (blue vector)

## C. Euler Angle Representation

As mentioned in previous studies [12], the angles of certain limbs encompass a wealth of information. This angle information is dependent on the subject's pose with respect to the sensor, so in order to render recordings pose invariant, as explained in the previous section, we choose to express the angles formed by the individual skeletal primitives, in terms of a torso based coordinate system.

The following figure illustrates the selected skeletal primitives corresponding to Humerous, Radius, Femur and Tibia bones from both sides of human body, considered to be the most informative regarding the body pose. Each pose is represented by a feature vector consisting of the Euler angles $\{\varphi, \theta\}$ denoting the orientation of each component in the sensor's coordinate system. Thus, the initial dimensionality of the pose data is being reduced from twenty 3D coordinates to eight pairs of angles.
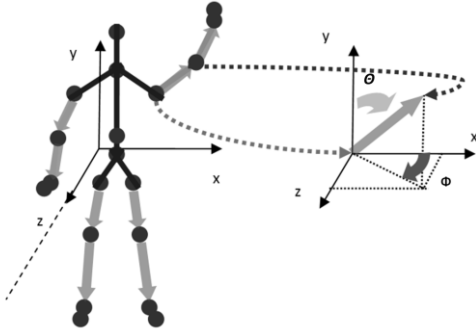


Figure 2.   Example of a figure caption. *(figure caption)*

Thus, following this stage, a sequence of poses corresponding to a walking person is represented by set of vectors into the 16-dimensional space of Euler angles.

## D. Mapping to the Dissimilarity space

We choose to map collected data into the dissimilarity space. Towards this goal the pair-wise distance matrix of the - angle sequences - vector data is formed. Besides representation efficiency, this mapping allows us to form a fixed length feature vector space [14].

Let $\mathbf{X} = \{o_1, o_2, ..., o_n\}$ be a training set of vectors of angle sequences. Given an appropriate dissimilarity function $D$, which in our case is the Mutual Nearest Point Distance (MNPD) [15] a mapping $D(\bullet, P): X \rightarrow \mathbb{R}^k$ can be defined, where $P = \{p_1, p_2, ..., p_k\}$ is, in the general case, a set of $k$ selected sequences, namely prototypes, and can be a subset of $X$. In the resulting space, called dissimilarity space, each dimension $D(\bullet, P_i)$ describes the dissimilarity to $i^{\text{th}}$ prototype. Our case matches the simplest scenario where $P := \mathbf{X}$ so that every angle sequence is represented by an n-dimensional vector of dissimilarities to all n training vector sequences $\mathbf{y} = D(o, \mathbf{X}) = \left[ d(o, o_1), d(o, o_2), ..., d(o, o_n) \right]^{\text{T}}$.

Thus, the representative vectors for the objects of $\mathbf{X}$ are simply the columns of the corresponding dissimilarity matrix $\mathbf{D} = D(\mathbf{X}, \mathbf{X})$. Thus the final descriptor for a gait sequence is given by the following formula:

$$\mathbf{y}_{gait} = D(o, \mathbf{X}) = \left[ d(o, o_1), d(o, o_2), ..., d(o, o_n) \right]^{\text{T}} \quad (1)$$

## E. Classification

Several classification techniques based on models of the class densities have been proposed. Most of these techniques differ in flexibility and the assumptions made about the underlying features' distribution.

As we explain in detail, LDA arises as a special case, where it is assumed that classes follow the Gaussian distribution and have a common covariance matrix. In particular, suppose that every class density is modeled according to the following formula. For simplicity we refer to the mapped gait sequence vector x.

$$p_c(\mathbf{y}_{gait}) = \frac{1}{(2\pi)^{d/2} |\Sigma_c|^{1/2}} e^{-1/2(\mathbf{y}_{gait} - \mu_c)^{\text{T}} \Sigma_c^{-1} (\mathbf{y}_{gait} - \mu_c)} \quad (2)$$

,where index c represents the class c ,and d is the dimensionality of the descriptors and x denotes a vector representing a gait sequence and $\mu_c$ denotes the center of class c. It is known that discriminant functions of the following form (equation (3)) can lead to a minimum-error-rate classification.

$$g_c(\mathbf{y}_{gait}) = \ln(p(\mathbf{y}_{gait} | w_c)) + \ln P(w_c) \quad (3)$$

The previous equation or similar formulations imply an equivalent decision rule whose role is to partition the feature space into $c$ regions.

Moreover, the previous expression can be evaluated if the densities $p(\mathbf{y}_{gait} | w_c)$ are multivariate normal distributions.

As we stated before, under the LDA framework we are assuming that the following equation holds: $\Sigma_c = \Sigma \forall c$ .The geometrical interpretation of this case, is when samples fall in hyperellipsoidal clusters of equal size and shape. This assumption simplifies our model and equation (3) takes the form of equation 4:

$$g_c(\mathbf{y}_{gait}) = \mathbf{y}_{gait}^{\text{T}} \Sigma^{-1} - \frac{1}{2} \mu_c \Sigma^{-1} \mu_c + \ln P(w_c) \quad (4)$$

According to the LDA framework a sample is assigned to class $c$ that maximizes $g_c(\mathbf{y}_{gait})$.

## III. Dataset and Methods

The dataset was collected at the Patras University Hospital, Greece, Department of Neurology where the prototype system is operated. Both MS patients as well as HC signed an approval form prior to data collection. Additionally to the two minute walking test all subjects completed the Modified Fatigue Impact Scale (MFIS). Furthermore for every patient the Expanded Disability Status Score was recorded. Skeleton tracked data as well as depth data were stored in a central computer. In total 8 Healthy Controls (3 males and 4 females, age mean±std: 31±7 years) and 9 MS patients (5 males and 4 females, age mean±std: 31±9 years) have been recorded. For the following experiments we kept only the portions of the recordings where the subjects were captured frontal view since as we noticed Kinect performance degrades significantly when the user is captured from the back. Also we didn't incorporate any filter for noise reduction prior to the analysis/processing of the data.

## IV. Experimental Results

In order to evaluate the efficiency of the proposed system we present experimental results both for the task of estimating the 2-minute walking test distance, as well as of the achieved classification performance of gait sequences into two categories namely HC and MS.

All computations were performed using the Matlab programming language and the IBM's SPSS version 21.

### A. Automated 2-minute walking test

In this paragraph we present results indicating the reliability of the proposed system. In particular, we test the reliability of our system in estimating the distance covered by the patients as the correlation between the estimated distance and the covered distance as measured by the physicians. The correlation between the estimated and the measured by the specialist distance was significant Pearson's Rho=0.72, p<0.001. Additionally the mean speed for HC was measured as (mean and ±std) 1.13 m/sec ±0.1 m/sec and for MS patients 0.95m/sec ± 0.13m/sec. Our system estimated as average speed for HC 1.16 m/sec ±0.2 m/sec and for MS patients 1.05m/sec±0.19m/sec.

### B. Gait based Classification

Here we focus into the classification of gait sequences into one of the two categories: HC and MS. Using the dissimilarity vectors we evaluated our system following the leave one out procedure. The dimensionality of our descriptor and of our classification space as well equals the number of training samples. This is due to the fact that during the mapping into the dissimilarity space, we used all available subjects as prototypes. Experimental results indicate that our method could differentiate between healthy and non-healthy subjects even for low EDSS values. Confusion matrix results are presented in the following table 1. It is worth to notice that only one sample from each category was misclassified thus the false positive and false negative percentage is in both cases

5.9%. The overall classification 88.2% rate proves the efficacy of our system.

TABLE I. CLASSIFICATION RESULTS

| Gait Sequences Classification Results using the Leave one out method | | |
|---|---|---|
| Correct Classification Rate | False Positive Rate | False Negative Rate |
| 88.20% | 5.9% | 5.9% |

These results while preliminary strengthen our initial belief that information encompassed in gait can be used to characterize a subject's health status. In particular the joint directions as represented via the two Euler angles seem to contain proper information for distinguishing between HC and MS patients.

## References

[1] Noseworthy JH, Lucchinetti C, Rodriguez M, Weinshenker BG: Multiple sclerosis. N Engl J Med 2000, 343(13):938–952.

[2] Pearson OR, Busse ME, van Deursen RWM, Wiles CM: Quantification of walking mobility in neurological disorders. QJM 2004, 97(8):463–475.

[3] Hausdorff JM, Gait variability: methods, modeling and meaning. J Neuroeng Rehabil. 2005 Jul 20;2:19.

[4] J. Shotton, et al., "Real-time human pose recognition in parts from single depth images," presented at the Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition, 2011.

[5] Richard Baker, Gait analysis methods in rehabilitation, J Neuroengineering Rehabil. 2006; 3: 4.

[6] Moshe Gabel, Erin Renshaw, Assaf Schuster, and Ran Gilad-Bachrach, Full Body Gait Analysis with Kinect, in Proceedings of EMBC 2012, Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), August 2012.

[7] Ilias Theodorakopoulos, Dimitris Kastaniotis, George Economou, Spiros Fotopoulos, Pose-based human action recognition via sparse representation in dissimilarity space, Journal of Visual Communication and Image Representation, Volume 25, Issue 1, January 2014, Pages 12-23

[8] Kastaniotis, D.; Theodorakopoulos, I; Economou, G.; Fotopoulos, S., "Gait-based gender recognition using pose information for real time applications," Digital Signal Processing (DSP), 2013 18th International Conference on , vol., no., pp.1,6, 1-3 July 2013

[9] Janina Behrens,Caspar Pfüller, Sebastian Mansow-Model, Karen Otte, Friedemann Paul,and Alexander U Brandt. Using perceptive computing in multiple sclerosis - the Short Maximum Speed Walk test, J Neuroeng Rehabil. 2014; 11: 899

[10] Sébastien Piérard, Samir Azrour, Rémy Phan-Ba and Marc Van Droogenbroeck, GAIMS: A Reliable Non-Intrusive Gait Measuring System, , on ERCIM news, October 201.

[11] Spiros Fotopoulos and Dimitris Kastaniotis, KAD - An Intelligent System for Categorizing and Assessing the State of Patients with Multiple Sclerosis, on ERCIM news, October 2013

[12] Michael J. Socie and Jacob J. Sosnoff, "Gait Variability and Multiple Sclerosis," Multiple Sclerosis International, vol. 2013, Article ID 645197, 7 pages, 2013. doi:10.1155/2013/645197

[13] Stone, E.; Skubic, M., "Fall Detection in Homes of Older Adults Using the Microsoft Kinect," Biomedical and Health Informatics, IEEE Journal of , vol.PP, no.99, pp.1,1

[14] S.-C. Fang and H.-L. Chan, "Human identification by quantifying similarity and dissimilarity in electrocardiogram phase space," Pattern Recogn., vol. 42, pp. 1824-1831, 2009.

[15] E. Pękalska and R. Duin, The Dissimilarity Representation for Pattern Recognition: Foundations And Applications: World Scientific Publishing Company, 2005.