34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Beyond the Standard Clinical Rating Scales: Fine-Grained Assessment of Post-Stroke Motor Functionality Using Wearable Inertial Sensors

Mi Zhang[1], Belinda Lange[2], Chien-Yen Chang[2], Alexander A. Sawchuk[1], and Albert A. Rizzo[2]

Abstract— **Accurate motor function assessment of post-stroke patients plays a critical role in their rehabilitation interventions. In this paper, we propose an approach to use wearable inertial sensing technology to quantitatively evaluate the patients' motor behavior. Different from existing wearable motor function assessment techniques that focus on building mapping functions that correlate sensed movement signals to the standard clinical rating scales, our approach provides a fine-grained assessment by capturing detailed patterns contained in the patients' movements. We collected data on three subjects including two post-stroke patients who have varying degrees of upper extremity hemiparesis. Our experimental results validate our approach and demonstrate that the captured patterns can be used to complement the standard clinical scores to provide fine-grained motor function assessment and help clinicians to track patients' gradual progress during rehabilitation.**

I. Introduction

Stroke is a leading cause of adult disability and death worldwide today [1]. While post-stroke patients often exhibit impaired balance, partial or full paralysis, and spasticity problems [2], many improve their mobility and coordination through comprehensive physical rehabilitation. Thus, it is important to assess accurately the patient's current motor functionality to design the most appropriate rehabilitation interventions. Clinicians and physical therapists typically supervise patients' movements and assess motor functions using standard clinical rating scales. However, since this approach relies on the clinicians' observational measures, the assessment accuracy can be influenced heavily by clinicians' subjective judgments. In addition, the rapid growth of the post-stroke patient population overburdens the clinicians such that they have little time to complete required assessment and treatment protocols.

Therefore, a tool that can quantitatively evaluate the patients' motor function would be valuable for both clinicians and patients. One popular technique uses video cameras and special markers attached to patients' bodies to capture and assess their movements [3]. Although this optical-based sensing technology tracks the movement accurately, it is typically expensive and complicated to set up. More importantly, optical sensors do not sense the strength of the patients' movement, which is an important metric for motor function assessment. Alternatively, wearable inertial sensing technology uses accelerometers and gyroscopes to directly measure the movement's acceleration and rotation. Compared to optical devices, inertial sensors are cheap, tiny, and can be worn on the body to track movement continuously and unobtrusively [4]. In addition, occlusion, noisy background, and viewpoint change problems of optical-based sensing technology are not a factor.

The use of wearable sensing technology for assessment of motor function has grown significantly in recent years. Hester *et al.* [5] used accelerometers and linear regression models based on their measurements to predict functional ability scores for stroke rehabilitation. Patel *et al.* [6] used clustering techniques to correlate accelerometer signals with the severity of dyskinesia in patients with Parkinson's disease. They demonstrated that patients with different severities can be represented by well separated clusters. Other studies cast the motor function assessment as a classification problem. For example, the authors in [7] used accelerometer, gyroscope, and magnetometer data to sense patient upper limb movements after neurological injury. A decision tree classifier inferred the functional ability score of the Wolf Motor Function Test (WMFT). Similarly, the authors in [8] apply the support vector machine (SVM) as the classifier to learn the mapping function between the functional ability scores and the severity of Parkinsonian motor fluctuations.

Previous techniques use linear regression, clustering, and classification to building mapping functions that correlate sensor signals with standard clinical rating scales. We feel that the clinical rating scales cannot record all details of motor behavior, thus failing to evaluate precisely the patients' performance during rehabilitation interventions. To bridge this gap, we describe a fine-grained motor function assessment approach that captures detailed patterns of the patients' motor behavior which standard clinical scores fail to acquire. We do not regard our approach as a replacement to the existing clinical score system. Instead, our approach should act as a significant complement to the standard clinical rating scales in the sense that combining the scores and the detailed patterns detected by our approach could produce a more accurate assessment of patients' motor behavior.

II. Our Method

Different from the existing motor function assessment methods in which the whole segment of the motor task is mapped to a single point in the feature space, the first step of our fine-grained approach is to divide the streaming sensor data sampled from each motor task segment into a

[1]Mi Zhang and Alexander A. Sawchuk are with the Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089, USA mizhang@usc.edu, sawchuk@sipi.usc.edu

[2]Belinda Lange, Chien-Yen Chang, and Albert A. Rizzo are with the Institute for Creative Technologies, University of Southern California, Playa Vista, CA 90094, USA lange@ict.usc.edu, kchang@ict.usc.edu, rizzo@ict.usc.edu

978-1-4577-1787-1/12/$26.00 ©2012 IEEE

sequence of fixed-length tiny windows whose length is much smaller than the duration of the motor task itself (In this study, the duration of motor task ranges from 2 seconds to 10 seconds. The length of the tiny window we use is 0.2 second). Then we extract a number of features which capture the intrinsic characteristics of the motor behavior from each tiny window and stack them together to form a local feature vector. As a consequence, each motor task segment is transformed into a sequence of local feature vectors which forms a motion trajectory in the feature space. Compared to the "single-point" representation used in the existing methods, this trajectory-based representation provides more information about the patients' motor behavior in the sense that it captures the local details of the motor tasks in a fine-grained manner. Moreover, we have developed a trajectory comparison algorithm on top of the fine-grained representation, which helps clinicians to quantitatively keep track of the patients' progress during rehabilitation. In the remainder of this section, we explain the features we extract and the details of the trajectory comparison algorithm.

A. Feature Extraction

The raw signals sampled from wearable inertial sensors are not only noisy but also difficult for clinicians who have little engineering background to understand and interpret. Therefore, we need to extract features which contain useful information about the patients' motor behavior and more importantly, are meaningful and interpretable to clinicians.

Below is the list of features we use in this study. These features are selected since they have clear meanings related to the physical movements and have been demonstrated to be able to represent the important characteristics of the movements in existing studies [9] [6].

- **Mean Value of Movement Intensity (MI)**: Motion Intensity (MI) is defined as

$$MI(t) = \sqrt{a_x(t)^2 + a_y(t)^2 + a_z(t)^2}, \tag{1}$$

  the Euclidean norm of the total acceleration vector, where $a_x(t)$, $a_y(t)$, and $a_z(t)$ represent the t^{th} acceleration sample of the x, y, and z axis in each window respectively. The value of $MI(t)$ can be seen as an indirect measure of the instantaneous intensity (strength) of the performed movement at sample t. We calculate the mean value of MI within the window and use it as one of our features.
- **Movement Intensity Variation (VI)**: VI is computed as the variation of MI defined above. It is intended to measure the strength variation (range) of the movements.
- **Smoothness of Movement Intensity (SI)**: SI is computed as the derivative values of MI. It is used in our study to measure the smoothness of the movements.
- **Averaged Acceleration Energy (AAE)**: AAE calculates the mean value of energy as the sum of the squared discrete FFT component magnitudes of the sensor signals over three accelerometer axes. It measures the total movement acceleration energy.
- **Averaged Rotation Energy (ARE)**: ARE calculates the mean value of energy over three gyroscope axes. It measures the total movement rotation energy.
- **Time Taken to Complete the Task (TIME)**: The time duration is also used as a metric to indirectly measure the degree of difficulty in completing the movement.

B. Trajectory Comparison

The constructed motion trajectory is expected to capture the intrinsic characteristics of the patients' motor behavior. Although clinicians can find critical patterns and compare pattern differences between trajectories to track patients' progress by just visual observation, it is helpful to compare the trajectories and measure the differences in a quantitative and objective manner. However, one of the biggest challenges for the comparison is that trajectories from any two motor task segments may have different lengths. In this work, we develop a trajectory comparison algorithm based on dynamic time warping (DTW) technique to resolve this issue. DTW is a nonlinear alignment technique for measuring similarity/difference between two signals (normally time series) which may have different lengths or durations [10]. One classical application of DTW is to accommodate different speaking speeds in the domain of automatic speech recognition. For our case, DTW is used to cope with different movement speeds when patients perform motor tasks. Specifically, let $\boldsymbol{X}$ and $\boldsymbol{Y}$ denote two trajectories constructed from two motor task segments of length M and N respectively:

$$\boldsymbol{X} = \boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_i}, \ldots, \boldsymbol{x_M} \tag{2}$$

$$\boldsymbol{Y} = \boldsymbol{y_1}, \boldsymbol{y_2}, \ldots, \boldsymbol{y_j}, \ldots, \boldsymbol{y_N} \tag{3}$$

where $\boldsymbol{x_i}$ and $\boldsymbol{y_j}$ represent the ith and jth local feature vector in $\boldsymbol{X}$ and $\boldsymbol{Y}$ respectively. DTW compensates for the length differences and finds the optimal alignment between $\boldsymbol{X}$ and $\boldsymbol{Y}$ by solving the following dynamic programming (DP) problem:

$$\begin{aligned} D(i,j) = &\min\{D(i-1,j-1), D(i-1,j), D(i,j-1)\} \\ &+ d(i,j) \end{aligned} \tag{4}$$

where $d(i,j)$ represents the distance function which measures the local difference between local feature vector $\boldsymbol{x_i}$ and $\boldsymbol{y_j}$ in the feature space, and $D(i,j)$ represents the cumulative (global) distance between sub-trajectory $\{\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_i}\}$ and $\{\boldsymbol{y_1}, \boldsymbol{y_2}, \ldots, \boldsymbol{y_j}\}$. The solution of this DP problem is the cumulative distance between the two trajectories $\boldsymbol{X}$ and $\boldsymbol{Y}$ which sits in $D(M,N)$ and a warp path $\boldsymbol{W}$ of length K

$$\boldsymbol{W} = w_1, w_2, \ldots, w_k, \ldots, w_K \tag{5}$$

which traces the mapping between $\boldsymbol{X}$ and $\boldsymbol{Y}$. Finally, since the cumulative distance $D(M,N)$ is dependent on the length of the warp path $\boldsymbol{W}$, we normalize $D(M,N)$ by dividing it by the warp path length K and use this averaged cumulative distance as the metric to measure the distance between trajectories $\boldsymbol{X}$ and $\boldsymbol{Y}$ as

$$Dist(\boldsymbol{X}, \boldsymbol{Y}) = \frac{D(M,N)}{K} \tag{6}$$

It should be noted that many forms of distance function $d(i,j)$ (e.g. Euclidean distance and Mahalanobis distance) can be used to calculate the local difference. In this work, we use the cosine distance as the local distance function defined as

$$d(i,j) = 1 - \frac{\boldsymbol{x_i}^T \cdot \boldsymbol{y_j}}{\|\boldsymbol{x_i}\| \cdot \|\boldsymbol{y_j}\|} \tag{7}$$

Compared to other distance functions, the benefit of using the cosine distance is that $d(i,j)$ is by nature in the range $[0,1]$. As a result, the averaged cumulative distance $Dist(\boldsymbol{X},\boldsymbol{Y})$ defined in Eq.(6) is also in the range $[0,1]$, and thus can be interpreted as the dissimilarity between $\boldsymbol{X}$ and $\boldsymbol{Y}$ in terms of percentile. Therefore, we can also define the corresponding similarity in percentile between $\boldsymbol{X}$ and $\boldsymbol{Y}$ as:

$$Sim(\boldsymbol{X},\boldsymbol{Y}) = 1 - Dist(\boldsymbol{X},\boldsymbol{Y}) \tag{8}$$

III. Evaluation

A. Experimental Setup

To evaluate the effectiveness of our approach, three subjects including one healthy subject (female) and two subjects (one male and one female) with different levels of upper limb hemiparesis from stroke were recruited at the Precision Rehabilitation Clinic[1] and Rancho Los Amigos National Rehabilitation Center[2] located in Los Angeles. The wearable inertial sensor we use for this study is called MotionNode[3](see Figure 1(a)). MotionNode is a high-performance inertial measurement unit (IMU) that can sense $\pm$6g acceleration and $\pm$500dps rotation rate. This is high enough to capture all the details of the patients' movements. In addition, the size of MotionNode is extremely small such that it can be attached to the patient's body nonintrusively.

During data collection, one MotionNode was attached to the forearm of the subjects (see Figure 1(b)). Each subject followed the instructions from a physical therapist and performed a subset of five upper limb motor tasks from the Fugl-Meyer Assessment (FMA). We choose FMA since it is well-known for its comprehensiveness as a measure of motor impairment after stroke and it is widely recommended for motor rehabilitation for post-stroke patients [11]. These five motor tasks include: (1) Flexor Synergy; (2) Hand to Lumbar Spine; (3) Shoulder Flexion; (4) Pronation; and (5) Supination (See Table I for detailed explanations). Each motor task was repeated five times by each subject for both affected and unaffected limbs and was assigned a functional ability score based on the FMA scale (0 = can not perform, 1 = performs partially, 2 = performs fully) by the therapist.

(a) MotionNode sensor

(b) The placement of MotionNode on the upper limb of the subject

Fig. 1. Wearable sensor and experimental setup

[1]http://www.precisionrehabilitation.com
[2]http://www.rancho.org/
[3]http://www.motionnode.com/

Motor Task	Description
Flexor Synergy	Fully supinate the forearm, flex the elbow, and bring the hand to the ear of the affected side
Hand to Lumbar Spine	Move the hand behind the back
Shoulder Flexion	Flex the shoulder to 90°, keeping the elbow extended
Pronation	Flex the elbow to 90° and pronate the forearm through the full available range of motion
Supination	Flex the elbow to 90° and supinate the forearm through the full available range of motion

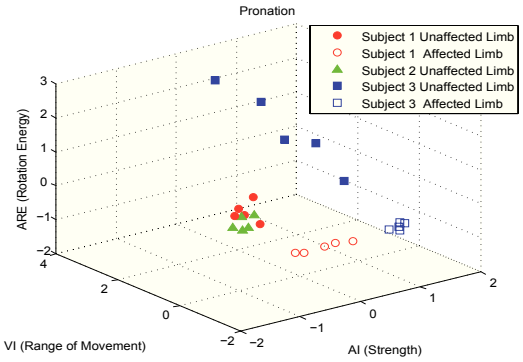
TABLE I

MOTOR TASKS CONSIDERED IN OUR STUDY

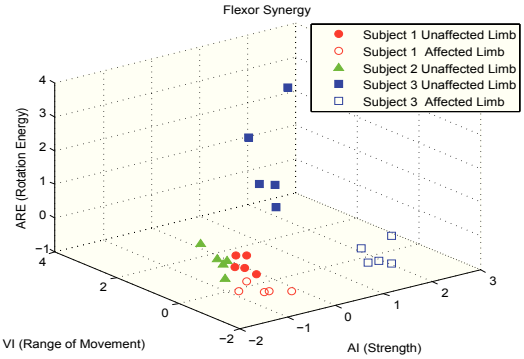
B. Experimental Results

1) Fine-Grained Trajectory-based Representation: In order to demonstrate better the benefits of our fine-grained trajectory-based approach, we first implemented the traditional motor function assessment method where each motor task segment is represented as a single point in a high dimensional feature space. Figure 2(a) and Figure 2(b) show the scatter plots in the 3D feature space for motor task Pronation and Flexor Synergy respectively. The three features used for the plots are AI, VI, and ARE. Here subject 1 is the female patient with upper limb hemiparesis, subject 2 is the healthy female, and subject 3 is the male patient with upper limb hemiparesis. As shown in Figure 2(a), data from different subjects and limbs forms compact clusters. Each cluster is well separated from others except one case where the data from the unaffected limbs of subject 1 and subject 2 overlaps. This observation can be explained by the fact that subject 1 and subject 2 are both female and the motor tasks are all performed by their unaffected limbs. For Figure 2(b), data is aggregated as compact clusters as previous example. However, the cluster formed by the data from subject 1's affected limb is very close in distance to the other two clusters formed by the data from subject 1 and subject 2's unaffected limbs. Although we can learn a classifier (e.g. nearest neighbor, SVM) to find a boundary to partition different clusters and then map the clusters to different clinical rating scores as in many existing research work, we argue that the distances between the data points in the feature space does not reflect their true differences.

In comparison, Figure 3 and Figure 4 illustrate our fine-grained trajectory-based solution in time-feature space to tackle the problem observed in the traditional method mentioned above. As illustrated, our fine-grained approach is capable of capturing the detailed patterns of the patients' motor behavior which traditional methods fail to acquire. As an example, Figure 3(a) and Figure 3(b) show the fine-grained trajectory representation of one segment of motor task Pronation in terms of feature AI and ARE respectively. The red curve represents the task performed by the



(a) Scatter plot of the motor task Pronation in 3D feature space



(b) Scatter plot of the motor task Flexor Synergy in 3D feature space

Fig. 2. The 3D scatter plots of the traditional automatic motor function assessment method. Subject 1 is the female patient with upper limb hemiparesis, Subject 2 is the healthy female, and Subject 3 is a male patient with upper limb hemiparesis. The three features used for the plots are AI, VI, and ARE.

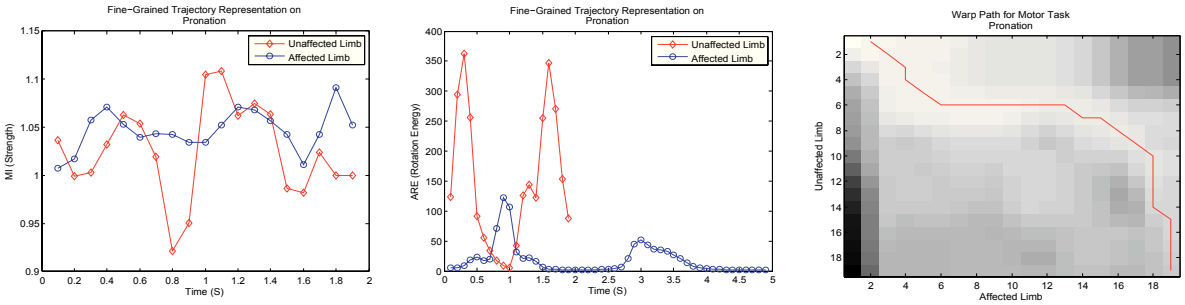
unaffected limb and the blue one for the affected limb. In Figure 3(a), our approach is able to capture the two troughs as patterns for the unaffected limb which does not appear for the affected limb. Similarly, two high peaks are captured for the unaffected limb in Figure 3(b) which represent the two key points during the movement of Pronation where significant rotation energy is exerted. As another example, Figure 4 shows the trajectories of motor task Flexor Synergy. In Figure 4(a), the curve of the unaffected limb is smooth with two peaks and two troughs detected while the curve of the affected limb shows less motion intensity but more fluctuations during the movement. This detailed difference is by no means reflected by the 1 point difference in the FMA score assigned by the clinician where the segment of unaffected limb is given 2 points and the segment of affected limb is given 1 point. Finally, similar conclusion can be drawn for Figure 4(b) as in Figure 3(b).

2) Trajectory Comparison Performance using DTW: As shown in the previous subsection, our trajectory-based approach captures the detailed pattern differences of the patients' motor behaviors. Here, we apply the DTW technique and the dissimilarity/similarity metric defined in Eq.(6)/Eq.(8) in percentile to quantitatively measure the differences/similarities between the fine-grained trajectories of the affected and unaffected limbs. Specifically, we first use DTW to extract the warp path of the two trajectories to be compared. Figure 3(c) and Figure 4(c) show two extracted warp path examples (the red curves overlaid on the grayscale similarity matrix, where dark entries indicate less similarity) for motor task Pronation and Flexion Synergy respectively. Based on the extracted warp paths, we then compute the similarity metric values in percentile between each motor task segment performed by the affected limb and the unaffected limb for each subject and all five motor tasks. The value of similarity indicates the recovery status of the affected limb. In other words, the higher the similarity value is, the better the performance of the affected limb. We also compare the trajectory similarity between segments from the same motor task of the unaffected limb. This is useful to validate whether our similarity metric is robust to noise or not. The accumulated results are listed in Table II. Since subject 2 is healthy, we only show the results for subject 1 and subject 3. For each subject, the first row in the table lists the averaged similarity between segments of affected limb and unaffected limb with the standard deviation in bracket. The second row shows the corresponding averaged FMA scores assigned by the professional physical therapist. The third row lists the similarity between segments from the same motor task of the unaffected limb with the corresponding FMA scores in the forth row.

For all five motor tasks, the trajectory similarity between segments from the same motor task of the unaffected limb is very high. This observation indicates that our similarity metric is robust to noise incurred from the patients' movements. The more interesting cases lie in the comparison between unaffected limb segment and affected limb segment. As shown in the table, their similarity values vary significantly across different motor tasks. More importantly, the calculated similarity values show a positive correlation to the FMA scores assigned by the therapist. For example, the two highest similarity values of subject 1 come from the motor task Flexor Synergy (70.55%) and Hand to Lumbar Spine (72.10%). These two motor tasks also have the highest FMA scores. Similarly, the motor task which gets the highest similarity value from subject 3 also has the highest FMA score. On the other hand, for all the motor tasks which get FMA score zero, the similarity values range from 33.30% to 49.39%. This indicates that our similarity metric can show more intermediate levels which would be extremely valuable for clinicians to track patients' gradual progress which are not reflected by the standard clinical scores.

IV. Conclusion and Future Work

In this paper, we propose a fine-grained approach that uses wearable inertial sensors to evaluate the motor function of the patients with stroke. The experimental results validate the effectiveness of our approach in capturing the detailed patterns that standard clinical scores fail to reflect. It should be noted that our approach is a general technique for fine-grained motor function assessment that can be used for many neurological injuries besides stroke. Therefore, based on the promising results reported in this paper, we plan to apply

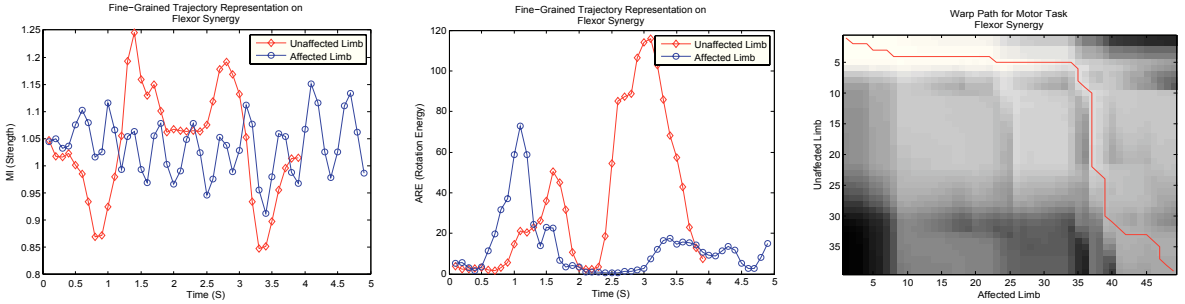


(a) The fine-grained trajectory representation of Pronation using feature AI

(b) The fine-grained trajectory representation of Pronation using feature ARE

(c) The warp path (red curve) of Pronation between affected and unaffected limbs overlaid on the grayscale similarity matrix

Fig. 3. The fine-grained trajectory representation and the warp path calculated from DTW of Pronation



(a) The fine-grained trajectory representation of Flexor Synergy using feature AI

(b) The fine-grained trajectory representation of Flexor Synergy using feature ARE

(c) The warp path (red curve) of Flexor Synergy between affected and unaffected limbs overlaid on the grayscale similarity matrix

Fig. 4. The fine-grained trajectory representation and the warp path calculated from DTW of Flexor Synergy

		Flexor Synergy	Hand to Lumbar Spine	Shoulder Flexion	Pronation	Supination
Subject 1	Similarity	70.55% (±8.20%)	72.10% (±6.72%)	62.10% (±8.06%)	33.30% (±7.42%)	41.17% (±4.27%)
	FMA	1.6	1.6	1	0	0
	Similarity	90.18% (±5.29%)	85.17% (±8.71%)	85.53% (±7.46%)	91.27% (±3.32%)	87.65% (±5.06%)
	FMA	2	2	2	2	2
Subject 3	Similarity	51.46% (±5.32%)	46.28% (±6.90%)	49.39% (±10.53%)	41.54% (±9.28%)	43.68% (±7.74%)
	FMA	0.6	0	0	0	0
	Similarity	83.37% (±7.38%)	79.04% (±10.51%)	85.13% (±9.27%)	91.89% (±5.64%)	90.66% (±6.93%)
	FMA	2	2	2	2	2

TABLE II

TRAJECTORY COMPARISON RESULTS USING DTW AND THE CORRESPONDING FMA SCORES FOR SUBJECT 1 AND SUBJECT 3

our approach to a larger sample of patients with neurological injuries such as spinal cord injury in the future.

REFERENCES

[1] American Heart Association. Heart disease and stroke statistics 2012 update: A report from the american heart association.

[2] N. Gordon et al. Physical activityand exercise recommendations for stroke survivors: an american heart association scientific statement from the council on clinical cardiology, subcommittee on exercise, cardiac rehabilitation, and prevention; the council on cardiovascular nursing; the council on nutrition, physical activity, and metabolism; and the stroke council. *Circulation*, 109:2031–2041, 2004.

[3] A. Mirelman et al. Effects of virtual reality training on gait biomechanics of individuals post-stroke. *Gait & Posture*, 31(4):433–437, 2010.

[4] M. Zhang and A.A. Sawchuk. A customizable framework of body area sensor network for rehabilitation. In *International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL)*, pages 1–6, Bratislava, Slovak Republic, November 2009.

[5] T. Hester et al. Using wearable sensors to measure motor abilities following stroke. In *International Workshop on Wearable and Implantable Body Sensor Networks*, pages 4–8, London, April 2006.

[6] S. Patel et al. Analysis of the severity of dyskinesia in patients with parkinson's disease via wearable sensors. In *International Workshop on Wearable and Implantable Body Sensor Networks*, pages 123–126, London, April 2006.

[7] V. Bento et al. Towards a movement quantification system capable of automatic evaluation of upper limb motor function after neurological injury. In *EMBC*, pages 5456–5460, Boston, September 2011.

[8] S. Patel et al. Monitoring motor fluctuations in patients with parkinson's disease using wearable sensors. *IEEE Transactions on Information Technology in Biomedicine*, 13(6):864–873, 2009.

[9] M. Zhang and A.A. Sawchuk. A feature selection-based framework for human activity recognition using wearable multimodal sensors. In *International Conference on Body Area Networks (BodyNets)*, Beijing, China, November 2011.

[10] E. Keogh and M. Pazzani. Derivative Dynamic Time Warping. In *SIAM International Conference on Data Mining*, Chicago, April 2001.

[11] D. Gladstone et al. The Fugl-Meyer Assessment of Motor Recovery after Stroke: A Critical Review of Its Measurement Properties. *Neurorehabilitation and Neural Repair*, 16(3):232–240, 2002.