Automated Administration of the Wolf Motor Function Test for Post-Stroke Assessment

Eric Wade†
Department of Biokinesiology and
Physical Therapy
ericwade@usc.edu

Avinash Rao Parnandi†
Department of Electrical
Engineering
parnandi@usc.edu

Maja J. Matarić†
Professor of Computer Science
and Neuroscience
mataric@usc.edu

†University of Southern California, Los Angeles, CA, USA

Abstract—The advent of new health sensing technologies has presented us with the opportunity to gain richer data from patients undergoing clinical interventions. Such technologies are particularly suited for applications requiring temporal accuracy. The Wolf Motor Function Test (WMFT) is one such application. This assessment is an instrument used to determine functional ability of the paretic and non-paretic limbs in individuals post-stroke . It consists of 17 tasks, 15 of which are scored according to both time and a functional ability scale. We propose a technique that uses wearable sensors and performance sensors to estimate the timing of seven of these tasks. We have developed a sensing framework and an algorithm to automatically detect total movement time. We have validated the system's accuracy on the seven selected WMFT tasks. We also suggest how this framework can be adapted to the remaining tasks.

I. INTRODUCTION

Occurrences of stroke are severely taxing on the population. In the U.S. alone, upwards of 800,000 strokes occur each year, with around 400,000 affected individuals surviving with some form of motor or cognitive deficit [6]. In many cases, this motor deficit takes the form of hemiparesis, or weakness or loss of function in one side of the body. It is known that, by working intensely with trained clinicians, those suffering from hemiparesis can often regain partial use of the affected limb. However, the numbers of people suffering from stroke is quickly outpacing the number of trained clinicians who can help this population to regain motor functionality. Thus, there is a large and growing gap between the quantity and duration of therapeutic interventions that should be administered, and that which we are able to provide. This has led to a push innovative approaches to physiological interventions. To evaluate the efficacy of these interventions, suitable tools are required. The tools used to do this are referred to as assessment instruments. A number of instruments have been developed to evaluate functional recovery after stroke. For instance, some examples of instruments for determining upper-extremity functional

ability are the Chedokee-McMaster (CM) assessment, the Fugl-Meyer (FM) assessment, and the Wolf Motor Function Test (WMFT) assessment [20], [21], [25]. These assessments can indicate the extent to which an intervention is effective in helping the patient to regain a dimension of functional motor capability.

Because assessment instruments utilize subjective measures, it is critical that these assessments are valid, reliable, and administered according to a standard. Within a study, there are numerous variables that can affect the administration of an assessment. For instance, in larger trials, there can be multiple sites and multiple examiners. Such variability has led to studies that seek to develop proper standardization, organization, and administration of large multi-site trials [9], [19].

Motivated by these concerns, we propose the technological solution of automating assessment instruments for motor task rehabilitation. We suggest using a wearable motion sensor and an overhead camera to administer the WMFT. The WMFT is ideal for a number of reasons. First, it is a reliable and valid instrument for determining upper extremity (UE) functional ability [25]. Second, it is scored according to time and a functional ability scale. An instruction manual indicates how the tasks should be measured and scored. According to this manual, the task movement time is obtained by a clinician with a stopwatch. Using our sensors, we can obtain this timing information for various WMFT tasks with a higher degree of accuracy and consistency.

An added benefit of automating assessments is the potential to save time for clinicians. The WMFT can take more than two hours to complete on both limbs. For large clinical trials, there may be many clinicians capable of administering assessments. However, smaller trials have limited resources and access to clinicians. By automating the assessment, we can shift time normally spent in the provision of the assessment to the provision of the therapeutic intervention.

In the following, we will describe our system's capabilities and the WMFT tasks on which it can be used. We will discuss the results of a technological feasibility study, the system accuracy, and the extension of our system to the remaining WMFT tasks. Finally, we will discuss the next steps for the technology.

1. Forearm to table (side)
2. Forearm to box (side)
3. Extend elbow (side)
4. Extend elbow (weight)
5. Hand to table (front)
6. Hand to box (front)
7. Weight to box*
8. Reach and retrieve
9. Lift can
10. Lift pencil
11. Lift paper clip
12. Stack checkers
13. Flip cards
14. Grip strength*
15. Turn key in lock
16. Fold towel
17. Lift basket

* Indicates tasks that are not timed.

II. THE WMFT ASSESSMENT

The Wolf Motor Function Test (WMFT) assessment is designed to evaluate upper extremity (UE) motor deficits for individuals post-stroke or traumatic brain injury (TBI) of mild to moderate severity. It was designed to help link treatment planning with the desired functional restitution for a patient. The assessment requires the participant to perform 17 functional tasks (see Table I). The WMFT utilizes time and a functional ability (FA) scale to score the tasks (15 of the tasks receive this score; the remaining 2 are strength tests). The tasks are arranged in order of increasing complexity, and from proximal to distal joints. The WMFT has been shown to be both valid and reliable [25].

The 17 tasks include gross motion (e.g. lifting the arm from the lap to the top of a box on a table), forearm pronation/supination (e.g. turning a key 180°), dexterity (e.g. flipping cards and stacking checkers), and grip strength (measured with a dynamometer). Most WMFT tasks require a table and chair, both of which are standard in dimension, and placed with a set orientation and distance with respect to each other. For tasks that require the tabletop, there is a template that is taped down to indicate object placement and success conditions (e.g. when extending the elbow, the thumb must pass a specific line). The template is symmetric, as the WMFT is done for both limbs. The test is also video taped, with video cameras placed in pre-determined, fixed locations.

It is important to note that, while the timing is measured in real-time, the FA score is often determined post-hoc after watching video data. Thus, automating the time measurement is effectively automating the assessment. In this paper, we describe how to obtain the time score for these tasks using our sensor framework. We note that there are aspects of the WMFT that are not accommodated by our system as yet (e.g. offering encouragement during task performance). We explore these, and the extension of this work to additional WMFT tasks, in Section VII.

III. BACKGROUND: AUTOMATED ASSESSMENT

Automated assessment is a sub-category of technological approaches to health care known as "telemedicine." The dominant approach to automation of assessments has been to use technological devices to quantitatively measure a physiological output and to determine its relationship to functional ability. Mazzoleni et al. developed a robotic apparatus through an iterative design process with clinicians, stroke patients, and engineers [17]. The robotic apparatus is designed to measure force and torque (F/T) applied by the user during the performance of 6 activities of daily living (ADL) tasks [15]. The system restricts users to isometric movements and makes use of forward models to determine patient-generated forces and torques. Balasubramanian et al. developed a new apparatus that also used F/T data [7]. Using a robotic device, they sought to quantify the required amount of assistance, motion smoothness, and movement synergy in participant motions. Both papers focus on the correlation of quantitative F/T measures with functional motor characteristics.

A large body of work in this field focuses on the correlation of quantitative physiological measures to existing assessment instruments, including those mentioned in Section I. Cinkelj and Van Dijck further Mazzoleni's work by describing the robot's software and by evaluating posterior probability profiles of the robotic system [8], [10]. They are able to show correlations between the quantitative F/T measures and the FM score.

An alternative approach to robot-aided assessments is to use simpler, sensor-based systems. One feature that is assessed in patients post-stroke is relative levels of proprioception in the affected limb. Leibowitz et al. focused on this issued by developing a system where sensors are worn on the hand; the participant is asked to move their hands to target locations indicated by a screen [16]. This system was able to measure proprioception with more accuracy and resolution than standard "up/down" tests administered in practice. Bonato et al. affixed tri-axial accelerometers to a subject's arm [14]. They then performed the WMFT and compared accelerometer readings for the participants. The goal of the study was to determine if different functional impairment levels could be distinguished in the raw accelerometer data. Hester et al. correlated accelerometer data from the trunk and arms to CM, FM, and WMFT scores using a linear regression technique [13]. They extracted a number of temporal and spectral features from the signals, and evaluated correlations with the WMFT FA score.

Our approach is to systematically evaluate the WMFT tasks with the goal of automation. As we have suggested, we focus on timing because timing is the only measure taken in real-time. Automating the timing of the WMFT accomplishes our stated goals of minimizing the burden on and need for trained clinicians. Subsequent evaluation of the video data can be used to determine the FA scores.

IV. AUTOMATED ASSESSMENT FRAMEWORK

We have devised an experimental setup to evaluate the efficacy of our system that uses a combination of wearable

sensors and performance sensors.

A. Wearable Sensors

In our experimental setup, the user wears a single sensor on the wrist of the limb under evaluation. The wearable sensor used in this study is an inertial measurement unit (IMU) developed in the USC Interaction Lab. This device has been validated in previous studies [22], [24]. The IMUs rely on inertia-based MEMS sensors to obtain motion information from the user. Each IMU contains a 3-axis accelerometer, three single-axis rate gyros, and one single- and one dual-axis magnetometer. This design is an updated version of that presented in our previous work [18].

The IMU sensor outputs are low-pass filtered (using analog filters) and sampled by a 10-bit ADC. An Atmel© ATmega324 microcontroller reads the digitized sampled data, arranges them into individual packets, and transmits each packet (along with a timestamp) to the central computer using the I^2C bus interface. Each IMU is enclosed in a custom plastic enclosure and secured to the user using a Velcro strip (Figure 1). The lightweight enclosure has a concave shape in order to contour to the human body.
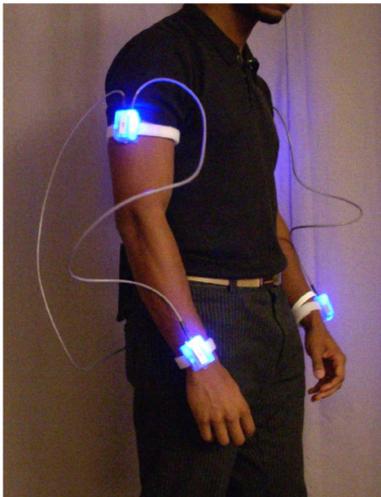

Fig. 1. The motion suit worn by the participants consists of the IMUs and wearable computer. In our experiments, only one IMU is used.

B. Wearable Computer

The central control unit of the motion suit contains a computer, a battery pack, a power converter, and other discrete components (see Figure 2). Our chosen hardware platform is the Gumstix©Wifi-Stix pack, a small Linux computer (80mm×20mm) with a 400MHz processor with 64MB of RAM. Wireless communication is facilitated by a Netwifi-MicroSD wireless transmitter. The transmitter, which includes an antenna, is FCC-certified and can use the 802.11(b) and 802.11(g) transmission protocols. The small size and form factor allow for easy packaging of the computer for wearable applications. It is powered by a polymer lithium ion cell phone battery pack. The Gumstix has a large developer community, and software for a number of interface devices

(including the I^2C and SPI data transmission protocols) is readily available. An advantage of the Gumstix platform is that it lends itself to development for various software suites, such as the open source Player/Stage development environment that we have employed [12]. The entire system is called the motion suit, and is pictured in Figure 1.

Fig. 2. The central controller worn on the participant's body contains a computer (Gumstix), battery, DC-to-DC converter, and a plastic housing.

C. Performance Sensors

The performance sensor used in this study is an overhead camera. By marking the work surface appropriately, an overhead camera can be calibrated to look for specific events (e.g. the arm crossing a line or the hand covering a strip). We used a Logitech QuickCam Pro 5000 USB camera for our vision system. This camera, when connected to a computer, incurs an average delay of 0.114s in capturing and processing 100 images (.001s/image).

D. Software Drivers

As described in our previous work, the sensors and controller are integrated using the Player/Stage framework. Player is an open source software suite that allows for the control and coordination of multiple devices using a server/client architecture. It is optimized for robot programming and control, and its design, which treats each driver as an input/output device, is suitable for our framework. By using the Player camera driver, and our own drivers for the IMUs and Gumstix, we can coordinate the data received from each device and feed it into a master controller. This controller maintains models of the user state (e.g. "arm is moving") and the trial state (e.g. "we are on task 3").

Real time image processing is done using the Open Computer Vision Library (OpenCV) [5], [4]. OpenCV is an open source computer vision library written in C/C++. It is independent of the operating system and hardware, and is optimized for real-time applications. In our system, for each WMFT task trial, the image processing code checks a region of interest (ROI) in each captured image to detect the success condition of a gesture. The ROI is a fixed section of the image which corresponds to markings we have made on the

table surface for the detection of success conditions for each task. For example, in the WMFT tasks that have a success condition denoted by *"the forearm touching the table and covering a line"* (Tasks 1, 2, 5, 6, and 8), the ROI corresponds to a marking on the table to which the participant is told to move their arm. The vision system looks for this specific cue by continually computing the total intensity (sum of intensity) of all the pixels in the ROI. This intensity value is compared to a predetermined threshold. Variations in this intensity indicate whether or not a gesture has been completed (we are making the reasonable assumption of standard lighting conditions.). For tasks which involve the thumb crossing a line (Tasks 3 and 4), it again looks at the intensity value of the ROI to detect when the line was occluded and when it becomes visible. We have developed tools for calibration and intensity threshold computation for this system.

V. METHODS

We designed a technological feasibility study to determine if our system could accurately measure timing of the various WMFT tasks. Our hypothesis was that the system would obtain more accurate timing scores than those obtained from a stopwatch.

The participants wore an IMU on the wrist of the arm being tested, and were monitored by an overhead camera (Figure 3). There are markings on the table, used by the monitoring system to determine when various task success conditions have been met. The participants performed tasks from the WMFT (Tasks 1, 2, 3, 4, 5, 6, and 8, listed in Table II. These tasks were chosen because they require only the IMU and overhead camera). Two participants were recruited for the study. The participants were healthy individuals with no apparent motor limitations.

Fig. 3. A participant performing the WMFT assessment. The participant is being monitored by an overhead camera and an IMU on the wrist.

The examiner administered the seven selected WMFT tasks, and timed them using a stopwatch. The examiner followed the protocol in the WMFT Instruction manual. These instructions specify the following: participant starting position (location with respect to the table, initial position of the limb being tested), task description, the timing procedure,

TABLE II
COMPARISON OF AUTOMATICALLY MEASURED TIME AND THE GROUND TRUTH TIME

Task	T_{auto} [s]	T_{gnd} [s]
1. Forearm to table (side)	2.43	2.50
2. Forearm to box (side)	1.65	1.65
3. Extend elbow (side)	2.35	2.50
4. Extend elbow, weight (side)	2.29	2.30
5. Hand to table (front)	2.60	2.55
6. Hand to box (front)	1.56	1.65
8. Reach and retrieve	2.03	2.15

verbal instructions, and the FA scoring metric. The tasks were administered on the participants' right arms (because these were healthy subjects, differences in the right and left arm scores were not expected to be significant). For each task, the automated time estimate is defined as T_{auto}. The time measured by the examiner's stopwatch is defined as T_{clock}. As specified in the manual, this is the time between the experimenter uttering *"Ready, set, go!"* and the completion of the task-specific success condition.

VI. RESULTS

A. Obtaining the Ground Truth Movement Time

Our first goal was to determine the temporal accuracy of our measurement system. To do so, we compared the measured movement time, T_{auto}, with the ground truth movement time, T_{gnd}, which we obtained from the IMU. To obtain T_{gnd}, we simultaneously recorded overhead camera and IMU data while a subject performed the WMFT tasks. We used a button on the experiment laptop to begin timing. The timer continued to run until the camera determined a task success condition. For instance, on Task 1, Forearm to Table (side), the WMFT instructions indicate that gesture is complete once *"the forearm and hand touch the table in the required fashion."* The camera determined that the gesture was complete once the table marking, specified *a priori*, was occluded by the participant's arm and hand. After the completion of these tasks, we hand-annotated the IMU data to determine the actual movement time. This was done by prototyping various motions, and observing the accelerometer and rate gyro signals. We found the start and end points of the gestures using an empirically determined threshold for the change in signal variance. We compared the IMU-obtained time to the automatically measured time (The results are shown in Table II.). By using the IMU data as a ground truth, we determined the RMS error of the automated timer to be 0.22s. Thus, the camera-based timing system is an accurate indicator of movement time.

B. WMFT Task Timing Results

Having determined the accuracy of the automated time measure, we next sought to compare the automated time to the time obtained when the examiner used a stopwatch. The examiner instructed the two participants to repeat the WMFT tasks. The automated and stopwatch-measured movement times for both participants are indicated in Table III. We

Task	T_{clock} [s]	T_{auto} [s]
Participant 1		
1. Forearm to table (side)	2.18	1.24
2. Forearm to box (side)	2.66	1.58
3. Extend elbow (side)	1.78	1.13
4. Extend elbow, weight (side)	1.81	1.01
5. Hand to table (front)	2.35	1.10
6. Hand to box (front)	2.03	1.11
8. Reach and retrieve	2.00	1.04
Participant 2		
1. Forearm to table (side)	2.03	1.03
2. Forearm to box (side)	2.75	1.72
3. Extend elbow (side)	2.03	1.28
4. Extend elbow, weight (side)	2.00	1.34
5. Hand to table (front)	2.31	1.32
6. Hand to box (front)	3.07	1.89
8. Reach and retrieve	1.60	0.60

originally hypothesized that T_{auto} would be shorter than T_{clock}. This is due to our observation that, when timing a movement using a stopwatch, there will always be a finite delay between the experimenter observing the completion of the task, and relaying that information to the finger to press the stopwatch button (this is a function of the examiner's reaction time). This delay is evident in the results presented in the table. A question that immediately comes to mind regards the extent to which this delay affects the measured time. For our examiner, the distribution of the error between T_{auto} and T_{clock} has mean $\mu = 0.94$ and variance $\sigma^2 = 0.03$. Thus, errors on the order of 1s were typical. To put this in perspective, the maximum time allowed for most WMFT tasks according to the manual is 120s. For those participants with lower functional abilities (and longer movement times), this error is of less significance. For those with higher functional abilities, the error must be taken into account.

VII. DISCUSSION

Our technological feasibility study was successful. During the preparation and administration of this experiment, we produced a number of interesting results and observations. In the following, we will discuss the issues of examiner bias, the difference between measured movement time and what we will define as gesture time, the requirements for full automation of the WMFT, extension of our system to the other WMFT tasks, and limitations of our system.

A. Examiner Bias

As we have discussed, the WMFT examiner's reaction time introduces an inherent delay that is included in the time score. For a trained examiner, we hypothesized that this error was a constant bias error. It has not affected the validity or reliability of the WMFT, presumably, because the relative movement times for stroke-affected individuals are typically much longer in duration than the examiner bias delay. Furthermore, if the same experimenter administers pre- and post-intervention WMFT assessments, the bias will

be present in both cases, and will be rendered negligible when the results are compared to each other. Each examiner will have a different bias – however, as long as this bias is relatively consistent, it does not pose a problem in measuring a patient's change in functionality. The error we empirically found between T_{auto} and T_{clock} follows from this assumption – the mean error was 0.94s, but had a relatively small variance of $\sigma^2 = 0.03$. In the future, we will validate this hypothesis by evaluating the error for different examiners.

B. Movement Time vs. Gesture Time

After performing the initial round of experiments, we realized that there can also be human-introduced error due to the variable delay between the experimenter saying *"Go!"* at the beginning of a bout, and the participant starting the movement. Our previously defined movement time T_{auto} is the elapsed time between the experimenter saying *"Go"* and the determination of the success condition. We define the gesture time T_{ges} as the elapsed time between the initiation of motion by the participant to the determination of the success condition. Thus, T_{ges} must always be less than or equal to T_{auto}. The delay between *"Go"* and the beginning of participant motion is equal to $T_{auto} - T_{ges}$. We conducted an initial exploration into the significance of this delay to determine how much it varied in our trials.

The participants repeated the seven WMFT tasks, but instead of using the button push on the laptop to determine the beginning of the gesture, we used the IMU data for the start condition (again, threshold values on the accelerometer and rate gyro signal variance were used for this condition). We then measured T_{ges} for the tasks. The original experiments, where a button-push was used to determine the start condition, will be defined as *vision-only*; the experiments that used the IMU for the start condition will be defined as *vision + IMU*. The results for the two participants are shown in Figures 4 and 5.
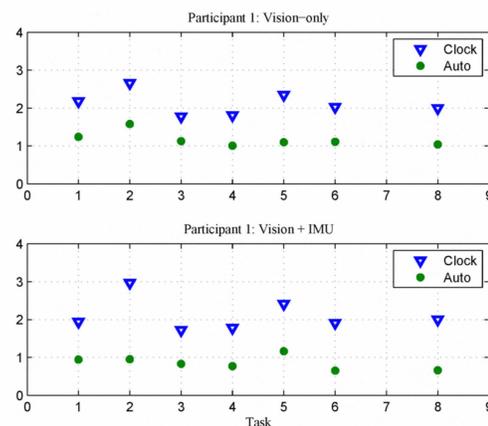

Fig. 4. A comparison of the time measured by the stopwatch and the automated system for Participant 1. The results for the *vision-only* are shown in plot 1, and the results for the *vision + IMU* are shown in plot 2.

As expected, in both cases, T_{auto} is always less than or equal to T_{clock}. It follows that the error between T_{auto} and
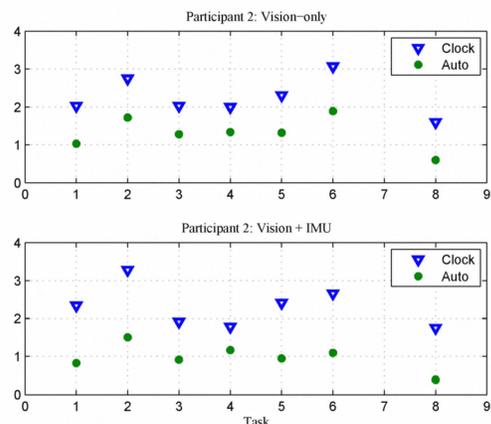
Fig. 5. A comparison of the time measured by the stopwatch and the automated system for Participant 2. The results for the *vision-only* are shown in plot 1, and the results for the *vision + IMU* are shown in plot 2.

TABLE IV
MEAN AND VARIANCE OF THE ERROR BETWEEN T_{auto} AND T_{clock} FOR *vision only* AND *vision + IMU* CONDITIONS

Condition	μ	σ^2
Participant 1: Vision only	0.94	0.04
Participant 1: Vision + IMU	1.25	0.14
Participant 2: Vision only	0.94	0.03
Participant 2: Vision + IMU	1.33	0.15

T_{clock} is larger in the *vision + IMU* condition. The mean and variance of the error between T_{auto} and T_{clock} for both participants are shown in Table IV.

C. Full Automation of the WMFT

To completely automate this system, a mechanism for delivering instructions to the participant must be automated. We have created such a system using a socially assistive robotic (SAR) device in our previous work [11]. SAR is the study of the use of non-contact robots to administer assistive therapeutic interventions to populations suffering from various deficits. To extend this work to be compatible with a SAR system, we can utilize the timing data to create a user state model, and a trial state model. The user state model is a constant estimate of the user's state. This consists of tags such as `user_resting`, `user_start`, `user_moving`, and `user_stop` that correspond to the user's state between tasks, the commencement of a task, motion, and the completion of a task (respectively). Meanwhile, the trial state model is a constant estimate of the task number, as well as the state of completion. The state of completion for each task includes tags including `task_pre`, `task_prog`, and `task_comp` corresponding to the time before the task is commenced, the task being progressed, and the task completion (respectively). Each state also includes different modes of operation. During pre-task, instructions are provided to the participant. For each task, once specific completion conditions are met (as indicated by the user model), the elapsed time is recorded as the task score.

With full automation these tags will correspond to the robot verbalizing or gesturing appropriately. This can include giving instruction, offering feedback, and providing encouragement. We have already piloted an assistive robotic agent with participants with Alzheimer's and stroke-related motor deficits [23], [22]. Thus, while we now require a human to administer the WMFT, we are actively working on the technological incorporation of the assistive agent.

D. Remaining WMFT Tasks

The seven WMFT tasks described here all make use of the same setup; a marked table, a wearable IMU, and an overhead camera. Other WMFT tasks would require a different combination of performance sensors. We have piloted the required technologies but have not fully validated them with the tasks. Here we will provide a brief outline on how the timing and administration of the remaining WMFT tasks can be automated.

Automation of Task 7, weight to box, can be accomplished by mounting a weight scale in the experiment table surface underneath the box. Task 9 involves lifting a can to the mouth. We are developing a C++ gesture recognition module to recognize and time the correct gesture with Hidden Markov Models (HMMs). For Tasks 10 and 11, we envision the use of the AcceleGlove motion capture glove to detect the desired finger motion [1]. Task 12 can be automated by using the aforementioned weight scales with appropriate threshold values. Automation of Task 13 can be accomplished with the existing setup by tracking the cards with augmented reality (AR) tags and detecting the required success condition using the overhead camera [2], [3]. AR tags allow for 3D tracking of objects in real-time. The tags are 2D symbols placed on objects in an environment. Once an overhead camera is calibrated with the locations of the tags, whenever they are moved, their position and orientation data is relayed to the controller. For Task 14, the dynamometer can be instrumented with a linear potentiometer, and the resulting resistance can be correlated with grip strength. For Task 15, a rotary potentiometer along with the wrist mounted IMU can be utilized to detect the start and end conditions. Finally, Task 16 can be administered by placing AR tags on the towel (in this case, the towel would have the AR tags calibrated while it is open. The subsequent occlusion of a tag would indicate folding of the towel.).

E. Limitations of our System

The methodical nature of the timing measure for assessments is a strong argument for automation. However, it goes without saying that automation removes the 'human element' of the assessment. One important characteristic of the human is recognizing failure scenarios. Specifically, these are motions or behaviors that would invalidate a score. For instance, most of the WMFT tasks are uni-manual. If a participant uses the opposite hand to aid in performing a task (typically, this occurs when the paretic limb is being assessed), the task is invalid, the participant is asked not to use the compensatory limb, and the task is repeated. For this

specific case, the use of the limb not under assessment can be captured by placing IMUs on both wrists, but we recognize that there are other scenarios that might occur. Through our interactions with our collaborators at the USC Health Sciences Campus, we are building models to determine the most relevant failure scenarios. We expect a pilot study, using approximately ten stroke-affected participants, to result in more insight into the additional sensing and processing required to capture failure scenarios.

As we outlined in the Introduction, with the work presented in this paper, we were focused on obtaining the timing score, and not the FA score (we mentioned other projects working in this arena). We are now in the preliminary stages of investigating the ability of our sensors to estimate the FA score. The problem is complicated because of the sometimes vague nature of the WMFT instructions for creating the FA score. For instance, on Task 10, the instructions state *"If participant slides the pencil away from the center of the table somewhat they can still get a "5" if it looks "normal" otherwise."* While this may become clear for a human receiving training from a clinician, translating such a standard to an automated system is challenging. Such a complication makes it easy to understand why the projects mentioned in Section III attempt to determine the FA score through correlation techniques. We intend to use pilot data from post-stroke individuals to begin formulating models for the determination of the FA score based on IMU and camera data.

Another characteristic of the WMFT is the provision of encouragement to the participant – during task performance, the examiner should offer encouraging and motivational phrases. Our SAR agent can offer encouragement to participants during the performance of various functional tasks [22]. In the future, we expect to extend these capabilities to the implementation of the WMFT.

VIII. CONCLUSION

We have described a technique for automating the timing of WMFT tasks. Specifically, we have completed a technological feasibility study with two participants. We have also presented a number of interesting technical and practical issues associated with the automation of motor task assessments. The next logical step is to validate with a larger number of participants, and with the full battery of WMFT tasks. We will also establish the robustness of our system by using more examiners.

REFERENCES

[1] Anthrotronix AcceleGlove sdk. *Available: http://www.acceleglove.com/products.asp.*
[2] AR Tag Applications. *Available: http://www.artag.net/.*
[3] AR Tag Toolkit. *Available: http://www.hitl.washington.edu/artoolkit/.*
[4] Introduction to programming with OpenCV. *Available: http://www.cs.iit.edu/~agam/cs512/lect-notes/opencv-intro/index.html.*
[5] OpenCV Library. *Available: http://opencv.willowgarage.com/wiki/.*
[6] American heart association: Heart disease and stroke statistics: 2009 update. 2009.
[7] S. Balasubramanian, R. Wei, R. Herman, and J. He. Robot-measured performance metrics in stroke rehabilitation. *Complex Medical Engineering*, pages 1 – 6, April 2009.
[8] J. Cinkelj, M. Mihelj, D. Bacciu, M. Jurak, E. Guglielmelli, A. Toth, J. D. LaFonteyne, J.-L. Vershelde, S. Mazzoleni, J. V. Vaerenbergh, S. de Ruijter, and M. Munih. Assessment of stroke patients by whole-body isometric force-torque measurements ii: software design of the alladin diagnostic device. *EMBC*, 2005.
[9] L. Criddle, C. Bonnono, and S. Fisher. Standardizing stroke assessment using the national institutes of health stroke scale. *J. Emerg. Nurs.*, 29(6):541 – 546, December 2003.
[10] G. V. Dijck, J. V. Vaerenbergh, and M. V. Hulle. Posterior probability profiles for the automated assessment of the recovery of stroke patients. *AAAI*, pages 347 – 353, 2007.
[11] D. Feil-Seifer and M. J. Matarić. Defining socially assistive robotics. *Proceedings, IEEE International Conference on Rehabilitation Robotics (ICORR-05)*, pages 465 – 468, June 2005.
[12] B. Gerkey, R. Vaughan, and A. Howard. The player/stage project: Tools for multi-robot and distributed sensor systems. *Proc. Int. Conference on Advanced Robotics*, pages 317 – 323, July 2003.
[13] T. Hester, R. Hughes, D. Sherrill, B. Knorr, M. Akay, J. Stein, and P. Bonato. Using wearable sensors to measure motor abilities following stroke. *Proc. of the Int. Workshop on Wearable and Implantable Body Sensor Networks*, pages 5 – 8, 2006.
[14] B. Knorr, R. Hughes, D. Sherrill, J. Stein, M. Akay, and P. Bonato. Quantitative measures of functional upper limb movement in persons after stroke. *Conference Proc. of IEEE EMBS, Neural Engineering*, pages 252 – 255, 2005.
[15] K. Krapp. *Activities of Daily Living Evaluation*. Gale Group, Inc., 2007.
[16] N. Leibowitz, N. Levy, S. Weingarten, Y. Grinberg, A. Karniel, Y. Sacher, C. Serfaty, and N. Soroker. Automated measurement of proprioception following stroke. *Disability and Rehabilitation*, 30(24):1829 – 1836, 2008.
[17] S. Mazzoleni, A. Toth, M. Munih, J. V. Vaerenbergh, G. Cavallo, S. Micera, P. Dario, and E. Guglielmelli. Whole-body isometric force/torque measurements for functional assessment in neuro-rehabilitation: platform design, development and verification. *J. Neuroeng. Rehabil.*, 6(38):1 – 15, Oct. 2009.
[18] N. Miller, O. Jenkins, M. Kallmann, and M. Matarić. Motion Capture from Inertial Sensing for Untethered Humanoid Teleoperation. *Proceeding of the IEEE-RAS International Conference on Humanoid Robotics (Humanoids)*, 2004.
[19] D. Morris, E. Taub, D. Macrina, E. Cook, and B. Geiger. A method for standardizing procedures in rehabilitation: Use in the extremity constraint induced therapy evaluation multisite randomized controlled trial. *Arch. of Phys. Medicine and Rehabilitation*, 90(4):663 – 668, April 2008.
[20] R. Perry. Chedoke-mcmaster stroke assessment development, validation and administration manual -. *Physiotherapy*, 82(3):211 – 357, March 1996.
[21] J. Sanford, J. Moreland, L. Swanson, P. Stratford, and C. Gowland. Reliability of the fugl-meyer assessment for testing motor performance in patients following stroke. *Phys. Ther.*, 73(7):447 – 454, July 1993.
[22] A. Tapus and M. Matarić. User personality matching with hands-off robot for post-stroke rehabilitation therapy. In *Proceedings, International Symposium on Experimental Robotics (ISER)*, Rio de Janeiro, Brazil, Jul 2006.
[23] A. Tapus, C. Tapus, and M. Matarić. Music therapist robot for people suffering from dementia: Longitudinal study. *Proceedings, International Conference on Alzheimer's Disease (ICAD 2009)*, July 2009.
[24] E. Wade and M. Matarić. Design and testing of lightweight inexpensive motion-capture devices with application to clinical gait analysis. *Proc. ICST 3rd International Conf. on Pervasive Healthcare*, pages 1 – 7, April 2009.
[25] S. Wolf, P. Catlin, M. Ellis, A. Archer, B. Morgan, and A. Piacentino. Assessing wolf motor function test as outcome measure for research in patients after stroke. *Stroke*, 32:1635 – 1639, 2001.