This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2017.2755667, IEEE Transactions on Neural Systems and Rehabilitation Engineering

Automated Evaluation of Upper-limb Motor Function Impairment using Fugl-Meyer Assessment

Seunghee Lee, *Student Member, IEEE*, Yang-Soo Lee, and Jonghyun Kim, *Member, IEEE*

***Abstract*— The Fugl-Meyer assessment (FMA) is the most popular instrument for evaluating upper extremity motor function in stroke patients. However, it is a labor-intensive and time-consuming method. This paper proposes a novel automated FMA system to overcome these limitations of the FMA. For automation, we used Kinect v2 and force sensing resistor sensors owing to their convenient installation as compared with body-worn sensors. Based on the linguistic guideline of the FMA, a rule-based binary logic classification algorithm was developed to assign FMA scores using the extracted features obtained from the sensors. The algorithm is appropriate for clinical use because it is not based on machine learning, which requires additional learning processes with a large amount of clinical data. The proposed system was able to automate 79% of the FMA tests because of optimized sensor selection and the classification algorithm. In clinical trials conducted with nine stroke patients, the proposed system exhibited high scoring accuracy (92%) and time efficiency (85% reduction in clinicians' required time).**

***Index Terms*—Stroke, Fugl-Meyer assessment, automated upper-limb assessment, rule-based binary logic classification**

I. Introduction

Stroke is one of the major disease and the second leading cause of death in humans [1], [2]. More than fifty percent of the patients with stroke experience decreased quality of life owing to motor function impairment [3]. Exercise-based interventions to restore motor function in stroke patients are accompanied by motor function assessment, and this assessment plays an important role in determining motor function status and intervention directions [4]–[6].

The Fugl-Meyer Assessment (FMA) is a representative evaluation tool for the upper extremities. It has been used to comprehensively evaluate upper extremity motor function, and it exhibited high inter-rater and intra-rater reliability [7]–[9]. However, the FMA is not performed frequently in clinical settings owing to the following problems: First, it is labor intensive; it includes several tests that should be conducted by clinicians [8], [9]. Second, its evaluation time is long (at least 30 minutes [8]), and thus, the clinicians' available time for intervention is reduced. These problems eventually lead to an increase in medical costs and decrease in the quality of rehabilitation services. Moreover, other clinical instruments, such as the Wolf Motor Function Test (WMFT) [10], Action Research Arm Test [11], NIH Stroke Scale [12], and Functional Ability Scale (FAS) [13], experience similar problems.

To solve these problems, a few studies have been conducted regarding automatic evaluation of the upper extremity motor function of stroke patients; however; they still have the following limitations: First, they primarily use body-worn sensors to acquire human motion data. For instance, accelerometers (or inertial measurement units (IMUs)) [14]–[16] and glove sensors [15], [16] were widely used; however, attaching them requires considerable time and they cause discomfort. Moreover, in our previous study [16], we found that the sensors were difficult to attach owing to the synergistic pattern and muscle contracture of stroke patients. Second, machine learning algorithms (such as random forests [14], extreme machine learning [15], support vector machines [16] and artificial neural networks [17]) are used to classify a clinical scale based on acquired data. However, these algorithms require a large amount of reliable data (sensor data and clinical score) and a learning process based on the data to derive a relationship between the features extracted from sensor data and clinical score [14]–[17]. Even though the quality and quantity of data directly affect classification accuracy [18], it is difficult and impractical to collect data from numerous patients with various motor abilities only for machine learning. Third, existing studies have automated only a small number of tests (7 [15] or 13 tests [17]) while the FMA consists of 33 upper extremity tests. It is clear that the higher the number of unautomated tests, the less effective it is to reduce evaluation time. Moreover, there is insufficient evidence that total FMA scores can be estimated using the scores of automated tests [15], [17]. It is noteworthy that several studies on automating different clinical instruments (FAS [19], WMFT [20], [21], and dystonia [22]) have similar limitations.

This paper proposes an automated upper extremity motor function assessment system that can be practically used in a

This work was supported by the DGIST R&D Program of the Ministry of Science, ICT and Future Planning (17-BD-0401), and by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2017R1C1B2010284). Corresponding author: Jonghyun Kim (e-mail: jhkim@dgist.ac.kr). S. Lee and J. Kim are with the Department of Robotics Engineering, DGIST (Daegu Gyeongbuk Institute of Science and Technology), Daegu, 42988 Republic of Korea (e-mail: leesh@dgist.ac.kr). YS. Lee is with the Department of Rehabilitation Medicine, KNU (Kyungpook National University School of Medicine), Daegu, 41944 Republic of Korea (e-mail: leeyangsoo@knu.ac.kr).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2017.2755667, IEEE Transactions on Neural Systems and Rehabilitation Engineering

clinical environment. We select 26 target FMA tests (79% of the total FMA tests) and two optimized sensors (no body-worn sensors) for effective automation by analyzing the linguistic FMA guideline and listing required sensor information. Then, the target tests are classified depending on the features of joint motions, and based on the FMA guideline, we develop a rule-based binary logic classification algorithm for each class to overcome the limitations of machine learning. The accuracy and time efficiency of the proposed system are evaluated through a clinical experiment involving nine stroke patients.

II. Automated FMA system

2.1. Target FMA tests and selected sensors

The FMA consists of 33 tests for shoulders/elbows/forearms, wrists, hands, and coordination/speed [7]. We classified the FMA tests into four groups based on a clinical evaluation method, as summarized in Table I. In several tests (G1: 23 tests), a clinician directs a patient to perform specific motions and evaluates the patient's performance through observation. The remaining tests were evaluated using the clinician's tug test (G2: 5 tests), by applying external force to the patient's specific joints for a reflex test (G3: 3 tests), and by sensing resistance and fine tremors during the motions (G4: 2 tests).

Even though it is necessary to automate as many FMA tests as possible for effective evaluation of upper limb motor function, clinicians' extensive involvement in a few evaluation methods makes this difficult. In this study, G1 was selected as the primary target test group because of including largest number of tests and the clinician's limited role in the tests. Among the other groups, G2 was selected as the secondary target because the tug test is replaced by measurement of grip strength when grasped objects are removed.

TABLE I
Target FMA tests

Group	FMA test	Target test symbol
G1	Shoulder retraction during hand to ear	N/A
	Shoulder elevation during hand to ear	T1
	Shoulder abduction during hand to ear	T2
	Shoulder external rotation during hand to ear	T3
	Elbow flexion during hand to ear	T4
	Forearm supination during hand to ear	T5
	Shoulder adduction/inward rotation during hand to knee	T6
	Elbow extension during hand to knee	T7
	Forearm pronation during hand to knee	T8
	Hand to lumbar spine	T9
	Shoulder flexion 0°- 90° with elbow extended	T10
	Forearm pronation/supination with elbow 90°	T11
	Shoulder abduction 0°- 90°	T12
	Shoulder flexion 90°- 180°	T13
	Forearm pronation/supination with elbow 0°	T14
	Wrist flexion/extension with elbow 90°	T15
	Wrist flexion/extension with elbow 0°	T16
	Wrist Circumduction	T17
	Finger mass flexion	T18
	Finger mass extension	T19
	Tremor during finger to nose	N/A
	Dysmetria during finger to nose	T20
	Movement time of finger to nose	T21
G2	Grasp A (Hook grasp)	T22
	Grasp B (Radial grasp)	T23
	Grasp C (Pincer grasp)	T24
	Grasp D (Cylindrical grasp)	T25
	Grasp E (Spherical grasp)	T26
G3	Biceps and finger flexors reflex activity	N/A
	Triceps reflex activity	
	Normal reflex activity	
G4	Wrist stability measurement with elbow 90°	
	Wrist stability measurement with elbow 0°	

The information required for evaluating G1 tests consists of joint angles, segment rotation, landmark position, and hand open/close, as shown in Table II. For obtaining most of this information, an IMU, which is a typical body-worn sensor, has been widely used [14]–[16], [19], [20]. However, after simulating the number of G1 tests that can be theoretically automated using the IMU, we concluded that the IMU is not an appropriate sensor, as shown in Table III. As an upper limb consists of several segments, multiple IMUs must be attached to these segments (Table III). Moreover, it is difficult to measure the landmark position using an IMU; even if eight IMUs are used, only 18 FMA tests can be automated (Table III). Existing studies have used glove sensors to obtain hand open/close information [15], [16]. However, practical issues, such as synergistic patterns and muscle contractures, made it difficult to wear a glove [23]. Hence, we selected Kinect v2 (Microsoft, Redmond, WA, USA) as the primary sensor (Table II). It is a widely used low-cost sensor in the rehabilitation field and does not require to be worn or attached to the body [16], [17], [21], [23]–[29].

In particular, Kinect v2 is an optimal sensor that can provide all required information, including hand open/close. Note that

TABLE II
Sensor selection based on required information to automate target FMA tests

Required information	FMA tests	Possible sensors	Selected sensors
Joint angle	T1 to T4, T6, T7, T10, T12, T13, T15, T16	Motion capture system, Kinect v1 and v2, IMU	Kinect v2
Segment rotation	T5,T8,T11,T14	Motion capture system, Kinect v2, IMU	
Landmark position	T9,T17,T20,T21	Motion capture system, Kinect v1 and v2	
Hand open/close	T18, T19	Motion capture system, Kinect v2, IMU	
Grip strength	T22 to T26	Force/torque sensor, FSR	FSR

TABLE III
Simulation of automated FMA system with IMU and Kinect

Sensors	Body segment attachment	# of required sensor	# of automatable FMA tests
IMU	Upper arm, forearm	2	2
	Upper arm, forearm, torso, hand thumb, index finger	6	16
	Upper arm, forearm, torso, hand thumb, index finger, shoulder, back	8	18
Kinect v1	N/A	1	19
Kinect v2	N/A	1	23 (*21)
Kinect v2 + FSR (proposed)	N/A	2	28(*26)

* is the number of automatable FMA tests except for the tests which are expected to be difficult to acquire sensor data by considering accuracy of required sensor.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2017.2755667, IEEE Transactions on Neural Systems and Rehabilitation Engineering

its older version, Kinect v1, cannot provide hand open/close information (Table II). After selecting the primary sensor, 21 out of 23 tests in G1 were selected as target FMA tests. This was because it is difficult to observe considerably fine motions, such as shoulder retraction and tremors, using Kinect v2 (Table I).

The secondary target group (G2) requires grip strength information for various grasped objects (Table II). As this information cannot be obtained using Kinect v2, a force-sensing resistor (FSR) (Interlink Electronics, Westlake Village, CA, USA) was selected as the secondary sensor for the automated FMA system (Table III). An FSR is a low-cost film type sensor that is easy to attach to various objects, and it is widely used in the rehabilitation field [30]–[32].

In summary, the proposed upper limb motor function assessment system automated 26 FMA tests (21 in G1 and 5 in G2) using only two sensors, i.e., Kinect v2 and the FSR.

2.2. Rule-based binary logic algorithm

2.2.1. Feature selection based on FMA guideline

In general, each FMA test is evaluated as a score using a 3-point scale (0: cannot be performed at all; 1: can be performed partly; 2: can be performed faultlessly), and a guideline to

TABLE IV
GROUPAGE FMA TESTS AND SELECTED FEATURES

Category	# of feature			FMA test	Feature		Sensor information (Used sensor)	FMA group
	F_{Va}	F_{Vb}	F_M		Type	Description		
C1	1	0	0	T1	F_{Va}	ROM* in shoulder elevation during the motion	Joint angle (Kinect v2)	G1
				T2	F_{Va}	ROM in shoulder abduction during the motion		
				T3	F_{Va}	ROM in shoulder outward rotation during the motion		
				T4	F_{Va}	ROM in elbow extension during the motion		
				T5	F_{Va}	ROM in forearm pronation during the motion		
				T7	F_{Va}	ROM in elbow extension during the motion	Segment rotation (Kinect v2)	
				T8	F_{Va}	ROM in forearm pronation during the motion		
				T17	F_{Va}	The maximum distance ratio of 'hand tip to wrist' and 'hand to wrist which was projected by perpendicular plane of elbow to wrist	Landmark position (Kinect v2)	
				T20	F_{Va}	The minimum distance ratio of 'nose to knee' and 'head to hand'		
				T21	F_{Va}	Motion execution time		
				T18	F_{Va}	Level of hand close	Hand open/close (Kinect v2)	
				T19	F_{Va}	Level of hand open		
C2	1	0	2	T15, T16	F_{Va}	ROM in wrist flexion during the motion	Joint angle (Kinect v2)	
					F_{M_1}	Mean of elbow flexion during the motion		
					F_{M_2}	SD* of elbow flexion during the motion		
C3	1	0	4	T11, T14	F_{Va}	ROM in forearm pronation during the motion	Segment rotation (Kinect v2)	
					F_{M_1}	Mean of elbow flexion during the motion	Joint angle (Kinect v2)	
					F_{M_2}	SD of elbow flexion during the motion		
					F_{M_3}	Mean of shoulder abduction during the motion		
					F_{M_4}	SD of shoulder abduction during the motion		
C4	1	0	9	T10, T13	F_{Va}	ROM in shoulder flexion during the motion	Joint angle (Kinect v2)	
					F_{M_1}	Mean of shoulder flexion at the onset		
					F_{M_2}	Mean of elbow flexion at the onset		
					F_{M_3}	SD of elbow flexion at the onset		
					F_{M_4}	Mean of shoulder abduction at the onset		
					F_{M_5}	SD of shoulder abduction at the onset		
					F_{M_6}	Mean of elbow flexion after the onset		
					F_{M_7}	SD of elbow flexion after the onset		
					F_{M_8}	Mean of forearm pronation after the onset		
					F_{M_9}	SD of forearm pronation after the onset		
				T12	F_{Va}	ROM in shoulder abduction during the motion		
					F_{M_1}	Mean of shoulder abduction at the onset		
					F_{M_2}	Mean of elbow flexion at the onset		
					F_{M_3}	SD of elbow flexion at the onset		
					F_{M_4}	Mean of shoulder flexion at the onset		
					F_{M_5}	SD of shoulder flexion at the onset		
					F_{M_6}	Mean of elbow flexion after the onset		
					F_{M_7}	SD of elbow flexion after the onset		
					F_{M_8}	Mean of forearm pronation after the onset		
					F_{M_9}	SD of forearm pronation after the onset		
C5	0	2	0	T6	F_{1_Vb}	ROM in shoulder abduction during the motion	Joint angle (Kinect v2)	
					F_{2_Vb}	ROM in shoulder inward rotation during the motion	Segment rotation (Kinect v2)	
				T9	F_{1_Vb}	The minimum distance ratio of 'hip to hip center' and 'hip to hand'	Landmark position (Kinect v2)	
					F_{2_Vb}	Position tracking confidence level of hand		
				T22 ~ T26	F_{1_Vb}	Valid grasp	Grip strength (FSR)	G2
					F_{2_Vb}	Grip strength		

The SD* represents the standard deviation. The onset is defined as a short period from 0.5 seconds before start of the motion to 0.5 seconds after the start.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2017.2755667, IEEE Transactions on Neural Systems and Rehabilitation Engineering

assign a score is provided as a linguistic expression for each test [7]. A clinician evaluates several key tasks decomposed from a patient's instructed movement and assigns a score after considering all evaluation results of the tasks. The features required to evaluate the tasks can be divided into three types. The first type is to evaluate the dominant joint motion task during movement in three degrees based on performance level (F_{Va}). This feature includes the range of motion (ROM) (T1 to T8, T11, T10 to T16), distance between two specific joints (T9, T17, T20), motion execution time (T21), and level of hand open/close (T18, T19). The second type is to evaluate grasp motion in two degrees (F_{Vb}). The third type is to evaluate a specific posture maintaining task during movement in two degrees (F_M). This includes the mean and standard deviation (SD) of a joint angle.

An FMA test has several F_{Va}, F_{Vb}, and F_M features. For instance, the linguistic guideline of T16 describes that based on a patient's wrist flexion movement (Table I), the 3-point scale is determined by the performance level of a wrist flexion task and the success of a 0° elbow flexion posture maintaining task, as follows [7]: "*The patient is instructed to perform alternating wrist movements from maximum dorsiflexion to maximum volar flexion. The elbow fully extended*." Hence, the features consist of one F_{Va} feature regarding wrist flexion, no F_{Vb} features, and two F_M features regarding elbow posture (flexion) (Table IV). In contrast, all G2 tests only use two F_{Vb} features to evaluate whether a patient is able to correctly grasp a particular shaped object and to evaluate two leveled grip strength (Table IV).

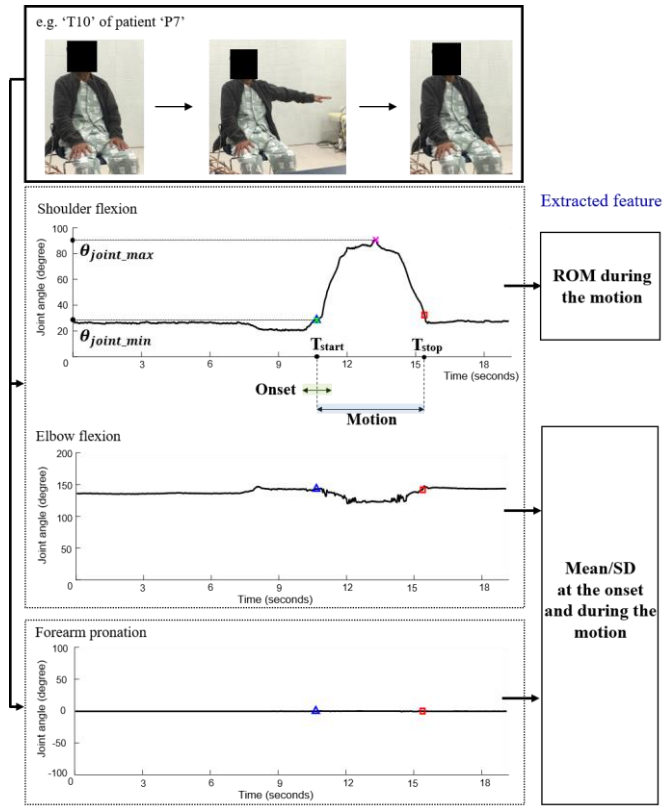


Fig. 1. Typical feature extraction in T10 test. Motion execution time was obtained as T_E-T_S, where the start time (T_{start}) and the stop time (T_{stop}) of the motion were determined when the moving SD of the joint angle with 10 frame window size was larger/smaller than 20% of the maximum moving SD obtained from entire joint angle data, respectively [42]. ROM was calculated as θ_{max}-θ_{min}, where θ_{max} (magenta X) and θ_{min} (green X) denote the maximum and minimum joint angles of the motion phase, respectively.

As mentioned above, we determined the features of each target test based on the FMA guideline. Then, all target FMA tests were classified into five categories according to the number of F_{Va}, F_{Vb}, and F_M features, as summarized in Table IV.

2.2.2. *Feature extraction*

As shown in Table IV, the information required to obtain the features of G1 tests consists of joint angles, segment rotation, landmark position, and hand open/close. Calculating joint angles and segment rotation requires pre-processing of raw data collected by Kinect v2, while the other types of information are directly provided by the Kinect v2 software development kit (SDK) (Microsoft, Redmond, WA, USA). A joint angle was calculated using a target landmark and its proximal/distal landmark position [23]. As the Kinect v2 SDK provides a quaternion form of segment rotation, we converted it to a roll angle for simplification [23].

Based on joint angles and segment rotation, several features (ROM, motion execution time, and mean/SD) can be calculated, as shown in Fig. 1. Note that the start and the stop times of the motion were determined to obtain the features. A feature of several tests (T9, T17, and T20), i.e., distance ratio ($D_{ratio} = \overrightarrow{P_1} / \overrightarrow{P_2}$), was calculated using the landmark position; for example, in T20, the extent to which a hand can approach the head was evaluated (dysmetria). Table V shows the manner in which position vectors ($\overrightarrow{P_1}$ and $\overrightarrow{P_2}$) are determined in each test to calculate the distance ratio. In addition, the level of hand open/close was calculated through the grasp information (open, mid-state, and close) provided by the Kinect v2 SDK. Note that for robust grasp detection, we used the information only when the tracking confidence level provided by Kinect v2 was high.

The feature of G2 tests, i.e., grip strength, was obtained using FSR sensors. The tests are used to evaluate five kinds of hand

TABLE V
LANDMARK POSITIONS TO OBTAIN DISTANCE RATIO FOR T9, T17 AND T20

FMA test	$\overrightarrow{p_1}$	$\overrightarrow{p_2}$
T9	Hip to hand	Hip to center of two hip joints
T17	Hand tip to wrist which was projected from perpendicular plane of elbow to wrist	Hand tip to wrist
T20	Head to hand	Head to knee

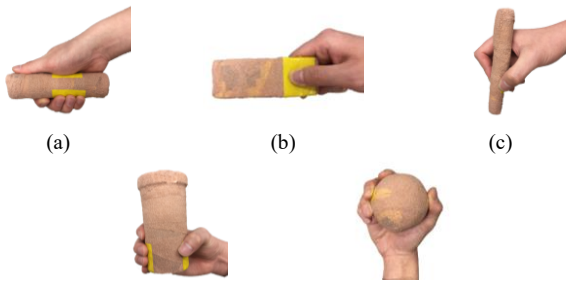
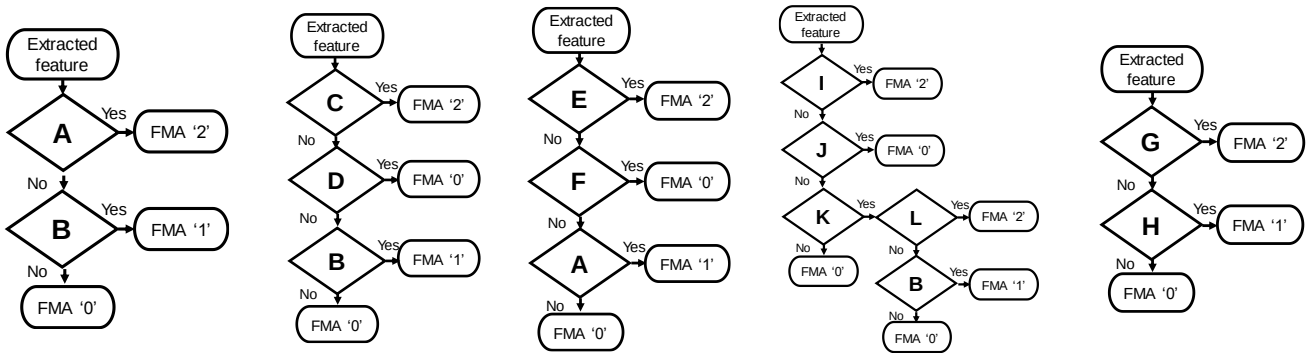


Fig. 2. Customized hand grip tools: (a) hook grasp (T22), (b) radial grasp (T23), (c) pincer grasp (T24), (d) cylindrical grasp (T25), and (e) spherical grasp (T26). Yellow parts indicate the position where FSR sensors are attached to the tools.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2017.2755667, IEEE Transactions on Neural Systems and Rehabilitation Engineering



A	x_{va_1}=1 AND x_{va_2}=1	**B**	x_{va_1}=1
C	x_{va_1}=1 AND x_{va_2}=1 AND x_{1M}=1 AND x_{2M}=1	**D**	$x_{1M} \cdot x_{2M}$=0
E	x_{1M}=1 AND x_{2M}=1 AND x_{3M}=1 AND x_{4M}=1 AND x_{va_1}=1 AND x_{va_2}=1	**F**	$x_{1M} \cdot x_{2M}$=0 AND $x_{3M} \cdot x_{4M}$=0
G	x_{1vb}=1 AND x_{2vb}=1	**H**	x_{1vb}=1
I	x_{1M}=1AND x_{2M}=1 AND x_{3M}=1 AND x_{4M}=1 AND x_{5M}=1 AND x_{6M}=1 AND x_{7M}=1 AND x_{8M}=1 AND x_{9M}=1 AND x_{va_1}=1 AND x_{1vb}=1	**J**	x_{1M}=0 OR x_{2M}=0 OR x_{3M}=0 OR x_{4M}=0 OR x_{5M}=0 OR x_{va_1}=0
K	$x_{6M} \cdot x_{7M}$=1 AND $x_{8M} \cdot x_{9M}$=1	**L**	x_{va_1}=1 AND x_{1vb}=1

Fig. 3. Binary logic-based classification. Categories (a) C1, (b) C2, (c) C3, (d) C4, and (e) C5.

grasp (hook, radial, pincer, cylindrical, and spherical grasp), and thus, the FMA guideline describes the shapes of each handheld object and grip methods [7]. Hence, we developed five customized hand grip tools by attaching two FSRs at locations where the object and finger would be in contact with each other during guided grip, as shown in Fig. 2.

2.2.3. Rule-based binary logic classification

Most existing studies on automating the FMA, which used classification algorithms based on machine learning, have a clear limitation that a learning process with a large amount of clinical data is essential [14]–[18]. In this study, we converted the linguistic guideline for the target FMA tests into binary logics and implemented a classification algorithm based on these logics to assign FMA scores using extracted features.

To obtain the logics, each feature was represented as binary variables using appropriate thresholds. According to the role of each feature type, as described in section 2.2.1, an F_{Va} feature is expressed using two binary variables (x_{va_1} and x_{va_2}) while F_{Vb} and F_M features are expressed using a single binary variable (x_{vb} and x_M). In general, x_{va_1} and x_{va_2}, x_{vb}, and x_M indicate the level of motion task performance, the validity of grasp motion, and the success of a posture maintaining task, respectively. All thresholds used to determine these binary variables based on extracted features were selected according to the FMA guideline and experienced (more than ten years of experience in the FMA) clinicians' advice. For instance, in T16, x_{va_1} and x_{va_2} are determined as (0, 0), (1, 0) or (1, 1) to represent the low, medium, or high levels of wrist flexion motion, respectively; x_{1M} and x_{2M} are determined as 0 or 1 to represent the success /failure of the 0° elbow flexion posture maintaining task. More details are provided in the appendix.

The binary logic-based classification of each test to rate the 3-point scale was obtained by analyzing the FMA guidelines and binary variables. For this purpose, experienced clinicians assisted us in interpreting the logic that was ambiguous owing to its insufficient linguistic explanation in the guideline. For instance, the classification of T16 was obtained as following. Since the guideline of T16, mentioned in section 2.2.1, clearly describes when FMA score 2 is rated, we firstly made the logic C in Fig. 3b to represent high level of wrist flexion motion with 0° elbow flexion posture. To distinguish score 0 and 1 came from the clinicians' advice that score 1 is clinically valid with maintaining the elbow posture. Hence we made the logics D and B to classify score 0 due to wrong elbow posture and low level of the wrist motion, respectively.

After the analysis, we found that the decision logic structure of all FMA tests corresponding to each category (from C1 to C5) is the same, as shown in Fig. 3. This result shows that the proposed automation system can be easily implemented using a few binary logic based algorithms, even though it is able to support most of the FMA tests.

2.3. Configuration of proposed system

The proposed FMA automation system follows the framework reported in our previous study [16], as shown in Fig. 4. The framework consists of a sensor part, a primary system, a classifier, and a user interface (Fig. 4). The sensor part includes

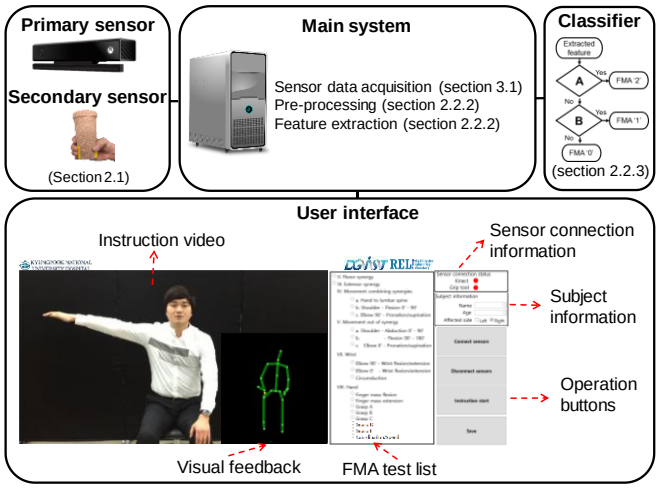


Fig. 4. Framework of proposed FMA automation system

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2017.2755667, IEEE Transactions on Neural Systems and Rehabilitation Engineering

Kinect v2 (primary sensor) and grasp tools with the FSR (secondary sensor). In the primary system, features are obtained through pre-processing and extraction of collected sensor data. Based on the features, the proposed rule-based binary logic classifier automatically assigns FMA scores.

To implement the proposed system, we developed a user interface as shown in Fig. 4. The primary purpose of the user interface was to provide accurate instructions and to enhance patients' motivation, even without clinician. For this purpose, we created an instruction video for each FMA test with a clinician who had sufficient clinical experience. In the video, the clinician conducted an instructed movement with detailed verbal explanation to enable a patient to clearly imitate the movement (Fig. 4). Moreover, the video could be replayed for sufficient understanding of the instructed movement. For motivating patients, the user interface provided the patients' skeleton images, which were obtained from the Kinect v2 SDK, as visual feedback that enabled the patients to observe their motion in real time, similar to a mirror (Fig. 4). Hence, the user interface enhanced patients' motivation, which led to the best motion performance. Note that a skeleton image was used instead of an RGB image based on the consideration of a few patients' psychological requirement to not observe their appearance. In addition, we added a stop button for atypical situations, such as an abrupt interruption by a patient during the automated FMA.

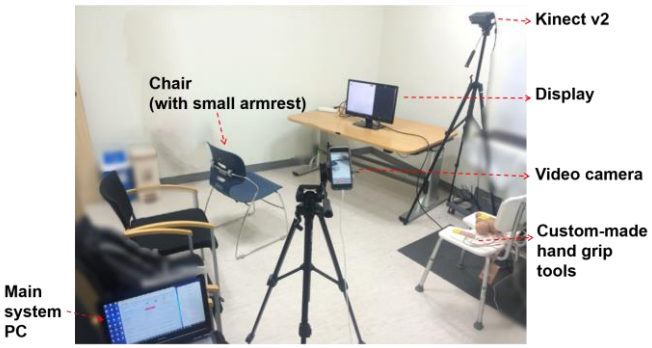


Fig. 5. Experimental setup of automated FMA system

TABLE VI
CHARACTERISTICS OF 10 STROKE PATIENTS (N=10)

Index	Age /Sex	Affected side	Etiology	MMSE	FMA
P1	33/M	Left	Intracerebral Hemorrhage	30	27
P2	39/F	Right	Left MCA infarction	25	48
P3	57/M	Right	Left MCA infarction	28	58
P4	64/M	Left	BS infarction, Right pons	28	34
P5	71/M	Left	Right MCA infarction	22	2
P6	45/F	Right	Left post putamen infarction	30	32
P7	86/M	Left	Right post putamen infarction	24	61
P8	44/M	Left	Right MCA infarction	27	12
P9	65/F	Left	Right CR infarction	30	16
P10	76/F	Left	Right MCA infarction	21	N/A

FMA score for the upper extremity of the affected side, assigned by clinicians during clinical experiment. MMSE score was measured before the experiment.

III. CLINICAL TRIALS

3.1. Experimental setup

A clinical experiment was conducted to evaluate the accuracy and time efficiency of the proposed system. Kinect v2 was installed at a distance of 1.5 m from a chair with a small armrest. The subjects would sit on this chair, and a display for the user interface was located at 1 m in front of the chair (Fig. 5). Five grasp tools with FSRs were placed approximately 20 cm away from a subject's affected arm (Fig. 5). We used an Arduino interface board (Adafruit Industries, Manhattan, NY, USA) to convert the analog signals from the FSRs to digital signals. Kinect v2 and the board were connected to the primary system PC via USB communication. While the patients were performing instructed motion, information about their motion and grip strength was collected at 30 Hz for Kinect and 120 Hz for the FSR, as shown in Table II. Before the FMA test, we adjusted the installation angle of Kinect v2 to enable it to track all landmarks of the subjects.

3.2. Participant

The participants consisted of ten stroke patients (six males, four females; age: 58±16.5 years) who did not have serious cognitive impairment (mini-mental state examination (MMSE) score > 15) and could maintain a sitting posture in a chair. The characteristics of the patients are shown in Table VI. All participants provided signed informed consent approved by Kyungpook National University Hospital institutional review board (KNUH 2016-08-034-001) prior to the experiment. In addition, two well-experienced clinicians (a medical doctor and an occupational therapist) participated in the experiment.

3.3. Protocol

The subjects sat on a chair by themselves or with a clinician's help. An instruction video was provided for each FMA test, and the subjects were asked to do their best to perform the instructed movement while watching the video. Whenever a subject was disappointed in his limited motor ability, the clinicians encouraged him/her. Based on the observation of a subject's movement, the clinicians individually determined the scores for each test and assigned a score based on mutual agreement to minimize the effect of imperfect inter-rater reliability of the FMA [8]. At the same time, the proposed system automatically scored each test; however, these scores were not visible to the clinicians. After all target FMA tests supported by the proposed system (Table I) were conducted, the remaining seven tests were directly performed by the clinicians without the proposed system. Experiment scenes were recorded using a video camera for measuring the time required for each experimental procedure and for analyzing the subjects' movements.

3.4. Data analysis

One out of ten subjects (P10) stopped during the experiment because of her sudden depression (Table VI), and thus, the data for this subject was excluded. First, we measured the installation time of the proposed system using the recorded video to evaluate its convenience. To evaluate the scoring accuracy of the proposed system, we calculated percent

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2017.2755667, IEEE Transactions on Neural Systems and Rehabilitation Engineering

TABLE VII
PERCENT AGREEMENTS OF FMA SCORES BETWEEN AUTOMATED ASSESSMENT AND IN-PERSON ASSESSMENT

Index	Classification Accuracy	Disagreed FMA tests
P1	96.2%	T3
P2	96.2%	T5
P3	92.3%	T18, T19
P4	80.8%	T6, T18, (T1, T2, T4)*
P5	100%	N/A
P6	80.8%	T1, T3, T4, T10, T24
P7	100%	N/A
P8	84.6%	T6, T14, T18, T23
P9	96.2%	T3

Note that the disagreements were caused by the subjects' unexpected violation of instructed movement.

TABLE VIII
COHEN'S KAPPA RESULT OF AUTOMATED FMA TESTS

		Automated			
		0	1	2	Total
In-person	0	85	1	3	89
	1	5	72	8	85
	2	0	2	58	60
	Total	90	75	69	234

Cohen's kappa coefficient = 0.877

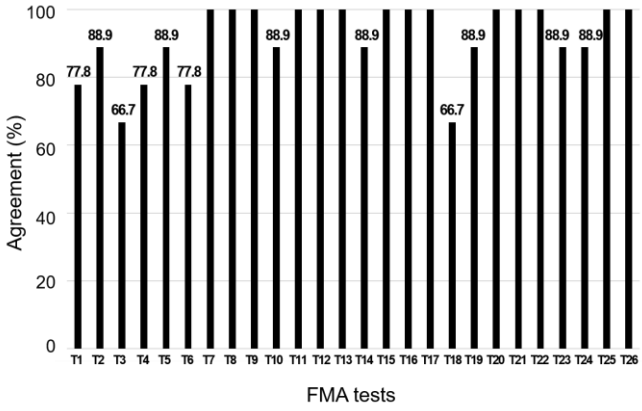


Fig. 6. Agreement of each automated FMA tests

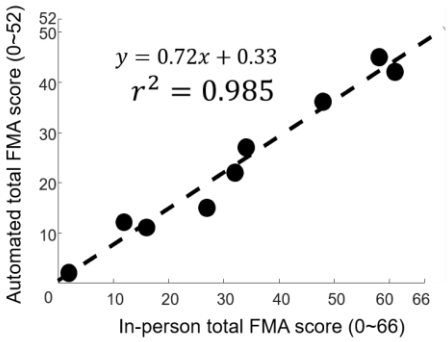


Fig. 7. Correlation analysis between automated FMA (26 tests) and in-person FMA (33 tests)

agreement and Cohen's kappa coefficient (k) by comparing the FMA scores obtained from the system and clinicians.

With the assumption that the proposed system is able to allow the clinician's absence, the clinician's administration time for FMA would be the time required for 7 unautomated tests. Hence, to estimate time efficiency due to the proposed system, we measured the required time separately using the recorded

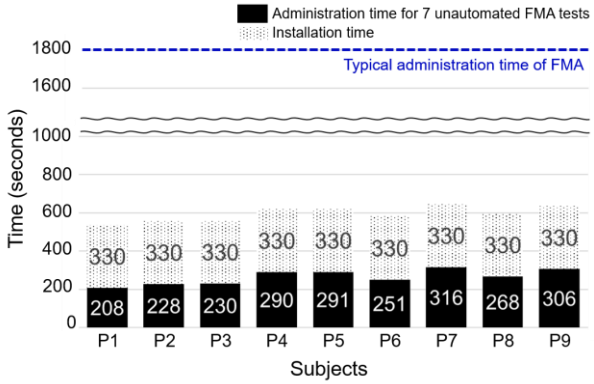


Fig. 8. Comparison of clinician's required time of the proposed system. Note that the required time for in-person FMA assessment was known as about 1800 seconds (30 minutes) [8].

video. In addition, to confirm whether the 26 automated FMA tests conducting using the proposed system could represent all 33 FMA tests, a correlation analysis was conducted by calculating the Pearson correlation coefficient (r) between the total score of all 33 tests assigned by the clinicians and the total score of 26 tests assigned by the proposed system. SPSS version 20 (IBM, Chicago, IL, USA) was used for all statistical analyses.

IV. RESULT

4.1. Installation time

The installation time of the proposed system was approximately 5.5 minutes (330 seconds).

4.2. Scoring accuracy

Table VII summarizes the results for the comparison of the FMA scores assigned by the proposed system and clinicians. In addition, the Cohen's kappa coefficients for 234 FMA tests (9 subjects and 26 automated tests per subject) are shown in Table VIII. The average percent agreement was approximately 92%, and for all subjects, the agreements were more than 80% (Table VII). A high kappa value (k = 0.877) indicates an almost perfect agreement between both scores (Table VIII) [33]. For each automated FMA test, the agreement was displayed in Fig. 6.

The result of the correlation analysis for 26 automated FMA tests and 33 in-person FMA tests is shown in Fig. 7. Pearson's correlation coefficient exhibits strong correlation (r = 0.985) [34].

4.3. Time efficiency

The average time required by a clinician for unautomated FMA tests was 265.3±36 seconds (black bar), as shown in Fig. 8. Assuming that the time required for conventional in-person FMA tests was approximately 1800 seconds (30 minutes) [8], the result showed that the proposed system could reduce the required time by approximately 85% (Fig. 8). Considering the installation time (white bar) as the worst case, a reduction of more than 64% was still observed in the required time (Fig. 8).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2017.2755667, IEEE Transactions on Neural Systems and Rehabilitation Engineering

V. Discussion

IMUs, glove sensors, and motion capture systems, which were used in existing studies, required considerable time, and it was inconvenient to wear a sensor or marker on the body [14]–[16], [19], [20], [22]. In contrast, the proposed system does not require time-consuming procedures; it requires only approximately five minutes for the initial setup of Kinect v2 (angle of view setting). Particularly, the proposed system can significantly improve a clinician's time efficiency because 1) it supports most FMA tests and 2) it is not necessary to perform its installation procedure for each test and subject. Moreover, compared with existing studies [14]–[17], which require a large amount of clinical data for learning processes, the proposed system is sufficiently practical for clinical applications.

The average accuracy of the FMA test scores assigned by the proposed system is considerably high (Table VII and VIII). This indicates that the proposed system can assess upper-limb motor function in a manner similar to that of in-person assessment by an experienced clinician. Despite high accuracy, as shown in Table VII and Fig. 6, a few tests exhibited disagreement between scores because of the following reasons: First, there were motion tracking errors in Kinect v2 caused by a few subjects' severe hand contracture. The tracking accuracy of the thumb strongly affects the detection of the segment rotation of Kinect, which results in extraction of inaccurate features on forearm pronation/supination (T5 and T14). Moreover, the contracture results in inaccurate measurement of the hand open/close state by Kinect (T18 and T19). For instance, subjects P3 and P8 performed a hand open movement with the thumb bent, and the proposed system with Kinect classified the movement as a fully performed task (2 points). However, the clinician judged it to be inadequate performance (1 point) considering the erroneous motion of the thumb. Note that this decision by the clinician was because the hand open/close movement includes thumb flexion/extension and the thumb performs the primary function of the hand [35]. Second, a few disagreements resulted from the threshold used to generate binary variables. In tug-related tests, the clinician's tug levels were not consistent for each subject, while the proposed system used the same grasp force threshold (T23 and T24). Third, shoulder internal/external rotation occurred occasionally because hospital gowns were loose (T3 and T6). Last, the clinician's variable scoring due to synergistic patterns led to errors in the tests, which included multiple features at the instructed movement (T1, T3, and T4). This was demonstrated by reviewing a subjects' recorded movement with the clinicians after the experiment. In addition, a subject (P4) did not follow the instruction video in a few tests (T1, T2, and T4) and used the ipsilateral arm for assistance while performing motions.

The above mentioned reasons for inaccurate scoring lead to the possibility that the accuracy of the proposed system can be improved, particularly in the case of relatively low-accuracy subjects (P6 and P8). If a subject wears tight clothes, the error in T6 for P8 can be removed. The error in T4 for P6 is not the critical reason of the decrease in accuracy, considering the clinician's scoring error (Table VII). Furthermore, excluding the cases in which a subject did not follow the instructed movement intentionally, the scoring accuracy of the proposed system would be at least 85% (Table VII). This level of accuracy is considerably higher than that obtained in existing studies (approximately 50%) [16], [17], which automate a significantly lower number of tests than the proposed system.

The extremely high correlation between the total score of the automated FMA tests and that of the in-person FMA tests indicates that the proposed system can adequately estimate complete FMA scores in spite of a few unautomated FMA tests. This is because the unautomated tests in the proposed system account for a relatively small proportion (21%) of all FMA tests. Another reason for this is that the unautomated tests do not affect the FMA significantly. In a previous study, the significance of each FMA test was investigated through Rasch analysis [36]. Based on the significance, six key tests were selected out of 33 FMA tests [36], [37]. These tests were automated in the proposed system (T1, T7, T11, T13, T15, and T18). This can be considered as indirect evidence that shows that the proposed system sufficiently automated the FMA.

The seven unautomated tests can be divided into the following two groups: five tests that require external force to be evaluated (reflex activity (G3) and wrist stability (G4)) and two tests in which measurement of fine motions (shoulder retraction and tremor) is required. The first group might be automated using robots and electromyography; however, this automation considerably increases system cost and installation time. Automation of the second group requires an apparatus that can precisely measure fine motions (i.e., a motion capture system); however, similar problems occur owing to high cost and inconvenient marker attachment.

This study proposed an automated FMA system for upper extremity motor function assessment in stroke patients. For developing a clinically relevant system, 79% of the FMA tests were automated through optimized sensor selection, and approximately 90% scoring accuracy was achieved by employing a rule-based binary logic classification algorithm without learning procedures. The proposed system can reduce a clinician's required time for the FMA by more than 85%, which would contribute to frequent evaluation of upper-limb motor function and improvement in upper-limb intervention for rehabilitation. However, this study has the following limitations: A few tests, such as T3 and T18, exhibited insufficient scoring accuracy (Fig. 6). To increase the accuracy, improvement in inaccurate tracking of thumb motion and in ambiguity in synergistic patterns is required. The performance evaluation of the proposed system was based on the participation of relatively few clinicians and patients. Future works will include clinical trials involving a larger population and verification of test-retest reliability. In addition, the proposed system has promising applications in telerehabilitation [38], [39], remote assessment, and in-home intervention.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2017.2755667, IEEE Transactions on Neural Systems and Rehabilitation Engineering

APPENDIX

The methods used for determining each binary variable used in this study are summarized in the following table:

Group	Type	Feature	Threshold	Binary value	
				x_{va_1}	x_{va_2}
G1	F_{Va}	Distance ratio (T17, T20)	If $F_{Va} < \boldsymbol{\lambda_1}$	0	0
			Else if $F_{Va} < \boldsymbol{\lambda_2}$	1	0
			Else	1	1
		Motion execution time (T21)	If $F_{Va} > \mathbf{T_{E1}}$	0	0
			Else if $F_{Va} > \mathbf{T_{E2}}$	1	0
			Else	1	1
		Level of hand open/close (T18, T19)	If F_{Va} = **'Hand close(open) state' not detected**	0	0
			Else if F_{Va} = **'Hand open(close) state' not detected**	1	0
			Else	1	1
		ROM (G1 tests except T17 to T21)	If $F_{Va} < \boldsymbol{\lambda_1 \cdot \theta_d}$	0	0
			Else if $F_{Va} < \boldsymbol{\lambda_2 \cdot \theta_d}$	1	0
			Else	1	1
					x_{1vb}
	F_{Vb}	Distance ratio (T9)	If $F_{1\_Vb} < \boldsymbol{\lambda_2}$		0
			Else		1
		Hand tracking confidence level (T9)	If F_{2_Vb} = **was not change from 'Tracked' to 'Not tracked'**		0
			Else		1
		ROM (T6)	If $F_{1\_Vb}(F_{2\_Vb}) < \boldsymbol{\lambda_2 \cdot \theta_d}$		0
			Else		1
					x_M
	F_M	Mean (G1 tests)	If $F_M < \boldsymbol{\lambda_1 \cdot \theta_f}$		0
			Else		1
		SD (G1 tests)	If $F_M > \boldsymbol{\theta_{SD}}$		0
			Else		1
					x_{1vb}
G2	F_{Vb}	Grip strength (G2 tests)	If $F_{1\_Vb} < \mathbf{P_v}$		0
			Else		1
					x_{2vb}
			If $F_{2\_Vb} < \mathbf{P_t}$		0
			Else		1

The bold text indicates the threshold. λ_1 and λ_2 are adjustable constants. In this study, λ_1 and λ_2 were set as 0.3 and 0.7, respectively, based on clinicians' advice. θ_d is the joint angle according to the instructed movement, and θ_f is the ROM of the normal joint [40]. P_v is the pressure detected when a grip tool is lightly grasped using the correct method. P_t represents the mean of the maximum pressure detected when an object is removed from the repeated in-person tug tests. θ_{SD} was set as ±20° considering the clinicians' advice and the body segment tracking error in Kinect v2 [41]. T_{E1} and T_{E2} were set as 16 and 12 seconds, respectively, according to the FMA guideline.

ACKNOWLEDGEMENT

The authors would like to thank all subjects and clinicians who participated in the study. Particularly, we would like to thank Dr. Min-gu Kang, Dr. Tae-woo Nam, Dr. Hyun-min Oh, Ja-won Yu, and Seok-hyun Noh, for volunteering and providing valuable advice.

REFERENCES

[1] V. Feigin, M. Forouzanfar, R. Krishnamurthi, G. Mensah, M. Connor, D. Bennett, A. Moran, R. Sacco, L. Anderson, T. Truelsen, M. O'Donnell, N. Venketasubramanian, S. Barker-Collo, C. Lawes, W. Wang, Y. Shinohara, E. Witt, M. Ezzati, M. Naghavi, and C. Murray, "Global and regional burden of stroke during 1990–2010: findings from the Global Burden of Disease Study 2010," *The Lancet*, vol. 383, no. 9913, pp. 245–255, 2014.

[2] C. Han, M. Arbib, and N. Schweighofer, "Stroke rehabilitation reaches a threshold," *PLoS Comput. Biol.*, vol. 4, no. 8, p. e1000133, 2008.

[3] T. Tarver, "Heart disease and stroke statistics–2014 update: A report from the American Heart Association," *J. Consum. Health Internet*, vol. 18, no. 2, pp. 209–209, 2014.

[4] P. Pantano, R. Formisano, M. Ricci, V. Piero, U. Sabatini, B. Pofi, R. Rossi, L. Bozzao, and G. Lenzi, "Motor recovery after stroke," *Brain*, vol. 119, no. 6, pp. 1849–1857, 1996.

[5] C. Caldwell, D. Wilson, and R. Braun, "Evaluation and treatment of the upper extremity in the hemiplegic stroke patient," *Clin. Orthop. Relat. Res.*, vol. 1, no. 63, pp. 69–93, 1969.

[6] R. Bonita and R. Beaglehole, "Recovery of motor function after stroke," *Stroke*, vol. 19, no. 12, pp. 1497–1500, 1988.

[7] A. R. Fugl-Meyer, L. Jaasko, I. Leyman, S. Olsson, and S. Steglind, "The post-stroke hemiplegic patient. 1. A method for evaluation of physical performance," *Scand. J. Rehabil. Med.*, vol. 7, pp. 13–31, 1975.

[8] D. Gladstone, C. Danells, and S. Black, "The Fugl-Meyer assessment of motor recovery after stroke: A critical review of its measurement properties," *Neurorehabil. Neural Repair*, vol. 16, no. 3, pp. 232–240, 2002.

[9] 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.*, vol. 73, no. 7, pp. 447–454, 1993.

[10] S. Wolf, D. Lecraw, L. Barton, and B. Jann, "Forced use of hemiplegic upper extremities to reverse the effect of learned nonuse among chronic stroke and head-injured patients," *Exp. Neurol.*, vol. 104, no. 2, pp. 125–132, 1989.

[11] R. Lyle, "A performance test for assessment of upper limb function in physical rehabilitation treatment and research," *Int. J. Rehabil. Res.*, vol. 4, no. 4, pp. 483–492, 1981.

[12] T. Brott, H. Adams, C. Olinger, J. Marler, W. Barsan, J. Biller, J. Spilker, R. Holleran, R. Eberle, and V. Hertzberg, "Measurements of acute cerebral infarction: A clinical examination scale," *Stroke*, vol. 20, no. 7, pp. 864–870, 1989.

[13] K. Avlund, S. Kreiner, and K. Schultz-Larsen, "Functional ability scales for the elderly: A validation study," *Eur. J. Pub. Health*, vol. 6, no. 1, pp. 35–42, 1996.

[14] S. Del Din, S. Patel, C. Cobelli, and P. Bonato, "Estimating Fugl–Meyer clinical scores in stroke survivors using wearable sensors," in *33rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Boston, MA, USA, 2011, pp. 5839–5842.

[15] L. Yu, D. Xiong, L. Guo, and J. Wang, "A remote quantitative Fugl-Meyer assessment framework for stroke patients based on wearable sensor networks," *Comput. Methods Programs Biomed.*, vol. 128, pp. 100–110, 2016.

[16] P. Otten, J. Kim, and S. Son, "A Framework to automate assessment of upper-limb motor function impairment: A feasibility study," *Sensors*, vol. 15, no. 8, pp. 20097–20114, 2015.

[17] W. Kim, S. Cho, D. Baek, H. Bang, and N. Paik, "Upper extremity functional evaluation by Fugl-Meyer assessment scoring using depth-sensing camera in hemiplegic stroke patients," *PLoS ONE*, vol. 11, no. 7, p. e0158640, 2016.

[18] S. Kotsiantis, I. Zaharakis and, P. Pintelas, "Machine learning: A review of classification and combining techniques," *Artificial Intelligence Review*, vol. 26, no. 3, pp. 159–190, 2006.

[19] S. Patel, R. Hughes, T. Hester, J. Stein, M. Akay, J. Dy, and P. Bonato, "A novel approach to monitor rehabilitation outcomes in stroke survivors using wearable technology," *Proceedings of the IEEE*, vol. 98, no. 3, pp. 450–461, 2010.

[20] A. Parnandi, E. Wade, and M. J. Mataric´, "Motor function assessment using wearable inertial sensor," in *32rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Buenos Aires, Argentina, 2010, pp. 86–89.

[21] A. Ozturk, A. Tartar, B. Ersoz Huseyinsinoglu, and A. Ertas, "A clinically feasible kinematic assessment method of upper extremity motor function impairment after stroke," *Measurement*, vol. 80, pp. 207–216, 2016.
[22] S. Kukke, L. Curatalo, A. de Campos, M. Hallett, K. Alter, and D. Damiano, "Coordination of reach-to-grasp kinematics in individuals with childhood-onset dystonia due to hemiplegic cerebral palsy," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 24, no. 5, pp. 582–590, 2016.
[23] S. Lee, M. Song, and J. Kim, "Towards clinically relevant automatic assessment of upper-limb motor function impairment," in *3rd IEEE EMBS Int. Conf. Biom. Heal. Info. (BHI)*, Las vegas, NV, USA, 2016, pp. 148–151.
[24] E. Olesh, S. Yakovenko, and V. Gritsenko, "Automated assessment of upper extremity movement impairment due to stroke," *PLoS ONE*, vol. 9, no. 8, p. e104487, 2014.
[25] M. Ye, C. Yang, V. Stankovic, L. Stankovic, and A. Kerr, "A depth camera motion analysis framework for tele-rehabilitation: Motion capture and person-centric kinematics analysis," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 5, pp. 877–887, 2016.
[26] K. Chun, D. Lim, C. Kim, H. Jung, and D. Jung, "Use of the Microsoft Kinect system to characterize balance ability during balance training," *Clin. Interv. Aging*, vol. 10, pp. 1077-1083, 2015.
[27] A. Da Gama, T. Chaves, L. Figueiredo, A. Baltar, M. Meng, N. Navab, V. Teichrieb, and P. Fallavollita, "MirrARbilitation: A clinically-related gesture recognition interactive tool for an AR rehabilitation system," *Comput. Methods Programs Biomed.*, vol. 135, pp. 105–114, 2016.
[28] M. Yates et al., "Virtual reality gaming in the rehabilitation of the upper extremities post-stroke," *Brain Inj.*, vol. 30, no. 7, pp. 855–863, 2016.
[29] C. Morrison, P. Culmer, H. Mentis, and T. Pincus, "Vision-based body tracking: Turning Kinect into a clinical tool," *Disabil. Rehabil. Assist. Technol.*, vol. 11, no. 6, pp. 516–520, 2014.
[30] Z. Abu-Faraj, G. Harris, J. Abler, J. Wertsch, and P. Smith, "A Holter-type microprocessor-based rehabilitation instrument for acquisition and storage of plantar pressure data in children with cerebral palsy," *IEEE Trans. Rehabil. Eng.*, vol. 4, no. 1, pp. 33–38, 1996.
[31] R. J. Goldman, "Capacitive biofeedback sensor with resilient polyurethane dielectric for rehabilitation," U.S. Patent 5 449 002, Sep 12, 1995.
[32] K. Kong, "A mobile gait monitoring system for abnormal gait diagnosis and rehabilitation: A pilot study for Parkinson disease patients," *J. Biomech. Eng.*, vol. 133, no. 4, p. 041005, 2011.
[33] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, no. 1, pp. 37–46, 1960.
[34] K. Pearson, "Note on regression and inheritance in the case of two parents," *Proc. R. Soc. Lond. (1854-1905)*, vol. 58, no. 1, pp. 240–242, 1895.
[35] D. Neumann and T. Bielefeld, "The carpometacarpal joint of the thumb: Stability, deformity, and therapeutic intervention," *J. Orthop. Sports Phys. Ther.*, vol. 33, no. 7, pp. 386–399, 2003.
[36] Y. Hsieh, I. Hsueh, Y. Chou, C. Sheu, C. Hsieh, and G. Kwakkel, "Development and validation of a short form of the Fugl-Meyer motor scale in patients with stroke," *Stroke*, vol. 38, no. 11, pp. 3052–3054, 2007.
[37] K. Chen, C. Chen, Y. Chou, C. Shih, C. Koh, and C. Hsieh, "Is the long form of the fugl-meyer motor scale more responsive than the short form in patients with stroke?," *Arch. Phys. Med. Rehabil.*, vol. 95, no. 5, pp. 941–949, 2014.
[38] H. Park, Q. Peng, and L. Zhang, "A portable telerehabilitation system for remote evaluations of impaired elbows in neurological disorders," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 16, no. 3, pp. 245–254, 2008.
[39] H. Park, J. Kim, and D. Damiano, "Development of a haptic elbow spasticity simulator (HESS) for improving accuracy and reliability of clinical assessment of spasticity," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 20, no. 3, pp. 361–370, 2012.
[40] D. Boone and S. Azen, "Normal range of motion of joints in male subjects," *J. Bone Joint Surg.*, vol. 61, no. 5, pp. 756–759, 1979.
[41] K. Otte, B. Kayser, S. Mansow-Model, J. Verrel, F. Paul, A. Brandt, and T. Schmitz-Hübsch, "Accuracy and reliability of the Kinect version 2 for clinical measurement of motor function," *PLoS ONE*, vol. 11, no. 11, p. e0166532, 2016.
[42] D. Giansanti and G. Maccioni, "Physiological motion monitoring : a wearable device and adaptive algorithm for sit-to-stand timing detection", *Physiological measurement*, vol. 27, pp. 713-723, 2006

Seunghee Lee (S'17) received the B.S. degree in electrical engineering from Chungbuk National University, Cheongju, Korea, in 2011. Since 2011, he has been the Ph. D. Candidate at the Department of Robotics Engineering in DGIST. His active research interests include rehabilitation robotic systems, sensor based medical assessment, algorithm design for automation.

Yang-Soo Lee received the M.D. degree from Kyungpook National University, Daegu, Korea at 1988. He completed residency at Hallym University Medical Center, Seoul, Korea at 1994. Since 1997, he has been assistant professor and professor at the department of rehabilitation medicine in Kyungpook National University School of Medicine. His current research interest include neurorehabilitation with task oriented training, rehabilitation robot, and sensor-based assessment.

Jonghyun Kim (S'05-M'09) received the B.S., M.S., and Ph.D. degrees in mechanical engineering from KAIST, Daejeon, Korea, in 2001, 2003, and 2010, respectively. He had been a postdoctodal fellow at the Rehabilitation Medicine Department, National Institutes of Health (NIH), Bethesda, MD, for 3 years. Since 2013, he has been an assistant professor at the Department of Robotics Engineering in DGIST, Daegu, Korea. His current research interests include neurorehabilitation with brain plasticity, telerehabilitation, rehabilitation robot, technology-driven sensorimotor assessment, and nonlinear control.