



Article

Proof-of-Concept of a Sensor-Based Evaluation Method for Better Sensitivity of Upper-Extremity Motor Function Assessment

Seung-Hee Lee ¹, Ye-Ji Hwang ², Hwang-Jae Lee ³, Yun-Hee Kim ³, Matjaž Ogrinc ^{4,5}, Etienne Burdet ⁴ and Jong-Hyun Kim ^{2,*}

¹ KEPCO Research Institute, Seoul 06732, Korea; s.lee@kepco.co.kr

² School of Mechanical Engineering, Sungkyunkwan University, Suwon 16419, Korea; yj.hwang@skku.edu

³ Center for Prevention & Rehabilitation, Heart Vascular and Stroke, Samsung Medical Center, Department of Physical and Rehabilitation Medicine, Sungkyunkwan University School of Medicine, Seoul 06351, Korea; goodptlee@skku.edu (H.-J.L.); yunkim@skku.edu (Y.-H.K.)

⁴ Department of Bioengineering, Imperial College London, London SW72AZ, UK; matjaz.ogrinc@imperial.ac.uk (M.O.); e.burdet@imperial.ac.uk (E.B.)

⁵ GripAble Limited, Thornton House, 39 Thornton Road, London, SW19 4NQ, UK

* Correspondence: jonghyunkim@skku.edu



Citation: Lee, S.-H.; Hwang, Y.-J.; Lee, H.-J.; Kim, Y.-H.; Ogrinc, M.; Burdet, E.; Kim, J.-H. Proof-of-Concept of a Sensor-Based Evaluation Method for Better Sensitivity of Upper-Extremity Motor Function Assessment. *Sensors* **2021**, *21*, 5926. <https://doi.org/10.3390/s21175926>

Academic Editor: Gregorij Kurillo

Received: 6 July 2021

Accepted: 27 August 2021

Published: 3 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: In rehabilitation, the Fugl–Meyer assessment (FMA) is a typical clinical instrument to assess upper-extremity motor function of stroke patients, but it cannot measure fine changes of motor function (both in recovery and deterioration) due to its limited sensitivity. This paper introduces a sensor-based automated FMA system that addresses this limitation with a continuous rating algorithm. The system consists of a depth sensor (Kinect V2) and an algorithm to rate the continuous FM scale based on fuzzy inference. Using a binary logic based classification method developed from a linguistic scoring guideline of FMA, we designed fuzzy input/output variables, fuzzy rules, membership functions, and a defuzzification method for several representative FMA tests. A pilot trial with nine stroke patients was performed to test the feasibility of the proposed approach. The continuous FM scale from the proposed algorithm exhibited a high correlation with the clinician rated scores and the results showed the possibility of more sensitive upper-extremity motor function assessment.

Keywords: sensor-based motor function assessment; depth sensor; fuzzy inference system; sensitivity; rehabilitation

1. Introduction

In rehabilitation, upper-extremity motor function evaluation for stroke survivors is important to plan effective rehabilitation intervention [1,2]. The most widely used in-person assessment in clinics is the Fugl–Meyer assessment (FMA) due to its validity and reliability [3–5]. Despite its popularity, FMA is (1) labor-intensive and time-consuming, and (2) not sensitive enough to fine changes in motor function ability due to the coarse three point grading scheme of the FM scale [4]. Although this grading scheme results in high inter/intra-rater reliability, it also has lower sensitivity than other clinical instruments, such as the medical research council muscle strength scale (six point scale) [4–6]. Many clinical studies have reported this limitation, meaning that it is not possible to track fine changes of a patient's motor function using the FM scale [4,7–10].

Thanks to recent advances in sensor technologies, several works had reported an automated FMA system to address labor-intensiveness and time consumption issues [11], these, however, did not attempt to propose a more sensitive FM scale using the virtue of sensor-based measurements to overcome the limitation on low sensitivity [12–18]. It might be because most work to automate FMA showed inadequate accuracy even though the

work focused on predicting the original three point FM scale. Another reason would be that the machine learning methods used in the existing works for FMA are not appropriate to handle this issue because (1) some of them (support vector machine [17,19] and Naive–Bayes classification [20]) cannot be used for regression and (2) the others (extreme machine learning [14], artificial neural network [15], and random forest [21,22]) require a large amount of dimension-reduced training data for regression, which could not be collected from numerous patients in practice.

A promising solution for a more sensitive FMA is to develop a continuous scoring algorithm for sensor-based automated FMA. Meanwhile, from our first attempt to apply a sensor-enabled body tracking to the FMA automation [17], we recently reported a sensor-based automated FMA system with a rule-based expert with binary logics originated from the linguistic grading guideline of FMA, which is different to the machine learning methods used in the other existing works [13]. Importantly, the binary logic was verified by high grading accuracy with the original FM scale [13] and could be applied to continuous scoring through the fuzzy logic approach, such as designing an appropriate fuzzy inference system (FIS) [23,24]. This FIS-based approach is promising because it (1) does not require collecting a large amount of patients' data, (2) can consider a clinician's ambiguous judgement mathematically [24,25] and (3) makes its reasoning process understandable [24].

The goal of this study was to check the feasibility of sensor-based continuous FM scale scoring. For that, we firstly chose three representative FMA tests and developed a novel scoring algorithm for the tests based on FIS defining the fuzzy variables and rules from the FMA guideline [23–26]. Then, a sensor-based automated FMA system that can provide continuous FM scale was implemented by using the scoring algorithm and a depth sensor (Kinect V2). After investigating the achievable number of grades under the system by considering the expected error of the sensor, we showed the feasibility of the proposed scoring method through a pilot trial with nine stroke patients.

2. Materials and Methods

2.1. Target FMA Tests and Sensor Selection

As the targets for developing a continuous FM scale scoring algorithm, we selected three tests out of the FMA tests automated in our previous study [13], which are listed in Table 1. These tests consist of volitional movements with synergies and without synergy [3], and all of them have quite different binary logics to prove the feasibility of the proposed approach [13].

Table 1. Target FMA tests.

FMA Test	Index
Elbow extension during hand to knee	T1
Shoulder abduction 0–90°	T2
Shoulder adduction/inward rotation during hand to knee	T3

2.2. Continuous FM Scale Scoring Algorithm

2.2.1. Fuzzy Variables

Based on the features used in our FMA studies [13,17], we determined the fuzzy input variables (X) for each target test as shown in Table 2. The fuzzy input variables can be divided into three types: F_{Va} and F_{Vb} for evaluating the selective/voluntary motor performance in three and two stages, respectively, and FM for evaluating the ability to maintain a specific constraint (posture) in two stages (Table 2). Therefore, the fuzzy set (R) of the fuzzy input variables (X) can be summarized as shown in Table 2. Note that we used the same feature extraction method [13] to obtain the variables.

Table 2. Fuzzy input variables and their fuzzy sets.

Fuzzy Input Variables (<i>X</i>)			
FMA Test	Type	Index	Description
T1	F_{Va}	F_{Va}	Elbow extension ROM
		F_{Va}	Shoulder abduction ROM
	F_M	F_{M_1}	Shoulder abduction angle (mean, onset)
		F_{M_2}	Elbow flexion angle (mean, onset)
		F_{M_3}	Shoulder flexion angle (mean, onset)
T2	F_M	F_{M_4}	Elbow flexion angle (mean, motion)
		F_{M_5}	Forearm pronation angle (mean, motion)
		F_{M_6}	Elbow flexion (SD, onset)
		F_{M_7}	Shoulder flexion angle (SD, onset)
		F_{M_8}	Elbow flexion angle (SD, motion)
		F_{M_9}	Forearm pronation angle (SD, motion)
T3	F_{Vb}	F_{Vb_1}	Shoulder abduction ROM
		F_{Vb_2}	Shoulder inward rotation ROM
Input Fuzzy Sets (<i>R</i>)			
Type	Index	Description	
F_{Va}	R_1	Performed fully	
	R_2	Performed partially	
	R_3	Not performed	
F_{Vb}	R_4	Performed	
	R_5	Not performed	
F_M	R_6	Maintained	
	R_7	Not maintained	

ROM denotes the range of motion; SD the standard deviation. ‘Onset’ and ‘motion’ denote the onset phase and the motion phase, respectively.

The fuzzy output variables (*Y*) are the FM scores for the target tests. Each FM score was assigned three levels (0: cannot be performed at all; 1: can be performed partially; and 2: can be performed fully) according to the degree of motor function judged by the clinician. This means that the fuzzy sets (*S*) of the fuzzy output variables were classified as shown in Table 3.

Table 3. Fuzzy output variables and their fuzzy sets.

Fuzzy Output Variables (<i>Y</i>)	Output Fuzzy Sets (<i>S</i>)	
	Index	Description
FM score	S_1	High motor function as FM score ‘2’
	S_2	Medium motor function as FM score ‘1’
	S_3	Low motor function as FM score ‘0’

2.2.2. Fuzzy Rules

The propositions expressed in words can be transformed into fuzzy rules for the scoring FM scale by using the explicitation process [24,25]. All established fuzzy rules can be defined as the following fuzzy canonical form:

$$\text{IF } X \text{ is } R \text{ THEN } Y \text{ is } S \quad (1)$$

From the viewpoint of the FMA, *X* is the clinician’s observation, *R* is the clinician’s judgment regarding the performance level of each feature, *Y* is the degree of the patient’s motor function, and *S* is the FM scale assigned by the clinician.

The fuzzy rules of the target tests are summarized in Table 4. T1 had the simplest logical structure with three rules; an F_{Va} is evaluated at three performance levels without maintaining a specific posture. As for T2, two rules were added to the rules of T1 in order

to consider the ability of posture constraint. T3 consisted of three rules, which were to evaluate two F_{Vb} in two levels without posture constraint. It should be noted that the weights among all fuzzy rules were set to 1.

Table 4. Fuzzy rules of each target FMA test.

FMA Test	# of Rules	Description
T1	3	(1) IF F_{Va} is R_1 THEN FM score is S_1 (2) IF F_{Va} is R_2 THEN FM score is S_2 (3) IF F_{Va} is R_3 THEN FM score is S_3
T2	5	(1, 2, 3) same as T1 fuzzy rules (4) IF F_{M_1} is R_7 OR F_{M_2} is R_7 OR F_{M_3} is R_7 OR F_{M_6} is R_7 THEN FM score is S_3 (5) IF F_{M_4} is R_7 OR F_{M_5} is R_7 OR F_{M_8} is R_8 OR F_{M_9} is R_7 THEN FM score is NOT S_1
T3	3	(1) IF F_{Vb_1} is R_4 AND F_{Vb_2} is R_4 THEN FM score is S_1 (2) IF F_{Vb_1} is R_5 OR F_{Vb_2} is R_5 THEN FM score is S_2 (3) IF F_{Vb_1} is R_5 AND F_{Vb_2} is R_5 THEN FM score is S_3

2.2.3. Fuzzy Inference System

We adopted the Mamdani method to implement FIS. This method has the characteristic that inference results can be easily transformed into linguistic forms [25].

The fuzzy set R of X in (1) can be defined as follows:

$$R = \{(x, \mu_R(x)) : x \in X, \mu_R(x) \in [0, 1]\} \text{ with } X: \alpha \leq x \leq \beta; \mu_R : X \rightarrow [0, 1] \quad (2)$$

where x denotes the feature measured by the sensor; μ_R the membership function (MF) for the fuzzification of X ; and α and β are the minimum and maximum values within the universe of discourse, respectively. Here, α was set to 0, and β was determined as the desired feature value in the instructed motion or as 30 for the SD features (F_{M_6} to F_{M_9}) [13].

Like (2), the fuzzy set S of Y in (1) was defined as shown below:

$$S = \{(y, \mu_S(y)) : y \in Y, \mu_S(y) \in [0, 1]\} \text{ with } Y: \gamma \leq x \leq \delta; \mu_S : Y \rightarrow [0, 1] \quad (3)$$

where y denotes the degree of a patient's motor function judged by a clinician; μ_S is the MF for the implication of Y ; γ and δ were 0 and 1, respectively.

For the MFs in (2) and (3), we adopted a triangular shaped function that is simple and widely used for modeling the human's reasoning process [27–29], as displayed in Figure 1. The MF in the fuzzy set R_1 for F_{Va} was designed to satisfy the fact that a clinician's judgment approaches 'performed fully' when the measured feature value is closer to the desired value in the instructed motion (Table 2). The opposite meaning, 'not performed', was implemented as the MF in R_3 using the complement of R_1 (Figure 1). We also designed the MF in R_2 with the assumption that the clinician's 'performed partially' judgement results in the highest degree of membership when the measured feature value was half of the desired feature value (Figure 1). The MFs of the fuzzy sets S for Y were also designed with the same shape. Moreover, for F_{Vb} and F_M , which have two stages in FMA, the MFs in R_4 (R_6) and R_5 (R_7) were designed as the same shape to the MFs of R_1 and R_3 , respectively.

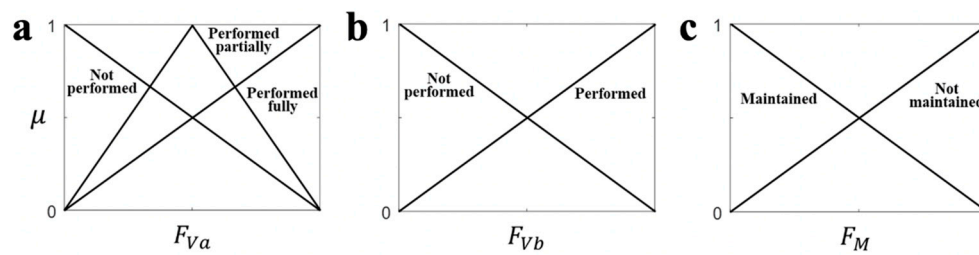


Figure 1. The membership functions for each type of feature. (a) F_{Va} , (b) F_{Vb} , (c) F_M .

The logic rules AND/OR were formulated by the minimum/maximum functions. The AND operation was applied for the truncate (implication) of S , and the maximum function was used for aggregation operation [30]. The result of the designed FIS, continuous FM scale, was provided as a constant value by using centroid defuzzification process [30,31]. The continuous FM scale scoring algorithm through FIS using the data acquired by the sensor is represented as shown in Figure 2.

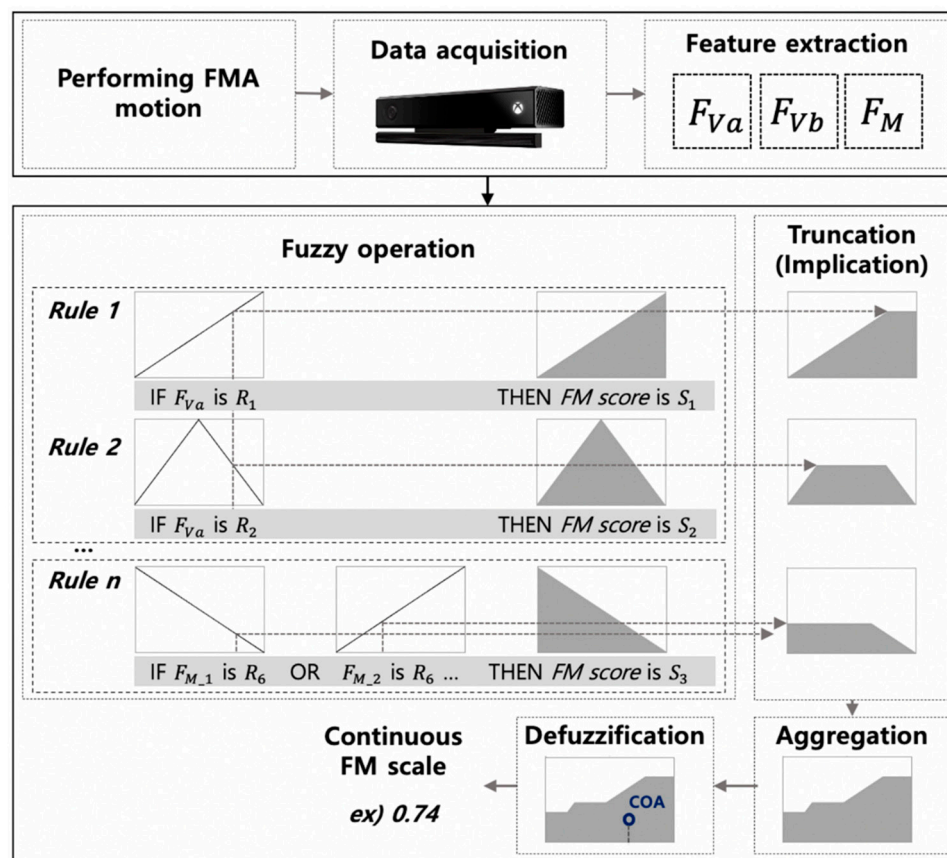


Figure 2. Schematic diagram of the continuous FM scale scoring algorithm; COA denotes the center of area in the centroid defuzzification method [31].

In summary, based on the rule-based logic obtained from the linguistic grading guidelines of FMA [13], we have developed a novel continuous FM scale scoring algorithm using a fuzzy logic approach. In contrast to a machine learning approach, our algorithm can easily rate a continuous FM scale without requiring any training data including a clinician's three point FM score.

2.2.4. The Sensor-Based Automated FMA System's Achievable Number of Grades

We implemented a sensor-based automated FMA system. Similar to our previous study [13], the system consisted of a sensor, user interface, and scoring algorithm. A depth

sensor, Kinect V2 (Microsoft, Redmond, WA, USA) was used to extract motion features (Figure 2), which is an inexpensive and easy to use sensor that has been widely applied in the rehabilitation area [13,15–17,32], and whose effectiveness was verified in our previous FMA studies [13,17]. The system has the same user interface as our previous automated FMA system [13], which provided an instruction video that was prerecorded by a well-experienced clinician. In contrast to our previous system, the system in this study could provide a continuous FM scale due to the proposed continuous FM scale scoring algorithm.

The continuous FM scale due to the scoring algorithm (FM_{CA}), which comes from the fuzzy result value of the FIS, would be a solution for better sensitivity of the FMA. However, the inaccurate input value due to measurement error of the depth sensor could restrict the achievable number of FM grades. To investigate this restriction, the confidence range of fuzzy result value was analyzed as follows. After simulating the expected maximum (worst) feature errors by using the reported tracking error of the depth sensor used (Kinect V2) [33] considering the effect of sensor position [34] and anthropometric data [35] (Table 5), we estimated the possible maximum fuzzy result errors and number of grades, as represented in Table 6. The results showed that FM_{CA} could be interpreted up to seven scale in all the target FMA tests (Table 6), which implies that motor function could be more sensitively evaluated than the conventional three point FM scale.

Table 5. Information used to estimate fuzzy result error due to sensor inaccuracy.

Joint Tracking Error of Kinect V2		
Kinect V2 Landmark	Mean Error (Standard Deviation)	FMA Tests Related
Shoulder	1 cm (1 cm)	T1, T2, T3
Elbow	2 cm (1 cm)	T1, T2, T3
Wrist	2 cm (1 cm)	T1
Spine shoulder	1 cm (1 cm)	T2, T3
Spine mid	1 cm (1 cm)	T2, T3
Shoulder internal rotation	16 deg	T3
Used Anthropometry		
Body segment	Length	FMA tests related
Upper arm (shoulder to elbow)	27.2 cm	T1, T2, T3
Lower arm (elbow to wrist)	29.4 cm	T1
Torso	90.1 cm	T2, T3

Note that anthropometry data were from a 40 year old Asian female [35].

Table 6. Estimated fuzzy result error and possible number of grades for each FMA test.

Target FMA Tests	Feature	Maximum Feature Error (Deg)	Maximum Fuzzy Result Error	Possible # of Grades
T1	F_{Va}	23.382	0.104	9
	F_{Va}	14.432		
	F_{M_1}	14.432		
	F_{M_2}	23.382		
	F_{M_3}	14.432		
T2	F_{M_4}	23.382	0.128	7
	F_{M_5}	7.81		
	F_{M_6}	6.222		
	F_{M_7}	2.934		
	F_{M_8}	6.222		
T3	F_{M_9}	4.52	0.081	12
	F_{Vb_1}	14.432		
	F_{Vb_2}	16		

Note that the features for each test can be found in Table 2.

2.3. Experiment

2.3.1. Experimental Setup

We conducted an experiment using the implemented system with stroke patients to test the feasibility of the proposed continuous FM scale scoring algorithm. Figure 3 shows the experimental setup. A Kinect V2 sensor was installed one meter in front of the subject (Figure 3), and data were recorded at a sampling rate of 30 Hz. The instructions were delivered to the subject through the instruction video. For the purpose of showing the video, a monitor (visual) and speaker (auditory) were installed near the sensor (Figure 3). The sensor data were automatically recorded after the start of the video, and this recording was finished when the subject's movement was completed.

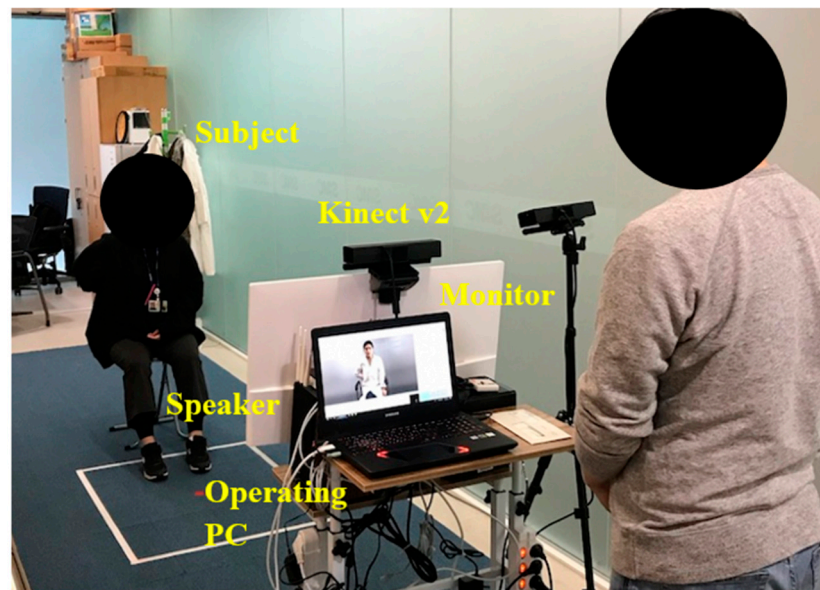


Figure 3. Experimental setup. Note that another Kinect that is located beside the monitor was just used to record depth images during FMA tests.

2.3.2. Protocol

Nine stroke patients (ages 49 to 77 years) participated in the experiment, whose biographical information is summarized in Table 7. This experiment was approved by the Samsung Medical Center institutional review board (SMC-2018-02-053), and all subjects gave their consent prior to the experiment.

The subjects sat in a chair without an armrest or wheelchair (Figure 3). They were asked to follow the motions in the instruction video, and thus they mimicked the motion while the video was being played. During the subjects' motion, a well-experienced (more than 10 years) clinician observed the motion, and rated the conventional three point FM scale (FM₃) as well as the following extended seven point FM scale (FM₇):

- 0: cannot be performed at all (same to FM₃);
- 0+: can be performed a little bit but close to level that cannot be performed at all;
- 1−: can be performed partially but close to level that cannot be performed;
- 1: can be performed partially (same to FM₃);
- 1+: can be performed partially but close to level that can be performed well;
- 2−: can be performed well but not perfectly;
- 2: can be performed perfectly (same to FM₃).

The reason why we used the FM₇ above was that it would not be enough to evaluate the proposed FM_{CA} using FM₃, considering the achievable number of grades estimated. In order to reduce the clinician's scoring difficulty, FM₇ was made as a straightforward extension by adding '0+', '1−', '1+' and '2−' to the FM₃.

Table 7. Biographical information of participated stroke patients.

	Age/Sex	Affected Side	Time Since Stroke (Month)	Etiology	MMSE	UE-FMA
S1	49/F	Left	15	Rt. FT ICH	30	12
S2	54/M	Left	18	Rt. LS inf.	28	27
S3	57/M	Right	32	Lt. Pons inf.	29	62
S4	69/F	Right	3	Lt. Pons inf.	27	63
S5	58/F	Right	50	Lt. BG ICH	25	23
S6	61/M	Left	1	Rt. MCA inf.	29	37
S7	77/F	Left	3	Rt. Thalamus ICH	24	9
S8	65/M	Left	1	Rt. Thalamus ICH	28	12
S9	68/M	Right	1	Lt. MCA territory	27	8

F and M denotes female and male, respectively; MMSE denotes Mini Mental State Examination score; UE-FMA is the total sum of conventional three point FM scale (FM₃). Rt = Right; Lt = Left; FT = Frontotemporal; LS = Lenticulostriate; BG = Basal ganglia; MCA = Middle cerebral artery; F = Frontal; ICA = Internal cerebral artery; CR = Corona radiata; ICH = Intracerebral hemorrhage; inf. = Infarction.

2.3.3. Data Analysis

Sensor data of 27 trials (three target FMA tests with nine patients) were recorded during the experiment. The high frequency spikes and jitters in the data were removed through a third-order low-pass Butterworth filter (10 Hz cutoff frequency) [16]. Then, two different FM scales were obtained from the data: (1) a three point FM scale using the automated FMA algorithm in our previous study (FM_{3A}) [13], and (2) FM_{CA} (continuous FM scale due to the proposed FIS-based scoring algorithm) (Figure 2). Here, the FIS was implemented by using a fuzzy logic toolbox in Matlab (Mathworks, Natick, MA, USA), and the fuzzy result value (FM_{CA}) was linearly normalized to have the range from 0 to 1.

In order to validate the quality of the collected data in this experiment, we investigated whether the data could provide an accurate three point FM scale. The agreement and Cohen's kappa were calculated between the FM_{3A} and FM₃. From this analysis, one can indirectly deduce whether the inaccurate FM_{CA} is due to low data quality or erroneous fuzzy rule/FIS design.

The proposed FM_{CA} needs to have the following clinical characteristics: FM_{CA} corresponds to the patients' degree of motor function evaluated by a clinician. For that, we calculated the Pearson's correlation coefficient between the FM_{CA} and FM₃/FM₇ [30]. Moreover, in clinic, since FMA used the total sum of the FM scale for each FMA test to evaluate a patient's overall motor function [4], the Pearson's correlation coefficient between the sum of FM_{CA} and sum of FM₃/FM₇ was calculated to check whether the proposed FM_{CA} could also follow the overall evaluation method [13–16]. The statistical analysis was performed using SPSS version 20 (IBM, Chicago, IL, USA).

3. Results

All the subjects' FM scales were automatically rated with three point level (FM_{3A}) and continuous level (FM_{CA}) by using our automated FMA system developed in [8], along with the manually scored FM₃ and FM₇ as summarized in Table 8.

Table 8. Rated FM scores by clinician and implemented sensor-based system.

	T1				T2				T3			
	FM ₃	FM _{3A}	FM ₇	FM _{CA}	FM ₃	FM _{3A}	FM ₇	FM _{CA}	FM ₃	FM _{3A}	FM ₇	FM _{CA}
S1	1	1	1−	0.490	1	1	1−	0.396	1	1	1	0.511
S2	1	1	1−	0.390	1	1	1+	0.585	1	1	1−	0.437
S3	2	2	2	0.864	2	2	2−	0.849	2	1	2	0.564
S4	1	1	1+	0.543	2	2	2	0.978	2	2	2	0.888
S5	2	2	2	0.791	1	1	1	0.484	2	2	2−	0.846
S6	2	2	2−	0.623	1	1	1−	0.367	2	2	2	0.970
S7	1	1	1−	0.360	1	1	1−	0.381	1	1	1	0.523
S8	0	0	0	0.061	1	1	1−	0.364	0	0	0	0.008
S9	0	0	0	0.245	0	0	0	0.259	0	0	0+	0.319

Note that the bold numbers mean the disagreement between FM₃ and FM_{3A}. FM₃, FM_{3A}, FM₇, and FM_{CA} denote clinician rated three point scale, system rated three point scale, extended seven point scale, and the proposed continuous scale, respectively.

3.1. Validity of Collected Data

The agreement between FM_{3A} and FM₃ was 96.3%, and Cohen's kappa was 0.940, as summarized in Table 9. This result was similar to our previous result for the automated FMA system (92% agreement; 0.877 Cohen's kappa) [13]. It supports the validity of the quality of data collected in this study. Disagreement only occurred in a T3 trial in which the clinician rated score '2' (Table 9).

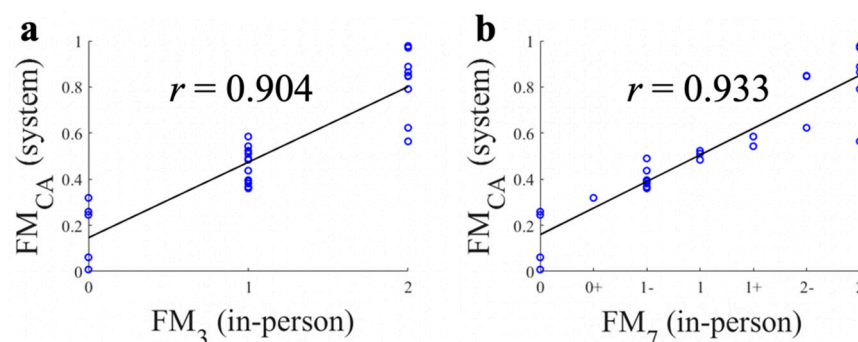
Table 9. Agreement of FM₃ and FM_{3A}.

		FM _{3A} (System)			
		0	1	2	Total
FM ₃ (in-person)	0	5	0	0	5
	1	0	13	0	13
	2	0	1	8	9
Total		5	14	8	27

Cohen's kappa = 0.940; 96.3% agreement; Note that the bold text represents the number of agreed FM trials. FM₃ and FM_{3A} denote clinician rated three point FM scale and system rated three point FM scale, respectively.

3.2. Continuous FM Scale

Figure 4 shows the correlations between FM_{CA} and FM₃/FM₇ in the 30 trials. Overall, a high Pearson's correlation coefficient was observed for FM₃ ($r = 0.904$) and FM₇ ($r = 0.933$) (Figure 4), and it was also valid for each FMA test (T1: $r = 0.930$ for FM₃, $r = 0.959$ for FM₇, T2: $r = 0.897$; for FM₃, $r = 0.966$; for FM₇; and T3: $r = 0.896$; for FM₃, $r = 0.903$; for FM₇). These results showed that the proposed FM_{CA} corresponded to the FM₇ rated by the clinician.

**Figure 4.** Correlations between continuous FM scale (FM_{CA}), three point FM scale (FM₃), and seven point FM scale (FM₇). (a) FM_{CA} and FM₃; (b) FM_{CA} and FM₇.

The correlations between the total sum of FM_{CA} and FM_3/FM_7 for each subject are shown in Figure 5. Note that there was no subject who had a total sum of FM_3 as two or four (Figure 5a). The high correlation results ($r = 0.940$; with FM_3 and $r = 0.976$; with FM_7) show that the proposed FM_{CA} system can estimate overall motor function accurately (Figure 5). Here, the total sum of FM_7 in Figure 5b was calculated through the conversion of FM_7 as '1−' to '2', '1' to '3', '1+' to '4', '2−' to '5' and '2' to '6'. It should be noted that all correlation analyses above showed significant correlation ($p < 0.001$).

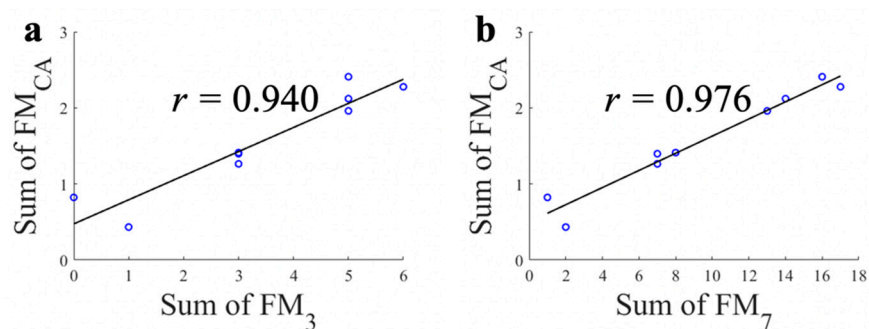


Figure 5. Correlations between the total sum of each FM scale. (a) Proposed continuous FM scale (FM_{CA}) and conventional three point FM scale (FM_3); (b) Proposed continuous FM scale (FM_{CA}) and extended seven point FM scale (FM_7).

4. Discussion

In this study, we used FM_7 along with FM_3 to evaluate the proposed FM_{CA} . The clinician reported that the rating of FM_7 was not difficult because FM_7 is a simple scale expansion of FM_3 . The extended scales in FM_7 (0+, 1−, 1+, and 2−) appeared in 51.9% of the total FMA tests (14 out of 27). This means that there is a clear demand in clinic for evaluating motor function by using a more sensitive FM scale than the existing FM_3 . It should be noted that FM_7 could not currently be regarded as a validated clinical tool.

The T3 FMA test resulted in lower correlation ($r = 0.903$) than T1 and T2, because of a disagreement in a trial between FM_{3A} and FM_3 highlighted in Table 8. The correlation becomes much higher ($r = 0.984$) when this trial is excluded. Since the FM_7 of the trial were '2', FM_{CA} had the greatest deviation for trials that belong to score '2' (Figure 4). We believe that the lower performance in T3 was caused by inaccurate tracking of the motion sensor used (Kinect V2). For T3, we extracted two F_{Vb} features when the subject moved the hand to the knee. Here, one of the features, shoulder inward rotation ROM, could not be precisely extracted because when the subject's distal segment of the upper limb was moving along the proximal direction, the subject's loose patient uniform, made the measurement of the angle unreliable (about a 16 degree error) [13]. If the proposed system was applied to 26 FMA tests, we expect that 22 of 26 tests would be free from the sensor inaccuracy problem above based on the characteristics of the inaccurate tracking investigated in [13], except the following tests: shoulder adduction/inward rotation during hand to knee (T3), shoulder external rotation during hand to ear, forearm supination during hand to ear, and forearm pronation/supination with elbow 0° .

As mentioned, this paper proposed a novel continuous FM_{CA} scoring algorithm based on the fuzzy logic derived from our previous rule-based expert (binary logics). One can expect that several existing studies on automated FMA could be extended for the continuous FM scale. For instance, a linearized model that is obtained from the correlation analysis between the extracted feature (i.e., range of motion) and original FM scale rated by clinician could enable the scoring of the continuous scale [16]. Those approaches, however, would suffer from inaccuracies due to the complexity of FMA (i.e., Pearson's correlation coefficient $r = 0.03$ in some tests [16]), as follows. Based on the Bobath concept [3,4], the instructions of FMA usually ask the patient to perform a certain joint motion while constraining the other joint motions for evaluating the selective/voluntary motor performance. Hence, the

FM scale is rated by clinician's comprehensive inference based on multiple features with different types: F_{Va} , F_{Vb} , and F_M , and thus it makes the dimension reduction used in those approaches (i.e., using principal component analysis [16]) difficult. Note that this statement is supported by the complex binary logic for automating some FMA tests that were shown in our previous work [13].

The aim of the proposed sensor-based continuous-scaled FMA system is to automate the evaluation of motor function more objectively and sensitively. From a clinical point of view, along with its convenience and time efficiency, the proposed system has the potential to improve the limited sensitivity of the conventional FM scale, which would be a novel instrument for better practice of rehabilitation. Moreover, the proposed system can contribute to effective robot-aided rehabilitation therapy due to its better sensitivity. For instance, thanks to FM_{CA} , the intensity and difficulty of the robotic therapy can be precisely chosen, and the fine monitoring of the motor function after the therapy could be used to accurately investigate its therapeutic effect. In addition, the proposed system is promising to be utilized as a key measure for achieving precise big data for upper-extremity motor function.

This study could still be improved. We only implemented three FMA tests for the proposed continuous FM scale so as to investigate its feasibility. Since the rule-based binary logic, the basis of FIS, for most FMA tests was already found in our previous work [13], it is promising that the unimplemented tests could be covered in a similar manner in the near future. As for the sensor system, the performance could be improved when we use a state-of-art depth sensor, such as RealSense (Intel, Santa Clara, CA, USA) or Leap motion controller (Leap Motion Inc., San Francisco, CA, USA) [36], both of which have better resolution than Kinect only. Moreover, the reliability (consistency) test of the proposed FMCA with repeated trials and various environment would be needed to confirm the feasibility of the proposed approach. In addition, the limited number of subjects in this study could be solved through an additional clinical trial with a larger population.

5. Conclusions

FMA, a well-known clinical instrument for stroke patients, still has low sensitivity, so it cannot evaluate fine changes of motor function. As a solution to this limitation, this study showed the possibility that sensor-based automated FMA system with the proposed FIS-based algorithm could provide a continuous FM scale (FM_{CA}), which is highly correlated with the conventional FM scale (FM_3) as well as the extended FM scale (FM_7). It means that the designed FIS in the system for scoring FMCA faithfully reflects the clinical knowledge of FMA. This is additionally supported by the high correlations between the total sum of FM_{CA} and FM_3/FM_7 . To our knowledge, this study is the first attempt (1) to develop the continuous FM scale and (2) to apply fuzzy logic approach (i.e., FIS) for automated and more sensitive FMA. Therefore, we expect that the proposed system could be a basis to improve the quality of motor function assessment for stroke patients significantly.

Author Contributions: J.-H.K. conceptualized and designed the study. Y.-H.K., E.B. and J.-H.K. provided the resources for the study. S.-H.L. and M.O. implemented the proposed system. S.-H.L., Y.-H.K. and J.-H.K. designed the experiments. H.-J.L. and Y.-H.K. recruited subjects and prepared IRB for the experiments. S.-H.L., Y.-J.H. and H.-J.L. acquired the data. S.-H.L., Y.-J.H. and J.-H.K. processed and analyzed the data from the experiments, and interpreted results from the data. S.-H.L. drafted the original manuscript. E.B. revised, and J.-H.K. finalized the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported in part by the MSIT (Ministry of Science and ICT), Korea, under the ICT Creative Consilience Program (IITP-2021-2020-0-01821) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation) and in part by the Translational Research Program for Care Robots funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HK20C0015).

Institutional Review Board Statement: The stroke patients signed an informed consent approved by the Samsung Medical Center IRB prior to the experiment (No. SMC-2018-02-053).

Informed Consent Statement: Not applicable.

Data Availability Statement: Data and materials can be made available upon request to the authors.

Acknowledgments: The authors would like to thank all subjects who volunteered for the study.

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

- Bonita, R.; Beaglehole, R. Recovery of motor function after stroke. *Stroke* **1988**, *19*, 1497–1500. [[CrossRef](#)]
- Barnes, M.P. Principles of neurological rehabilitation. *J. Neurol. Neurosurg. Psychiatry* **2003**, *74*, iv3–iv7. [[CrossRef](#)]
- Fugl-Meyer, A.R.; Jääskö, L.; Leyman, I.; Olsson, S.; Steglind, S. The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. *Scand. J. Rehabil. Med.* **1975**, *7*, 13–31.
- Gladstone, D.J.; Danells, C.J.; Black, S.E. The Fugl-Meyer assessment of motor recovery after stroke: A critical review of its measurement properties. *Neurorehabil. Neural Repair* **2002**, *16*, 232–240. [[CrossRef](#)]
- Duncan, P.W.; Propst, M.; Nelson, S.G. Reliability of the Fugl-Meyer assessment of sensorimotor recovery following cerebrovascular accident. *Phys. Ther.* **1983**, *63*, 1606–1610. [[CrossRef](#)]
- Craig, P.; Dieppe, P.; Macintyre, S.; Michie, S.; Nazareth, I.; Petticrew, M. Developing and evaluating complex interventions: The new Medical Research Council guidance. *BMJ* **2008**, *337*, 259–275. [[CrossRef](#)]
- van der Lee, J.H.; Beckerman, H.; Lankhorst, G.J.; Bouter, L.M. The responsiveness of the Action Research Arm test and the Fugl-Meyer Assessment scale in chronic stroke patients. *J. Rehabil. Med.* **2001**, *33*, 110–113.
- Barnes, M.; Dobkin, B.; Bogousslavsky, J. *Recovery after Stroke*; Cambridge University Press: Cambridge, UK, 2005. [[CrossRef](#)]
- Lin, J.H.; Hsu, M.J.; Sheu, C.F.; Wu, T.S.; Lin, R.T.; Chen, C.H.; Hsieh, C.L. Psychometric comparisons of 4 measures for assessing upper-extremity function in people with stroke. *Phys. Ther.* **2009**, *89*, 840–850. [[CrossRef](#)]
- Prabhakaran, S.; Zarah, E.; Riley, C.; Speizer, A.; Chong, J.Y.; Lazar, R.M.; Marshall, R.S.; Krakauer, J.W. Inter-individual variability in the capacity for motor recovery after ischemic stroke. *Neurorehabil. Neural Repair* **2008**, *22*, 64–71. [[CrossRef](#)]
- Simbaña, E.O.; Baeza, P.S.-H.; Huete, A.J.; Balaguer, C. Review of automated systems for upper limbs functional assessment in neurorehabilitation. *IEEE Access* **2019**, *7*, 32352–32367. [[CrossRef](#)]
- Song, X.; Chen, S.; Jia, J.; Shull, P.B. Cellphone-based automated Fugl-Meyer assessment to evaluate upper extremity motor function after stroke. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2019**, *27*, 2186–2195. [[CrossRef](#)]
- Lee, S.; Lee, Y.; Kim, J. Automated evaluation of upper-limb motor function impairment using Fugl-Meyer assessment. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2018**, *26*, 125–134. [[CrossRef](#)] [[PubMed](#)]
- Yu, L.; Xiong, D.; Guo, L.; Wang, J. A remote quantitative Fugl-Meyer assessment framework for stroke patients based on wearable sensor networks. *Comput. Meth. Programs Biomed.* **2016**, *128*, 100–110. [[CrossRef](#)]
- Kim, W.S.; Cho, S.; Baek, D.; Bang, H.; Paik, N.J. Upper extremity functional evaluation by Fugl-Meyer assessment scoring using depth-sensing camera in hemiplegic stroke patients. *PLoS ONE* **2016**, *17*, e0158640. [[CrossRef](#)] [[PubMed](#)]
- Olesh, E.V.; Yakovenko, S.; Gritsenko, V. Automated assessment of upper extremity movement impairment due to stroke. *PLoS ONE* **2014**, *9*, e104487. [[CrossRef](#)]
- Otten, P.; Kim, J.; Son, S.H. A framework to automate assessment of upper-limb motor function impairment: A feasibility study. *Sensors* **2015**, *15*, 20097–20114. [[CrossRef](#)]
- Li, Y.; Zhang, X.; Gong, Y.; Cheng, Y.; Gao, X.; Chen, X. Motor function evaluation of hemiplegic upper-extremities using data fusion from wearable inertial and surface EMG sensors. *Sensors* **2017**, *17*, 582. [[CrossRef](#)]
- Wang, J.; Yu, L.; Wanf, J.; Guo, L.; Gu, X.; Fang, Q. Automated Fugl-Meyer assessment using SVR model. In Proceedings of the 2014 IEEE International Symposium on Bioelectronics and Bioinformatics (IEEE ISBB 2014), Chung Li, Taiwan, 11–14 April 2014. [[CrossRef](#)]
- Wade, E.; Parnandi, A.R.; Mataric, M.J. Automated administration of the Wolf Motor Function Test for post-stroke assessment. In Proceedings of the 2010 4th International Conference on Pervasive Computing Technologies for Healthcare, Munich, Germany, 22–25 March 2010. [[CrossRef](#)]
- Del Din, S.; Patel, S.; Cobelli, C.; Bonato, P. Estimating Fugl-Meyer clinical scores in stroke survivors using wearable sensors. In Proceedings of the 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Boston, MA, USA, 30 August–3 September 2011. [[CrossRef](#)]
- Patel, S.; Hughes, R.; Hester, T.; Stein, J.; Akay, M.; Dy, J.G.; Bonato, P. A novel approach to monitor rehabilitation outcomes in stroke survivors using wearable technology. *Proc. IEEE* **2010**, *98*, 450–461. [[CrossRef](#)]
- Zadeh, L.A. The concept of a linguistic variable and its application to approximate reasoning-II. *Inf. Sci.* **1975**, *8*, 301–357. [[CrossRef](#)]
- Zadeh, L.A. From computing with numbers to computing with words—from manipulation of measurements to manipulation of perceptions. *IEEE Trans. Circuits Syst. I Fundam. Theor. Appl.* **1999**, *45*, 105–119. [[CrossRef](#)]

25. Zadeh, L.A. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans. Syst. Man Cybern.* **1973**, SMC-3, 28–44. [[CrossRef](#)]
26. Zadeh, L.A. Fuzzy sets. *Inf. Control* **1965**, *8*, 338–353. [[CrossRef](#)]
27. Wu, Y.; Zhang, B.; Lu, J.; Du, K.L. Fuzzy logic and neuro-fuzzy systems: A systematic introduction. *Int. J. Artif. Intell. Expert Syst.* **2011**, *2*, 47–80.
28. Ajiboye, A.B.; Weir, R.F. A heuristic fuzzy logic approach to EMG pattern recognition for multifunctional prosthesis control. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2005**, *13*, 280–291. [[CrossRef](#)]
29. Zheng, H.; Zhu, K.Y. A fuzzy controller-based multiple-model adaptive control system for blood pressure control. In Proceedings of the ICARCV 2004 8th Control, Automation, Robotics and Vision Conference, Kunming, China, 6–9 December 2004. [[CrossRef](#)]
30. Mamdani, E.H.; Assilian, S. An experiment in linguistic synthesis with a fuzzy logic controller. *Int. J. Man Mach. Stud.* **1975**, *7*, 1–13. [[CrossRef](#)]
31. Farhoud, A.; Erfanian, A. Fully automatic control of paraplegic FES pedaling using higher-order sliding mode and fuzzy logic control. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2014**, *22*, 533–542. [[CrossRef](#)]
32. Hwang, Y.; Lee, S.; Hong, J.; Kim, J. A novel end-effector robot system enabling to monitor upper-extremity posture during robot-aided planar reaching movements. *IEEE Robot. Autom. Lett.* **2020**, *5*, 3035–3041. [[CrossRef](#)]
33. Otte, K.; Kayser, B.; Mansow-Model, S.; Verrel, J.; Paul, F.; Brandt, A.U.; Schmitz-Hübsch, T. Accuracy and reliability of the kinect version 2 for clinical measurement of motor function. *PLoS ONE* **2016**, *11*, e0166532. [[CrossRef](#)]
34. Huber, M.E.; Seitz, A.L.; Leeser, M.; Sternad, D. Validity and reliability of Kinect skeleton for measuring shoulder joint angles: A feasibility study. *Physiotherapy* **2015**, *101*, 389–393. [[CrossRef](#)]
35. McConville, J.T.; Clauser, C.E.; Churchill, T.D.; Cuzzi, J.; Kaleps, I. *Anthropometric Relationships of Body and Body Segment Moments of Inertia*; Tech Report, AFAMRL-TR-80-119; Anthropol Res Proj Inc.: Yellow Springs, OH, USA, 1980.
36. Jeong, H.; Kim, J. Development of a guidance system for motor imagery enhancement using the virtual hand illusion. *Sensors* **2021**, *21*, 2197. [[CrossRef](#)]