

Sensor Abstracted Extremity Representation for Automatic Fugl-Meyer Assessment

Patrick Heyer¹(✉), Felipe Orihuela-Espina¹, Luis R. Castrejón²,
Jorge Hernández-Franco³, and Luis Enrique Sucar¹

¹ Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Puebla, Mexico
`patrickhey@prodigy.net.mx`

² Hospital Universitario de la Benemérita Universidad Autónoma de Puebla
(HU-BUAP), Puebla, Mexico

³ Instituto Nacional de Neurología y Neurocirugía (INNN), Mexico City, Mexico

Abstract. Given its virtually algorithmic process, the Fugl-Meyer Assessment (FMA) of motor recovery is prone to automatization reducing subjectivity, alleviating therapists' burden and collaterally reducing costs. Several attempts have been recently reported to achieve such automatization of the FMA. However, a cost-effective solution matching expert criteria is still unfulfilled, perhaps because these attempts are sensor-specific representation of the limb or have thus far rely on a trial and error strategy for building the underpinning computational model. Here, we propose a sensor abstracted representation. In particular, we improve previously reported results in the automatization of FMA by classifying a manifold embedded representation capitalizing on quaternions, and explore a wider range of classifiers. By enhancing the modeling, overall classification accuracy is boosted to 87% (mean: $82\% \pm 4.53$;) well over the maximum reported in literature thus far 51.03% (mean: $48.72 \pm \text{std}: 2.10$). The improved model brings automatic FMA closer to practical usage with implications for rehabilitation programs both in ward and at home.

Keywords: Automatic motor dexterity assessment · Gesture classification · Gesture representation · Sensor independent representation · Automatic Fugl-Meyer

1 Introduction

The economic burden of motor rehabilitation programs for patients with motor disability due to stroke or traumatic brain injury among others to public health systems as well as families is untenable [8]. Obvious measures to contain those costs include reducing the continuous demand of expert supervision during the rehabilitation therapy sessions. Robotic rehabilitation [12], virtual rehabilitation [1] and telerehabilitation [2] are among a new generation of rehabilitation therapeutic modalities which, with current status, already can match classical

occupational therapy (moderate [3]) success on motor recovery, but which alleviate the need for continuous supervision. Moreover, the later two can easily be delivered at patient’s home further reducing costs without compromising the recovery with still margin for improvement.

Whether in ward or at home, these innovative therapeutic alternatives still rely on an expert for something as routinary as the assessment of motor recovery using a clinically validated scale such as the Fugl-Meyer Assessment (FMA) [4]. Since the FMA is applied routinely to monitor patient progress and its application is almost algorithmic, it is no surprise that several attempts have been made to automatize the assessment procedure [6]. The automatization of the FMA can free therapist time, reduce any remainings of subjective appreciation [5], and also afford the aforementioned therapeutic alternatives even greater independence and wider home applicability. Ultimately, automatic assessment of the patient motor recovery may proved to be the enabling element for home based rehabilitation.

Despite the obvious interest to have an automatized version of the FMA, having a definitive solution sufficiently reliable that can help the clinicians remains unsolved. Even though we have suggested above that the assessment procedure proceeds almost algorithmically, matching the human expert criterion with a cost-effective solution is challenging. Differences in sensing strategies i.e. selection of the appropriate sensors, and their positioning in the assessment stage whether on-body or off-body, differences in signal processing and analysis strategies, and small variations in clinical application of the assessment among experts are likely candidates to explain this current mild success of the computational models developed for the task.

If the above hypothesized candidacies are preventing higher success of the computational models, then it is likely that the combination of (a) developing a body mechanics representation weakly dependent on the sensing strategy and (b) optimizing the subsequent modeling decisions shall boost the accuracy and success rates of the automatic assessment model. Focusing only on the upper limb, our contribution here is a new abstract representation of the arm mechanics that reduces the commonly tight dependency of the representation on the acquiring sensor. This is achieved by projecting different sensing configurations to a common space capitalizing on quaternions. Then, the classification stage is flexibilized by testing a wide number of combinations between processing and decision making stages. Preliminary results of this research (only 6 patients and only 2 classification models) have been published in [22].

2 Related Work

Given the obvious benefits of having an automated version of the motor assessment procedures, it is unsurprising that a number of solutions have been proposed in the literature across different clinical scales e.g. the Arm Motor Ability Test (AMAT) [9], the Wolf Motor Function Test (WMFT) [10, 19], the Chedoke-McMaster Hand Stage [11], and of course the Fugl-Meyer score [6, 7]. These

solutions either use expensive sensing geometries or their classification rate is still far from satisfying. Moreover, although those methods with expensive sensing setup have managed to obtain good results in controlled settings but they tend to have an obtrusive factor in the sensor arrangement that makes them unfit for the very purpose they were developed!. Additionally, these methods also use the data from the sensors directly as features for their evaluation which makes the solution sensor specific. Finally, several solutions for assessment of motor dexterity following impairment have been further suggested with unobtrusive low-cost sensing geometries e.g. [18,20], but since they do not rely on clinical standard scales they are naturally of limited interest. In summary, a cost-efficient solution capable of matching expert evaluation on clinically validated schemes is still elusive.

3 Methods

3.1 Experiment Setup

Following consent, 9 patients with motor impairment from different origin underwent FMA agreed to participate from two hospitals in Mexico; Hospital Universitario de la Benemérita Universidad Autónoma de Puebla (HU-BUAP) and Hospital General Sur de Puebla (HGSP). The patients present different types of neurological damage including: stroke, and traumatic brain injury, but all require FMA as part of their rehabilitation process. Blinded pictures of the hospital sessions at both hospitals are shown in Fig. 1. The participants were monitored during the assessment performed by a trained clinician whilst their upper limb kinematics were concurrently being monitored using two sensing geometries (Fig. 1); (a) two Inertial Measuring Units (IMU) -one within an *ad-hoc* controller of a virtual rehabilitation platform developed previously by our group [13] and (b) one Microsoft KinectTM. Additionally, the experimental session was video recorded for visual inspection purposes. Table 1 summarises the 10 items of the Fugl-Meyer score for the upper limb.

To compensate for this small sample size¹ additional data was further collected from 15 healthy volunteers recruited among the faculty and students of the Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE) in Mexico executing the Fugl-Meyer exercises using the same setup and protocol as used in the clinical data capture. The healthy volunteers carried out five repetitions of the exercises defined in the upper extremity subsection of the FMA, simulating all three levels of motor dexterity (at their own interpretation after a brief description by the experimenter). Data was segmented in a bespoke software developed by our group was used to separate the samples corresponding to each of the five repetitions, giving us a total of 750 synthetic samples (10 FMA exercises \times 5 repetitions \times 15 subjects) for each of the 3 levels of FMA in addition to the 60 samples obtained from patients.

¹ We are currently in the process of collecting further clinical data.

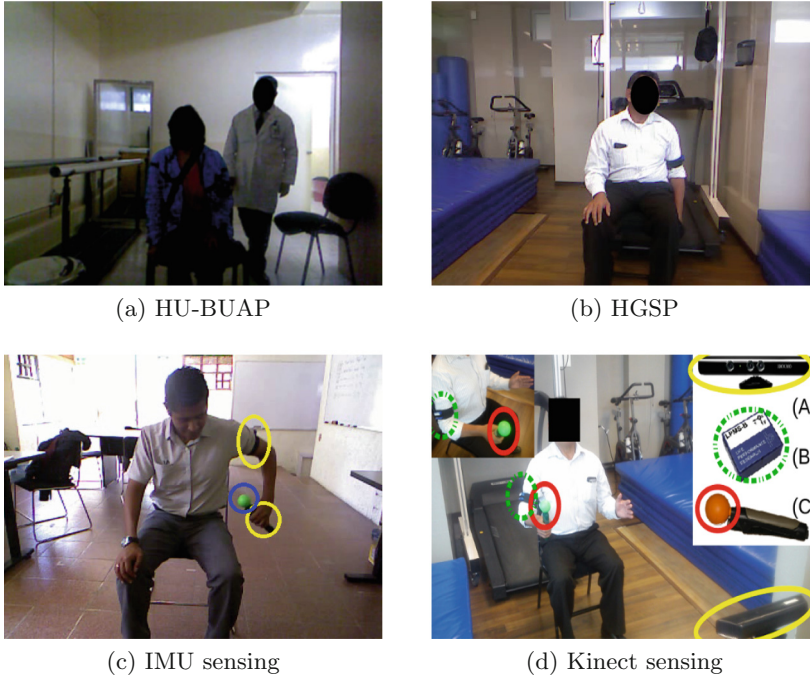


Fig. 1. Experimental setup and sensing geometry. (a-b) Sessions of motor assessment of patients at the two participating hospitals. (c) The placement of IMU's *yellow*, and color tracking reference *Blue*. (d) Location of the arm joints and relationship as established by the Kinect sensor. (Color figure online)

Table 1. The 10 items Fugl-Meyer subscale for the upper limb. Upon execution, each different exercise is scored 0 (no movement), 1 (clearly impaired movement) or 2 (normal or close to normal movement).

| Exercise | Description |
|----------|--|
| 1 | Move hand from knee to same side ear |
| 2 | Move hand from knee to opposite side knee |
| 3 | Move hand from knee to lumbar spine |
| 4 | Raise hand from knee to 90° (pointing at horizon) |
| 5 | With elbow touching body rotate hand |
| 6 | Raise hand from knee to 90° sideways |
| 7 | Raise hand from pointing at horizon to straight up |
| 8 | With elbow touching body flex and extend hand |
| 9 | With elbow touching body rotate hand clockwise |
| 10 | Move hand from knee to nose five times as fast as possible |

3.2 Abstract Representation of the Arm Mechanics

To partially address our hypothesis, we propose a representation R that is non-specific across a family of motion amenable sensors² to maintain the classification problem independent from the sensing technologies available at the rehabilitation centers. This representation R is the result of a composition of functions. First, transforming each sample of the sensors' output S to an orientation space represented by quaternions f where the dimensions are dictated by the limb segments upper arm (*brachium*), forearm (*antebrachium*) and hand (*manus*). Then, the dimensional differences in variance are nullified using some normalization g . Finally, by considering the movement during each FMA exercise as a trajectory in the space, a (manifold) projection to a space of salient components is used to reduce dimensionality while maintaining the most significant features h . The full representation is illustrated in Fig. 2 and formally given by Eq. 1:

$$R = h \circ g \circ f(S) \quad (1)$$

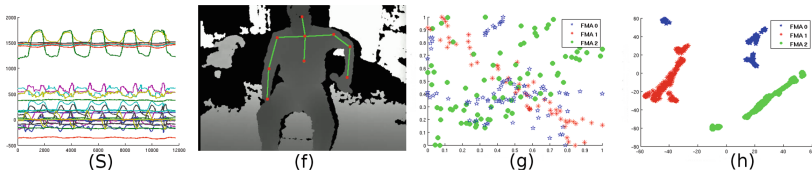


Fig. 2. Plot of the data during the different stages of composition. (S) Sensor data acquired by either, (f) Limb segment orientation (illustrative), (g) Normalized data from different FMA scores, (h) Projection of salient components using t-SNE.

The next subsections detail the last two transformations.

Normalization. To nullify dimensional variance one of three different normalization schemes were used, namely: classical *normalization* to a unitary range, *regularization*, and *quaternion* normalization. The first two consist in scaling the data to avoid overfitting when training machine learning algorithms. The third is avoids floating-point precession errors that will cause a quaternion not to have a unit length. The particular choice of each of these possible transformations lead to different representations, all sharing their detachment of the sensing geometry.

Feature Extraction; Projection to Salient Component Space. A large number of manifold embedding approaches (and its corollary dimensionality

² Full abstraction from the sensing geometry is beyond the scope of this work. For instance, we do not aim at being capable of achieving our goal of automatic motor assessment from say thermal sensors.

reduction application) exist, ranging from the classical Principal Component Analysis (PCA) to the sophisticated Isomap. For a review of the topic, the reader is directed to [21]. They all involve two steps; whether implicit or explicit; (1) imposing a distance function defining the topology of the space, and (2) projecting to a “different” space, often with less dimensionality, either by choosing a different view, i.e. a different coordinate set, and/or removing those dimension of less interest³. Previously we have also explore the rotational effect of PCA [22]. Here, considering the dynamics of the exercises that we aim to decode we opted for a projection with t-distributed stochastic neighbor embedding (t-SNE) [14]. t-SNE is a nonlinear probabilistic embedding which favours a similarity definition that abstracts the dynamics of the process as opposed to alternatives with a similarity purely based on the manifold shape e.g. Isomap or Locally Linear Embedding, as exemplified by its impressive separation of the handwriting MNIST dataset in the original publication. Specifically, t-SNE models each high-dimensional object into a (normally) two or three dimensional point Fig. 3 by converting the ambient Euclidean distances between datapoints into conditional probabilities representing similarities among objects as per Eq. 2.

$$dist(x_i, x_j) = p_{x_i, x_j} = \frac{\exp(-\|x_i - x_j\|^2 / 2\sigma_i^2)}{\sum_{k \neq j} \exp(-\|x_i - x_k\|^2 / 2\sigma_i^2)} \quad (2)$$

where $x \in X$ are the datapoints and σ_i is the variance of the Gaussian that is centered on datapoint x_i .

The projection itself minimizes the sum of Kullback-Leibler divergences over all datapoints using gradient descent.

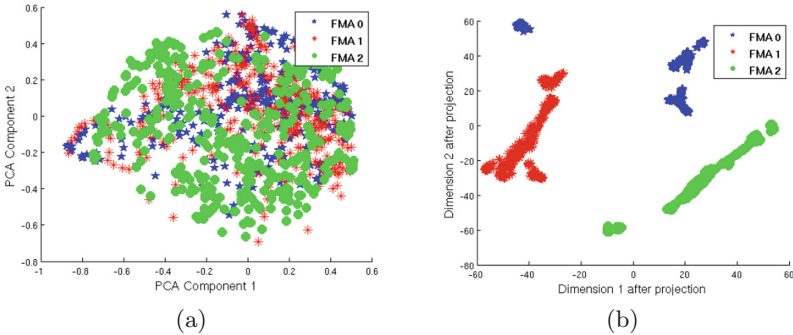


Fig. 3. Visual comparison of two different projection techniques showing clear advantage for using t-SNE in this particular domain. In both cases, the example corresponds to data from exercise seven from the FMA, and the colors index the labelled assessment score in that particular exercise. (a) First two salient dimensions using PCA. There is no evident separability among the different classes. (b) First two salient dimensions using t-SNE. The high separability among different classes is self-evident.

³ The definition of what is an interesting view of the dataset correspond to the domain demands.

3.3 Classification

Given the representation achieved with Eq. 1, the assessment itself consists in labeling the observation with a score 0, 1 or 3 according to the Fugl-Meyer scale. This is addressed here as a classical supervised classification problem. The classification model partitions the space into subregions each one assigned a class label. Figure 4 illustrates several partitioning possibilities. The problem is then not so much to build a particular model, but to choose the best classifier in some sense, often in terms of their capacity for generalization.

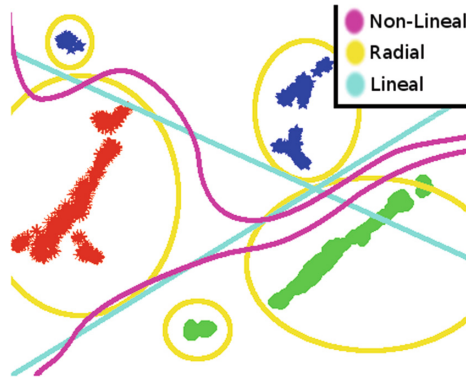


Fig. 4. Different separation possibilities for the same representation. Although the depicted possibilities are all correct, they differ in their capacity to generalize.

To determine the best classification strategy different classifiers were trained to compare using standard metrics:

- $Accuracy = (TP + TN) / (TP + FP + TN + FN)$
- $Precision = TP / (TP + FP)$
- $Sensitivity = TP / (TP + FN)$
- $Specificity = TN / (FP + TN)$
- $F - score = 2 * TP / (2 * TP + FP + FN)$

Where TP = true positives, TN=true negatives, FP = false positives and FN = false negatives. Since there are 3 classes for each exercise, the confusion matrices are summarized to the binary hit or miss labeling. For presenting the results, the different metrics are later averaged across all ten FMA exercises.

Validation was attempted by means of classical cross-folding experimental replication. A total of 2400 ($= 2 \times 3 \times 4 \times 10 \times 10$) classification exercises were carried out using;

- 2 sensing geometries (IMU-based or Kinect-based),
- 3 different normalization methods (see section on normalization above),

- 4 different classifiers; Naive Bayes classifier [15], Random Forest (RF) [16], and Support vector machines (SVM) [17] using linear and radial function based kernels,
- 10 FMA exercises (see Table 1), and
- 10 fold repetition for cross-folding based assessment of internal (reproducibility) and external (generalizability) validity.

ANOVA at 5% significance was used to determine statistical significance. Mann-Whitney-U pairwise comparisons when ANOVA detected significant differences in at least one treatment.

A leave one out classification comparison was made comparing results of classifying healthy participants and patients using a mixed training set and classifying all exercises for one participant at a time.

4 Results

Table 2 summarizes the classification rates across the different metrics. The specific combination of quaternion normalization and SVM with radial basis function kernel systematically affords the higher means (assumed to be associated to generalizability) and lower standard deviations (std) (assumed to be associated to reproducibility).

Figure 2 shows the average classification rates among the different classifiers using quaternion normalization with three different sensing setups, exhibiting similar results independent of the sensing technique used as hypothesized (Fig. 5).

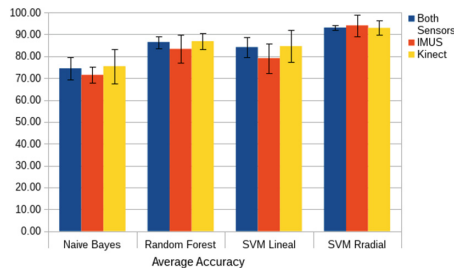


Fig. 5. Average classification for the different sensing setups. Three possible data inflow combinations are shown: Using both IMU’s and Kinect data, using only the data from the IMU’s, and that only of the Kinect sensor. (mean \pm std)

The effect of including data from healthy participants is dissected in Table 3. Understandably higher rates of classification are found for healthy subjects. This is not to be misunderstood as inflated rates. Although it is tempting to quickly argue that only data from patient should be accounted for, that is true only for the evaluation, but not for the training. In other words, only classification of

Table 2. Summary of classification rates achieved by the different approaches. In all cases mean \pm std is indicated. Top most results are highlighted in gray.

| Accuracy | | | | |
|--------------------------|------------------|------------------|-------------------|------------------|
| | {Naive Bayes} | {Random Forest} | {SVM linear} | {SVM radial} |
| Normalization | 76.96 \pm 9.32 | 79.43 \pm 6.29 | 63.81 \pm 12.54 | 71.05 \pm 5.23 |
| Regularization | 71.02 \pm 8.56 | 73.43 \pm 5.98 | 69.26 \pm 11.03 | 80.23 \pm 5.62 |
| Quaternion Normalization | 74.48 \pm 5.23 | 86.49 \pm 2.74 | 84.18 \pm 4.62 | 93.12 \pm 1.09 |
| Precision | | | | |
| | {Naive Bayes} | {Random Forest} | {SVM linear} | {SVM radial} |
| Normalization | 69.05 \pm 8.34 | 72.18 \pm 7.30 | 54.20 \pm 12.61 | 62.07 \pm 4.96 |
| Regularization | 62.12 \pm 8.25 | 64.97 \pm 5.71 | 60.17 \pm 11.00 | 73.15 \pm 5.25 |
| Quaternion Normalization | 66.01 \pm 4.89 | 81.04 \pm 3.29 | 78.25 \pm 4.83 | 90.19 \pm 0.87 |
| Sensitivity | | | | |
| | {Naive Bayes} | {Random Forest} | {SVM linear} | {SVM radial} |
| Normalization | 69.13 \pm 9.67 | 72.15 \pm 6.31 | 54.32 \pm 12.56 | 62.15 \pm 4.60 |
| Regularization | 62.06 \pm 9.01 | 64.85 \pm 6.68 | 60.27 \pm 12.05 | 73.14 \pm 5.76 |
| Quaternion Normalization | 66.04 \pm 5.01 | 80.96 \pm 2.64 | 78.05 \pm 5.33 | 90.01 \pm 1.32 |
| Specificity | | | | |
| | {Naive Bayes} | {Random Forest} | {SVM linear} | {SVM radial} |
| Normalization | 81.67 \pm 8.43 | 83.78 \pm 6.92 | 70.24 \pm 12.49 | 76.59 \pm 5.12 |
| Regularization | 76.50 \pm 8.31 | 78.64 \pm 6.74 | 75.08 \pm 11.87 | 84.44 \pm 6.03 |
| Quaternion Normalization | 79.55 \pm 4.75 | 89.48 \pm 3.15 | 87.69 \pm 4.72 | 94.75 \pm 0.61 |
| F-score | | | | |
| | {Naive Bayes} | {Random Forest} | {SVM linear} | {SVM radial} |
| Normalization | 68.96 \pm 8.58 | 71.96 \pm 7.25 | 54.04 \pm 12.33 | 62.08 \pm 6.03 |
| Regularization | 61.82 \pm 8.53 | 64.80 \pm 5.62 | 60.04 \pm 11.76 | 73.00 \pm 5.33 |
| Quaternion Normalization | 66.02 \pm 5.92 | 80.78 \pm 2.34 | 78.00 \pm 4.67 | 90.06 \pm 0.45 |
| AUC | | | | |
| | {Naive Bayes} | {Random Forest} | {SVM linear} | {SVM radial} |
| Normalization | 56.42 \pm 8.71 | 60.46 \pm 6.83 | 38.17 \pm 12.74 | 0.476 \pm 7.09 |
| Regularization | 47.43 \pm 8.47 | 50.91 \pm 5.97 | 45.20 \pm 11.58 | 61.75 \pm 5.84 |
| Quaternion Normalization | 52.57 \pm 5.45 | 72.46 \pm 2.85 | 68.42 \pm 4.86 | 85.32 \pm 0.31 |

patient observations has to be accounted, but the training of the model should benefit of whatever information can be given. Dissociation of the benefit of including these data in the training of the classifiers, but not counting them in the classification rates is still pending (Fig. 6).

5 Discussion and Conclusions

The proposed representation affords high classification rates⁴ regardless of the classifier and without depending on a specific sensing technology. The representation benefits from the so called quaternion normalization. Its low dependency on the sensing geometry, suggested by the small differences in classification across the tested geometries, suggests that the approach proposed facilitates its use in

⁴ Previously reported values were well below these figures.

Table 3. Classification rates by type of participant. Dissociation of the effect in the classification rates due to the participant status.

| Exercise | Patients | Healthy | Patients and Healthy |
|----------|-------------------|------------------|----------------------|
| 1 | 45.09 \pm 7.89 | 95.75 \pm 0.92 | 93.19 \pm 2.66 |
| 2 | 44.84 \pm 8.65 | 95.64 \pm 1.92 | 92.65 \pm 0.8 |
| 3 | 42.17 \pm 3.41 | 98.83 \pm 2.6 | 95.74 \pm 0.34 |
| 4 | 45.39 \pm 9.78 | 95.25 \pm 0.26 | 93.11 \pm 2.66 |
| 5 | 42.47 \pm 2.63 | 96.34 \pm 1.79 | 94.15 \pm 1.52 |
| 6 | 39.14 \pm 10.46 | 96.49 \pm 2.99 | 93.19 \pm 2.85 |
| 7 | 49.18 \pm 8.9 | 95.61 \pm 0.62 | 92.63 \pm 2.43 |
| 8 | 45.93 \pm 3.39 | 94.2 \pm 0.96 | 92.05 \pm 2.54 |
| 9 | 33.68 \pm 3.48 | 94.81 \pm 1.5 | 92.17 \pm 2.36 |
| 10 | 50.36 \pm 5.48 | 96.17 \pm 0 | 92.43 \pm 2.01 |

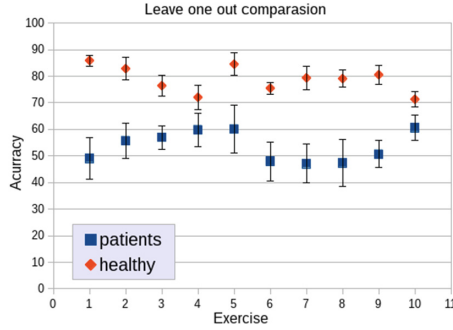


Fig. 6. Average classification for every exercise using leave one out classification results are shown as mean \pm std. *The similarity of the classification seems to be directly related to the level in the recovery hierarchy, during rehabilitation some movements are recovered sooner than others in a predefined sequence, further analysis is required.*

different rehabilitation settings including those where non-intrusive assessment is required.

The incorporation of data from healthy participants, unfortunately limits the generalization of these findings. Our previous results suggests these data may be easier to classify than patient data [22], and thus the classification rates may look like higher than they might have look shall only data from patient may have been used. In this sense, we consider that to get a stronger validation of our solution a bigger dataset only from patients data is necessary.

Although the present effort has enhanced the classification by exploring several possible combinations, this is still open to mathematical optimization. We are currently working to achieve optimal modeling by means of full model selection techniques.

Apart from the use of the proposed method for upper extremity assessment our proposal should be suitable for other kinds of movement analysis the possibilities being currently being discussed are directly related to the FMA scale using whole body information. Other areas could be gait analysis and evaluation of Parkinson syndrome tremors.

Acknowledgment. The leading author has received a scholarship No. 339981 from CONACYT.

References

1. Adamovich, S.V., Fluet, G.G., Tunik, E., Merians, A.S.: Sensorimotor training in virtual reality: a review. *NeuroRehabilitation* **25**, 29 (2009)
2. Reinkensmeyer, D.J., Pang, C.T., Nessler, J.A., Painter, C.C.: Web-based telerehabilitation for the upper extremity after stroke. *IEEE Trans. Neural Syst. Rehabil. Eng.* **10**, 102–108 (2002)
3. Krakauer, J.W., Carmichael, S.T., Corbett, D., Wittenberg, G.F.: Getting neurorehabilitation right: what can be learned from animal models? *Neurorehabilitation Neural Repair* **26**, 923–931 (2012)
4. 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.* **7**, 13–31 (1975)
5. Duncan, P.W., Propst, M., Nelson, S.G.: Reliability of the Fugl-Meyer assessment of sensorimotor recovery following cerebrovascular accident. *Phys. Ther.* **63**, 1606–1610 (1983)
6. Quintana, G.E., et al.: Qualification of arm gestures using hidden markov models. In: 8th IEEE International Conference on Automatic Face & Gesture Recognition, FG 2008, pp. 1–6. IEEE (2008)
7. Hou, W.-H., Shih, C.-L., Chou, Y.-T., Sheu, C.-F., Lin, J.-H., Wu, H.-C., Hsueh, I.-P., Hsieh, C.-L.: Development of a computerized adaptive testing system of the Fugl-Meyer motor scale in stroke patients. *Arch. Phys. Med. Rehabil.* **93**, 1014–1020 (2012)
8. Ma, V.Y., Chan, L., Carruthers, K.J.: The incidence, prevalence, costs and impact on disability of common conditions requiring rehabilitation in the US: stroke, spinal cord injury, traumatic brain injury, multiple sclerosis, osteoarthritis, rheumatoid arthritis, limb loss, and back pain. *Arch. Phys. Med. Rehabil.* **95**(5), 986–995.e1 (2014)
9. Allin, S., Ramanan, D.: Assessment of post-stroke functioning using machine vision. In: MVA2007 IAPR Conference on Machine Vision Applications, 16-18 May, Tokyo, Japan, pp. 8–18 (2007)
10. Virgilio, F.B., Cruz, V.T., Ribeiro, D.D., Cunha, J.P.: Towards a movement quantification system capable of automatic evaluation of upper limb motor function after neurological injury. In: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC, pp. 5456–5460. IEEE (2011)
11. Hester, T., Hughes, R., Sherrill, D.M., Knorr, B., Akay, M., Stein, J., Bonato, P.: Using wearable sensors to measure motor abilities following stroke. In: International Workshop on Wearable and Implantable Body Sensor Networks, BSN 2006, p. 4. IEEE (2006)

12. Balasubramanian, S., Wei, R., Perez, M., Shepard, B., Koeneman, J., Koeneman, E., He, J.: RUPERT: an exoskeleton robot for assisting rehabilitation of arm functions. In: *Virtual Rehabilitation*, 163–167. IEEE (2008)
13. Sucar, L.E., Orihuela-Espina, F., Velazquez, R.L., Reinkensmeyer, D.J., Leder, R., Hernández Franco, J.: Gesture therapy: an upper limb virtual reality-based motor rehabilitation platform. *IEEE Trans. Neural Syst. Rehabil. Eng.* **22**(3), 634–643 (2014)
14. der Maaten, V.L., Hinton, G.: Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008)
15. Murphy, K.P.: *Naive Bayes classifiers*. University of British Columbia (2006)
16. Svetnik, V., Liaw, A., Tong, C., Culberson, J.C., Sheridan, R.P., Feuston, B.P.: Random forest: a classification and regression tool for compound classification and QSAR modeling. *J. Chem. Inf. Comput. Sci.* **43**, 1947–1958 (2003)
17. Hearst, M.A., Dumais, S.T., Osman, E., Platt, J., Scholkopf, B.: Support vector machines. *IEEE Intell. Syst. Appl.* **13**, 18–28 (1998)
18. Olesh, E.V., Yakovenko, S., Gritsenko, V.: Automated assessment of upper extremity movement impairment due to stroke. *PLoS ONE* **9**(8), e104487 (2014)
19. Wade, E., Parnandi, A.R., Matarić, M.J.: Automated administration of the Wolf Motor Function test for post-stroke assessment. In: *4th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, Munich, Germany, pp. 1–7 (2010)
20. Hondori, H.M., Ling, S.-F.: A method for measuring human arm’s mechanical impedance for assessment of motor rehabilitation. In: *3rd International Convention on Rehabilitation Engineering & Assistive Technology (i-CREATE 2009)*, Singapore, p. 4 (2009)
21. Carreira-Perpiñán, M.A.: *A review of dimension reduction techniques* University of Sheffield, University of Sheffield, Technical report, CS-96-09 (1997)
22. Heyer, P., Felipe, O.-E., Castrejón, L.R., Hernández-Franco, J., Sucar, L.E.: Sensor adequacy and arm movement encoding for automatic assessment of motor dexterity for virtual rehabilitation. Accepted at 9th World Congress for NeuroRehabilitation