2020 17th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)
Mexico City, Mexico. November 11-13, 2020

Clustering of Data that Quantify the Degree of Impairment of the Upper Limb in Patients with Alterations of the Central Nervous System

Leonardo Anaya
Movement Analysis Lab
Instituto Nacional de Rehabilitacion
CDMX, Mexico
ORCID 0000-0002-3254-4243

Ivett Quinones
Movement Analysis Lab
Instituto Nacional de Rehabilitacion
CDMX, Mexico
ORCID 0000-0003-3522-5085

Yannick Quijano
Movement Analysis Lab
Instituto Nacional de Rehabilitacion
CDMX, Mexico
ORCID 0000-0003-4121-0576

Virginia Bueyes
Movement Analysis Lab
Instituto Nacional de Rehabilitacion
CDMX, Mexico
ORCID 0000-0002-8877-358X

Enrique Chong
Mechatronics Department
ITESM CEM
Atizapan, Estado de Mexico
ORCID 0000-0002-3192-1238

Victor Ponce
Computer Research Center
Instituto Politecnico Nacional
CDMX, Mexico
ORCID 0000-0001-5699-0478

***Abstract*—Previous studies have considered improving the classification procedures for motor impairment of the upper limb in patients with Central Nervous System alterations. This work compares two classification methods to be able to group the SSULF scale into five classes to have a better assessment, results showed that with the K-Means more than the 95% of the control group SALM values were correctly classified in SSULF 1 and with the Fuzzy C-means the 92%, so we can assume that the K-means method did a better classification for our purpose.**

***Index Terms*—clustering (unsupervised) algorithm, data classification, smoothness of movement.**

I. Introduction

The central nervous system (CNS) is vulnerable to various disorders caused by trauma, vascular interruptions, degeneration, structural defects, autoimmune disorders and tumors. These alterations can cause local and systemic deterioration and death, which makes them one of the main causes of multidisciplinary care and follow-up in the Luis Guillermo Ibarra Ibarra, National Institute of Rehabiitation (INR-LGII) for functional evaluation of the upper limb in patients with pathologies in the nervous system. such as spinal cord injury (SCI) and stroke (CVD).

These data were obtained by carrying out the Sorting Block Box (SBB) protocol, which consists of specific activities for filling and emptying an instrumented board [1]; The SBB together with an inertial measurement unit (IMU) sensor assess the patient's motor ability after processing data obtained and a specialized software, a smoothness of movement value is given for its interpretation and medical assessment. [2] [3], see Figure 1.

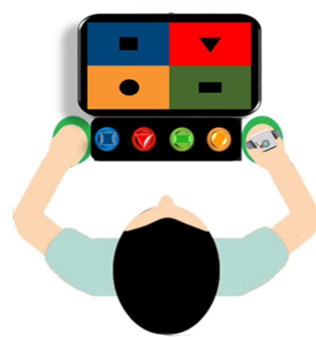

Figure 1: Measurement protocol with the SBB. [1]

II. Materials and Methods

A. Smoothness and evaluation

Various studies of smoothness of movement have created; new smoothness quantification metrics and compared among them to see which one obtained a better smoothness results. Smoothness of movement is shown as an effective, reliable and adequate measure to evaluate the upper limb, smooth and coordinated movements are an indication of healthy human motor control and learning [1].

These studies on the smoothness of movements, concluded in the realization of a new smoothness metric called Spectral Arc Length Metric ($SALM$) defined [3] as the arc length of the amplitude and the frequency normalized Fourier magnitude spectrum of the speed profile, considering a movement with a velocity profile $v(t)$ [5]:

$$n_{sal} = -\int_0^{\omega_c} \sqrt{(\frac{1}{\omega_c})^2 + (\frac{dV(\omega)}{d\omega})^2 * d\omega} \tag{1}$$

$$V(\omega) = \frac{V(\omega)}{V(0)} \quad (2)$$

Where in the Eq. 1 and Eq. 2 [5], ω_c = 20 Hz, is the upper cut frecuency that covers normal and abnormal human movements [2] and $V(\omega)$ is the Fourier magnitude spectrum of $v(t)$.

The spectral arc length metric assesses and evaluates a Fourier magnitude spectrum of the motion velocity profile to quantify smoothness of motion. This metric is tested systematically and compared with other measures of uniformity, using experimental data from impaired and healthy subjects, as well as simulated movement data. The results [3] indicate that the arc length spectral metric is a valid and consistent measure of smoothness of motion, which is sensitive to both changes in motor behavior and measurement noise. The Figure 2 is an example of the smoothness of two movements; on the left there is two speed profiles and on the right is the signal transformation of both with SALM.

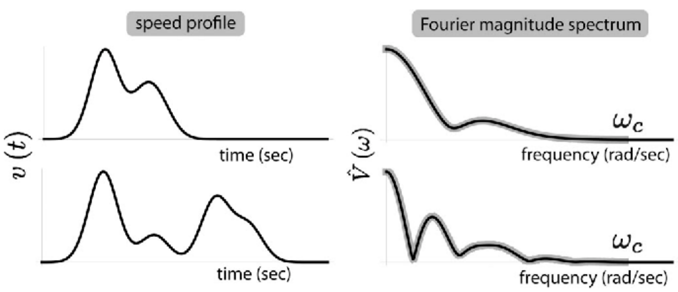

Figure 2: The motion speed profile with a smoother motion (top) is less complex than that of a non-smooth motion (bottom) [3].

Taking into account that this SALM metric uses kinematic data such as speed, it serves to detect the differences between motor control skills, learning and recovery, among all the metrics under investigation it is the chosen candidate to use in the experimental data of this work [1] [2] [3]. Figure 3 is an example of a movement smoothness quanitfied in the SALM scale with a our MATLAB software shown in Figure 4.

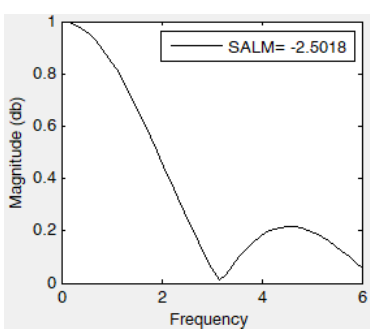

Figure 3: Result of the quantification of a task using the SALM metric.

The Smoothness Scale for Upper Limb Function (SSULF) [1] [3] is the proposed scale to be used based on the results obtained with the SALM.

B. Patients

For the experimentation, two groups were formed. One from healthy patients and one from patients with some motor impairment. The criteria for choosing healthy subjects and patients are shown below.

1) Inclusion criteria

- Healthy subjects 15-64 years of age who agree to participate in the study.
- Patients aged 15-64 years without total restriction of movement of the pelvic limb with diagnosis of:
  - CVD ischemic, transient ischemic or hemorrhagic.
  - Congenital, traumatic, infectious, tumor or secondary to systemic disease.

2) Elimination criteria

- Subjects who do not cooperate with the study.
- Subjects who do not wish to participate or do not wish to sign the informed consent.
- Subjects with severe cognitive, hearing or visual problems and / or not obeying orders.

3) Exclusion criteria

- Subjects without trunk control.
- Subjects taking sedation.

This procedure explained above was applied to the 28 subjects (control group, SCI, CVD); 2688 SALM results (28 patients x 6 attempts x 16 subtasks) were obtained. The patient data is shown below in table I and table II, SBBGC stands for control group, SBBEVC for CVD patients and SBBLM for SCI.

Table I: Control group data.

ID	HEALTHY SUBJECTS		
	Dominant	Age	Gender
SBBGC001	RIGHT	27	M
SBBGC002	RIGHT	28	M
SBBGC003	RIGHT	22	F
SBBGC004	RIGHT	50	F
SBBGC005	LEFT	21	F
SBBGC006	RIGHT	22	M
SBBGC007	RIGHT	27	F
SBBGC008	LEFT	37	M
SBBGC009	RIGHT	23	F
SBBGC010	RIGHT	47	F
SBBGC011	LEFT	53	F
SBBGC012	RIGHT	26	M
SBBGC013	RIGHT	23	M
SBBGC014	RIGHT	23	M
SBBGC015	RIGHT	23	F
SBBGC016	RIGHT	40	M

C. Data adquisition

For data acquisition, the Sorting Block Box (SBB) instrument was used, that consist of a board with holes in different geometric shapes (box, triangle, circle, rectangle) and 4 pieces of wood that must be placed in their respective spaces [1]; said board has various sensors inside that detect the force of the placement and whether or not the piece is in place, in addition to this, the information is being transfered by a data acquisition card (DAQ, National Instrument NI USB-6009),

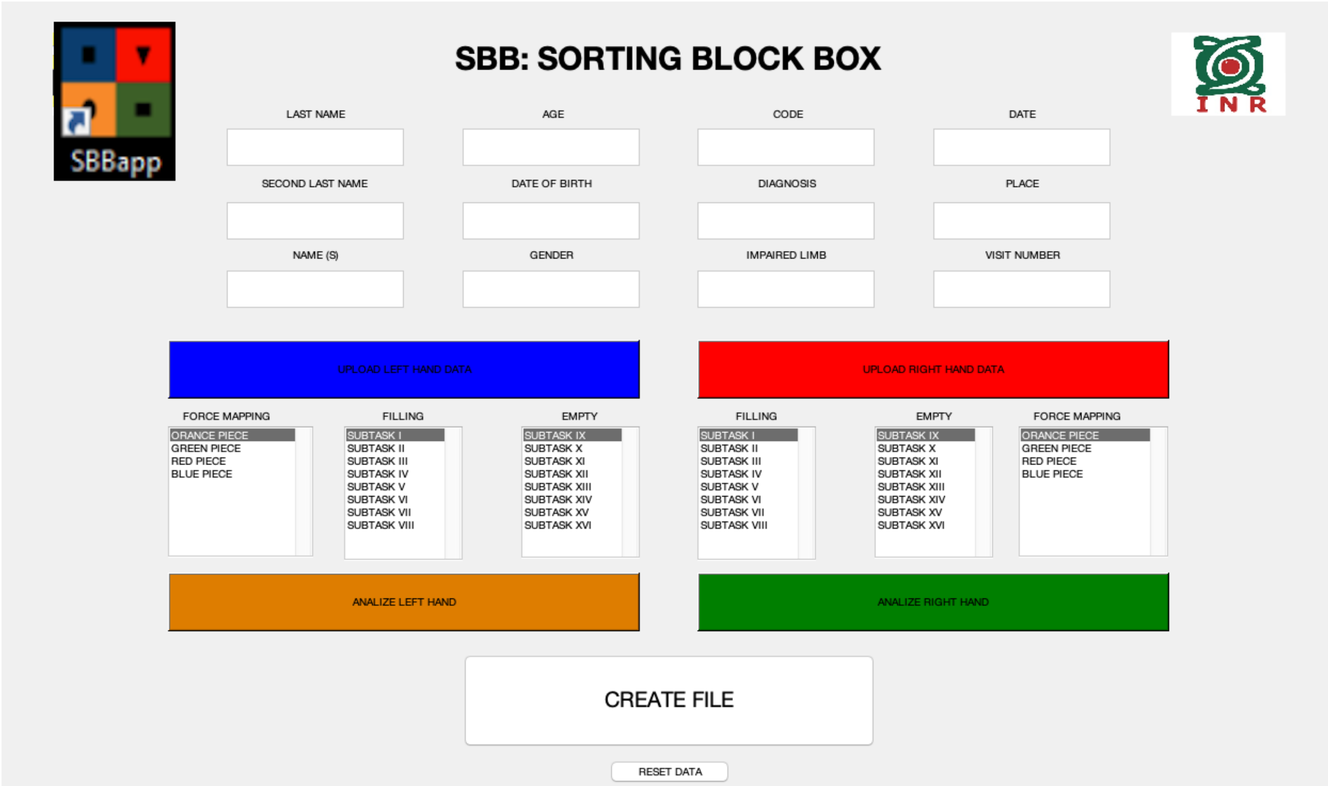

Figure 4: SBBapp is a software made in MATLAB that process the data from the IMU and SBB to obtain the SALM.

Table II: Data of the group of patients with diagnosed motor impairment.

ID	PATIENTS		
	Affected	Age	Gender
SBBEVC01	RIGHT	46	M
SBBEVC02	LEFT	42	M
SBBLM001	RIGHT	26	M
SBBLM002	RIGHT	22	M
SBBLM003	LEFT	20	M
SBBLM004	RIGHT	39	M
SBBLM005	LEFT	27	F
SBBLM006	RIGHT	76	F
SBBLM007	RIGHT	30	M
SBBLM008	RIGHT	30	M
SBBLM009	RIGHT	30	M
SBBLM010	RIGHT	23	M

to a computer [3]. After completing a part filling or emptying task, a program developed in LabVIEW (National Instruments, USA), generates two data files with the information collected from the various sensors involved, to be analyzed later. In addition to the SBB, an IMU sensor (Shimmer, Shimmer3), with sampling rate of 102.4 Hz, is placed on the back of the hand in order to quantify the trajectories [2], as shown in Figure 1; an observer was there all the time giving the instructions and checking that the system worked correctly during the trials.

The two files obtained are introduced to the software developed in MATLAB, see Fig. 4, which automatically segments the 16 tasks performed. This software allows to qualify each attempt of the patient and to obtain a SALM value for his motor skills of each one of his submoves, as shown in Figure 3.

With the 2688 SALM values, two classification methods were chosen for their relevance within the computational area are presented below [6]. They are among the most widely used for the classification of any data cloud. Both mentioned methods are able to significantly separate between the number of classes that are indicated, so they are reliable tools.

D. K-Means

K-means [7] is an unsupervised classification algorithm that groups objects into k groups based on their characteristics. Grouping is performed by minimizing the sum of distances between each object and the centroid of its group or cluster. Quadratic distance is often used. Therefore, given a set of observations N, the encoder C that assigns these observations to the clusters K such that, within each grouping, [8] the mean measure of difference of the assigned observations from the group mean is minimize.

1) The K-means algorithm is computationally efficient, since its complexity is linear in the number of clusters.
2) When clusters are tightly distributed in the data space, the algorithm faithfully recovers them.

E. C-Means

Fuzzy logic classification [9] is a class of grouping algorithms where each element has a diffuse degree of group membership. This type of algorithms arises from the need to solve a deficiency of the exclusive grouping, which considers that each element can be unequivocally grouped with the elements of its cluster and, therefore, does not resemble the rest of the elements.

Fuzzy C-Means algorithms [10] are some of the main algorithms used in fuzzy clustering and belong to a class of algorithms based on objective functions. These algorithms define a grouping criterion in the form of an objective function that depends on the fuzzy partition.

III. Classification of Data

After the experimentation process, the data obtained from a total of 2688 SBB repeats, the results obtained from the SALM of the control group and the diagnosed patients were used to construct a validation for extreme groups using two different methods (K-means and Fuzzy C-means) of clustering using the same data to compare and determine which one best classifies the data.

The following scatter diagram shows the distribution of the data of the 16 different subtasks from Sorting Block Box with the use of Shimmer IMU in the comparison of SALM vs SALM average.

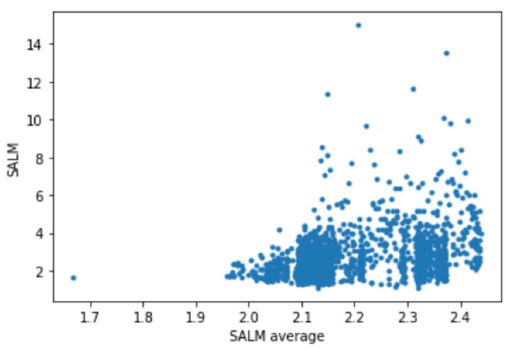

Figure 5: Clouded data.

A. Clustering with both methods

Using the K-Means grouping method, we obtained the results shown in Figure 6, where each different color represents a different class of the classification, with a total of 5 clusters.

Now using these results, we obtained for the SSULF classification the Parameter limits of each SSULF class:

$$SSULF = \begin{cases} 1. & 0 \leq SALM \leq 2.19 \\ 2. & 2.19 < SALM \leq 3.14 \\ 3. & 3.14 < SALM \leq 4.71 \\ 4. & 4.71 < SALM \leq 7.62 \\ 5. & 7.62 < SALM \end{cases} \tag{3}$$

Using the Fuzzy C-means grouping method we also obtained the results shown in figure 7; where each different color represents a different class of the classification, with a total of 5 clusters.

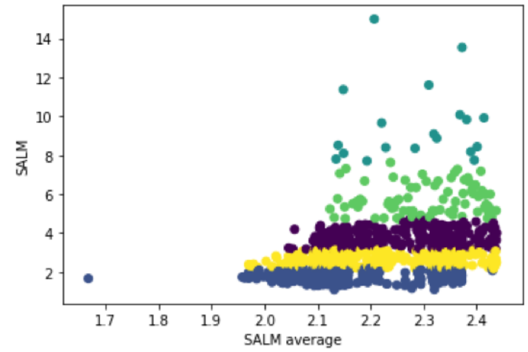

Figure 6: Clustering using Fuzzy Kmeans.

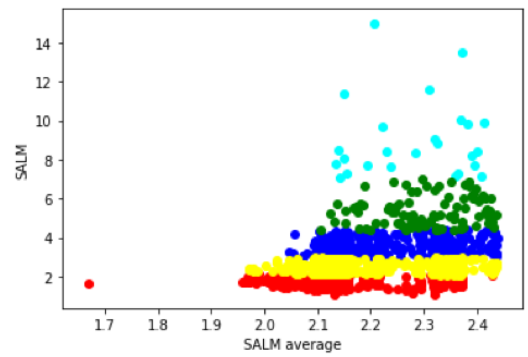

Figure 7: Clustering using Fuzzy Cmeans.

Now using these results, we obtained for the SSULF classification the parameter limits of each SSULF class:

$$SSULF = \begin{cases} 1. & 0 \leq SALM \leq 2.10 \\ 2. & 2.10 < SALM \leq 2.95 \\ 3. & 2.95 < SALM \leq 4.37 \\ 4. & 4.37 < SALM \leq 7.02 \\ 5. & 7.02 < SALM \end{cases} \tag{4}$$

Now, compared to the two methods proposed in this study, the table III was created to compare the results obtained with the different grouping methods.

Table III: Clustering method comparison

		SSULF				
No.	limits	1	2	3	4	5
(1)	MIN	0	2.20	3.15	4.72	7.63
	MAX	2.19	3.14	4.71	7.62	max
(2)	MIN	0	2.11	2.96	4.38	7.03
	MAX	2.10	2.95	4.37	7.02	max

1. K-Means
2. Fuzzy C-Means

You can see the difference between the limits of each class for the new SSULF metric proposed in this study. We can see that the Fuzzy C-means method is the one that has the classes with the limits less separated than the other technique; With the data obtained from the control group, we analyzed how many of the 1536 results were well classified using a classifier based on the artificial neural network approach with *backpropagation*, showing that with the K-Means more than the 95% of the control group SALM values were correctly classified in SSULF 1 and with the Fuzzy C-means the 92%; so we can assume that the K-means method did a better classification for our porpuse.

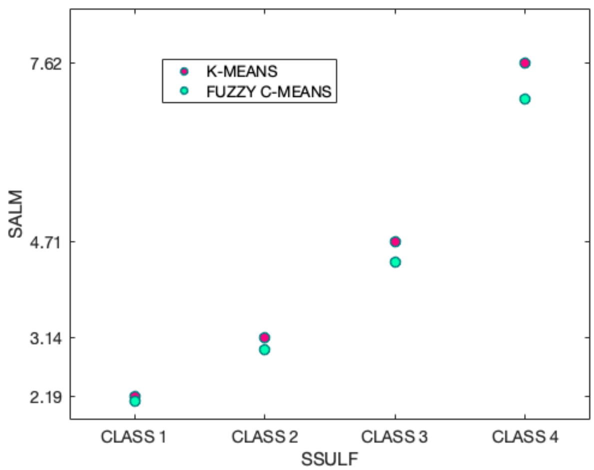

Figure 8: Graphical comparison of the two clustering methods.

IV. Conclusion

As a conclusion of this analysis, we can determine that with the data obtained and the classification made with the two methods (as can be seen in Figure 6 and Figure 7), we decided that the classification made by The K-means method is more consistent because it has a greater sensitivity when classifying SALM results into the SSULF class 1, that is, the classes are more separated from each other in their values, which makes them have a greater capacity when determining if a data is of one kind or another.

However, this does not mean that the classification made by the fuzzy-c or k-means method is incorrect, as you can see similarities between the classes of the two different classifications. Perhaps, with a modification of the classification algorithm, it is possible to improve the sensitivity and obtain a better grouping.

As this study is not definitive, since there are more and different grouping methods, future work could cover these doubts and clarify which method could achieve a better classification of the data presented, although with the approach made on this work we were able to prove that they do work and classify the information as expected.

There are several applications where decision making and exploratory pattern analysis must be performed on large data sets.

Studies of fuzzy or neuro-diffuse clustering, weighted arithmetic average, unweighted arithmetic average, room method, neural network methods are proposed, knowing that clustering is a subjective process; the same set of data elements often needs to be partitioned differently for different applications. This subjectivity makes the grouping process difficult.

References

[1] Quijano-Gonzalez, Y.; Melendez-Calderon, A.; Burdet, E.; Chong-Quero, J.E.; Villanueva-Ayala, D.; Perez-Moreno, J.C., ”Upper limb functional assessment of children with cerebral palsy using a sorting box” Engineering in Medicine and Biology Society (EMBC), 2014, pp.2330,2333.

[2] Quijano Gonzalez, Y., Quiñones Uriostegui Ivett, Chong-Quero J.E., Alessi Montero, A., Ulacia Flores, P., León Garmendia, A.. “IMU placement in the assessment of the upper limb using smoothness metrics”, 2016 GCMAS Annual Conference.

[3] Anaya, L., Quijano, Y., Quiñones, I., Alessi, A., Bueyes, V., Chong, E., ”Segmentacion de Tareas en el Dispositivo Sorting Block Bock, para la Valoración Funcional del Miembro Superior de Niños con Parálisis Cerebral”, Memorias del XXXIX Congreso Nacional de Ingenieria Biomedica, 2016.

[4] Balasubramanian, S., Melendez-Calderon, A., Roby-Brami, A. et al. On the analysis of movement smoothness. J NeuroEngineering Rehabil 12, 112 (2015).

[5] Selim, S.Z., and Ismail, M.A., ”K-Means-Type Algorithms: A Generalized Convergence Theorem and Characterization of Local Optimality” PAMI(6), No. 1, January 1984, pp. 81-87.

[6] Hartigan, J.A., Wong, M.A., ”A k-means clustering algorithm” AppStat(28), 1979, pp. 100-108.

[7] Navarro, A., Allen, C.R., ”Adaptive Classifier Based on K-Means Clustering and Dynamic Programming” OptEng(36), No. 1, 1997, pp. 31-38.

[8] Mc Neill, F.; Thro, E. (1994). ”Fuzzy Logic, a practical approach. (1ª Ed.)” Cambridge Mass.: Academic Press Ltd. 21.

[9] Mordeson, J.; Premchand S. (2001). ”Fuzzy Mathematics. (Studies in Fuzziness and soft computing No. 20)” New York: Physica-Verlag Heidelberg.