Journal of Innovations in Computer Science and Engineering
Shahid Beheshti Unversity

June 2025, Volume 3, Issue 1

Unlocking individual motor signatures using feature-based clustering of a graphomotor task

Zinat Zarandi✉, **Code Orcide:** 0000-0002-2488-4950
INSERM UMR1093-CAPS, UFR des Sciences du Sport, Université Bourgogne Franche-Comté, Dijon, France.
Amirreza Behmanesh, **Code Orcide:** 0009-0005-0783-1442
Department of Computer Engineering, Artificial Intelligent and Robotics, Amirkabir Kabir University of Technology, Tehran, Iran
Mohammad Mehdi Ebadzadeh, **Code Orcide:** 0000-0001-6466-5229
Department of Computer Engineering, Artificial Intelligent and Robotics, Amirkabir Kabir University of Technology, Tehran, Iran
Thierry Pozzo, **Code Orcide:** 0000-0002-0585-7965
INSERM UMR1093-CAPS, UFR des Sciences du Sport, Université Bourgogne Franche-Comté, Dijon, France.

Abstract**—**Understanding individual motor signatures (IMS) is essential for personalized treatment and performance optimization. This study investigates the effectiveness of Fuzzy C-Means (FCM) clustering for identifying individual motor signatures from graphomotor tasks. We analyze various kinematic and geometric features, such as movement duration, velocity, and trajectory length, to reveal which aspects of motor behavior are most effective in distinguishing individuals. The results show that features like length of movement are particularly discriminative, while others, such as beta and velocity, offer weaker clustering outcomes.

Keywords—Motor behavior, Fuzzy C-Means clustering, hand-drawing tasks, motor signatures, feature selection.

INTRODUCTION

Understanding individual motor signatures (IMS) is essential in neurorehabilitation, biomechanics, and motor learning. IMS refers to distinct movement patterns shaped by neural, muscular, and cognitive factors. Accurately capturing IMS offers valuable insights for personalized treatment and performance optimization. However, capturing these unique motor patterns is challenging due to variability in motor behavior across individuals. Statistical analyses often fail to capture subtle differences, making clustering techniques like Fuzzy C-Means (FCM) a potential solution for distinguishing meaningful motor differences. This paper investigates the use of FCM clustering to capture IMS based on kinematic and geometric features from hand-drawing tasks. In particular, the study seeks to identify which features are most effective for distinguishing motor patterns, as well as the conditions under which certain features outperform others. We hypothesize that while some features will excel in clustering by capturing strong individual characteristics, others may offer weaker discriminative power due to their sensitivity to finer, less idiosyncratic aspects of movement [1], [2].

METHOD

Subjects

Nine healthy subjects (1 male, 8 female) aged 18–25 (2 left-handed, 7 right-handed) participated in the study. All subjects were free from conditions affecting cognitive or motor abilities. The Edinburgh Inventory was used to assess handedness. The study was approved by an ethical review board, and informed consent was obtained from all participants.

Experiment Procedure

Subjects were instructed to draw cloverleaf and eight-shaped patterns with their dominant hand in a counterclockwise direction on a paper sheet for 5 trials, each including 20 continuous repetitions. A Wacom tablet recorded the coordinates of each movement (temporal resolution: 200 samples/s; resolution: 1748 by 2551; active area: 210 × 297 mm). In total, each subject performed 100 repetitions.

Feature Extraction

The extracted kinematic and geometric features included: **Length**—the sum of the distance between successive points (cm), **Movement duration** (s), **Maximum velocity (max V)**—peak velocity during the movement (cm/s), **Beta**—the exponent in the relationship between velocity and curvature of the trajectory, **Correlation coefficient**—covariation between velocity and trajectory curvature. Data were preprocessed with a low-pass filter (0.07 Hz cut-off). Successive repetitions were separated using local velocity minima.

Data Analysis and Clustering

The dataset was normalized using min-max normalization. We applied FCM clustering, which allows data points to belong to multiple clusters with varying degrees of membership. This clustering method is beneficial for analyzing motor data that may overlap between distinct motor patterns. The FCM algorithm minimizes an objective function iteratively based on membership values for each data point.

Submit Date: 2025-03-22
Accept Date: 2025-09-28
✉ Corresponding author

To evaluate the clustering performance of the FCM (Fuzzy C-Means) model, we used an entropy-based measure to assess the degree of "fuzziness" in the clustering results, with lower entropy (fuzziness) indicating better clustering quality. Let $U = [u_{ij}]$ represent the membership matrix, where u_{ij} is the degree of membership of data point i in cluster j. In FCM, u_{ij} values lie in the interval $[0, 1]$ and satisfy the constraint:

$$\sum_{j=1}^{c} u_{ij} = 1 \tag{1}$$

for each data point i, where c is the total number of clusters.

The entropy E for the clustering result is calculated as follows:

$$E = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{c} u_{ij}\log(u_{ij}) \tag{2}$$

where:

- N is the total number of data points,
- c is the number of clusters,
- $u_{ij}\log(u_{ij})$ represents the contribution of data point i to the entropy for cluster j.

If entropy values are low, it suggests that data points have high membership values for specific clusters and are assigned to clusters with greater certainty. Conversely, high entropy values imply greater uncertainty, with data points more equally distributed across multiple clusters, indicating a fuzzier clustering outcome. Thus, this entropy-based metric allowed us to quantitatively assess the FCM clustering quality, capturing the degree of overlap in cluster memberships and providing insights into the clarity of the clusters formed.

RESULTS

The effect of the number of clusters (single feature clustering)

Fig. 1.A, presents the entropy values obtained from clustering the hand-drawing data into two and three groups, based on different extracted features. Entropy here is used as a measure of clustering efficiency, where lower entropy indicates more distinct, well-separated clusters.

The results show that the feature "length of shape" yields the lowest entropy, suggesting it is the most effective feature for clustering hand-drawing data. In contrast, beta results in the highest entropy, making it the least useful feature for this task. Therefore, length emerges as the best discriminative feature, while beta performs the worst in distinguishing between different behaviors.

The figure also demonstrates how features affect the number of clusters that can be effectively discriminated. For example, max V and duration achieve better entropy values when clustering into three groups, indicating that these features provide more granularity and are better suited for finer distinctions. On the other hand, length, beta, and the correlation coefficient perform better when clustering into two groups, suggesting that these features are more effective at separating the data into broader categories.

Fig. 1, B to D further illustrate the distribution of clusters for the two extreme features: length (the best-performing feature) and beta (the worst-performing feature). The clusters, colored according to membership values, reveal notable differences in performance between these features.

In the case of length, although it clearly separates the data into more distinct clusters (with greater inter-cluster distance), the cluster sizes are imbalanced, meaning that one cluster contains significantly more members than the others. This suggests that while length is highly discriminative, it may overfit certain behaviors.

On the other hand, beta produces more uniform clusters, with relatively balanced sizes, but the separation between clusters is less distinct, implying poorer discriminative power. Beta's ability to cluster data is limited, which aligns with its higher entropy score.

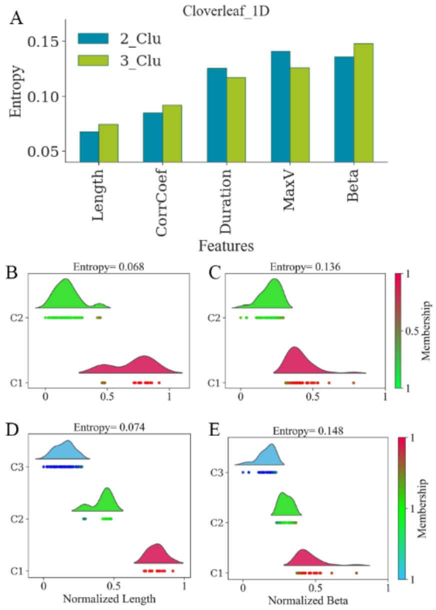

Fig. 1- **Clustering Performance for cloverleaf task.** A) Entropy values for all clustering cases using different features, shown for both 2 and 3 clusters. B) Cluster centers for the length feature in the 2-cluster case. C) Cluster centers for the beta feature in the 2-cluster case. D) Cluster centers for the length feature in the 3-cluster case. E) Cluster centers for the beta feature in the 3-cluster case. The size of each circle reflects the number of members in each cluster, highlighting differences in cluster size.

In Fig. 2, A, the results of the clustering of extracted data from the eight-shape task. Similar to the previous results, length had the best entropy for discriminating data in 3 clusters and beta had the worst value. In the second task, we applied the same clustering analysis to hand-drawing movements involving the eight-shape. Fig. 2, shows the entropy values for clustering into 2 and 3 groups.

As with the cloverleaf task, the length again showed the best clustering performance, yielding the idea that length is a highly effective feature for distinguishing between hand-drawing behaviors, regardless of the specific shape being drawn. In contrast, beta showed the worst clustering performance, resulting in the highest entropy values, indicating that it struggles to effectively differentiate behaviors in both tasks.

Interestingly, in this task, duration performed better when clustering into 3 groups, similar to the findings from the cloverleaf task. This suggests that this feature may capture more subtle differences in drawing dynamics, which are better revealed when the data is split into finer-grained categories. For max V, the entropy decreased in the eight-shape compared to the cloverleaf task. In this task, the temporal features showed better performance.

Fig. 2, B to E illustrate the clustering distributions for length (best feature) and beta (worst feature). As observed in the cloverleaf task, length produces more distinct clusters with greater separation between them, though the clusters are unevenly sized. Beta, on the other hand, forms more balanced clusters in terms of size, but with much less separation between the groups, which indicates poorer overall discrimination.

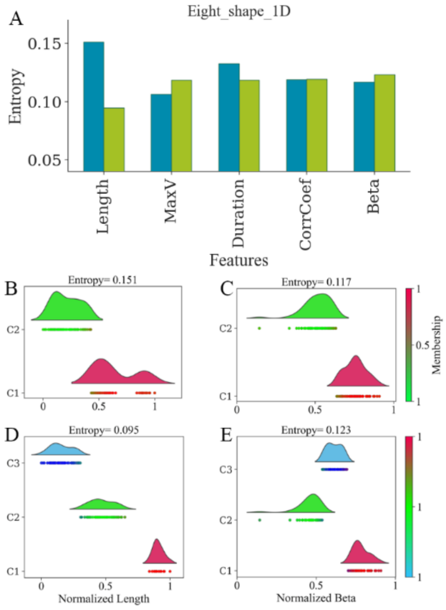

Fig. 2- **Clustering Performance for eight-shape task.** A) Entropy values for all clustering cases using different features, shown for both 2 and 3 clusters. B) Cluster centers for the length feature in the 2-cluster case. C) Cluster centers for the beta feature in the 2-cluster case. D) Cluster centers for the length feature in the 3-cluster case. E) Cluster centers for the beta

feature in the 3-cluster case. The size of each circle reflects the number of members in each cluster, highlighting differences in cluster size.

The effect of the feature dimensions (multiple features clustering)

Fig. 3, shows the mean entropy values obtained from clustering with different combinations of features, ranging from single-feature clustering (1 dimension) to using all five features together (5 dimensions). The results demonstrate a clear trend: as more features are combined for clustering, the performance deteriorates, as indicated by increasing entropy values. This trend is consistent across both tasks (cloverleaf and eight-shape) and suggests that adding more features leads to poorer clustering performance, possibly due to redundant or less discriminative information being introduced.

These findings highlight the importance of careful feature selection in clustering analyses of hand-drawing data, as using more features does not necessarily result in better discrimination. Instead, focusing on key individual features may yield more precise clustering and better identification of motor signatures.

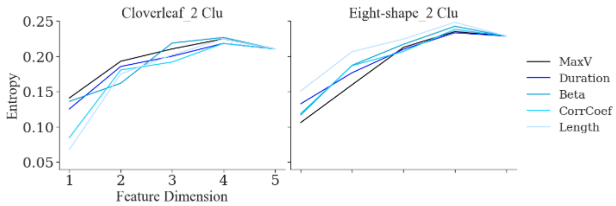

Fig. 3- Entropy values for clustering with different combinations of features (from 1 to 5 dimensions) across both tasks (cloverleaf and eight-shape) for 2-cluster discrimination.

Comparison with other methods

In this section, we compare the performance of FCM with two other clustering methods: **K-Means** [3], [4] and **Hierarchical Clustering (HC)** [5]. K-Means is a hard clustering algorithm that assigns each data point to the nearest cluster centroid. It is computationally efficient and performs well for datasets with spherical clusters. In contrast, HC is an agglomerative or divisive clustering method that constructs a hierarchy of clusters using a bottom-up (agglomerative, in this case) or top-down (divisive) approach. Unlike FCM and K-Means, HC does not require the number of clusters to be specified a priori. However, to ensure a fair comparison, we truncate the HC dendrogram at a specific level to obtain the same number of clusters as used in FCM and K-Means. This allows us to evaluate the performance of all three methods under identical conditions.

Since we used the fuzziness metric to evaluate the performance of FCM, this metric cannot be applied to K-Means and HC. This is because both K-Means and HC are hard clustering methods, assigning each data point to a single cluster, resulting in a fuzziness value of **0**. To enable a fair comparison across all methods, we employ the **Silhouette Index** [6], a metric that quantifies cluster quality by measuring how similar a data point is to its own cluster compared to other clusters. However, since the traditional Silhouette Score is designed for hard clustering and is not directly applicable to FCM, we use a variant called the **Fuzzy Silhouette Index** [7] with $\alpha = 1$. This metric incorporates fuzzy membership degrees while retaining the same interpretability as the traditional Silhouette Score, allowing us to compare FCM, K-Means, and HC on a common scale.

In Table 1, we compare the average cluster quality of FCM, K-Means, and HC using the Silhouette Score (Fuzzy Silhouette Index for FCM). For 3-cluster solutions, FCM achieves the highest average score (0.48), followed by HC (0.47) and K-Means (0.45). While FCM outperforms the other methods in this comparison, the small difference between FCM and HC suggests that both algorithms may be viable choices depending on the application context. Future work should validate these results with statistical significance testing.

Table 1- Average Silhouette Scores (Fuzzy Silhouette Index for FCM) Across All Feature Combinations for 3-Cluster Solutions. Higher values indicate better cluster quality, with a maximum possible score of +1.

	FCM	K-Means	HC
3 Clusters	**0.48**	0.45	0.47

DISCUSSION

The present study employed Fuzzy C-Means (FCM) clustering to analyze motor performance using hand-drawing tasks, aiming to understand how kinematic and geometrical parameters can reveal individual motor patterns and their underlying dynamics. By examining clustering efficiency across two distinct tasks—drawing cloverleaf and eight-shape patterns—we identified specific features that excel at distinguishing between motor behaviors, providing valuable insights into the feature-based clustering of complex motor actions.

Longer lengths may indicate fluid, continuous movements, while shorter lengths might suggest more controlled, segmented actions. The variability in length across individuals likely reflects their unique movement characteristics, such as drawing speed, force, and control stability [8], [9]. This variability makes length a highly idiosyncratic feature, able to reveal distinct motor patterns between individuals. Aligning with our results, research has demonstrated that spatial consistency often serves as a marker of individual motor behaviors [10], [11]. Moreover, length as a measure is minimally influenced by the shape type (cloverleaf or eight-shape), allowing it to act as a robust indicator of underlying motor behavior regardless of task specifics.

Conversely, beta—an exponent in the power-law relationship between velocity and curvature—was less effective in distinguishing motor patterns, likely due to its reliance on more abstract relationships sensitive to noise [12], [13]. Additionally, beta captures subtleties in the coordination between speed and path shape, but these subtleties may not vary significantly across individuals in a way that clearly distinguishes their motor behaviors. As a result, beta yielded clusters with higher intra-cluster ambiguity, illustrating that it is a less reliable indicator of distinct motor profiles in these tasks. These finding highlights that not all kinematic parameters are equally useful for IMS; features like length, which capture direct spatial and motor control aspects, provide clearer clustering outcomes [14].

Temporal features such as duration and maximum velocity also performed well in three-cluster configurations, capturing dynamic aspects of drawing reflective of individual pacing and motor control nuances [15], [16]. These results underscore the importance of dynamic control markers for differentiating motor behaviors, consistent with prior research on pacing and motor coordination [17].

Moreover, combining multiple features generally led to higher entropy, suggesting that redundant or conflicting information introduced noise that weakened clustering performance [18], [19]. This supports the notion that selective feature inclusion is essential in clustering, as excess dimensions can obscure rather than clarify individual motor signatures [20].

CONCLUSION AND IMPLICATIONS

In summary, this study underscores the potential of targeted feature selection in clustering motor behaviors and highlights the differential utility of specific kinematic and temporal features in distinguishing individual motor patterns. Length, with its ability to capture spatial control and idiosyncratic motor traits, emerged as a particularly valuable feature, while beta was found to be less effective due to its sensitivity to nuanced, less variable aspects of motor behavior. Temporal features like duration and maximum velocity offer additional insights, particularly in more granular clustering scenarios where finer distinctions in pacing and control are relevant.

These findings contribute to the methodological approach of using feature-based clustering to capture individual motor signatures. In future research, this framework could be applied to diverse motor tasks or clinical populations, where unique motor signatures may help identify early markers of neurological conditions or tailor rehabilitation protocols. Overall, this study provides a foundation for refining clustering techniques in motor analysis, demonstrating that careful feature selection can enhance the interpretability and effectiveness of clustering in capturing distinct motor behaviors.

REFERENCE

[1] T. Flash and B. Hochner, "Motor primitives in vertebrates and invertebrates," *Curr. Opin. Neurobiol.*, vol. 15, no. 6, pp. 660–666, 2005.

[2] Z. Zarandi, A. N. Stucchi, L. Fadiga, and T. Pozzo, "The effect of the preferred hand on drawing movement," *Sci. Rep.*, vol. 13, no. 1, p. 8264, May 2023. doi: 10.1038/s41598-023-34861-x.

[3] S. Lloyd, "Least squares quantization in PCM," *IEEE transactions on information theory,* vol. 28, no. 2, pp. 129-137, 1982.

[4] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, 1967, vol. 1, no. 14: Oakland, CA, USA, pp. 281-297.

[5] F. Nielsen and F. Nielsen, "Hierarchical clustering," *Introduction to HPC with MPI for Data Science,* pp. 195-211, 2016.

[6] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of computational and applied mathematics,* vol. 20, pp. 53-65, 1987.

[7] R. J. Campello and E. R. Hruschka, "A fuzzy extension of the silhouette width criterion for cluster analysis," *Fuzzy Sets and Systems,* vol. 157, no. 21, pp. 2858-2875, 2006.

[8] S. Mitra, M. A. Riley, and M. T. Turvey, "Variation and variability in motor behavior: A dynamical systems perspective," *Psychol. Rev.*, vol. 124, no. 2, pp. 182–203, 2017.

[9] H. Haken, J. A. S. Kelso, and H. Bunz, "A theoretical model of phase transitions in human hand movements," *Biol. Cybern.*, vol. 51, no. 5, pp. 347–356, 1985.

[10] E. Todorov, "Cosine tuning minimizes motor errors," *Nat. Neurosci.*, vol. 7, no. 4, pp. 384–390, 2004.

[11] M. L. Latash, J. P. Scholz, and G. Schöner, "Toward a new theory of motor synergies," *Motor Control*, vol. 6, no. 1, pp. 8–45, 2002.

[12] K. Schneider and D. Sternad, "Variability, stability, and the role of noise in motor learning," *Exerc. Sport Sci. Rev.*, vol. 42, no. 1, pp. 24–31, 2014.

[13] N. Hogan and D. Sternad, "Dynamic primitives of motor behavior," *Biol. Cybern.*, vol. 108, no. 6, pp. 645–663, 2013.

[14] K. M. Newell, "Constraints on the development of coordination," in *Motor Development in Children: Aspects of Coordination and Control*, M. G. Wade and H. T. A. Whiting, Eds. Springer, 1986.

[15] J. A. S. Kelso, *Dynamic Patterns: The Self-Organization of Brain and Behavior*. MIT Press, 1995.

[16] D. Sternad, "It's not (only) the mean that matters: Variability, noise and exploration in skill learning," *Curr. Opin. Behav. Sci.*, vol. 20, pp. 183–195, 2018.

[17] D. A. Rosenbaum, R. G. Carlson, and R. R. Gilmore, *Cognitive Foundations of Action: A Research Synthesis*. Erlbaum, 2001.

[18] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2009.

[19] B. Schölkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, 2002.

[20] M. L. Latash, *Neurophysiological Basis of Movement*. Human Kinetics, 2008.