Cluster Analysis of Movement Patterns in Multiarticular Actions: A Tutorial

Robert Rein, Chris Button, Keith Davids, and Jeffery Summers

The present paper proposes a technical analysis method for extracting information about movement patterning in studies of motor control, based on a cluster analysis of movement kinematics. In a tutorial fashion, data from three different experiments are presented to exemplify and validate the technical method. When applied to three different basketball-shooting techniques, the method clearly distinguished between the different patterns. When applied to a cyclical wrist supination-pronation task, the cluster analysis provided the same results as an analysis using the conventional discrete relative phase measure. Finally, when analyzing throwing performance constrained by distance to target, the method grouped movement patterns together according to throwing distance. In conclusion, the proposed technical method provides a valuable tool to improve understanding of coordination and control in different movement models, including multiarticular actions.

Keywords: Cluster analysis, coordination, movement pattern, control

A major aim of human movement science is to develop and test robust models of neurobiological processes that underpin movement coordination and control. Developments in computer technology and motion capturing devices now support collection of vast amounts of kinematic data with much greater precision than ever before. For example, Chow, Davids, Button and Koh (2008) collected kinematic data from a full body marker set in every trial while learners practiced a soccer kicking task over 12 weeks. Hence, analysis of increasingly complex movement models involving multiple biomechanical degrees of freedom is becoming more feasible (see Chen, Liu, Mayer-Kress, & Newell, 2005; Hung, Kaminski, Fineman, Monroe, & Gentile, 2008 for similar examples). However, with the availability of enhanced data collection tools there is also an increased demand for sophisticated analysis techniques to process the information contained in these datasets. As proposed in this paper, one candidate technique exists in the application of cluster analysis for investigating multiarticular actions. Although cluster analysis has previously appeared in some studies of motor control, knowledge about its application appears limited to only a small group of researchers. Providing more detailed information...
about its application and execution may support further applications in the study of multiarticular actions. Cluster analysis has been used to analyze movement patterns in a variety of tasks involving walking (O’Byrne, Jenkinson, & O’Brien, 1998; Schöllhorn, Nigg, Stefanshyn, & Liu, 2002; Schöllhorn, Stefanyshyn, Nigg, & Liu, 1999; Toro, Nester, & Farren, 2007), swimming (Wilson & Howard, 1983), long jump (Jaitner, Mendoza, & Schöllhorn, 2001), handball (Schorer, Fath, Joseph, & Jaitner, 2007), golf (Ball & Best, 2007; Lames, 1992), reaching (d’Avella, Portone, Fernandez, & Lacquaniti, 2006), and soccer kicking (Chow et al. 2008). The present article aims to provide a tutorial style overview of cluster analysis of kinematic data thereby presenting a general framework for the analysis of multiarticular actions. To maintain the scope of the topic treatment, the present paper focuses only on one type of cluster analysis techniques, the agglomerative hierarchical cluster analysis, which represents the most common type used in previous research (Blashfield & Aldenderfer, 1978; Gnanadesikan, Kettenring, & Srinivas, 2007). Accordingly, in this paper the term cluster analysis is used to refer to this specific type.

The paper is divided into four main parts. A general introduction of cluster analysis is followed by an overview of past applications in studies of motor control. Subsequently, a discussion of cluster analysis and its relation to alternative analysis techniques like principal component analysis (PCA) and the uncontrolled manifold technique (UCM) is provided. Finally, data from three experiments are briefly presented to highlight some of the benefits associated with the application of cluster analysis to the study of movement coordination data.

An Introduction to Cluster Analysis

Historically, cluster analysis was developed to deal with problems in data mining when investigators needed to identify patterns in high-dimensional datasets. Typically a dataset subjected to cluster analysis consists of a collection of objects of interest measured on several characteristic dimensions. To be able to identify shared patterns between the objects a common metric among the different characteristics needs to be defined. Subsequently, the identified metric might afford an estimation of the degree of (dis)similarity between the objects. Cluster analysis uses similarity information to quantitatively group objects into clusters in an iterative step-wise manner. Therefore, when adopting hierarchical agglomerative methods, initially all objects belong to separate clusters and the cluster algorithm progressively groups objects into clusters of increasing size until eventually all objects are grouped into a single cluster (Blashfield & Aldenderfer, 1978; Everitt, Landau, & Leese, 2001; Kaufmann & Rousseeuw, 1990). The final result of the procedure yields a hierarchical tree (termed a dendrogram). From this dendrogram, it is possible to identify groupings in the dataset. The height of the merger between clusters indicates the grade of similarity or dissimilarity between the objects. The higher the merger, the smaller the similarity between clusters. Thus, well-separated clusters are characterized by a great merging height which can be used to obtain the number of clusters that best characterizes a specific dataset. The unit of the height scale of the dendrogram represents the distances of the common metric as previously calculated. Common metrics of cluster distances, therefore, depend on their units. If for example joint angle data were used as input data, the unit of the scale is angular degree.
One common criticism of cluster analysis methods concerns the fact that a cluster analysis algorithm always groups objects into clusters even when no real underlying structure might exist in the dataset. Together with the, sometimes, arbitrary choice of number of clusters, this fact has contributed somewhat to a view that cluster analysis may be discredited as a serious movement analysis technique (Ball & Best, 2007; Ketchen & Shook, 1996; Toro, et al., 2007). To overcome these perceived limitations of cluster analysis, much research has been devoted to the development of cluster validation techniques, which are exemplified later. Acknowledgment of these perceived limitations of cluster analysis techniques mandates the need to validate results obtained by use of cluster analysis to avoid naïve acceptance of research results.

To summarize, a typical application of cluster analysis involves three main steps: 1) Data preprocessing 2) Cluster analysis 3) Cluster validation and choice of number of clusters (Handl, Knowles, & Kell, 2005; Ketchen & Shook, 1996; Kettenring, 2006). In the following sections of this paper, the three steps are elucidated in more detail.

**Preprocessing**

During data preprocessing the researcher must choose which variables should be included in the cluster analysis similar to the choice of elementary variables required by UCM (Latash, Scholz, & Schöner, 2007). The choice of input variables should be driven by theoretical knowledge since the inclusion of unrelated variables might have a negative impact on the quality of the analysis. The number of different input variables defines the number of input dimensions of the dataset. In principle, input variables can stem from any theoretical domain involving interval, ordinal, nominal, or ratio-scaled scales. In extant research all types have been successfully included in cluster analyses (Everitt, et al., 2001). When mixing different scale levels of analysis, however, it becomes more difficult to define a common distance measure. In studies of motor control, typically interval scaled variables are used which circumvent this problem. If the dataset contains combinations of variables with different units, e.g., position and velocity data, the scales must be made commensurate by a normalization scheme (e.g., range, z-scores, etc). However, normalization always introduces a weighting of the different input variables. For example, using z-scores with joint angular data yields a stronger weighting of joints with small ranges of motion as these variations are up-scaled, whereas variables with greater ranges of motion are down-scaled. Adopting an inappropriate scaling scheme can lead to spurious results. Therefore, naïve application of z-scores, also called “autoscaling”, is inappropriate (Gnanadesikan, et al., 2007; Kettenring, 2006).

The inclusion of all input variables results in a data matrix with \( n \) objects measured on \( M \) variable dimensions for each of \( N \) trials, where the \( M \) input variables describe the specific properties of each individual object. For example, one object could be a particular joint angle and then \( M \) would specify the according number of frames collected. Next, the differences between the objects must be calculated. When using interval-scaled variables, a common dissimilarity measure is the canonical Euclidean distance (Han & Kamber, 2006).

\[
d(i, j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \cdots + (x_{iM} - x_{jM})^2}
\]
Other possible distance measures include the Manhattan (city block), or the general Minkowski distance. The choice of distance measure depends on properties of the variables included in the analysis (Everitt, et al., 2001; Han & Kamber, 2006; Kaufmann & Rousseeuw, 1990). By calculating the distances between all pairs of $N$ trials the resulting set of distances can be arranged into a so-called distance matrix $D$. The entries $(d_{ij})$ thereby represent the distances between the trails $i$ and $j$. $D$ forms the basic input for cluster analysis algorithms. $D$ is symmetric with zeros on the main diagonal, as distances between objects are symmetric $(d_{ij} = d_{ji})$ and the distance $d_{ii}$ between the object and itself is zero.

Cluster Analysis

After calculation of distances, $D$ is submitted to the cluster analysis algorithm. At present, several different hierarchical clustering algorithms are available, unfortunately without any clear-cut criteria to determine optimality of an algorithm. The most widely-used algorithms include single linkage, complete linkage, average linkage, centroid linkage, median linkage, and Ward’s method (Everitt, et al., 2001). These methods vary in terms of how they determine which objects are merged into clusters.

For example, let $Q1$ and $Q2$ be two clusters which are about to be merged. Using the single linkage method the dissimilarity score is calculated according to the minimum of all pairwise distances, or more formally:

$$d(Q1, Q2) = \min_{i \in Q1, j \in Q2} d(i, j)$$

For this reason the single linkage methods is also often referred to as nearest neighbor method. Unfortunately, choosing the minimum distance often leads to a so-called ‘chaining effect’ where poorly separated clusters are grouped together (Kaufmann & Rousseeuw, 1990; Theodoridis & Koutroumbas, 2006). The group average method is based on the average distance between all objects of two clusters:

$$d(Q1, Q2) = \frac{1}{|Q1||Q2|} \sum_{i \in Q1, j \in Q2} d(i, j)$$

$|Q1|$ and $|Q2|$ describe the number of items for each respective cluster. The Ward algorithm, also called the minimum-variance method, is based on another variation and follows from a least-squares argument (see Kaufmann & Rousseeuw, 1990; Ward, 1963 for details).

let $\bar{x}(Q) = (\bar{x}_1(Q), \bar{x}_2(Q), \ldots, \bar{x}_m(Q))$

and $\bar{x}_i(Q) = \frac{1}{|Q|} \sum_{j \in Q} x_{ij}$

$$d(Q1, Q2) = \frac{2|Q1||Q2|}{|Q1| + |Q2|} \|\bar{x}(Q1) - \bar{x}(Q2)\|$$
Of these methods, the group average and the Ward’s algorithm have been shown to be particularly robust, yielding valid results in several simulation studies (Breckenridge, 2000; Everitt, et al., 2001; Hands & Everitt, 1987; Kaufmann & Rousseeuw, 1990; Kuiper & Fisher, 1975; Milligan, 1981; Scheibler & Schneider, 1985). The former method has also demonstrated some robustness against outliers (Rencher, 2001), whereas the Ward algorithm appears to be somewhat biased to equal group sizes, which has to be taken into consideration with regards to characteristics of a specific dataset.

During each clustering step performed by the algorithm, the minimum entry in the distance matrix is determined and according objects are grouped together into one cluster. Subsequently, the distance matrix is updated according to the choice of clustering algorithm and distance measure and the steps are repeated. One limitation concerning hierarchical agglomerative methods is that grouping decisions made by the algorithm during one step cannot be undone in subsequent steps, which can lead to suboptimal clustering results (Han & Kamber, 2006; Kaufmann & Rousseeuw, 1990; Rencher, 2001). This problem further highlights the need to validate cluster analysis results. Finally, all objects are grouped into a single cluster and the cluster algorithm terminates.

Cluster Validation

Having obtained the results from the cluster analysis, the researcher must validate the results and determine the correct number of clusters that capture the relationship between objects of interest. Following the work of Handl et al. (2005), the different validation schemes available can be classified into two main groups: 1) External measures and 2) Internal measures (see also Halkidi, Batistakis, & Vazirgiannis, 2002a; Jain, 1988). External measures can be applied when knowledge about “true” clustering or a so-called gold-standard is available. For example, when performing a Monte Carlo test and sampling from known populations, the true grouping of the data are known a priori and can be used to test the concordance between the obtained clustering and the a priori known clustering. Unfortunately, the known clustering measure is seldom available in experimental data. In contrast, internal measures assess the quality of the clustering solution by using only information contained in the cluster analysis and the distance matrix itself. Once more, several different internal methods are currently available without definitive criteria about technique optimality (Handl, et al., 2005). Internal measures all suffer from the fact that it is not possible to provide an absolute assessment of whether the cluster analysis results provide meaningful groupings. This type of information would require complete knowledge about the “true” nature of clustering and would make the cluster analysis superfluous. However, this limitation is no different for any other analysis technique in inferential statistics. Therefore, common use of cluster validation procedures involves combining different techniques to verify the results. Cluster validation serves two main purposes. It provides an assessment of the degree of “real” clustering in the data and it provides an indication about the appropriate number of clusters underlying the dataset. In the following section we briefly present two validation techniques, which will be discussed subsequently during the exemplar experimental analyses in more detail.
The first validation measurement, the normalized Hubert-\( \Gamma \) statistic tests the compactness of clusters by measuring the compliance between the distance matrix and the grouping (Halkidi, Batistakis, & Vazirgiannis, 2002b; Handl, et al., 2005; Jain, 1988; Zhao, Liang, & Hu, 2006). The Hubert-\( \Gamma \) statistic is defined as:

\[
\hat{\Gamma} = \frac{1}{M} \sum_{i=1}^{N-1} \sum_{j=1}^{N} \left[ D(i, j) - \mu_X \right] \left[ Y(i, j) - \mu_Y \right] \sum_{i=1}^{N-1} \sum_{j=1}^{N} \left[ D(i, j) - \mu_X \right] \left[ Y(i, j) - \mu_Y \right] \]

\( D \) and \( Y \) represent the distance matrix and the group matrix accordingly and \( \mu \) and \( \sigma \) the according mean and variance values. A high relative value for the score indicates that a particular cluster partition fits the data well and should be preferred over other partitions. Typically, the different Hubert-\( \Gamma \) values are plotted against the according number of clusters and the experimenter searches for a significant “knee” (peak) in the plot (Zhao, et al., 2006). This peak determines the optimum number of clusters as indicated by the Hubert-\( \Gamma \) score.

A second measure based on a completely different approach can be found in the approximately unbiased (AU) test which uses a multiscale bootstrapping procedure and tests the stability of the cluster analysis results (Shimodaira, 2002, 2004, 2008; Suzuki & Shimodaira, 2004, 2006a). By forming repeatedly scaled bootstrap samples from the original sample the procedure estimates the stability of the clustering yielding a p-value between 0 and 1 which gives an indication about the probability of true clustering. The p-values can be interpreted as they are in common inferential statistical procedures and provide an indication of confidence in the results with a specific data sample. The necessary number of bootstrapping repetitions can be checked through the standard deviations of the p-values. Typical numbers of repetitions lie between 10,000 and 20,000. As this validation procedure was initially developed to deal with problems in DNA microarray clustering, the procedure copes especially well with datasets containing a great number of input dimensions (Suzuki & Shimodaira, 2004).

Obviously, this discussion of two validation methods is somewhat arbitrary and, as already mentioned, there are a vast number of different validation techniques available. Readers are encouraged to consult the wider literature to extend understanding of different methods (for example Gan, Ma, & Jianhon, 2007; Theodoridis & Koutroumbas, 2006). This small selection should be seen as encouragement for the use of validating procedures rather than a comprehensive overview. The reason for the choice in favor of these specific techniques was based on the fact that they test different properties of cluster results and, therefore, are complementary introducing some robustness against the biases inherent in different validation schemes (Ball & Best, 2007; Handl, et al., 2005). In the following section several applications of cluster analysis in the area of motor control are presented.

### Past Applications of Cluster Analysis to Movement Analysis Data

One research group led by Wolfgang Schöllhorn has commonly adopted cluster analysis approaches for studying various sports movements (Bauer & Schöllhorn,
Cluster Analysis of Movement Patterns in Multiarticular Actions

1997; Jaitner, et al., 2001; Schmidt, Schöllhorn, & Bauer, 1997; Schöllhorn, 1998, 2003; Schöllhorn, et al., 2002; Schöllhorn, et al., 1999). For example, Schorer et al. (2007) investigated movement patterning in five handball players. The authors analyzed shots to different goal locations and under two conditions (trick vs. normal shot). Displacement data of the shoulder, elbow, wrist and the hip of the throwing side were recorded and the data were time normalized and z-score transformed. The authors adopted a two-stage strategy with the first stage applying a single-linkage cluster analysis to identify outliers followed by a cluster analysis using the Ward-algorithm. The results indicated a grouping consisting of three major clusters based on the skill level of the participants. Investigating the subgrouping within the clusters, the authors were able to show that the player with the greatest skill (an Olympic athlete) exhibited the greatest levels of movement pattern variability. In contrast, the novice player used very few movement patterns indicated by a smaller number of movement clusters (Schorer, et al., 2007). Although the data are convincing, the researchers did not validate the cluster analysis results, which might have influenced interpretation of the findings.

Ball and Best (2007) applied cluster analysis to study force plate data recorded during golf swings. They analyzed the movement of the center of pressure among 62 participants at specific time events during the swing using a combination of hierarchical and nonhierarchical cluster analysis methods. To secure the results against cluster algorithm bias, the authors applied a combination of several validation methods. The authors were able to distinguish between two different weight transfer strategies. The first strategy involved continuous movement of the center of pressure in the forward direction during and after club-ball contact, whereas the second strategy was indicated by a reverse of the direction during club-ball contact (Ball & Best, 2007).

Toro, Nester, and Farren (2007) investigated movement patterning of gait with healthy children in comparison to children with cerebral palsy. The authors used a hierarchical cluster analysis based on a furthest-distance measure together with Euclidean distance for establishing the distance matrix. The authors chose the angle data of the knee and hip joint during initial contact, maximum stance extension, and maximum swing flexion as input data. To validate the clusters the authors specified a priori that movement patterns of healthy children should be grouped together into a single cluster. Further, it was expected that the standard deviations of the joint angle data within each clusters should remain stable over a range of different clusters and that movement kinematics should exhibit distinct differences. Applying the prespecified criteria the authors were able to identify a 13-cluster solution. Inspection of the kinematics within each cluster enabled the identification of distinct movement patterns between the clusters (Toro, et al., 2007).

In summary, these applications of cluster analysis to movement patterning data show how high-dimensional datasets can be successfully compressed enabling subsequent identification of different movement strategies. However, all these approaches show some limitations. In the study by Schorer et al. (2007) no validation was used, therefore it is not clear whether the results from the cluster analysis were reliably interpreted. The study by Ball and Best (2007) required detailed knowledge about critical events during the golf swing movement pattern, which is not always available, and might even bias the analysis (Post, Daffertshofer, & Beek, 2000). Finally, the study by Toro et al. (2007) needed a-priori knowledge
of groupings which is often not available. Thus, a more general framework for the application of cluster analysis might be necessary.

A question of interest concerns the relationship between cluster analysis procedures to other emerging analysis techniques which, on the surface, appear to have similar aims. For example, recently, several applications of PCA have appeared in the literature (Balasubramaniam & Turvey, 2004; Chau, 2001; Chen, et al., 2005; Corder, Levin, Li, & Swinnen, 2005; Daffertshofer, Lamoth, Meijer, & Beek, 2004; Post, et al., 2000). From first impressions, one might assume that cluster analysis and PCA perform the same function since the latter also acts as an information-compression method. Typically, PCA is applied in movement studies to extract certain features of the movement and it is assumed that these features provide information about underlying control processes used by the central nervous system (CNS) to coordinate movement (see for example Post, et al., 2000). However, in this regard, cluster analysis is different to PCA since different cluster groupings do not necessarily imply different control strategies, only that different movement patterns are preferred by participants. Alternatively, the mere statement that different numbers of eigenvalues are identified as necessary to control a movement does not necessarily imply the presence of different movement patterns. However, this information is available in the orientation of the eigenvectors which typically does not get reported in studies using PCA methods. Further, interpreting this orientation information leaves one with the problem of how to quantify group differences between different orientation vectors. Thus, the problem of identification similarity between movements is only postponed. With regards to the results obtained from the clusters, the information can subsequently be used to subject the data to further analysis to obtain functional interpretations of clusters. For example, when using joint angles as input data, angle-plots can subsequently be compared between clusters (see Toro, et al., 2007 for an example). A similar argument exists regarding the uncontrolled manifold hypothesis (UCM) and its implicit control hypothesis (de Freitas, Scholz, & Stehman, 2007; Latash, Danion, Scholz, Zatsiorsky, & Schöner, 2003; Latash, et al., 2007; Scholz & Schöner, 1999). Changes in the variability structure of the movement do not imply the presence of different movement patterns per se only that certain elementary variables are controlled.

Taken together, cluster analysis provides different information from both methods discussed here and it analyses different aspects of movement control data. Potentially, a two-stage application involving cluster analysis followed by either PCA or UCM could circumvent some of the key issues raised here. For example, during the application of UCM the experimenter has to linearize the Jacobian (Scholz & Schöner, 1999). This becomes problematic when vastly different movement patterns are present in the dataset as the linearization approximation might not be valid anymore. Using the cluster analysis approach though, the experimenter could separate the trials into different groups and calculate the UCM within each group. Thus, such an approach would provide the best of “both worlds”. For example, Wei, Scholz, Zatsiorsky, and Latash (2008) combined UCM and PCA using a two-stage process with initial calculation of the UCM, followed by projecting force modes onto the UCM and applying PCA to the projected forces. Through analysis of factor loading the authors were able to obtain information about control strategies (Wei, et al., 2008). Thus, by combining different complex approaches, novel information becomes accessible. Further work is needed using
such a combinatorial methodology to elucidate mechanisms of control. In addition, the relationships between cluster analysis and the range of techniques used to study stability and variability in performance of complex actions needs to be clarified.

**General Framework for Cluster Analysis in Studies of Motor Control**

In the following section we will outline a general framework for the application of cluster analysis methods to movement coordination data. Following the procedures popularized by some researchers (Jaitner, et al., 2001; Schöllhorn, 1998; Schorer, et al., 2007; Wilson & Howard, 1983) we propose using joint angle displacement or force time-series data as input variables.

Adopting this approach justifies a simple Euclidean distance-based distance measure which can be seen as a generalization of the normalized root mean square (NoRMS) procedure (Sidaway, Heise, & Schoenfelder-Zohdi, 1995) with variants adopted in numerous studies (Hong & Newell, 2006; Magill, 1989; Mullineaux, Bartlett, & Bennett, 2001). The clustering technique does not include amplitude normalization, which can be justified on the assumption that, because only joint angles are involved, typically similar potential ranges of motion are present. A potential problem can be identified with joints which hold their position constant during a trial. However, joints and their attached segments generate inertial forces which act on the movement system during motion. Therefore, it can be argued that a constant joint position should not be treated differently from a moving joint as a constant joint angle does not necessarily imply a lack of control by the CNS (Schöner & Scholz, 2007). Further, humans have many muscles which span across several joints (e.g., biarticular muscles) and which afford deliberate “control” and coordination by the CNS (Latash, et al., 2003; Latash, et al., 2007; Yang, Scholz, & Latash, 2007). Nevertheless, the issue of normalization of inputs must eventually be determined on the basis of the specific constraints of each movement task. No universally valid approach can be identified and instead it depends on the specific research question (Gnanadesikan, et al., 2007). The cluster analysis per se is neutral against different normalization schemes. However, as normalization affects the distances between items it is possible to obtain different results depending on the normalization scheme (see Experiment 1).

Having obtained the raw representation of the joint displacement time-series, different trials must be time-normalized to obtain equal numbers of time slices for each trial. These time-normalized trials can be arranged in a matrix for each trial. Let $T_1$ and $T_2$ be the matrices representing the angular data of two trials, where rows represent the $M$ different input variables and columns represent $N$ successive time points leading to the following representation.

\[
T_1 = \begin{bmatrix}
a_{111} & a_{112} & \cdots & a_{11n} \\
a_{121} & \ddots & & \\
\vdots & & \ddots & \\
a_{1M1} & \cdots & a_{M1n} \\
\end{bmatrix},
\]

\[
T_2 = \begin{bmatrix}
a_{211} & a_{212} & \cdots & a_{21n} \\
a_{221} & \ddots & & \\
\vdots & & \ddots & \\
a_{2M1} & \cdots & a_{2Mn} \\
\end{bmatrix}
\]
The dissimilarity between these two matrices can now be calculated by subtracting the matrices from each other and calculating the Frobenius norm of the resulting difference matrix which yields the dissimilarity score \( (d_{12}) \).

\[
d_{12} = \| T_1 - T_2 \|_F = \sqrt{ \sum_{i=1, j=1}^{M \times N} |a_{1ij} - a_{2ij}|^2 }
\]

When two trials are performed in exactly the same way the score is zero, and accordingly, the more the kinematics between two trials differ from each other, the higher the \( (d_{ij}) \) value yielded. Calculating all possible permutations of trial pairings, the obtained \( (d_{ij}) \) scores can be arranged in the distance matrix \( D \). Through this approach amplitude and timing information contained in the joint displacement data are considered at the same time. The distance matrix can then be subjected to the cluster algorithm. Based on data from the literature reported earlier in this paper, the average distance method is recommended since it is relatively robust against outliers. To secure the results the previously discussed combination of the Hubert-Γ scale and the approximately unbiased test can be recommended. Taken together, time-normalization, calculation of the Euclidean distance, cluster analysis and cluster validation represent a general framework which can be tailored to various contexts in studies of motor control.

**Special Consideration with Movement Data: Error Influences**

Because the calculated dissimilarity score forms the input for the clustering algorithm, an analysis of possible error sources, especially with regards to time normalization procedures, seems necessary. Although time normalization is commonly applied in studies of motor control, unfortunately the implicit assumptions behind the method are seldom elucidated. Basically, time normalization implies that the CNS specifies at each time slice across trials the same postural state (Cordero, Koopman, & van der Helm, 2006; Scholz & Schöner, 1999). This assumption needs to be investigated for each specific trial as it justifies the comparison of performance across different trials. To highlight the entailing problems which might lead to spurious cluster analysis results, a simple model with only two linear curves will be used.

Let \( f_1 \) be the reference graph with \( f_1 := a^*t_i \ (i = 1, \ldots , M) \), where \( a \) represents a constant slope, \( M \) is the number of time normalized frames and \( f_2 \) is the testing graph. Three different influences will be investigated:

(I.) Different intercepts, (II.) different slopes, and (III.) different offsets between the graphs (see Figure 1).

(I.) Different intercepts: Let \( f_2 := a^*t_i + A \)

\[
D = \sqrt{ \sum_{i=1}^{M} (at_i - at_i - A)^2 } = \sqrt{MA}
\]
(II.) Different slopes: Let \( f_3 := c^* t_i \), with \( c := a + B \).

\[
D = \sqrt{\sum_{i=1}^{M} (a t_i - c t_i)^2} = B \sqrt{\sum_{i=1}^{M} (t_i)^2}
\]

(III.) Different offset: Let \( f_4 := a(t_i + C) \).

\[
D = \sqrt{\sum_{i=1}^{M} (a t_i - a(t_i + C))^2} = \sqrt{M} aC
\]

In the case of different intercepts, the error term consists of two parts. The first square root term affects all trials in the same way since all trials should have the same length after time normalization. Therefore, only the second part is of interest. With regards to movement coordination, the different intercepts represent differences in actual movement execution, from where it follows that the obtained score is actually desirable (compare Figure 1.a). In case (II.) two cases have to be distinguished. In the case of a true difference, the former argument applies. In the case of an erroneous difference when the true slopes are equal, differences in the slope can be introduced either if one of the cutting points is not the same in the two trials or when the trial lengths differ. In this case, the dissimilarity is erroneous and gets multiplied by the squared time points (compare Figure 1.b). Similarly, in case (III.) the difference can be introduced when wrong cutting points are being used. The error term gets inflated, not only by the offset error, but it is also multiplied by the slope and the square root of the number of points which will always be higher than one (compare Figure 1.c). Thus, to obtain meaningful data, both cases (II) and (III) highlight the importance of using trials that cannot differ much in length before time normalization. Further, cutting points must be used which can be reliably identified in all observed trials otherwise spurious differences will be introduced.

---

**Figure 1** — Measurement error influences. a) Different intercepts b) Different slopes c) Different offsets
which will affect all subsequent analysis. Therefore, special consideration is needed when using cutting points based, for example, on acceleration data. For instance, when investigating a tennis movement, participants might use follow through or stop techniques. When using the acceleration of the racquet as cutting points, all trials with a follow through will be much longer. Although in principle, up to the point of the ball impact, the movements can be completely equal because of the later cutting point during follow through, the movement patterns will actually be identified as different. Similarly, the case of a different intercept can also purely result from the type of input data observed. For example, when spatial data are used, participants might shift between trials to different starting positions. Thus, differences between trials might purely result from this fact. Again, whether these differences should be considered part of the experiment or should be accommodated through normalization depends on the nature of the question the experimenter is seeking to address.

**Experimental Case Studies**

In the following three exemplar experiments, the cluster analysis procedure will be presented in a tutorial fashion. All studies will use some a priori knowledge about the movement patterning which enables a better validation of the cluster analysis. The experimental case studies we discuss follow the strategies proposed by Button and colleagues (2006; 2003) adopting a multiple single-participant design.

**Experiment 1**

Four professional basketball players (age = 22 ± 4 yrs, height = 1.91 ± 0.8m) were instructed to score baskets using three different shooting techniques: free throws (4.6m) and three point shots (6.5m), and hook shots from 4m distance which are typical distances for these shooting techniques. Joint centers of participants were identified using passive reflective markers to enable recording of joint trajectory data with a 12-digital-camera system (Motion Analysis Corp., USA) set at a sampling frequency of 100Hz. A mobile basket fixed at a conventional height of 3.05 m with a Plexiglas backboard and a standard sized basketball (Size 7, FIBA approved) was used for all shooting trials. From the marker trajectories the ankle, knee, hip, included shoulder, and elbow angle displacements were derived. Trials were cut from the instance of maximum knee flexion before throwing to the instance of maximum wrist height after ball release. Both points could be reliably identified and are meaningful in terms of the specific task investigated. As only joint angle data were used further error influences like spurious intercept differences could be prevented. As these events can be unambiguously identified in all trials irrespective of the technique used, it can be assumed that all differences between trials represent true differences. All trials were normalized to 50 data points and the number of bootstrapping samples was chosen according to the standard errors of the p-values ranging between 10,000 and 20,000 iterations. Once the range of all p-values ± 2*standard errors fell between 0 and 1, the number of iterations was deemed sufficient.

Based on a priori knowledge a three-cluster solution, representing the three movement patterns, was expected whereby the cluster representing the free throws
and the three point shots should be more similar compared with the cluster of patterns for the hook shots since the former techniques are closely related.

**Results**

Figure 2 shows the time courses of the joint angles of one trial from each technique from Participant 3. Only joint angles of the right arm and the left knee are depicted. Visual inspection of the data suggested that the three-point shot and the free throw appeared more similar regarding movements of the throwing arm, unlike the hook shot which seem to exhibit some greater differences.

The plot for the left knee shows that the starting position for the normalization procedure can be reliably identified for each technique (see Figure 2.c), as could the maximum height of the throwing hand (not shown). Further, the number of frames representing each trial did not considerably differ between the techniques (maximum difference was 30%). Thus, possible error influences, because of erroneous cutting of trials and comparison of trials which differ too much in length, can be ignored.

Figure 3 shows the cluster solutions for the different participants with the three predicted main clusters for all four participants. The p-values obtained from the bootstrapping procedure supported the reliability of the three-cluster solution. The performance of Participants 1–3 corroborated the perception of greater similarity between the free throw and the three-point shot, although some individual variability in the extent of this similarity was present in the data. For Participant 1, the height of the merger between free throws and three-point shots on the right indicates greater separation between these two techniques compared, for example, with Participants 2 and 3 where the merging between the clusters occurred at a much lower dissimilarity value. For Participant 4, the cluster analysis grouped

---

**Figure 2** — Joint angle data from one free throw (---), one hook shot (---, light gray), and one 3 point shot (---, dark gray) from Participant 3. Cutting points for the start of the trial are depicted with arrows.
the three-point shots and the hook shots together. Inspection of the kinematics for Participant 4 revealed a special technique for the hook shot, a so-called “jump-hook” which is much more similar to a jump shot (Martin, 1992), and explains the outcome of the cluster analysis. Variations in dissimilarity values, when comparing the different techniques for the individual participants, was supported by the Hubert-Γ values in Figure 4.

For Participants 1 and 4, the three-cluster solution showed the highest score whereas Participants 2 and 3 demonstrated a peak at two. However, for the latter the decrease to the 3-cluster solution was relatively small indicating a possible divide using three clusters. The difference between the Hubert-Γ and bootstrapping procedure showed that some interpretation in their use is needed. To exemplify the influence of data normalization, the same analysis was performed with normalized angle data. Angle data were centered at zero and range-normalized yielding standard deviations equal to 1 (z-scores). For Participants 1, 3, and 4 the results were

**Figure 3** — Individual dendrograms for each participant. ▲: free throw; ●: 3 point throw, ■: hook shot.
equal to those obtained with raw angle data. However, for Participant 2 a different clustering was obtained (compare Figure 5).

The dendrogram shows some distinct differences from that shown in Figure 3. Both Hubert-Γ scores and the bootstrapping procedure favored a four-cluster solution. Applying this clustering, free throws and three-point shots are grouped into a single cluster together with two hook-shot trials. The remaining hook shots are separated into two separate clusters and a single trial. When investigating angle-angle plots of the right arm it becomes apparent that the normalization procedure decreased the differences between the right arm kinematics compared with those observations in nonnormalized trials (see Figure 6). Therefore, the cluster analysis picked up differences from other segments which were not as influential when using nonnormalized trials.

This observation again highlights the importance of proper data preparation when analyzing results using cluster analysis. Investigators must be sure that the data are adequate for the intended analysis. For example if important differences between trials stem from small variations in segments kinematics which might be masked by greater variations from less important segments, normalization procedures might be appropriate.

In summary, the cluster analysis applied to three different shooting techniques performed by skilled performers led to entirely feasible results. The cluster analysis was able to reproduce a priori known differences between the diverse movement patterns. Furthermore, the technique was also able to identify some interesting

**Figure 4:** Hubert-Γ for each individual participant. Participant 1: □, Participant 2: ●, Participant 3: ▲, Participant 4: ♦.
Figure 5 — Dendrogram from Participant 2 based on normalized angle data. ▲: free throw, ●: 3 point throw, ■: hook shot.

individual differences in relation to the amount of dissimilarity between the three basketball shooting techniques. Additional analysis of normalized data highlighted the importance of appropriate data-preparation procedures.

Experimental Case Study 2

Next, a wrist supination-pronation task was investigated in a single participant design which highlights how cluster analysis can be used with movement models typically associated with different analysis techniques (i.e., discrete relative phasing). A participant was seated in front of two manipulanda and instructed to synchronize an antiphase movement of the wrists with a metronome. Movement frequency was scaled from 1 Hz to 2.5 Hz in steps of 0.25 Hz and each step lasted for 8 s. As is the convention in many studies of phase relations in cyclical movements, the participant was instructed that, if the movement patterning were to become unstable, he should adopt the most comfortable pattern (Kelso & Schöner, 1988). The recording frequency of accelerometers attached to the manipulanda was set at 500Hz and 12 trials in total were collected. Single movement trials were determined according to the occurrences of maximum pronation of the right hand and time normalized to 100 data points. This procedure ensured that no spurious cutting
Figure 6 — Mean angle-angle plots of the right shoulder and the elbow joint angle from normalized and non-normalized trials for each cluster.

Shaded hulls indicate ± 1 standard deviation.
points were used since the position of the maximum excursion could be reliably identified as a meaningful event for data analysis and unwanted error influences on the movement patterning data were prevented. Discrete relative phase (DRP) was calculated according to Kelso (1995). For the bootstrapping procedure 10,000 iterations were performed and the standard errors of the p-values were investigated.

In the following section only a single representative trial will be reported as the results for the individual scaling trials were all very similar. However, the trials varied in their apparent transitions frequencies, thus averaging across trials was not deemed appropriate.

Results

Figure 7 shows the mean DRP for each movement frequency bin from all trials. As expected from previous research, the antiphase pattern was stable for the lower movement frequencies and a shift to an in-phase pattern was apparent at a critical frequency of 2.0 Hz. At 2Hz one more antiphase pattern was maintained, followed by a drop in DRP and subsequent transition to an in-phase pattern.

Figure 8.a shows the dendrogram of the cluster analysis clearly indicating a three-cluster solution. The three clusters are well separated as indicated by the high merger compared with the much smaller values within the two clusters. The p-values obtained from the bootstrapping procedure supported this interpretation (all clusters $p < .05$) providing evidence for a true clustering. Figure 8.b shows the resultant Hubert-$\Gamma$ statistic which also supported a three-cluster solution. Figure 6.a indicates that the cluster solution consists of two main clusters with an additional single trial which is in complete agreement with the analysis using the DRP measure.

Figure 9 shows the cluster distribution as a function of frequency with the associated mean DRP superimposed. The plot provides clear evidence for a relation

![Discrete Relative Phase](image)

**Figure 7** — Discrete relative phase data for each individual trial
between the clustering and the movement frequency. For the lower frequencies the participant used Cluster 1, which represented the antiphase movement. At 2Hz, a switch from Cluster 1 to Cluster 3 took place. Two single occurrences of Cluster 1 and 2 were present, which represented the first two trials at 2.0hz visible in Figure 7. In the current study, the occurrence of a cluster containing only a single trial marks transient behavior. Thus, cluster analysis is able to identify singular events varying from main clusters which can be interpreted as transient or outlier behavior depending on the specific context.

In summary, the results of the clustering analysis closely resembled the results obtained from the discrete relative phase analysis, which represents the conventional way to analyze data from this movement model. No investigation into the variation of the transition frequency between trials has been reported in this paper because it is a tutorial about the application of cluster analysis and not an in-depth analysis about variability in phase transition behavior. Nevertheless, the cluster analysis approach proved valuable and the data suggested that it may be used as a substitute for other analysis approaches.

**Experimental Case Study 3**

The third validation study also used a scaling methodology as a strategy to analyze a multiarticular movement: this time a basketball throwing task. Two professional basketball players (age = 20 and 21 yrs, height = 1.93 and 2.00m) were instructed to score baskets using a hook shot technique from distances between 2m to 9m from the basket. Both participants typically played as forward and were experienced at the hook shot technique. The changing distance imposed differential strength

![Figure 8](image_url) — a) Dendrogram of the cluster solution for a single trial. Cluster numbering as indicated at the bottom b) Hubert-Γ statistic.
requirements on participants and was assumed to induce changes in movement patterning (Southard, 1998, 2002). The study was performed in a gymnasium selected for its ceiling height of only 4.5m. This physical location characteristic imposed a spatial structural constraint on the shooting technique because the specific ball trajectory arcs normally observed during hook shots from further distances could not be performed by the participants. This set-up was chosen, after initial pilot work, to facilitate a strong division between techniques used for close- and far-range shots. A two-cluster solution was expected to result from this particular experimental manipulation providing some external framework for assessing the sensitivity of the cluster analysis. During shots from closer distances participants could perform their typical techniques since the ceiling did not limit selected throwing arcs. At greater distances it was believed that the structural constraint would lead to an adaptation in movement patterning, which would be identified by the cluster analysis.

Participants were prepared in the same way as in the Experimental Case Study 1 and the same recording set-up was used. After a 15-min warm-up period each participant performed eight trials from each distance starting at 2m and ending at 9m in increments of one meter. Afterward the same procedure was applied backward starting from 9 m. Data analysis was performed as in Experimental Case Study 1 which insured minimal error influences.

Results

Figure 10.a shows the cluster solution obtained from the hook shots performed by Participant 1. The dendrogram presents a clear two-cluster solution with high dissimilarities between the two clusters indicated by the great height of the tree.
Figure 10 — Dendrograms for Participant 1 and Participant 2. Cluster indices are indicated at bottom of the plot.
branches. This visual impression was also supported by the bootstrapping procedure where both p-values for the according cluster were \( p < .05 \). The Hubert-\( \Gamma \) scores also supported the two-cluster solution attaining a maximum value for two clusters (compare Figure 11).

Plotting clusters against throwing distances (see Figure 12), it becomes apparent that the cluster 2 contained trials from further distances and the cluster 1 captured trials from closer distances.

This observation indicates that Participant 1 used different movement patterns according to throwing distance. When increasing throwing distance, Participant 1 used movement patterns belonging to cluster 1 until 6m, where a switch to cluster 2 took place with a transition to a different functional movement pattern (see Figure 12). This different movement pattern was retained until the final distance. However, when starting to shoot from 9m, Participant 1 used movement patterns belonging to cluster 2 until the distance dropped to 4m, and only switched to movement pattern one at 3 m. Thus, the cluster analysis was able to identify that the participant used two distinct movement patterns according to the throwing distance as hypothesized a priori.

Figure 10.b shows the cluster solution for the hook shots performed by Participant 2. Except for weak separation into leftmost and rightmost clusters, the dendrogram shows no clear separation between the different clusters which was also indicated by a smaller height of the dendrogram compared with Participant 1. The solution appeared to be a result of a “chaining effect” where subsequent clusters are rather merged into a single ever-growing cluster, opposed to differ-

![Figure 11](image_url) — Hubert-\( \Gamma \) values for Participant 1 (■) and Participant (●).
ent well-separated clusters being later merged into a single cluster. These results indicated that differences between the trials followed more of a gradual adaptation of a single movement pattern rather than a transition between completely different patterns observed in Participant 1. This finding also impacts on the stability of the cluster results, mirrored in the data obtained from the bootstrapping procedure which exhibited generally lower values. Using a higher p-value of $p < .2$, the resulting clustering consists of 6 clusters as highlighted by the rectangles in the dendrogram (compare Figure 10.b). The use of a higher p-value can be justified by the exploratory structure of the current study and the fact that $p < .2$ was the lowest value which yielded interpretable results. The Hubert-$\Gamma$ scores indicated possible solutions for 5–8 clusters, therefore supporting the findings from the bootstrapping procedure. The absolute values of the Hubert-$\Gamma$ scores were much lower compared with those obtained in Participant 1 which is indicative of a weaker clustering as well (compare Figure 11). Accordingly, a clustering scheme with 6 clusters was used for further analysis.

Plotting the cluster solution against distance to targets in Figure 13 had a differential effect for the two conditions apparent. During the increasing distance condition, Participant 2 mainly used movement patterns belonging to three different clusters: clusters 1, 6, and 4. Cluster 1 was used for lower distances (2m-4m), cluster 6 for middle distances (5m-7m), and cluster 4 at greater throwing distances (8m-9m). Comparing the clustering in the dendrogram, clusters 6 and 4 are close together compared with cluster 1. During the decreasing condition a slightly different
patterning was apparent. Here, Participant 2 used in general the same movement clusters at similar distances. However during throws from middle distances, the patterning was much less stable and the participant seemed to switch back and forth between clusters 6 and 4, although with decreasing frequency as throwing distance decreased. Interestingly, at the smallest throwing distances Participant 2 used a movement pattern different from those used during the increasing condition. Inspecting the dendrogram again, it indicates that cluster 1 and cluster 3 are more similar compared with cluster 3 and cluster 4 and cluster 6. Thus, it appears that Participant 2 used a different strategy from Participant 1 and rather adapted a single movement pattern according to the distance constraints which was counter to the initial hypothesis. This strategy can also be observed when plotting angle-angle plots for the right arm for both participants as (see Figure 14).

For Participant 1, the two movement clusters exhibit clearly distinct patterns in the right arm with no intermediate patterns, whereas for Participant 2 movements of the right arm are much more similar and are only rescaled and slightly shifted versions of each other. This approach exemplifies also how to proceed after application of cluster analysis to obtain functional interpretations of movement clusters and further interpret the data. Nevertheless, the cluster analysis proved sensitive enough to unravel these interindividual differences in movement control strategies.

In conclusion, the results of the scaling experiment in a constrained physical environment justified the application of the cluster analysis. Particularly for Participant 1 the results showed strong indications of usage of two distinct patterns. For Participant 2 the movement patterning showed distinctively fewer differences

![Figure 13 — Cluster distribution for Participant 2. A small amount of noise has been added to improve visibility of individual points.](image-url)
as distance varied. Possibly, the constrained environment influenced the movement patterning in a different way for this participant as expected. The bootstrapping procedure and the Hubert-$\Gamma$ values showed that, when interpreting results of a cluster analysis, a validation approach is compulsory, without which outcomes could be open to naïve interpretation.

**Summary and Outlook**

In the present paper, a general framework for the analysis of movement coordination data using cluster analysis was introduced. The paper explained the basic steps necessary for conducting a cluster analysis and showed some of the important decisions necessary for a successful application of cluster analysis in studies of motor control. Using three different case studies it was shown how the cluster analysis framework can be adapted to examine movement patterning in specific tasks and how to extract important information from movement data. Through the adoption of a cluster analysis approach it is possible to identify and quantify patterns contained in complex datasets. Thereby, few a priori assumptions about the structure of the data set are required.

With continuing advances in movement measurement technologies, evermore complex datasets will be obtainable and cluster analysis provides one possible way to identify the existence of patterns within these large datasets. The approach seems especially appropriate to be used together with other analysis tools like PCA or UCM to obtain more detailed information about a dataset under scrutiny. For example, the results from Experiment 1 could be further analyzed using separate
UCMs for each cluster. Subsequently, adopting the approach introduced by Wei et al. (2008), PCA of joint angles applied to each UCM can be analyzed yielding factor loadings for each cluster. Fluctuations of factors loading than provide information about different control strategies observed across and within clusters. Adopting such an approach, the different analysis tools can be seen as part of a whole tool-chain with each tool responsible for different aspects of the data.

We hope that we succeeded in showing that the cluster analysis framework is general and can be applied to various task and contexts. Potentially cluster analysis creates a new, exciting approach to the analysis of degenerate, multiarticular movement models with no a-priori information about the underlying structure of such actions. The overarching aim of this tutorial paper was to help researchers in the area of motor control to incorporate this method into their canon of analysis techniques yielding new insights on behavior in neurobiological systems.

Acknowledgment

The authors would like to thank Mike Garry, PhD, School of Psychology, University of Tasmania, for collecting the bimanual data for Experimental Case Study 2.

References


