OPEN

A data-driven approach to decompose motion data into task-relevant and task-irrelevant components in categorical outcome

Daisuke Furuki & Ken Takiyama*

Decomposition of motion data into task-relevant and task-irrelevant components is an effective way to clarify the diverse features involved in motor control and learning. Several previous methods have succeeded in this type of decomposition while focusing on the clear relation of motion to both a specific goal and a continuous outcome, such as a 10 mm deviation from a target or 1 m/s hand velocity. In daily life, it is vital to quantify not only continuous but also categorical outcomes. For example, in baseball, batters must judge whether the opposing pitcher will throw a fastball or a breaking ball; tennis players must decide whether an opposing player will serve out wide or down the middle. However, few methods have focused on quantifying categorical outcome; thus, how to decompose motion data into task-relevant and task-irrelevant components when the outcome is categorical rather than continuous remains unclear. Here, we propose a data-driven method to decompose motion data into task-relevant and task-irrelevant components when the outcome takes categorical values. We applied our method to experimental data where subjects were required to throw fastballs or breaking balls with a similar form. Our data-driven approach can be applied to the unclear relation between motion and outcome, and the relation can be estimated in a data-driven manner. Furthermore, our method can successfully evaluate how the task-relevant components are modulated depending on the task requirements.

Diverse features of motor control have been reported. One striking feature is that motion data can be decomposed into at least two components: task-relevant and task-irrelevant components. Variability in task-relevant components is smaller than in task-irrelevant components[1]. Quantifying task-relevant and task-irrelevant components can thus reveal the salient features of motor control.

Several methods can quantify task-relevant and task-irrelevant motion components. The uncontrolled manifold (UCM) method quantifies the components by focusing on forward kinematics, such as hand position in sit-to-stand[2]. While the UCM focuses on kinematic outcome (e.g., the position of hand or center of mass), the goal equivalent manifold (GEM)[3] and noise-tolerant-covariance (TNC)[4] methods quantify the task-relevant and task-irrelevant motion components by explicitly defining the relations between kinematic parameters and task outcome (e.g., in ball throwing, the height of the ball flight can be approximated by parabolic motion). In addition to the GEM and the TNC, our recent methods detect the task-relevant and task-irrelevant motion components after estimating the unknown relation between time-varying motion and task outcome in a data-driven manner[5,6].

Although these methods succeeded at quantifying task-relevant and task-irrelevant components, they all focused on continuous task outcomes, such as the error between desired and actual outcome[3,4,6] or hand position[2]. In contrast, task-relevant and task-irrelevant motion components are more difficult to quantify when the outcomes are categorical, such as throwing a fastball or breaking ball, whether the action results in success or failure, serving wide or down the line in tennis, whether the subject is a healthy control or a patient, and whether the player is an amateur or professional. Despite the fact that investigating the influences of these categorical outcomes can be central topics in biomechanics, motion science or related research areas, how to decompose motion data into task-relevant and task-irrelevant components remains unclear.

Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology, Koganei-shi, Tokyo, 184-8588, Japan. *email: t.j.ken.takiyama@gmail.com
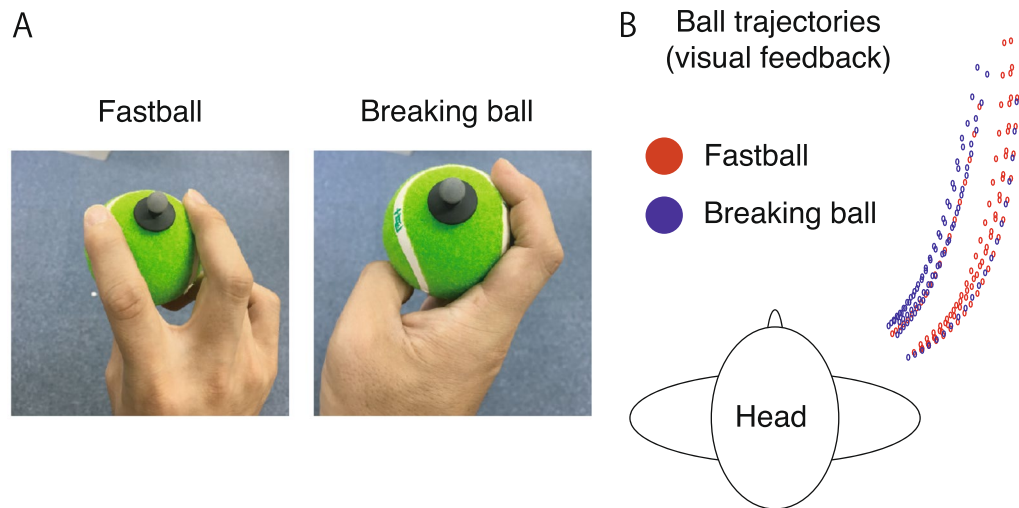
Figure 1. Protocol of our experiments. (**A**) The instructions for subjects described how to grasp the ball when throwing a fastball or breaking ball. (**B**) On the second day, the subjects were instructed to throw the fastballs and breaking balls using a similar form. The subjects could confirm the trajectory of the marker attached to the ball only on the second day.

In this paper, we propose a data-driven technique to detect task-relevant and task-irrelevant motion components with categorical task outcomes. Our method relies on a linear regression technique for solving classification problems, such as logistic regression[7]. For example, logistic regression enables classification of whether the current motion data are associated with throwing a fastball or breaking ball. Our data-driven method can be applied even when the relation between motion data and outcome is unknown: the relation can still be estimated in a data-driven manner. Along with recent data-driven approaches in biomechanics that focused on unsupervised methods[8,9], we rely on supervised methods to address the task-relevant and task-irrelevant components with categorical outcomes in a data-driven manner. Notably, our current mathematical framework is the same as that used in our previous methods[5,6], which focused on continuous task outcomes by applying a linear regression technique. Along with our previous methods, we propose a unified data-driven approach to detect task-relevant and task-irrelevant motion components for multiple kinds of task outcomes.

Methods

Participants. Eight healthy volunteers (aged 18-21 years, four females) participated in our experiment for two days (not consecutively). All the participants were informed of the experimental procedures and their conformance with the Declaration of Helsinki, and all participants provided written informed consent before the start of the experiments. All the procedures were approved by the Ethics Committee of the Tokyo University of Agriculture and Technology.

Experimental protocol. The subjects were seated on a chair and instructed to throw a ball towards a target 2.8 meters from the chair. We used one target location throughout this study to focus on whether the subjects threw either a fastball or breaking ball. On the first and the second days, the subjects were instructed to throw either a fastball or breaking ball pseudorandomly (i.e., they threw each once in every two trials) towards the target following their preferred motion form. Because all the participants were naive to baseball pitching, how they grasped the ball was the only difference between throwing the two types of pitches (Fig. 1A). On the second day, the subjects were additionally instructed to throw a fastball and a breaking ball using a similar motion. The purpose of the current study is to investigate the influence of task instruction on task-relevant and task-irrelevant motion components. To facilitate similar forms, we provided feedback about the motion trajectories by displaying the ball trajectories on a monitor only on the second day (Fig. 1B, 20 frames before the release).

At the beginning of each trial, three beeps sounded at a one-second interval. The subjects were instructed to throw either a fastball or breaking ball at the timing of the third beep. On each day, subjects performed 100 main trials (50 for fastball and 50 for breaking ball) after 20 practice trials (10 for fastball and 10 for breaking ball). We analyzed the 100 main trials.

Data acquisition and processing. Subjects' motions were recorded at 120 Hz using nine cameras (Optitrack Flex 13, NaturalPoint Inc., Corvallis, Oregon). Markers were attached to the ball, neck, right clavicle, right shoulder, right elbow, right wrist, right index finger, and right ilium of the participants. The marker position data were filtered in MATLAB 2016a (Mathworks, Inc) using a 12th-order, 10 Hz, and zero-phase Butterworth filter. The joint angles of the shoulder (three degrees of freedom), elbow (one degree of freedom), and wrist (three degrees of freedom) were analyzed. The release timing was determined based on the moment at which the distance between the markers attached to the ball and index finger exceeded 0.01 meter. We did not analyze the trajectory of the marker attached to the ball except to detect the release time.
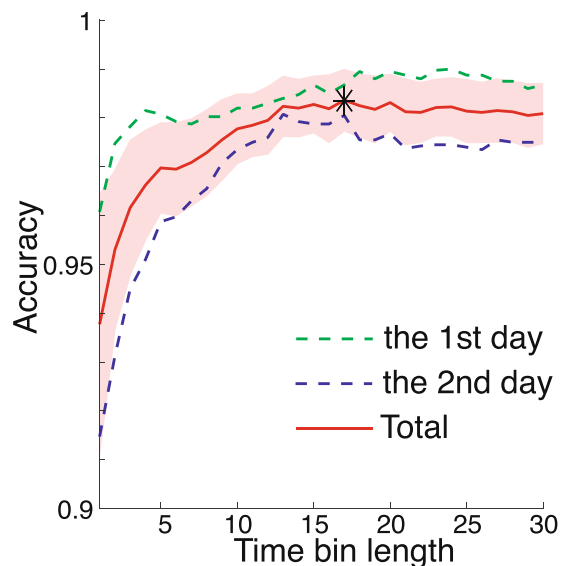
Figure 2. Classification outcome using logistic regression. The horizontal axis denotes the number of time frames to define time-varying motion data X, and the vertical axis denotes the classification accuracy in cross-validation. The green and blue dotted lines denote the classification accuracy averaged across all subjects using the motion data on the first and second days, respectively. The red solid line and the shaded area indicate the mean and standard error of the classification accuracy for all subjects and two days, respectively. We used joint angles with 17 time frames to determine the motion sequence data X for the analysis throughout this paper because the mean classification accuracy was the best for the logistic regression with 17 time frames.

Logistic regression. We applied logistic regression to classify whether the subjects threw a fastball or breaking ball based on time-varying motion data $X \in \mathbf{R}^{T \times D}$, where T and D denoted the number of trials to be analyzed and the dimensions of the time-varying motion data. In the current study, X corresponded to the vectorized temporal sequences of the joint angles[6]. We defined the target value $d \in \mathbf{R}^{T \times 1}$ as $d_t = 0$ when the subjects threw the fastball; otherwise, $d_t = 1$, $(t = 1, \ldots, T)$. In the logistic regression, appropriate weight values $w \in \mathbf{R}^{D \times 1}$ were estimated to predict the target value based on $p(d_t = 1) = f(w_0 + X_t w)$, where $p(d_t = 1)$ denotes the estimated probability that $d_t = 1$ (i.e., $p(d_t = 0) = 1 - p(d_t = 1)$), $f(z) = \frac{1}{1 + \exp(-z)}$ denotes a sigmoid function, and w_0 is a bias term to be estimated. The motion sequence data X were normalized so that the mean and standard deviation of each component across the trials were 0 and 1, respectively, following a standard cross-validation procedure[10] that did not affect the results[11].

Appropriate w_0 and w values were estimated to minimize the following cost function E:

$$E(w_0, w) = -\frac{1}{T}\sum\nolimits_{t=1}^{T}(d_t \log f(w_0 + X_t w) + (1 - d_t)\log(1 - f(w_0 + X_t w))) + \frac{\lambda}{2}w^T w, \qquad (1)$$

where w^T indicates the transposition of w. The first and second terms on the right-hand side indicate that the cross entropy equaled 0 when logistic regression provided a perfect prediction and log2 when the prediction failed (i.e., f = 0.5 when d = 0 and d = 1). The third term is a regularization term using the parameter λ. An appropriate λ value was determined based on 10-fold cross-validation with five iterations. We used glmnet[11] throughout this study.

Because the relation between the time-varying motion data and task outcome is unclear in general, we confirmed whether logistic regression was able to discriminate between the motions used to throw a fastball or breaking ball. A correct classification means that the logistic regression for the motion to throw a fastball results in $p(d_t = 0) > 0.5$ (i.e., $p(d_t = 1) < 0.5$), while the motion used to throw a breaking ball results in $p(d_t = 1) > 0.5$ (i.e., $p(d_t = 0) < 0.5$). Fig. 2 shows the percentage of accurate classifications (i.e., accuracy). Because the best classification accuracy was achieved when we used 17 time frames before the release to determine X: 98.37 $\pm$ 0.0063 % (mean $\pm$ standard error of mean [s.e.m.]), we used these 17 time frames in the following analysis. Another measure to evaluate the classification accuracy is area under the curve, or AUC. The AUC was 0.9983 $\pm$ 0.0035, indicating high classification accuracy. We focused on the time series of the seven joint angles; thus, we analyzed 119 data points for each trial.

Statistical analysis. Repeated measures analysis of variance (ANOVA) were conducted when there were no assumptions about the statistical tests, followed by Tukey's post hoc comparisons. The current study considered two within-subject factors, including 'Part' (former or latter part of the same day) and 'Day' (day 1 or 2). All the statistical analyses were performed using MATLAB 2016a.

Results

Eight subjects threw fastballs and breaking balls towards a fixed target in 100 trials (50 trials for each pitch type) while seated on two days. On the first day, the subjects were instructed to throw the ball towards the target. On the second day, the subjects were additionally instructed to throw the fastball or the breaking ball with a similar motion. We analyzed seven joint angles (three shoulder joint angles, one elbow joint angle, and three wrist angles) to investigate how the task-relevant components were modulated depending on the task requirement. Throughout this study, we relied on logistic regression, although the method to extract task-relevant and task-irrelevant motion components was invariant across linear classification methods as follows.

Decomposition into task-relevant and task-irrelevant differences in categorical outcome.

Using linear classification methods such as logistic regression, we can obtain weight values w to determine the probability that the analyzed data will be classified into the first or second category (i.e., $p(d_t = 1) = f(w_0 + X_t w)$, where d_t is the target value and $d_t = 1$ indicates that the category associated with motion data at the tth trial, X_t, is the breaking ball, f is a sigmoid function, and w_0 is the bias to be estimated). Based on these results, we can discuss how each joint angle at each time point contributes to the categorical task outcome (i.e., fastball or breaking ball in the current study)[5]. The purpose of this study was to focus on task-relevant and task-irrelevant components inherent in time-varying motion data when the task outcome was categorical. We thus derived task-relevant and task-irrelevant components within the framework of linear classification methods, including logistic regression, as follows.

After estimating w, we can obtain the task-relevant components by minimizing the following cost function

$$E_{\text{rel}}(X_{\text{rel}}) = \frac{1}{2}(Xw - X_{\text{rel}}w)^T(Xw - X_{\text{rel}}w) \tag{2}$$

for all trials because the task-relevant components provide the same classification probability as $p(d_t = 1) = f(w_0 + X_t w) = f(w_0 + X_{\text{rel},t} w)$. In other words, the task-relevant components provide all the task-relevant information included in the motion data. By solving $\frac{\partial E_{\text{rel}}(X_{\text{rel}})}{\partial X_{\text{rel}}} = 0$ while avoiding the self-evident answer (i.e., $X_{\text{rel}} \neq X$), we obtain

$$X_{\text{rel}} = X\frac{ww^T}{|w|^2}, \tag{3}$$

where $|w|^2$ denotes the L2-norm of $|w|$. The derived $X_{\text{rel}} = X\frac{ww^T}{|w|^2}$ equaled X projected on the vector w because $X_{\text{rel}}w = Xw$ and $X_{\text{rel}}w^\perp = 0$ ($w^\perp$ is a vector orthogonal to w).

After obtaining X_{rel}, we can calculate the task-irrelevant components as follows:

$$X_{\text{irr}} = X - X_{\text{rel}}, \tag{4}$$

because $X_{\text{irr}}w = Xw - X_{\text{rel}}w = 0$. In other words, the task-irrelevant components do not provide any task-relevant information included in the motion data.

Notably, the above-mentioned derivations and interpretations of task-relevant and task-irrelevant components are the same as those used in our previous method, which focused on these components in continuous task outcomes[6]. Our data-driven technique thus forms a unified framework for extracting task-relevant and task-irrelevant motion components with either categorical or continuous task outcomes.

To investigate the properties of X_{rel} and X_{irr}, we defined $X_{\text{rel},0} \in \mathbf{R}^{T_0 \times D}$ as the task-relevant components in the trials estimated as throwing a fastball, where T_0 indicates the number of these trials. Similarly, we defined $X_{\text{rel},1} \in \mathbf{R}^{T_1 \times D}$, $X_{\text{irr},0} \in \mathbf{R}^{T_0 \times D}$, and $X_{\text{irr},1} \in \mathbf{R}^{T_1 \times D}$ (see the Supplementary Material for the details of the task-irrelevant components). Our method enabled an analysis of how the statistics of the components were modulated when subjects threw a fastball or breaking ball with or without the task requirement to use a similar throwing motion.

Modulation of task-relevant components.

Because the subjects were instructed to throw fastballs and breaking balls with a similar motion on the second day, we needed to evaluate the similarity between the throwing motions for fastballs and breaking balls. A possible and common measure of the similarity was the sensitivity index, or d', previously proposed in the framework of signal detection theory. d' indicates the classification accuracy of two signals generated from different probability distributions. When μ_1, μ_2, and σ are the mean of the first distribution, that of the second distribution, and the common standard deviation of both distributions, d' $= \frac{\mu_1 - \mu_2}{\sigma}$. A smaller d' indicates a more difficult classification; thus, d' should be smaller when throwing a fastball and breaking ball using similar motions. There are two ways for d' to become small: a small $\mu_1 - \mu_2$ or a large σ.

Two possibilities thus exist for the modulation of the task-relevant components depending on the task requirements. One possibility is modulation in the means of $X_{\text{rel},0}$ and $X_{\text{rel},1}$ (Fig. 3A,B,D,E). When subjects threw the fastballs and breaking balls in the trials where they had been instructed to use similar motion, the difference in the means of $X_{\text{rel},0}$ and $X_{\text{rel},1}$ should be reduced compared to the trials without the similar motion task requirement. We refer to this difference as the task-relevant difference. In addition, we refer to the possibility of the reduction of the task-relevant difference as mean modulation in the following. The second possibility is modulation in the variance of $X_{\text{rel},0}$ and $X_{\text{rel},1}$ (Fig. 3A,C–F); these variances could be increased with the task requirement, and this modulation enabled similar forms among conditions. We refer to the variance of $X_{\text{rel},0}$ and $X_{\text{rel},1}$ as task-relevant variability. In addition, we refer to the possibility of the increase of the task-relevant variability as variability modulation. We investigated these possibilities by conducting an experiment in which subjects threw fastballs or breaking balls.
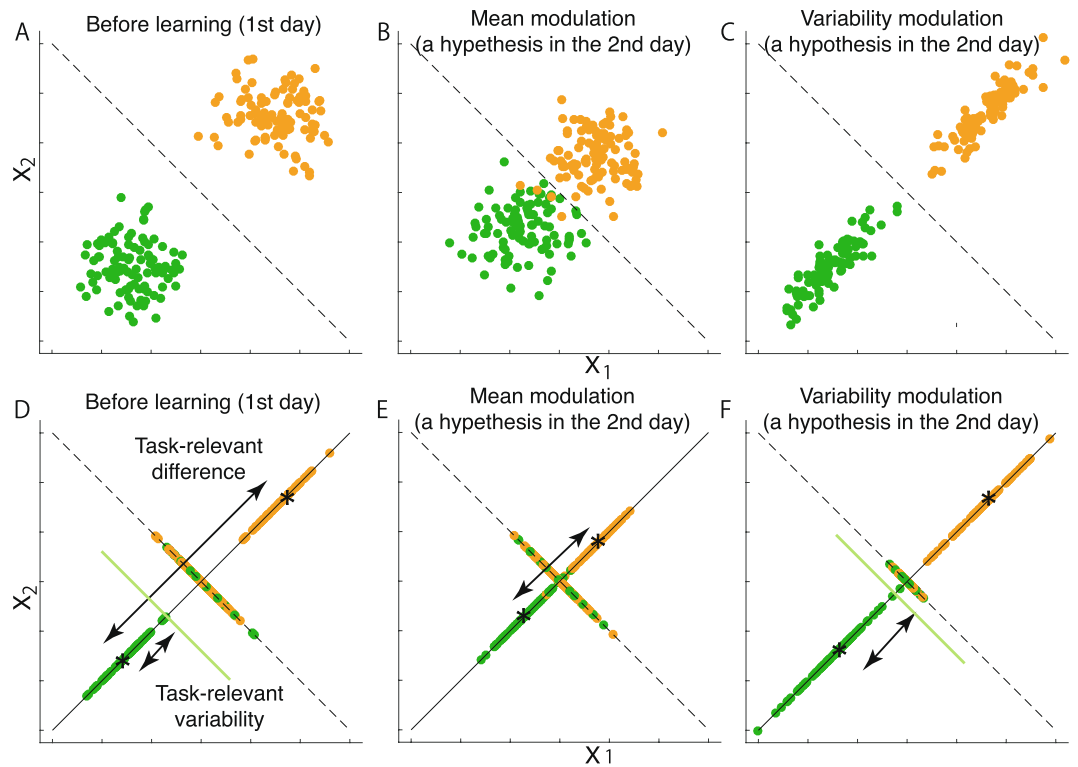
Figure 3. Simulated data. (**A**,**D**) Simulated data in a simple two-dimensional case $X = (X_1, X_2)$. The green and orange dots indicate simulated data for different movement patterns (e.g., the green dots represent motion data for throwing a fastball, and the orange dots represent motion data for throwing a breaking ball). In this case, the classification boundary is drawn as a black dotted line. Panel (D) shows the task-relevant components (the green and orange dots on the solid black line) and the task-irrelevant components (the green and orange dots on the black dotted line) of the simulated data in panel (A). The task-relevant difference is defined as the distance between the means of task-relevant components, as indicated by the black asterisks. The task-relevant variability is defined as the variability in the task-relevant components, as shown by the variability between the black asterisk and the solid green line. The task-relevant variabilities can also be calculated for the data in a similar manner as denoted by orange dots. (**B**,**E**) A simulated mean modulation. In the mean-modulation hypothesis, the task-relevant differences are reduced, as shown in panel (E). (**C**,**F**) A simulated variability modulation. In the variability-modulation hypothesis, the task-relevant variance is increased, as shown in panel (F).

On the first day, the subjects were instructed to throw either a fastball or a breaking ball towards a target. On the second day, the subjects were asked to throw the two types of balls with a similar form. No significant difference in ball trajectories and motion data was observed between the two days (see the Supplementary Materials).

However, a significant modulation was observed in the task-relevant difference between the two days (Fig. 4A, $p = 0.0029$, paired t-test). In contrast to the task-relevant difference, there was no modulation of the task-relevant variability (Fig. 4B, $p = 0.4777$, paired t-test). The task-relevant variabilities were calculated by averaging the variance of $X_{rel,0}$ and that of $X_{rel,1}$. Although these results appeared to support mean modulation, there are two possibilities: the modulation may have been induced via learning or may have been produced due to the task requirement. If the first possibility were correct, we would expect to observe smaller task-relevant differences in the last 50 trials than in the first 50 trials within the first day. We divided the task-relevant components into those for the former and latter trials on each day. No significant interaction was observed between 'Part' and 'Day' ($p = 0.890$, see Methods for details), and no significant difference was observed in the task-relevant difference between the former and latter trials on the first day ($p = 0.3276$) or second day ($p = 0.2667$). These results contradicted the first possibility; therefore, the modulation of the task-relevant difference was not induced by learning within each day. In contrast, a significant difference was observed in the task-relevant difference between the former trials on the first day and second day ($p = 0.00130$) and between the latter trials on the first day and second day ($p = 0.01001$). Therefore, the task requirement to throw the fastballs and breaking balls with similar forms induced the mean modulation.

Discussion

We proposed a data-driven method to detect task-relevant and task-irrelevant components when task outcomes are categorical using a machine learning technique. Our data-driven method can be applied even when the relationship between the motion data and the outcome is unknown; the relation can be estimated in a data-driven manner. Because the mathematical form between the current method and our earlier method focusing on continuous outcome[5,6] is invariant, our data-driven approach forms a unified approach for detecting task-relevant and
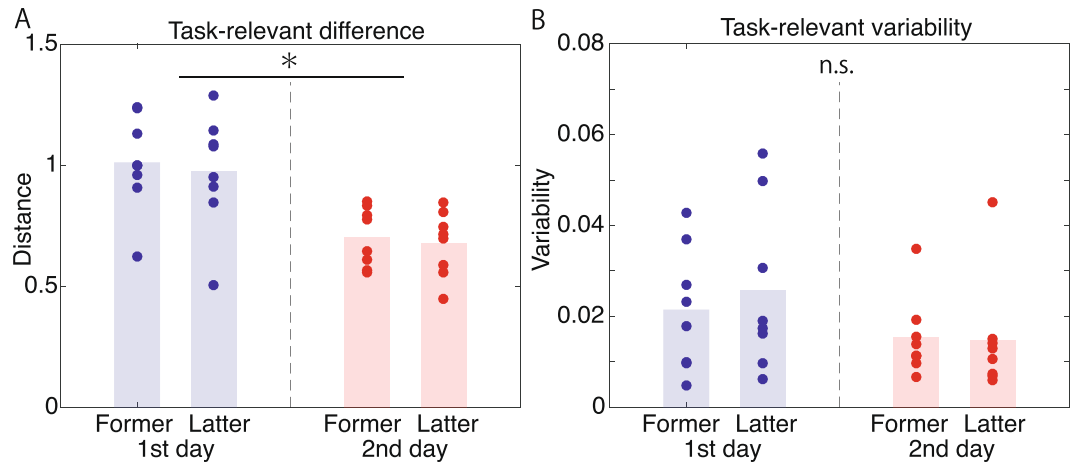
Figure 4. Experimental results. (**A,B**) The task-relevant difference and the variability calculated for the former and latter trials on the first and second day. A significant difference (denoted by a black asterisk) was observed only for the task-relevant difference between the first and the second day.

task-irrelevant motion components independent of the properties of the outcome. Our approach detected mean modulation in the task-relevant difference under the task requirement to throw fastballs and breaking balls with similar forms (Fig. 4A). The task requirement thus affected the mean rather than the variability in task-relevant motion components.

Although we relied on simple logistic regression, more complicated methods, such as mixture models[12–15], kernel techniques[16], and deep learning[17], could also be applied. Linear regression has several advantages, including that it is related to motor primitive, a conventional model of motor control and learning[18–23]. Although most conventional models of motor learning focus on continuous outcome values, our method can also be applied to motor learning with categorical outcomes.

One limitation inherent to our method is that we assume a linearly separable case in the current study. When nonlinear transformation of X is indispensable for classification, kernel methods can work in many cases[12]. Using a kernel method, it is impossible to quantify the task-relevant differences except in the case of a linear kernel. The kernel method relies on nonlinear transformations of X, $\phi(X)$. Although the explicit function of $\phi(X)$ is not necessary, the second-order statistics of $\phi(X)$ (i.e., kernel $K(X, X') = \phi(X)\phi(X')^T$) should be defined. Thus, we can possibly evaluate task-relevant variabilities (i.e., the second-order statistics) using $K(X, X')$ but cannot evaluate task-relevant differences (i.e., first-order statistics). Because the data in the current study were linearly separable without nonlinear transformation of X (Fig. 2), we were able to quantify the modulation of task-relevant differences (Fig. 4).

Because the current study is the first step to apply a data-driven approach for discussing categorical motion outcomes, it is still unclear whether the current data-driven method is useful in other cases. The present study validated the effectiveness of a data-driven approach to discussing the throwing motions of fastball and breaking ball in naive participants. It thus remains unclear whether our method is practical in other contexts, such as the throwing motions of curveball and slider in naive participants, neurological patients, expert baseball pitchers, or baseball pitchers with a shoulder injury. In these and other contexts, some context-dependent modifications of a data-driven approach are possibly necessary and practical. To propose effective data-driven approaches to discuss diverse situations, such as detecting success-relevant components, patient-specific components, or professional-specific components, would be not only open questions but also promising future works.

One alternative to our method is to use Fisher linear discriminant analysis (LDA)[12]. Because LDA is a linear classification method whose classification performance is almost equivalent to logistic regression[24], Eqs. (3) and (4) can still be applied. LDA could thus be a possible alternative to logistic regression. Other nonlinear classification techniques, such as quadratic discriminant analysis (QDA), are also possible alternatives to logistic regression when sufficient data are available. Given a sufficient quantity of data, nonlinear classification techniques can provide better classification accuracy than linear classification techniques. However, Eqs. (3) and (4) can be applied only to linear classification methods. To discuss task-relevant and task-irrelevant components, we must rely on linear classification methods, such as logistic regression or LDA.

It is possible to apply our method to some neural network models with some constraints. In general, a neural network model with L layers can be written as $y = g(h(XW))$, where g and h denote the nonlinear functions generated from the second to the Lth layer and the nonlinear function in the first layer, respectively. In this form, we can still expect to be able to apply Eqs. (3) and (4) to the neural network. When h is a rectified linear function ($h(x) = ax$ when $x > c$ and $h(x) = 0$ otherwise, where c is a threshold), which is a common nonlinear function used in recent deep neural network models, many patterns of XW will generate identical 0 values. Because we assumed a one-to-one relation between X and d in Eq. (2), it cannot be appropriate for deriving task-relevant components under the rectified linear function. In contrast, when h is a sigmoid function, which is a common nonlinear function used in conventional neural network models, Eqs. (3) and (4) can be applied to extract task-relevant and task-irrelevant components.

Although it is difficult to define the relation between time-varying motion data and categorical task outcomes using mathematical equations such as parabolas, UCM focuses on kinematic parameters. UCM can thus be an alternative to our data-driven method in the current situation. A key point in UCM is that it requires linear approximations around averaged kinematics. For example, to apply UCM to our data, we need to evaluate the task-relevant and task-irrelevant components around averaged joint angles. In the simple case shown in Fig. 3, the averaged joint angles can be located in a blank region without any motion data. In addition, the linear approximation assumes that all the data are located close to the averaged values. In discussing the task-relevant and task-irrelevant components, this assumption will be invalid in several situations, such as in Fig. 3.

Another possible use case for UCM is to apply it separately to the motion data for throwing fastballs and breaking balls. Although this analysis is possible practically, it is beyond the assumption of the linear approximation theoretically. The linear approximation enables calculation of the task-relevant and task-irrelevant components around averaged joint angles by projecting the motion data onto the vectors corresponding to the directions that affect kinematic performance and to the directions that do not affect the performance, respectively. In addition, the directions of the vectors depend on the averaged joint angles for the linear approximation. Consequently, to analyze the current data separately in each condition, we need to compare task-relevant components using different vectors for each condition. For an imaginary and simplified example, the task-relevant directions can be proportional to $2\theta_e + \theta_s$ when throwing a fastball, and proportional to $\theta_e + 2\theta_s$ when throwing a breaking ball, where θ_e and θ_s are the elbow and shoulder joint angles, respectively. Even when the mean and variance of the task-relevant components differ among the two conditions, it is difficult to judge whether the difference originates from the difference of task condition or task-relevant directions. In sum, although the UCM is sophisticated for discussing task-relevant and task-irrelevant components within the same condition, it is unsuitable for comparing those components, especially in situations that have multiple and categorical task outcomes.

Data availability

The datasets analyzed in the current study are available from the corresponding author upon reasonable request.

References

1. Bernstein, N. A. *The co-ordination and regulation of movements.* Pergamon, London (1967).
2. Scholz, J. P. & Schöner, G. The uncontrolled manifold concept: identifying control variables for a functional task. *Exp. Brain Res.* **126**(3), 289–306 (1999).
3. Cusumano, J. P. & Cesari, P. Body-goal variability mapping in an aiming task. *Biol. Cybern.* **94**(5), 367–379 (2006).
4. Müller, H. & Sternad, D. A randomization method for the calculation of covariation in multiple nonlinear relations: illustrated with the example of goal-directed movements. *Biol. Cybern.* **89**(1), 22–33 (2003).
5. Furuki, D. & Takiyama, K. Detecting the relevance to outcome of whole-body movements. *Sci. Rep.* **7**(1), 15659 (2017).
6. Furuki, D. & Takiyama, K. Decomposing motion that changes over time into task-relevant and task-irrelevant components in a data-driven manner: application to motor adaptation in whole-body movements. *Sci. Rep.* **9**(1), 7246 (2019).
7. David, W. H., Lemeshow, S. & Cook, E. D. Applied Logistic Regression, Second Edition: Book and Solutions Manual Set. Wiley-Interscience (2001).
8. Phinyomark, A., Osis, S., Hettinga, B. A. & Ferber, R. Kinematic gait patterns in healthy runners: A hierarchical cluster analysis. *J. Biomech.* **48**(14), 3897–3904 (2015).
9. Halilaj, E. *et al.* Machine learning in human movement biomechanics: Best practices, common pitfalls, and new opportunities. *J. Biomech.* **81**, 1–11 (2018).
10. Raschka, S. & Mirjalilili, V. *Python Machine Learning.* Packt Publishing Ltd. (2017)
11. Qian, J., Hastie, T., Friedman, J., Tibshirani, R. & Simon, N. Glmnet for Matlab (2013). http://www.stanford.edu/~hastie/glmnet_matlab/
12. Bishop, C. M. *Pattern Recognition and Machine Learning.* Springer Verlag. (2006).
13. Naruse, Y., Takiyama, K., Okada, M. & Umehara, H. Statistical method for detecting phase shifts in alpha rhythm from human electroencephalogram data. *Phys. Rev. E* **87**(4), 042708 (2013).
14. Takiyama, K., Katahira, K. & Okada, M. Exact inference in discontinuous firing rate estimation using belief propagation. *J. Phys. Soc. Jpn* **78**(6), 4003 (2009).
15. Takiyama, K. & Okada, M. Detection of hidden structures in nonstationary spike trains. *Neural. Comput.* **23**(5), 1205–1233 (2011).
16. Rasmussen, C. E. & Williams, C. K. I. *Gaussian Processes for Machine Learning.* MIT Press. (2006).
17. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**(7553), 436–444 (2015).
18. Thoroughman, K. A. & Shadmehr, R. Learning of action through adaptive combination of motor primitives. *Nature* **407**, 742–747 (2000).
19. Takiyama, K. Context-dependent memory decay is evidence of effort minimization in motor learning: a computational study. *Front in Comput Neurosci* **9**(4), 1–10 (2015).
20. Takiyama, K., Hirashima, M. & Nozaki, D. Prospective errors determine motor learning. *Nat. Comm.* **6**, 5925 (2015).
21. Takiyama, K. & Sakai, Y. Balanced motor primitive can explain generalization of motor learning effects between unimanual and bimanual movements. *Sci. Rep.* **6**, 23331 (2016).
22. Takiyama, K. & Shinya, M. Development of Portable Motor Learning Laboratory (PoMLab). *PLoS One* **11**, e0157588 (2016).
23. Ishii, K., Hayashi, T. & Takiyama, K. Influence of switching rule on motor learning. *Sci. Rep.* **8**(1), 13559 (2018).
24. Pohar, M., Bias, M. & Turk, S. Comparison of logistic regression and linear discriminant analysis: a simulation study. *Metodološki zvezki* **1**(1), 143–161 (2004).

Acknowledgements

Author contributions

D.F. and K.T. designed the experiments and D.F. performed the experiments. D.F. and K.T. performed the analyses and wrote the manuscript.

Competing interests

The authors declare no competing financial or nonfinancial interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-59257-z.

Correspondence and requests for materials should be addressed to K.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.