Aim of this Supporting Information file

As is widely known, the performance of the machine learning model highly depends on the hyper parameter settings. Although effect of some parameters has been verified in the original paper of NGC (neural granger causality) [19], since this is the first study to use machine learning model for analyzing the whole-body human movement in a sports context within our best knowledge, we consider it is essential to sincerely provide how each hyperparameter affects the results and how we tuned those parameters for avoiding the overvaluation of our method. Specifically, in this supporting information file, we first describe the criteria for hyper parameter settings in this study, then, verified the relation between the changes of each hyper parameters and results of the analysis.

Criteria for hyper parameter setting

For setting the hyper parameters, variable usage rate, which indicates the ratio of candidate combinations of different body joints showing a positive NGC, was used for tuning of them. Figure S1 shows the effect of different variable usage rate on the causal graph obtained from the NGC analysis. If the value of variable usage rate is 0, there is no NGC among body joints and all joints' movement depends on the past information of themselves (i.e., autoregressive model). Also, if it shows 1.0, all joints have NGC to all other joints. Obviously, these two cases should be avoided. Specifically, we tuned the hyper parameters so that the mean variable usage rate among 16 pairs becomes around 0.25-0.50 based on the following insights. First, if the variable usage rate is over 0.50, there always inter-body joint combinations that have positive NGC, and this may lead the overestimation of the number and size of inter-body NGC. Also, to avoid the underestimation of the number of joint combinations that have NGC, we set the lower limit at 0.25 to capture about 50% of the NGC between inter-body joints (0.25/2+0.25/2). If there is any body joint that has large NGC to opponent's joints, we can also identify through this criterion. As a result of parameter tuning, mean variable usage rate among 16 pairs is 0.30 ± 0.03 in this study.

(a)   (b)

(Variable usage rate: 0.16)   (Variable usage rate: 0.33)

(c)   (d)

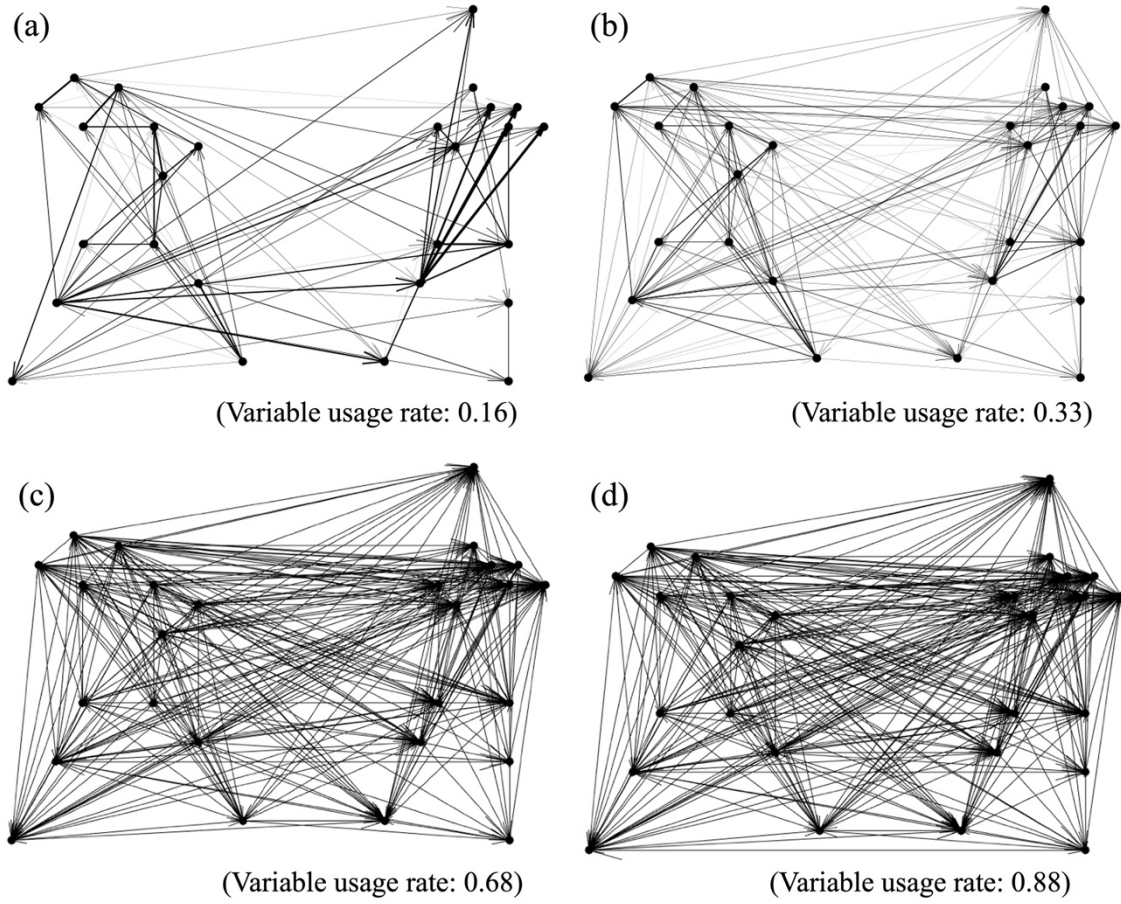(Variable usage rate: 0.68)   (Variable usage rate: 0.88)

Figure S1. The relations between differences variable usage rate and obtained causal graph from NGC analysis (a) 0.16, (b) 0.33, (c), 0.68, (d) 0.88.

Testing effect of hyper parameters on the result

In this section, we verify the effect of hyper parameters on the result of the analysis of the paper. Specifically, we verify the relations between the changes of hyper parameters and variable usage rate, and the ration among four causal indexes (ngc_{pp}, ngc_{bb}, ngc_{pb}, ngc_{bp}) of the Analysis1 of the paper. Since the numerical value of NGC itself is not meaningful, we converted it to a ratio and examined how their relative size changes with parameter setting. The following are explanations of each parameter.

penalty term λ: this parameter penalizes the total amount of weights included in the trained model in equation (3). If we set smaller λ value, the variable usage rate will be larger. We tested the impact of this parameter with different values of 0.0001, 0.001, 0.003 (our setting), 0.005, 0.01, 0.1 while keeping other parameters constant.

Sampling rate (down sampling rate): this parameter controls the time interval between the points

included in the input time series data. Since penalty term in equation (3) penalizes the sum of weights for avoiding to include too many lagged causalities, if we try to capture the characteristics of the IC with long time scale (several hundred milliseconds ~ several seconds), we need to down sample the data from the original motion capture data with high frequency. We tested the impact of this parameter with different values of 250 Hz (1.0), 125 Hz (2.0), 50 Hz (5.0, our setting), 25 Hz (10.0).

Maximum lag: this parameter controls the maximum time lag that is considered in the calculation. If there is any hypothesis about this parameter (e.g., near the ball travel time 0.50 sec for the baseball), it should be set as enough large to include it. We tested the impact of this parameter with different values of 2.0 sec, 1.0 sec (our setting), 0.5 sec, 0.1 sec. Note that although we only change the value of a single parameter in this file, the sampling rate should also be adjusted to keep the range of maximum time-lag const (e.g., 100 Hz with 50 points (0.50 sec) and 200 Hz with 100 points (0.50 sec)).

Number of hidden units: this parameter indicates the number of hidden units included in the layer of multi-layer perceptron. While the number of features that can be considered in the model increases as the number of hidden units increases, the complexly of the model and required time for the training process also increases. We tested the impact of this parameter with different values of 128, 64, 32 (our setting), 16, 8.

Learning rate: this parameter controls the step size for updating the weight parameters at each iteration based on the value of loss function. While if too small value of it increases the training time, too large value makes it difficult to find the optimal solution. We tested the impact of this parameter with different values of 0.1, 0.07, 0.05 (our setting), 0.03, 0.01.

Type of input variable: this parameter indicates the type of input variable to the model. Although this study tried to capture the characteristics of the interaction between two players from the aspects of the "velocity interaction", the form of the interaction may change depends on the type. Therefore, we also calculated the resultant acceleration data of each joint and verified this viewpoint.

Results and discussions
Table S1-S6 shows the effect of changing the hyper parameter on the mean value of variable usage rate and the ratio of four causal indexes (ngc_{pp}, ngc_{bb}, ngc_{pb}, ngc_{bp}) to sum of them. First, as shown in the Table S1, the size of λ parameter affects the variable usage rate, i.e., if the penalty for the weight parameter increases, the number of causal relations detected through the NGC analysis will decrease. However, they are consistent with respect to their relative ratios to the sum of the weight

parameters. Hence, the change of the variable usage rate could be due to an increase in the detection of small causalities. Therefore, if criteria are set to eliminate combinations of joints with such a small causality, as in this study, we believe that a robust analysis is possible with respect to the increase or decrease of this parameter except for the case that the variable usage rate is 1.0 or 0.0. On the contrary, the size of the sampling rate affects the ratio of four causal indexes, i.e., the ratio of inter-personal NGC (ngc_{pb} and ngc_{bp}) decreases as this parameter increases (Table S2). As mentioned above, this seems to be caused by the difficulty for detecting the long-time scale interaction (such as 0.50 sec delay) with high sampling rate. This is evidenced by the fact that the same trend is observed for the results of maximum time lag, i.e., if we allow only few time-lag, the inter-personal NGC decreases.

Further, from the results of number of hidden units (Table S4), this parameter has small effect on the analysis. The slight difference of the variable usage rate is thought to be caused by differences in convergence speed, i.e., larger number of hidden units require larger training time. However, if the model tries to identify the causal structure in more complex interaction, this parameter should be carefully tuned. Also, the learning rate has similar effect of the λ parameter, i.e., it has large effect on the variable usage rate but not on the ratio of causal indexes (Table S5). The reason for increasement of variable usage rate in 0.1 is that it cannot converge to the optimal solution due to the large step size. Finally, slightly different results were observed by using resultant acceleration as the input data (Table S6). Namely, by using the acceleration data, the ratio of NGC among pitcher's body joints increased while that of batter's body joints decreased. Therefore, the type of input data may have some influence on the results even if the overall trend is unchanged. To account for the effects of amplified noise in acceleration data, this study adopts resultant velocity data as the input variable.

In summary of these results, although there are some sensitivities to the hyper parameter settings, we consider that the criteria and analytical procedures defined in this study have yielded robust results. Therefore, we believe that the setting of this study has a certain level of validity. However, when it applies to other type of interaction or human movement, the sensitive hyper parameters such as the down-sampling rate, maximum lag and type of input data should be carefully tuned.

Table S1. The impact of changing the value of λ parameter on the results. NGC_pp, NGC_bb, NGC_pb, NGC_bp indicates the ratio of four causal indexes canulated in Analysis1 to the some of them.

λ value	Variable usage rate	NGC_pp	NGC_bb	NGC_pb	NGC_bp
0.0001	1.00±0.00	0.25±0.00	0.25±0.00	0.25±0.00	0.25±0.00
0.001	0.69±0.03	0.47±0.01	0.30±0.10	0.21±0.09	0.02±0.01
0.003 (ours)	0.30±0.03	0.46±0.03	0.33±0.09	0.20±0.09	0.01±0.01
0.005	0.21±0.02	0.44±0.04	0.36±0.09	0.19±0.09	0.01±0.01
0.01	0.15±0.01	0.43±0.05	0.36±0.09	0.20±0.10	0.01±0.01
0.1	0.00±0.00				

Table S2. The impact of changing the value of sampling rate (down sampling rate) parameter on the results.

Sampling rate (Hz)	Variable usage rate	NGC_pp	NGC_bb	NGC_pb	NGC_bp
250	0.15±0.01	0.41±0.05	0.55±0.05	0.03±0.01	0.00±0.00
125	0.19±0.01	0.43±0.04	0.48±0.05	0.08±0.03	0.01±0.01
50 (ours)	0.30±0.03	0.46±0.03	0.33±0.09	0.20±0.09	0.01±0.01
25	0.33±0.01	0.36±0.16	0.33±0.09	0.30±0.16	0.00±0.00

Table S3. The impact of changing the value of maximum lag parameter on the results.

Maximum lag (sec)	Variable usage rate	NGC_pp	NGC_bb	NGC_pb	NGC_bp
2.0	0.51±0.05	0.30±0.02	0.39±0.15	0.29±0.15	0.01±0.01
1.0 (ours)	0.30±0.03	0.46±0.03	0.33±0.09	0.20±0.09	0.01±0.01
0.5	0.27±0.02	0.45±0.04	0.36±0.07	0.17±0.07	0.01±0.01
0.1	0.25±0.01	0.47±0.03	0.48±0.03	0.03±0.01	0.01±0.01

Table S4. The impact of changing the number of hidden units of the multi-layer perceptron on the results.

Number of hidden units	Variable usage rate	NGC_pp	NGC_bb	NGC_pb	NGC_bp
128	0.35 ± 0.04	0.46 ± 0.03	0.33 ± 0.09	0.20 ± 0.09	0.01 ± 0.01
64	0.31 ± 0.03	0.46 ± 0.04	0.33 ± 0.09	0.20 ± 0.09	0.01 ± 0.01
32 (ours)	0.30 ± 0.03	0.46 ± 0.03	0.33 ± 0.09	0.20 ± 0.09	0.01 ± 0.01
16	0.29 ± 0.03	0.46 ± 0.04	0.34 ± 0.09	0.20 ± 0.09	0.01 ± 0.01
8	0.28 ± 0.03	0.46 ± 0.04	0.34 ± 0.09	0.19 ± 0.09	0.01 ± 0.01

Table S5. The impact of changing the value of learning rate parameter on the results.

Learning rate	Variable usage rate	NGC_pp	NGC_bb	NGC_pb	NGC_bp
0.1	0.50 ± 0.10	0.44 ± 0.04	0.35 ± 0.09	0.19 ± 0.09	0.01 ± 0.01
0.07	0.21 ± 0.02	0.44 ± 0.04	0.34 ± 0.09	0.21 ± 0.09	0.01 ± 0.01
0.05 (ours)	0.30 ± 0.03	0.46 ± 0.03	0.33 ± 0.09	0.20 ± 0.09	0.01 ± 0.01
0.03	0.40 ± 0.04	0.46 ± 0.03	0.31 ± 0.10	0.21 ± 0.10	0.01 ± 0.02
0.01	1.00 ± 0.00	0.26 ± 0.00	0.25 ± 0.00	0.25 ± 0.00	0.24 ± 0.00

Table S6. The impact of changing the type of input variable on the results.

Type of input variable	Variable usage rate	NGC_pp	NGC_bb	NGC_pb	NGC_bp
Resultant velocity (ours)	0.30 ± 0.03	0.46 ± 0.03	0.33 ± 0.09	0.20 ± 0.09	0.01 ± 0.01
Resultant acceleration	0.52 ± 0.10	0.55 ± 0.07	0.22 ± 0.11	0.22 ± 0.11	0.01 ± 0.01