The Twenty-Sixth International Symposium on Artificial Life and Robotics 2021 (AROB 26th 2021),
The Sixth International Symposium on BioComplexity 2021 (ISBC 6th 2021),
ONLINE, January 21-23, 2021

Animacy perception of a pair of movements under quantitative control of its temporal contingency: a preliminary study

Kenichi Oguma[1], Takuma Torii[1], and Shohei Hidaka[1]

[1] Japan Advanced Institute of Science and Technology, Japan

[1] {s1910055, tak.torii, shhidaka}@jaist.ac.jp

Abstract: People perceive life-likeness from the movements of even geometric shapes. This phenomenon is called animacy perception. In this study, we examine our hypothesis that the degree to which one movement contributes to the predictability of the other movement characterizes the degree of animacy. To quantify temporal contingency in this sense, we define temporal contingency by Granger causality in multivariate animacy. To test our hypothesis, we created various movie clips of the two moving circles and asked human participants to rate the degree of animacy, intention, and contingency in each movie. The result of this experiment showed that the degrees of contingency and intension of one object were correlated to those of animacy, only if there was low Granger causality from the other to the object. This result suggests that perceived autonomy, measurable by Granger causality, would play a major role in deciding factors correlated to animacy perception.

Keywords: Animacy perception, intention, contingency, and Granger causality

1 PERCEPTION OF LIFE-LIKENESS

People perceive some degree of life-likeness of a given moving object even if it is a non-living thing such as a geometric shape [1]. When people observe a geometric shape changing its position continuously over time, people often perceive that the geometric shape 'itself' moves as if it were a living thing. This is known as animacy perception that the observer interprets the object looks like animate or life-like. The human observer can recognize animacy or life-likeness to the movement of a simple geometric shape with few superficial clues indicating its life-likeness (such as eyes, face, etc.). This is considered to be essentially related to the ability to read the intentions of others and the social relationships of multiple agents from their actions [2].

In psychology, many experimental studies since Heider and Simmel [1] have explored the factors by which subjects (observers) perceive animacy using animation films (movies) in which geometric shapes 'plays their roles' of the human actors in their story [2, 3, 4, 5]. Bassili [3] suggested that humans can perceive life-likeness even for simple "one follows the other" movements. Bassili [3] defined "temporal contingency", as the key concept of his hypothesis on animacy perception in social interaction, by "moves by A will be followed within a short lag by compensatory moves by B". Ueda [6] reviewed that many empirical studies on animacy perception are related to temporal contingencies between multiple moving objects. In recent years, Takahashi and Watanabe [7] examined the degree of "synchrony" in the movements among dozens of objects, and showed that the degree of animacy perception was lower, if the degree of synchronization of the observed object with the surrounding objects is higher. In a certain definition of temporal contingency, a high degree of synchrony between moving objects is supposed associated with a high degree of temporal contingency of them. If one employs this definition of temporal contingency, it is inconsistent with Bassili [3]'s hypothesis, saying that the degree of temporal contingency correlates with the degree of perception of animacy. Thus, this naive or vague idea of temporal contingency hypothesis needs to be refined.

In this study, we considered a sort of temporal dependency between two interactively moving objects and explore the effects of temporal contingency on animacy perception of these objects. While many studies have considered temporal structure of interacting objects, to the best of our knowledge, many researchers have not employed a quantitative characterization of their hypothesized temporal dependency of the observed/generated movements (instead, mostly they only manipulated parameters of movement-generating algorithms). To quantitatively define the concept of temporal contingency, we employ the idea of conditional predictability of one's future movement given the one's and other's past movements. Given two moving geometric shapes in a movie clip, we hypothesized that conditional predictability of a pair of movements correlates with the degree of perception of animacy. Intuitively, when knowing the past movement of object B improves the prediction of the future movement of object A, the observer of both objects in a movie clip may attribute animacy to those objects. According to this hypothesis, a pair of perfectly synchronized movements has no contribution to improving the predictability of each other, so it is expected to have a low animacy evaluation. This is conceptually consistent with the finding by Takahashi and Watanabe [7], showing that animacy perception is unlikely to occur for a given set of perfectly synchronized movements. In this study, we conducted a preliminary experiment to see if the degree of conditional predictability characterizes the degree of animacy perception.

2 METHODS

2.1 Generation of pairs of movements

In order to construct a simple visual stimulus that displays movements of a pair of two objects with or without interaction, we used a vector auto-regressive (VAR) model as a statistical model.

We specifically used a first-order bivariate VAR model. We call two geometric shapes in the movie clips R (depicted by a red circle) and B (a black circle). The one-dimensional positions of these circles at time $t = 0, 1, \dots$ are denote by R_t and B_t. The random variables R_t and B_t evolve using the following VAR model equation for $t \geq 0$:

$$\begin{pmatrix} R_{t+1} \\ B_{t+1} \end{pmatrix} = \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix} \begin{pmatrix} R_t \\ B_t \end{pmatrix} + \begin{pmatrix} \epsilon_{R,t} \\ \epsilon_{B,t} \end{pmatrix} , \quad (1)$$

where $R_0 = B_0 = 0$, $A = (a_{ij}) \in \mathbb{R}^{2\times 2}$ is the coefficient matrix with the real entries, and both random variables $\epsilon_{R,t}$ and $\epsilon_{B,t}$ for each t have the normal distribution $N(\epsilon_{R,t}|0, \sigma^2)$ and $N(\epsilon_{B,t}|0, \sigma^2)$ with mean 0 and variance σ^2 as its probability density function.

2.2 Quantification of conditional predictability

In this study, the degree to which one object contributes to the predictability of the other object is measured by the statistic known as Granger causality [8]. Granger causality was originally proposed to measure the conditional predictability of a univariate time-series, generated by a VAR model, given its past or/and other past states of random variables.

Granger causality is a descriptive statistic that shows how much better the future R_t can be estimated than the one random variable R_{t-1} alone by using two random variables R_{t-1} and B_{t-1}. Thus, Granger causality is directional. Denote a Granger causality from R to B by $G_{R\to B}$ and the opposite directional one by $G_{B\to R}$. Granger causality can be estimated from the sums of squared residuals, Eq. (2)–(5), of the ordinary least-square procedure.

$$E_R = \min_{\alpha,\gamma} \sum_{t=1}^{T} (R_t - \alpha R_{t-1} - \gamma)^2 \quad (2)$$

$$E_B = \min_{\alpha,\gamma} \sum_{t=1}^{T} (B_t - \alpha B_{t-1} - \gamma)^2 \quad (3)$$

$$E_{R,B} = \min_{\alpha,\beta,\gamma} \sum_{t=1}^{T} (R_t - \alpha R_{t-1} - \beta B_{t-1} - \gamma)^2 \quad (4)$$

$$E_{B,R} = \min_{\alpha,\beta,\gamma} \sum_{t=1}^{T} (B_t - \alpha B_{t-1} - \beta R_{t-1} - \gamma)^2 \quad (5)$$

With this notation, the Granger causality $G_{R\to B}$ and $G_{B\to R}$ for the sample of bivariate timeseries $((R_t, B_t))_{t=1,\dots,T}$ can be estimated by Eq. (6) and Eq. (7), respectively, as follows.

$$G_{R\to B} = \log\left(\frac{E_B}{E_{B,R}}\right) \quad (6)$$

$$G_{B\to R} = \log\left(\frac{E_R}{E_{R,B}}\right) \quad (7)$$

Intuitively, $G_{R\to B}$ is interpreted as a measure of the degree to which the future value of B is easy to predict by using R, and $G_{B\to R}$ is interpreted as a measure of the degree to which the future value of R is easy to predict by using B.

2.3 Participants

Seven Japanese graduate students (6 males and 1 female) in Japan Advanced Institute of Science and Technology participated in the experiment.

2.4 Apparatus and stimuli

Fig. 1 shows a schematized situation of the experimental setting with an example of stimuli displayed. Using the VAR model, Eq (1), we generated the movements of two circles, denoted by R the red and B the black one, and created movie clips as the experimental stimuli. Each object is depicted by a colored circle in order not to give visual cues indicating biological characteristics. The position of the red circle R, R_t, was displayed as the position in the vertical axis to the display, and the trajectory of the red circle's 5 past positions was displayed along the horizontal axis on the target red circle. The k^{th} previous position of the red circle was displayed at the position of $-(10\,k + 20)$ pixels relative to the center of the display for $k = 0, 1, \dots, 5$. For the black circle B, the position in the vertical direction is B_t, and the k^{th} previous position in the horizontal direction is displayed at the position of $+(10\,k + 20)$ pixels relative to the center of the display. In other words, the two circles move vertically while extending their history horizontally. The position of the past red circle displayed earlier was displayed in red with a higher degree of transparency. In this experiment, we set the framerate at 6 Hz in consideration of the visibility of the movement in the computer display.

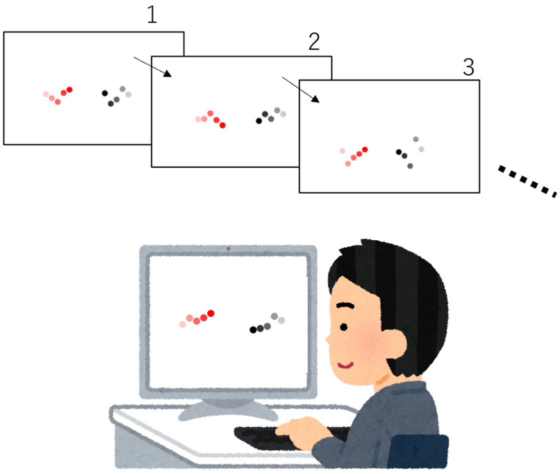

Fig. 1. A schematized experimental setting. Each participant watched a movie clip (animation) of the red and black circles and then evaluate the degree of animany, intention, and contingency of the red circle in each trial. Repeat this procedure for 90 trials with different movie clips. Numbers in this figure indicate frame numbers of a movie clip (not exposed as this form to participants).

According to the hypothesis of this study, a stimulus with a larger Granger causality would be expected to have a higher degree of animacy. To compare $G_{R\to B}$ and $G_{B\to R}$, we created the experimental stimuli under the three conditions as follows. The first condition is called "Red→Black", under which each stimulus has high $G_{R\to B}$ and low $G_{B\to R}$. The second condition called "Black→Red", under which each stimulus has low $G_{R\to B}$ and high $G_{B\to R}$. The third condition is called "Low GC" (stands for Granger Causality), under which each stimulus has low $G_{R\to B}$ and low $G_{B\to R}$.

In general, there are multiple sets of the parameters of the VAR model, Eq. (1), that give the same pairs of Granger causality. Therefore, in the preliminary experiment of this study, we defined three types of the coefficient matrices A for each of the three conditions above. The six types of coefficient matrices out of all the nine types are shown in the Eq. (8)–(10). For the other three types, we used the matrices in which the signs of the non-diagonal components of $A_{\mathrm{HL}}, A_{\mathrm{LH}}, A_{\mathrm{LL}}$ were negative (i.e., $A = \left(\begin{smallmatrix} a & -b \\ -a & b \end{smallmatrix}\right)$ for the "negative" type corresponding to A_{HL}). For each stimulus, the coefficient $a \gg b > c$ in the coefficient matrix was sampled with the uniform probability in the range of $a \in [0.9, 0.99]$, $b, c \in [0.01, 0.1]$. In addition, a coefficient matrix of $|a| + |b| > 1$ or $|b| + |c| > 1$ was removed from stimulus generation, to let the values of the random variables R_t and B_t have a proper limit in a finite value. The variance of R_t, B_t at each step was fixed at $\sigma^2 = 10$.

Red→Black condition:

$$A_{\mathrm{HL}} = \begin{pmatrix} a & b \\ a & b \end{pmatrix}, A_{\mathrm{H0}} = \begin{pmatrix} a & 0 \\ a & b \end{pmatrix} \tag{8}$$

Black→Red condition:

$$A_{\mathrm{LH}} = \begin{pmatrix} b & a \\ b & a \end{pmatrix}, A_{\mathrm{0H}} = \begin{pmatrix} b & a \\ a & 0 \end{pmatrix} \tag{9}$$

Independent condition:

$$A_{\mathrm{LL}} = \begin{pmatrix} b & c \\ b & c \end{pmatrix}, A_{\mathrm{00}} = \begin{pmatrix} b & 0 \\ 0 & c \end{pmatrix} \tag{10}$$

Time-series generated by the coefficient matrices A_{HL}, A_{LH}, and A_{LL} for 10 seconds (60 frames) are shown in Fig. 2. In the Red→Black and Black→Red conditions (e.g., A_{HL} and A_{LH}), the red circle R and the black circle B move in the same direction with a relatively high probability. In the Low GC condition (e.g., A_{LL}), they move almost statistically independently.

In this study, we created several movies for each of these nine types of coefficient matrices. Then, we selected the top 10 movies with the largest $G_{R\to B}$ for the Red→Black condition, with the largest $G_{B\to R}$ for the Black→Red condition, with the smallest $G_{R\to B}$ and $G_{B\to R}$ for the Low GC condition. The selected ones vary $G_{R\to B} \in [0.51, 0.61]$ in the Red→Black condition, $G_{B\to R} \in [0.53, 0.58]$ in the Black→Red condition, $G_{R\to B}, G_{B\to R} \in [1.34 \times 10^{-6}, 5.71 \times 10^{-13}]$ in the Low GC condition.

2.5 Procedure

Participants evaluated 90 movies in total (3 conditions × 3 types × 10 clips) in sequence. Each of the participants sat on a chair and watched a movie clip displayed on the computer display on the desk, which is approximately 50cm away from the participant. Each of them could watch the repeating movie clip as long as they wanted until they started the evaluation phase. The presentation order of the stimulus movies was randomized with all the 90 stimuli under the three conditions among each participant. Although the red and black circles are displayed in the movie clip, each participant was asked to evaluate only the movement of the *red* circle. In each trial, the participants answered the three questions regarding "Animacy", "Intention", and "Contingency" on a five-point scale (1. Do not feel, 2. Do not feel much, 3. Neither, 4.

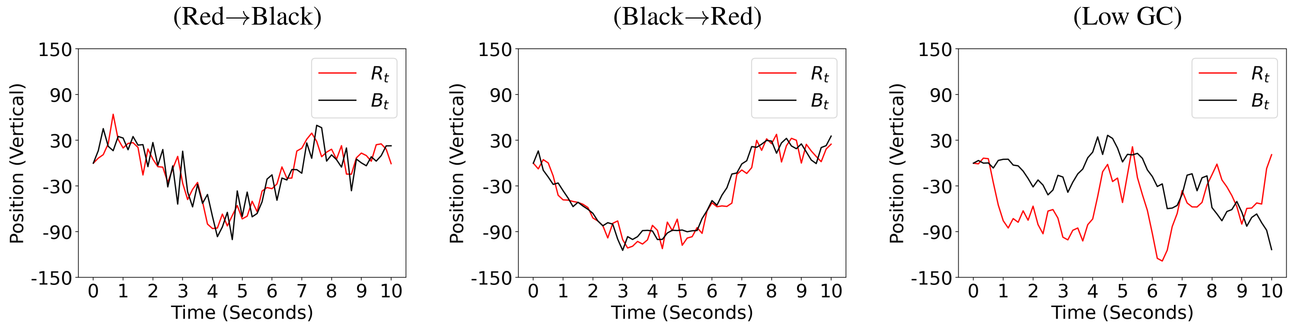

Fig. 2. Examples of a pair of movements (experimental stimulus) for the three conditions

Feel a little , 5. Feel). The actual question text is as follows (translated from the original text in Japanese).

- Animacy: Did you feel that the movement of the red circle was like a living thing?
- Intention: Did you feel that the red circle was moving with purpose?
- Contingency: Did you feel that the movement of the red circle responded to the movement of the black circle?

3 ANALYSIS

Each participant evaluated 10 movie clips generated by a VAR model with each type of coefficient matrix. We adopted the average of the 10 ratings as the participant's rating for the movements generated with the type of coefficient matrix. Since the animacy, intention, and contingency ratings across participants of the two types of the coefficient matrix in the same condition generally showed similar values (e.g., the pair A_{HL} and A_{H0} for the Red→Black condition), we treated these pairs to compare among the three conditions. Although in this preliminary study we also investigated the movements generated from the coefficient matrices whose non-diagonal components are negative, we later noticed that those matrices may produce unexpectedly unnatural movements. To avoid including other factors than Granger causality considered to affect the rating, we excluded these from the analysis of this paper [1]. Eventually, the analyzed data for each condition consists of the 14 average ratings from the 7 participants for the movements with the two types of the coefficient matrix.

4 RESULTS

The results of the mean values (M) and standard deviations (SD) of ratings are depicted in Fig. 3. The average of rating in animacy is $M = 3.529$ and $SD = .487$ for the Red→Black, $M = 2.921$ and $SD = .813$ for the Black→Red, $M = 3.564$ and $SD = .410$ for the Low GC condition. The average of rating in intention is $M = 3.614$ and $SD = .630$ for the Red→Black, $M = 3.543$ and $SD = .951$ for the Black→Red, $M = 2.743$ and $SD = .851$ for the Low GC condition. The average of rating in contingency is $M = 4.036$ and $SD = .760$ for the Red→Black, $M = 4.157$ and $SD = .816$ for the Black→Red, $M = 2.250$ and $SD = .573$ for the Low GC condition.

There was a significant difference between the animacy ratings of the Red→Black and the Black→Red condition ($t(13) = 2.272,\ p = .032$). There was also a significant difference between the animacy ratings of the Black→Red and the Low GC condition ($t(13) = 2.571,\ p = .017$).

There was a significant difference between the intention ratings of the Red→Black and the Low GC condition ($t(13) = -3.287,\ p = .003$). There was also a significant difference between the intention ratings of the Black→Red and the Low GC condition ($t(13) = -2.596,\ p = .015$).

There was a significant difference between the contingency ratings of the Red→Black and Low GC condition ($t(13) = -6.033,\ p = 2.313 \times 10^{-6}$). There was also a significant difference between the contingency ratings of the Black→Red and the Low GC condition ($t(13) = -5.951,\ p = 2.870 \times 10^{-6}$).

We performed correlation analysis among the ratings of animacy, intention, and contingency for each of the three conditions. The results are summarized in Table 1–4.

Table 1 shows the results of the correlation analysis of animacy, intention, and contingency ratings over all the three conditions. [3] reported a positive correlation between animacy and contingency. However, we found no significant correlation ($r = -.111,\ p = .482$). We found a significantly positive correlation between intention and contingency ($r = .647,\ p = 3.598 \times 10^{-6}$).

Table 2 shows the results of the correlation analysis for the

[1]The ratings for the movements with these matrices were quite similar to those of A_{LL} and A_{00}.

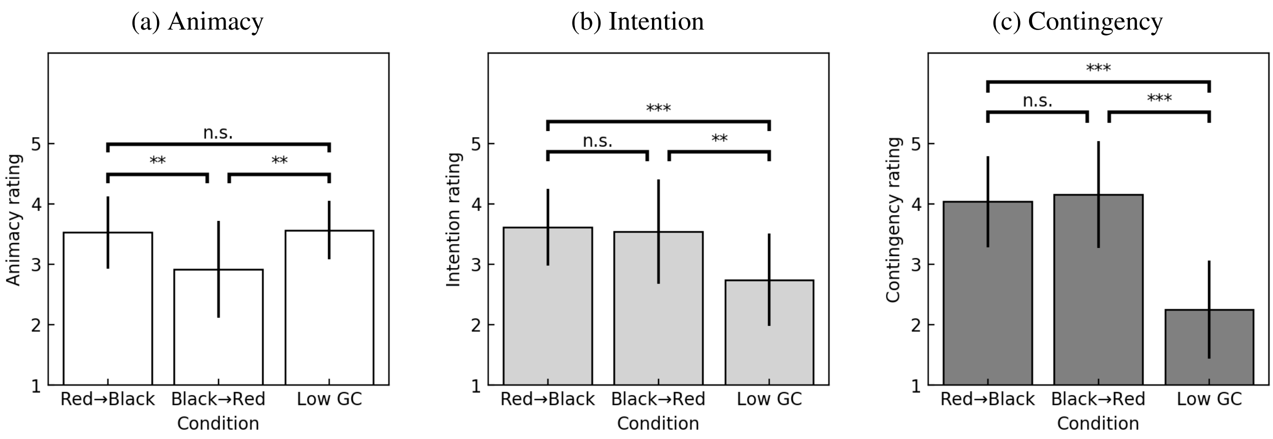

Fig. 3. Ratings for Animacy, Intention, and Contingency questions. Results of Welch's t test were indicated in the figures. n.s.: not significant, *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$.

Red→Black condition. We found significantly positive correlation between animacy and contingency ($r = .683,\ p = .007$), animacy and intention ($r = .628,\ p = .016$), and intention and contingency ($r = .717,\ p = .004$).

Table 3 shows the results of the correlation analysis for the Black→Red condition. The correlation between animacy and contingency ($r = -.024,\ p = .936$), animacy and intention ($r = -.292,\ p = .310$), and intention and contingency ($r = .441,\ p = .114$) were not significant.

Table 4 shows the correlation analysis for the Low GC condition. We did not find any significantly correlation in all the pairs, i.e., animacy and contingency ($r = -.235,\ p = .419$), animacy and intention ($r = .215,\ p = .461$), and intention and contingency ($r = .424,\ p = .131$).

5 DISCUSSION

In this study, we hypothesized that the degree to which one movement contributes to the predictability of the other movement characterizes the degree of animacy. The results of this preliminary study relating to animacy and contingency are summarized in Table 5. The rows $G_{R\to B}$ and $G_{B\to R}$ are regarding the experimental stimuli, the rows "Animacy rating" and "Contingency rating" are based on the mean ratings in Fig. 3, and the row "Significant correlation" is based on the correlations in Table 2–4. From Table 5, when $G_{B\to R}$ is high, the average animacy rating perceived by the observer to R is low, and when $G_{R\to B}$ is low, the average animacy rating is high. This suggests that there is a negative correlation between $G_{B\to R}$ and the average animacy rating that the observer perceives about R. For Granger causality, when $G_{B\to R}$ is low, the degree of prediction of R's movement from B's movement is small, and conversely, when $G_{B\to R}$ is high, the degree of prediction of R's movement from B's movement is large. Considering this, it was suggested that the degree to which the movement of other objects contributes to the predictability of the movement of the observed object characterizes the animacy perceived by the observer.

We treated Granger causality as the degree to which one movement contributes to the predictability of the other movement. We questioned whether Granger causality can be a measure of temporal contingency. From Table 5, the participants highly rated the average contingency rating in the Red→Black and Black→Red condition, which have high Granger causality in one direction. No significant difference was observed between these conditions. On the other hand, in the Low GC condition, the participants rated the average contingency rating low. From these results, it was found that the observer perceives contingency if $G_{R\to B}$ or $G_{B\to R}$ was high regardless of the direction of Granger causality. Therefore, we think that some contingency can be measured by Granger causality.

In this paper, we asked about the relationship between animacy and contingency. From Table 5, it was found that there is a significant correlation between animacy and contingency only when $G_{B\to R}$ is low and either $G_{R\to B}$ or $G_{B\to R}$ is high. No correlation was found between animacy and contingency in the Black→Red condition with high contingency. From these results, it is considered that there is a significant correlation between animacy and contingency when the average contingency rating is high and the average animacy rating is high.

Based on the above, when $G_{B\to R}$ is low, the average animacy rating that the observer perceived in R was high, so

Table 1. Correlation analysis among ratings for Animacy, Intention, Contingency. Pearson's correlation coefficient r with the statistical test against the null hypothesis $r = 0$. Results for linear regression $y = \alpha\, x + \beta$ were also shown. n.s.: not significant, *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$.

y	x	α	β	r	p
Animacy	Contingency	-0.06	3.56	-0.11	n.s.
Animacy	Intention	0.15	2.83	0.18	n.s.
Intention	Contingency	0.46	1.71	0.65	***

Table 2. Correlation analysis for the Red→Black condition. n.s.: not significant, *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$.

y	x	α	β	r	p
Animacy	Contingency	0.54	1.34	0.68	***
Animacy	Intention	0.59	1.37	0.63	**
Intention	Contingency	0.60	1.19	0.72	***

Table 3. Correlation analysis for the Black→Red condition. n.s.: not significant, *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$.

y	x	α	β	r	p
Animacy	Contingency	-0.02	3.01	-0.02	n.s.
Animacy	Intention	0.27	1.95	0.30	n.s.
Intention	Contingency	0.43	1.75	0.44	n.s.

Table 4. Correlation analysis for the Low GC condition. n.s.: not significant, *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$.

y	x	α	β	r	p
Animacy	Contingency	-0.14	3.87	-0.23	n.s.
Animacy	Intention	0.13	3.19	0.21	n.s.
Intention	Contingency	0.40	1.84	0.42	n.s.

it is considered that when there are few clues to predict the movement of the observation target R, the observer attributes the explanation of the movement to the observation target R and perceives a high degree of animacy. Regarding this, even in the case of $G_{R\to B}$ where the average animacy rating perceived by the observer to R was low in the preliminary experiment, it is considered that the average animacy rating perceived by the observer to R becomes higher by hiding the movement of B, which is useful for predicting the movement of R. In our future work, to test this refined hypothesis, we are planning to conduct a new experiment including the case where the black circle B is not displayed in the movie clips and only the red circle R is displayed (apparently an isolated movement of the single object).

Table 5. Summary of our preliminary experiment results related to our hypothesis and animacy and contingency.

	Red→Black	Black→Red	Low GC
$G_{R\to B}$	High	Low	Low
$G_{B\to R}$	Low	High	Low
Animacy rating	High	Low	High
Contingency rating	High	High	Low
Significant correlation	Yes	No	No

REFERENCES

[1] Heider F, Simmel M (1994), An experimental study of apparent behavior. *The American Journal of Psychology*, 57(2):243–259

[2] Scholl BJ, Tremoulet PD (2000), Perceptual causality and animacy. *Trends in Cognitive Sciences*, 4(8):299–308

[3] Bassili JN (1976), Temporal and spatial contingencies in the perception of social events. *Journal of Personality and Social Psychology*, 33(6):680–685

[4] Dittrich WH, Lea SEG (1994), Visual perception of intentional motion. *Perception*, 23(3):253–268

[5] Tremoulet PD, Feldman J (2000), Perception of animacy from the motion of a single object. *Perception*, 29(8): 943–951

[6] Ueda K (2013), Animacy perception: Lifelikeness that we perceive from artifacts. *Journal of the Robotics Society of Japan*, 31(9):833–835

[7] Takahashi K, Watanabe K (2015), Synchronous motion modulates animacy perception. *Journal of Vision*, 15(8): 1–17

[8] Granger CWJ (1969), Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438