Part XIV
Free Talk Session 8

A NEW PROBABILISTIC APPROACH TO DESCRIBING THE STREAM OF DECISION-MAKING EVENTS NEAR PERCEPTION THRESHOLD

Ihor Lubashevsky and Kosuke Hijikata

University of Aizu

Ikki-machi, Aizu-Wakamatsu, Fukushima 965-8580, Japan

IL: `<i-lubash@u-aizu.ac.jp>`

Abstract

Within the paradigm of human intermittent control over unstable systems human behavior admits the interpretation as a sequence of point-like moments when the operator makes decision on activating or halting the control. These decision-making events are assumed to be governed by the information about the state of system under control which the operator accumulates continuously. In the present work we propose the concept of reinforcement learning with decision inertia (the status quo bias) that opens a gate to applying the formalism of reinforcement learning to describing human intermittent control. The basic feature of such reinforcement learning is that human behavior in a sequence of selecting available options exhibits quasi-continuous dynamics. Numerical simulation based on a fairly simple model demonstrates that the proposed formalism does possess the required properties of quasi-continuous behavior.

Model Background

Nowadays human intermittent control has become a dominant paradigm of describing human behavior in controlling various unstable systems (for a review see, e.g., Loram et al., 2011). It assumes a human operator to activate and, then, halt the control alternately instead of keeping it active continuously though the whole course of actions. The event-driven mechanism of human intermittent control is now accepted to be the main mechanism governing such human actions. According to the event-driving scenario the operator activates the control when the state of controlled system deviates from the desired one substantially and halts the control when this difference becomes rather small and the operator cannot recognize or affect it with a required accuracy. The threshold model of control activation is rather popular in mathematical description of human intermittent control. It assumes an operator to activate the control when the difference between the current and desired states exceeds a certain threshold and to halt it in the opposite case. However, recently (Zgonnikov et al., 2014) based on experiments on balancing overdamped pendulums, we proposed a novel concept of control activation. It is called the noise-induced activation and implies the control activation to be probabilistic in nature. In particular, the noise-induced activation represents a certain interplay between the necessity of starting the control process and the possibility of postponing it until this necessity becomes absolutely clear.

The formalism of reinforcement learning allows for such aspects in adopting to changing environment and finding an optimal choice via the trial-error strategy (see, e.g., Frank and Claus, 2006). Unfortunately, this approach on its own cannot be applied directly to describing human intermittent control. The matter is that on time scales much larger than the elementary step the produced decision-making process is strongly discontinuous. Conversely, the desired decision-making process must demonstrate the choice of the same option on scales comprising many elementary steps of decision-making. This

human behavior can be captured turning to the notion of decision inertia. Decision inertia can arise via some mechanism by which a human operator (agent) prefers not to change a chosen option for a relatively long time interval comprising a large number of elementary steps. In particular, it is the status quo bias well-known in economics (Samuelson and Zeckhauser, 1988). Recently Akaishi et al. (2014) found this type phenomenon in direct experiments with human decision-making in a situation similar to human intermittent control.

The introduction of decision inertia opens a gate to describing human intermittent control as (*i*) continuous accumulation of information about the state of controlled system and (*ii*) the sequence of events when the subject makes decision on changing the control state in response to the accumulated information.

The goal of presented work is developing a reinforcement learning model that allows for decision inertia. Actually, the present work proposes a fairly simple model for reinforcement learning with status quo bias and demonstrates that it does meet the desired property of quasi-continuous dynamics. The gist of this model is the concept of multichannel information processing used previously (Zgonnikov and Lubashevsky, 2014) to describe human learning affected by novelty-seeking (intrinsic motivation).

Model

An agent is assumed to make repeated choice between finite number of options $i \in [1, N]$ and accumulate the information about them, which in turn affects their choice probability. The processing of information is considered to be implemented through two independent mental channels. One of them is the deliberate analysis of obtained rewards. The other is irrational and exhibits status quo bias which can be explained turning to a reason like this: "If just now I have chosen a new option it is not reasonable to choose another option immediately, it could be better to wait some time until the quality of the chosen option becomes clear." The two channels interact via their cumulative effect on the option selection (Fig. 1).
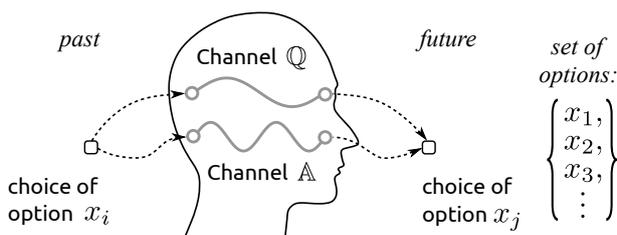

Fig. 1. Two channels of information processing: the channel $\mathbb{Q}$ via which the information about the rewards is processed and the channel $\mathbb{A}$ dealing with the information not directly related to the rewards.

Within the channel $\mathbb{Q}$ each option i is related to the corresponding reward r_i the agent receives each time it has chosen the given option. The preference of choosing option i is quantified by a value q_i; the value q_i results from the experience the agent gains each time it chooses the corresponding option. Namely, at every time step $t_k = k\tau$ $(k \in \mathbb{N})$ the preference values q_i are updated with currently received rewards $r_{q,i}$ and, besides, subjected to the memory loss:

$$q_i(t_{k+1}) = q_i(t_k) + \delta_{ii_k} r_{q,i} - \epsilon_q q_i(t_k) \,, \tag{1}$$

where the index i_k points to the option chosen at the given time step t_k and the value $0 < \epsilon_q < 1$ quantifies the agent memory capacity $\propto 1/\epsilon_q$. The second term on the right-

hand side of Eq. (1) containing the Kronecker delta δ_{ii_k} ($\delta_{ii_k} = 1$ if $i = i_k$ and 0 if $i \neq i_k$) reflects our assumption that only the preference value of chosen options is increased by the obtained rewards. In the present analysis the effects of forgone payoffs as well as different evaluation of rarely or frequently chosen options by humans can be omitted. Below time will be measured in units of the elementary time step τ between the neighboring decision-making events.

The channel $\mathbb{A}$ allows for the effect of status quo bias on choosing the same option at the next time moment. This effect is taken into account through some additional preference value $\Delta_i > 0$ to choose the same option i. This value decreases if the same option i has been chosen again, otherwise, it increases with a certain saturation, namely:

$$\Delta_i(t_k) = \Delta_i(t_{k-1})(1 - \epsilon_a) + (1 - \delta_{ii_k}\delta_{ii_{k-1}})r_{a,i}, \tag{2}$$

where the coefficient $r_{a,i}$ specifies the restoration of the status qua bias measure Δ_i for the option i when it has not been chosen. The given model imitates the human preference to wait for a certain time $T_a \sim 1/\epsilon_a$ to recognize the quality of the made choice.

The proposed model assumes the channels $\mathbb{Q}$ and $\mathbb{A}$ to be mutually independent, which implies that the resulting probability of choosing any option i is just the product of their individual contributions. Besides, we consider that only the difference between the preference values has physical sense. Under such conditions the probability of choosing option i is

$$p_i = \frac{1}{Z} \exp\left\{ \beta \left[q_i + \Delta_i \delta_{ii_{k-1}} \right] \right\}, \tag{3}$$

where the normalization coefficient Z is determined by the expression

$$Z = \sum_{i=1}^{N} \exp\left\{ \beta \left[q_i + \Delta_i \delta_{ii_{k-1}} \right] \right\}. \tag{4}$$

Expression (3) is the mathematical implementation of the concept of fuzzy perception threshold for the given system; for example, the choice between two options becomes practically equiprobable when the difference in their cumulative preference quantities is comparable with the perception threshold $1/\beta$.

It should be noted that the obtained expressions admit the interpretation in terms of transitions to another state j provided currently the state i have been chosen. Namely, the probability of this transition can be written as

$$\mathcal{P}_{i \to j} = \frac{1}{Z_i'} \exp\left\{ \beta \left[q_j - q_i \right] \right\},$$

and the probability of choosing the same option i given by the expression

$$\mathcal{P}_{i \to i} = \frac{1}{Z_i'} \exp\left\{ \beta \Delta_i \right\}.$$

Here the normalization factor Z_i' now depending on the current state i is related to the previous one, Ex. (4), as $Z_i' = Z \cdot \exp\{\beta q_i\}$.

Results and Conclusion

To illustrate the characteristic features of the proposed model the system with two options $i = 1, 2$ symmetric in properties ($r_{q,1} = r_{q_2} = r_q$ and $r_{a,1} = r_{a_2} = r_a$) was studied

numerically. In the numerical simulation time is measured in units of τ and the preference values are measured in units of the perception threshold $1/\beta$. The particular values of system parameter were set equal to $T_q = 1/\epsilon_q = 30$, $T_a = 1/\epsilon_a = 30$, and $Q_{\max} = r_q T_q = 2.5$, $D_{\max} = r_a T_a = 10$. In this case the agent actions should exhibit substantial status quo bias.

Figure 2 demonstrates the obtained results. We see that after selecting an option the agent keeps the made choice for a relatively long time. As a consequence the shown time patterns of difference in the option priorities $(q_1 - q_2)/2$ and, e.g., the status quo bias measure Δ_2 do admit the interpretation as piece-wise continuous dynamics. Finally, the distribution of option life time—the time during which the same option is chosen continuously—mainly comprises values much larger than the elementary time step τ. In particular, its maximum for the analyzed system is about $\tau_{\text{life}} > 40$.
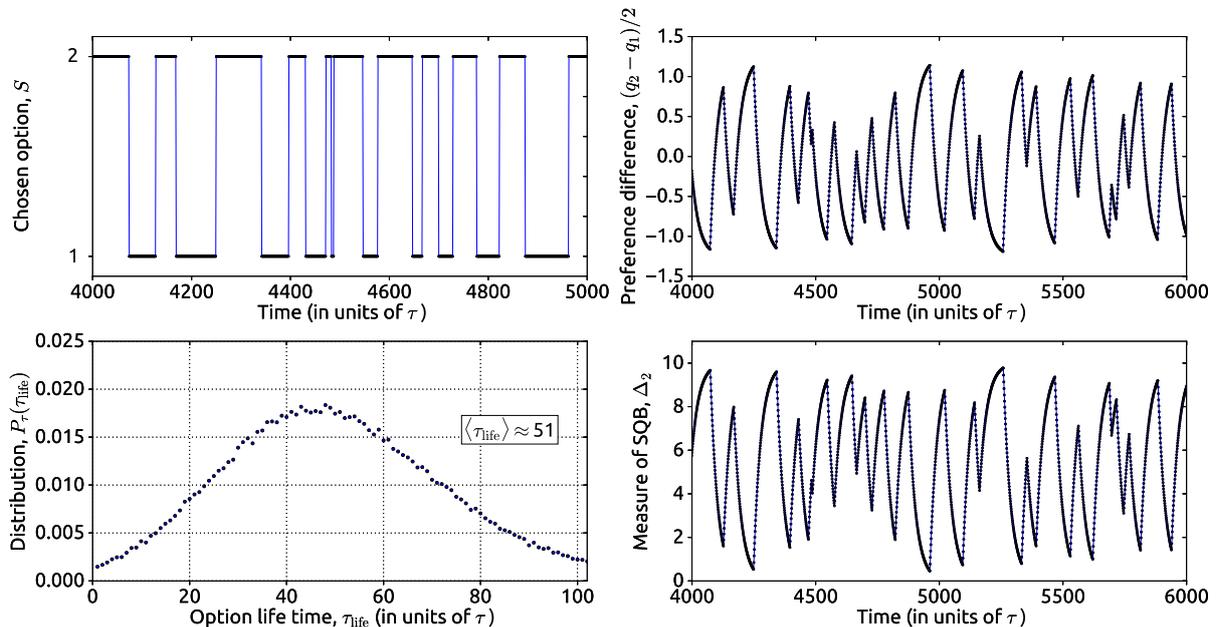

Fig. 2. Time patterns and statistical properties of reinforcement learning dynamics for the system with two equivalent options whose choice is affected by the status quo bias.

In summary, the proposed formalism of reinforcement learning with decision inertia opens a gate toward constructing theories of human intermittent control dealing with (*i*) intentional actions of operators activating or halting the control depending on the system state and (*ii*) allowing for the bounded capacity of human cognition. The main premise of this formalism is the existence of two independent channels of the information processing in the human mind. They are

- a channel responsible for the deliberate analysis of the obtained rewards and ordering them according to their preference; it is the standard component of all the models for reinforcement learning;

- a channel stimulating a subject not to change the current choice and keep it for a certain time interval.

As demonstrated, under such conditions the reinforcement learning with decision inertia can generate time patterns of subject's actions admitting the required smoothness of the

discontinuous stochastic process mimicking, in particular, human actions in controlling unstable mechanical objects.

References

Akaishi, R., Umeda, K., Nagase, A. and Sakai, K.: 2014, Autonomous Mechanism of Internal Choice Estimate Underlies Decision Inertia, *Neuron* **81**(1), 195–206.

Frank, M. J. and Claus, E. D.: 2006, Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal., *Psychological Review* **113**(2), 300–326.

Loram, I., Gollee, H., Lakie, M. and Gawthrop, P.: 2011, Human control of an inverted pendulum: is continuous control necessary? Is intermittent control effective? Is intermittent control physiological?, *The Journal of Physiology* **589**(2), 307–324.

Samuelson, W. and Zeckhauser, R.: 1988, Status quo bias in decision making, *Journal of Risk and Uncertainty* **1**(1), 7–59.

Zgonnikov, A. and Lubashevsky, I.: 2014, Unstable Dynamics of Adaptation in Unknown Environment due to Novelty Seeking, *Advances in Complex Systems* **17**(3 & 4), 1450013 (17 pages).

Zgonnikov, A., Lubashevsky, I., Kanemoto, S., Miyazawa, T. and Suzuki, T.: 2014, To react or not to react? Intrinsic stochasticity of human control in virtual stick balancing, *Journal of The Royal Society Interface* **11**, 20140636.

AN OBJECT-FREE ALGORITHM FOR GENERATING BIOLOGICAL MOTION STIMULI FROM NATURAL MOVIES

Wataru Suzuki and Noritaka Ichinohe

Department of Ultrastructural Research, National Institute of Neuroscience
National Center of Neurology and Psychiatry
4-1-1 Ogawa-Higashi, Kodaira, Tokyo 187-8502, Japan;
Ichinohe Neural System Group, Laboratory for Molecular Analysis of Higher Brain
Functions, RIKEN Brain Science Institute, RIKEN
2-1 Hirosawa, Wako, Saitama 351-0198, Japan

Hiroshige Takeichi

Computational Engineering Applications Unit
Advanced Center for Computing and Communication (ACCC), RIKEN
2-1 Hirosawa, Wako, Saitama 351-0198, Japan;
Department of Developmental Disorders, National Institute of Mental Health
National Center of Neurology and Psychiatry
4-1-1 Ogawa-Higashi, Kodaira, Tokyo 187-8553, Japan

Atsuko Gunji

Department of Developmental Disorders, National Institute of Mental Health
National Center of Neurology and Psychiatry
4-1-1 Ogawa-Higashi, Kodaira, Tokyo 187-8553, Japan;
Department of Advanced Neuroimaging, Integrative Brain Imaging Center
National Center of Neurology and Psychiatry
4-1-1 Ogawa-Higashi, Kodaira, Tokyo 187-8551, Japan;
College of Education, Yokohama National University
79-2 Tokiwadai, Hodogaya, Yokohama 240-8501, Japan

Masumi Inagaki

Department of Developmental Disorders, National Institute of Mental Health
National Center of Neurology and Psychiatry
4-1-1 Ogawa-Higashi, Kodaira, Tokyo 187-8553, Japan

Abstract

We have developed an object-free algorithm for extracting motion information from natural movies by estimating the dynamics of the local normal vectors of the image intensity projected onto the x-y plane. Here, we used the algorithm to generate random dot movies from 60 movies of moving animals. We then tested whether six healthy adult male observers could identify which of two similar random dot movies matched the original movie. The participants made forced-choice judgments on the correspondence between the original movie and the two random dot movies, which were presented normally, temporally reversed, and vertically inverted. The participants showed better-than-chance performance. The results provide a basis for broader application of the algorithm, which could also be used to expand the paradigms used in biological motion studies, most of which use point light walkers.

Motion in the visual environment provides a rich source of information for biological systems (Nakayama, 1985). Biological motion perception (Johansson, 1973) is a good

example. When several lights are attached to a person's body as markers and the person moves in total darkness, the time series of the images of the lighting markers, referred to as a point light walker, provides rich information about the age, gender, and even emotions of the marked person to human observers. This type of stimulus is called a biological motion stimulus. The perception of biological motion stimuli seems to be associated with social functions such as face- or voice-specific responses, which can be impaired in children with developmental disorders (e.g., Annaz, et al., 2012; Gunji et al., 2013; Hirai et al., 2014; Aglieri, et al., 2017).

Most studies of biological motion perception use point light walker stimuli because they enable robust extraction of rich information from complex motion. However, it is not easy to generate a comparable stimulus using a non-human animal or inanimate object as the actor. The use of more general or broader classes of stimuli is favorable for clarification of the neural substrates underlying intact and impaired biological motion perception in humans as well as animal models.

Thus, we have developed a novel algorithm for extracting motion information from natural movies to generate random dot movies (Suzuki et al., in press). In the previous study, we described the algorithm and a neurophysiological experiment on common marmoset (Callithrix jacchus). In this study, we conducted a psychophysical experiment to demonstrate a further application of the algorithm.

Methods

Ethics Statement. This study was approved in advance by the Ethical Committee of the National Institute of Neuroscience and Psychiatry (Kodaira, Japan). All participants gave written informed consent before participating in the experiment.

Participants. Six healthy men (mean age 38) with normal or corrected-to-normal visual acuity volunteered to participate in the experiment.

Stimuli. Details of stimulus generation are described by Suzuki et al. (in press). We made 60 random dot movies from 60 animal movies and divided them into 30 pairs with similar lower-order-motion properties. The pairs were based on six indices that were calculated for each random dot movie: the integrated moving distances in the x- and y-coordinates for all dots, the centroids of the moving distances for all dots during their appearance in x- and y-coordinates, and the distribution of the dots with a large moving distance around the centroid in the x- and y-coordinates. The six indices for all 60 random dot movies were normalized by calculating their z-scores. The Euclidean distances $D_{u,v}$ between two movies, u and v, were used to create pairs. Each pair of random dot movies was selected such that $D_{u,v}$ was minimized. Namely, the first pair comprised the movies with the lowest $D_{u,v}$, the second pair comprised the movies with the next lowest $D_{u,v}$, and so forth until all 30 pairs were formed.

Procedure. The participants sat 57 cm in front of a 19-inch (48 cm) CRT monitor. Head movement was minimized by a chinrest. The spatial resolution of the CRT monitor was 600 pixels in height and 800 pixels in width, and the refresh rate was 60 Hz. The size of the stimulus was 28.5 × 20.5 cm. The task was controlled using MATLAB (Mathworks, Natick, MA, USA) with the Psychophysics Toolbox. The experiment was carried out in a dark room. The 30 pairs of random dot movies were divided into three sets—stimulus

sets 1, 2, and 3—each consisting of 10 pairs. There were three sessions. The participants received one stimulus set in each session. Participants 1 and 2 received stimulus sets 1, 2, and 3, in Sessions 1, 2, and 3, respectively. Participants 3 and 4 received stimulus sets 2, 3, and 1, in Sessions 1, 2, and 3, respectively. Participants 5 and 6 received stimulus sets 3, 1, and 2, in Sessions 1, 2, and 3, respectively.

Session 1. The participants were asked to match a sample animal movie with a corresponding random dot movie in a temporal two-alternative forced-choice task. A trial started upon pressing the keyboard. A sample animal movie appeared for 1 s. After a 2-s delay in which a random-dot static mask was displayed, the first random dot movie appeared for 1 s. After another 2-s delay with the random dot static mask displayed, the second random dot movie appeared for 1 s. One of the random dot movies corresponded to the animal movie and the other corresponded to the paired movie. The participant pressed button "1" or "2" on the keyboard, depending on whether he judged the dot motion in the first or second random dot movie as corresponding to the animal motion in the sample animal movie. The trial ended when the participant responded. There were 10 trials in each session and each pair appeared only once, with the order of presentation pseudorandomized across participants. The participants were instructed to fixate on a cross presented at the center of the screen throughout the trial.

Session 2. The procedure was the same as that in Session 1, except that the frames of the two random dot movies were shown in reverse order (time-reversal condition). The participants were informed that the random dot movies were reversed.

Session 3. The procedure was the same as that in Session 1, except that the two random dot movies were spatially inverted so that they were upside-down (inverted condition). The participants were informed that the random dot movies were inverted.

Results

Session 1. We examined whether it was possible to match the original animal movie with the random dot movie using the motion information extracted by our algorithm. The participants' performance was $93.3 \pm 8.17\%$ (mean $\pm$ standard deviation, $n = 6$), which was significantly different from chance (50%) ($p < 0.05$, two-tailed binominal test).

Session 2. We examined whether it was possible to match the original animal movie with the random dot movie when it was presented in reverse order. The performance for the time-reversal condition was $93.3 \pm 8.17\%$ (mean $\pm$ standard deviation, $n = 6$), which was significantly different from chance (50%) ($p < 0.05$, two-tailed binominal test)

Session 3. We examined whether it was possible to match the original animal movie with when the random dot movie when it was presented upside-down. The performance for the inverted condition was $88.3 \pm 11.7\%$ (mean $\pm$ standard deviation, $n = 6$), which was significantly different from chance (50%) ($p < 0.05$, two-tailed binominal test).

Discussion

In this article, we describe our new algorithm for extracting motion information from natural movies. Our experimental results confirmed that human observers were able to match the original movie with the extracted motion presented as a random dot movie. Their performance was not impaired under the time-reversal condition, in which the velocity of the random dot motion was reversed, or the inverted condition, in which the spatial distribution of the random dot motion differed from that in the original movies; therefore, the results did not seem to reflect the early visual processing stage of motion detection (DeAngelis et al., 1995). Nor did the results exhibit the deterioration in performance observed under time-reversal and inverted conditions using point light walker stimuli (Sumi, 1984; Pavlova and Sokolov, 2000).

It could be argued that the matching performance in the psychophysical experiment was generally good simply because the motion patterns in the paired movies was sufficiently dissimilar that any motion extraction algorithm could have generated discriminable random dot movies; in other words, the high matching performance might not reflect the high performance of our algorithm in extracting object motion. However, it should be emphasized that the similarity indices for the pairs were calculated from the generated random dot movies, not the original animal movies, and that the pairs with high similarity indices could still be discriminated.

The proposed motion extraction algorithm can be applied to natural movies to generate visual stimuli. The algorithm can provide a variety of complex visual motion stimuli because it does not require parameterized geometric shapes. By applying the algorithm to various natural movies, it is possible to generate random dot motion stimuli for humans, non-human animals, and man-made moving objects. It is also possible to dissociate dynamic and static information because the visual motion stimuli do not contain static shape information and the objects cannot be identified unless they move. This algorithm should be useful for clarifying the critical factors and neural mechanisms underlying biological motion perception and social cognition through the visual modality.

In conclusion, we have described a psychophysical application of a new algorithm. The algorithm can be used to extract motion information from any natural movie and generate a random dot movie. The global motion of an object emerge from the computation of normalized local normal vectors. The participants showed better-than-chance performance, providing a basis for broader applications of the algorithm. Our technique could also expand the paradigms used in biological motion studies, most of which use point light walkers.

Acknowledgements

References

Aglieri, V., Watson, R., Pernet, C., Latinus, M., Garrido, L., Belin, P. (2017). The Glasgow Voice Memory Test: Assessing the ability to memorize and recognize unfamiliar voices. *Behavior Research Methods* 49, 97-110.

Annaz, D., Campbell, R., Coleman, M., Milne, E., Swettenham, J. (2012). Young children with autism spectrum disorder do not preferentially attend to biological motion. *Journal of Autism and Developmental Disorders,* 42, 401-408.

DeAngelis, G.C., Ohzawa, I., Freeman, R.D. (1995). Receptive-field dynamics in the central visual pathways. *Trends in Neurosciences,* 18, 451-458.

Gunji, A., Goto, T., Kita, Y., Sakuma, R., Kokubo, N., Koike, T., Sakihara, K., Kaga, M., Inagaki, M. (2013). Facial identity recognition in children with autism spectrum disorders revealed by P300 analysis: a preliminary study. *Brain & Development,* 35, 293-298.

Hirai, M., Gunji, A., Inoue, Y., Kita, Y., Hayashi, T., Nishimaki, K., Nakamura, M., Kakigi, R., Inagaki, M. (2014). Differential electrophysiological responses to biological motion in children and adults with and without autism spectrum disorders. *Research in Autism Spectrum Disorders,* 8, 1623-1634.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics,* 14, 201-211.

Nakayama K. (1985). Biological image motion processing: a review. *Vision Research,* 25, 625-660.

Pavlova, M. and Sokolov, A. (2000). Orientation specificity in biological motion perception. *Perception & Psychophysics,* 62, 889-899.

Sumi, S. (1984). Upside-down presentation of the Johansson moving light-spot pattern. *Perception,* 13, 283-286.

Suzuki, W., Ichinohe, N., Tani, T., Hayami, T., Miyakawa, N., Watanabe, S., Takeichi, H. (in press). Novel method of extracting motion from natural movies. *Journal of Neuroscience Methods.* http://dx.doi.org/10.1016/j.jneumeth.2017.08.006