1H3-OS-02a-4

# サンバ運動の力学的位相構造による不変的表現
## Dynamical Invariant Representations for Samba Performances

日高　昇平[*1]
Shohei Hidaka

藤波　努[*2]
Tsutomu Fujinami

[*1] 北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology

[*2] 北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology

人の運動を表現する基礎的な特徴量は、多くの場合、課題固有的にアドホックに設定され、その妥当性には議論の余地がある。本研究では、物理的な特徴量(速度、間接角など)に代わり、その高次空間に埋め込まれた力学系の位相構造を基礎として、運動を記述する手法を提案する。本手法をサンバ演奏時の運動データに適用する事で、演奏者、演奏速度などによらないサンバ固有の不変構造を抽出できる事を示した。

## 1. Introduction

Recognition of motion is vitally important to any animal. Detection of another animal, whether predator or prey, or a conspecific, and subsequent detailed identification of the other and how it may behave is essential to taking any emergent actions (Johnson, Bolhuis, & Horn, 1985). Not surprisingly, our visual system is highly specialized to recognize others' actions. How do we recognize bodily movements?

The past experimental literature has explored capacity of motion perception using point-light displays (Johansson, 1973) in which the point-lights attached in major joints are only visible in the dark background. The available information is point-wise kinematic motion in multiple body parts. Despite of the limited information, people can recognize identity (Troje, Westhoff, & Lavrov, 2005), gender (Kozlowski & Cutting, 1977; Troje, 2002), emotions (Pollick et al., 2001; Atkinson; 2009; Hobson & Lee, 1999), dynamics such as the weight of a lifted object (Bingham, 1987) of actions from point-light displays. Accumulating empirical studies on action perception have suggested that velocity and its higher order derivatives in single or multiple body parts characterize actions: duration of action (Pollick et al., 2001), velocity (DeMeijer, 1989), acceleration (force or the second order time derivatives) (Chang & Troje, 2008; 2009), jerk or the third order time derivatives (Cook, Saygin, Swain, & Blakemore, 2009), and pairwise counter-phase oscillation (Chang & Troje, 2008; 2009). Although conventional biomechanical quantities mentioned above have been traditionally promoted as the likely variables of motor control and action perception, there are doubts about their appropriateness for the task (Turvey, 1998).

Therefore, instead of such biomechanical quantities, we hypothesize that an appropriate set of representations for action recognition is "dynamical invariances" under smooth transformation. This hypothesis views motor control underlying human movements as a set of dynamical systems, that is, a sequence of interactions between elements involved in controlling movements such as body joints, muscles, neural systems, etc. The properties retained in dynamical systems for long term can be captured with invariant measures such as attractor dimension or Lyapnov exponent (Kantz & Schreiber, 1997).

We define a higher dimensional space, i.e., phase space, within which all possible combinations between elements involved in controlling movements can be found. An action is then defined as a trajectory on the space. Trajectories can be projected onto lower dimensional spaces, e.g., actual movements observable from outside. In our study, we collect motion data to reconstruct the dynamical systems by embedding the time series in a higher dimensional space. For graphical examples, Figure 1 illustrates *attractors*, or the state space which the system may take in the three theoretical dynamical systems, the Hennon map, Rossler system, and Lorenz system (Figure 1 (a-1), (a-2), (a-3)). A univariate time series (as imperfect observation of the system) is shown in Figure 1 (b) for each of these systems. Since the original systems live in two or more dimensions, these univariate time series do not have full information due to missing dimensions. Thus, we need to "reconstruct" the phase space instead of studying the degenerated patterns (Takens, 1981). By taking the time delay vector (e.g., $\{x(t), x(t+\delta)\}$), the topological nature of the phase space is reconstructed (Figure 1c). In Figure 1 (c-1)-(c-3), the time-delay embedding (a map from low to high dimensional space) successfully recovers similar topological structure shown in Figure 1 (a-1)-(a-3) only from the degenerated data Figure 1 (b-1)-(b-3). Although the original phase space is unknown for empirical bodily movements, we expect the intrinsic topological nature can be reconstructed in the same way as the theoretical dynamical systems (Figure 1 (b-4) and (c-4)). See Kantz and Schreiber (1997) for a detailed description of these procedures for nonlinear time series analysis.

To study dynamical invariances, we investigated topological similarities of motor coordination. The

連絡先：日高昇平，北陸先端科学技術大学院大学，石川県能美市旭台 1-1，0761-51-1717，shhidaka@jaist.ac.jp

rationale for the approach is found in observations such that one can mimic other's behaviors no matter how different their individual appearances. Topology abstracts over physical particulars such as distance, speed, etc., to extract some dynamical invariances independent of these physical properties. Specifically, we examined the dynamical properties of rhythmic movements for two main reasons. First, rhythmic movements are not just a period but with fluctuating accents, and this is expected to show complexity to some extent neither too simple nor too complex. Second, actions which an actor can maintain continuously and produce a substantial amount of datasets are necessary for characterization of dynamical invariances.

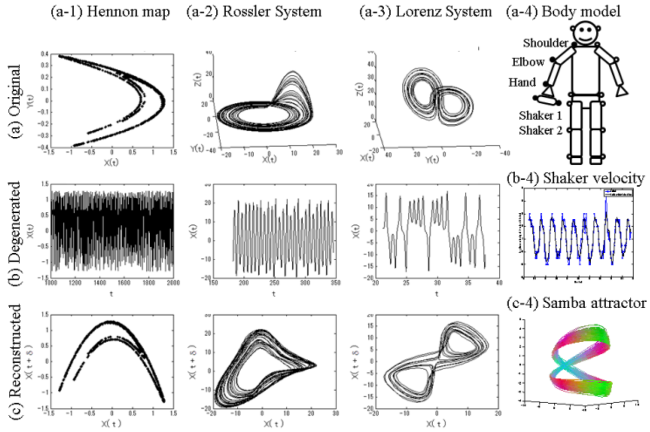


Figure 1: Phase space of (a-1) the Hennon map, (a-2) the Rossler system, (a-3) Lorenz System, and (a-4) the body model and attached markers (filled circles: analyzed, open circles: attached but not reported in this study). (b1-3) A univariate time series from the original phase space in (a1-3) (b-4) An x-axis phase of the Shaker 1 in the expert player in the 60-bpm trial (blue circles) with the estimated noise-reduced time series (black dots). (c1-4) The reconstructed phase space from the low dimensional observed time series in (b1-4).

## 2. Characterizing Complex Rhythmical Actions

The data was originally obtained in order to analyze the levels of expertise in the samba music plays (Yamamoto, Ishikawa, & Fujinami, 2006). The dataset consists of five players, and each player performed basic samba shaking actions in five different tempos (60, 75, 90, 105, and 120 beats per minute, and each trial lasted 97.4 seconds on average) by being cued with a metronome. While playing, three dimensional motions of 18 markers, attached on body parts and musical instruments, were recorded at 86.1Hz of sampling rate (Figure 1a-4). As well as the original study, here we aim to find the relationship between dynamical properties among bodily actions. For simplicity, we limited ourselves to analyze a subset of the original datasets, 3190 samples (74.1 seconds long) of four markers attached on right wrist, right elbow, and two sides of the musical instrument (shaker), having the right shoulder asreference point (Figure 1a-4). These were the essential parts of the samba actions making sounds directly, and we expected that dynamic coordination among them would be crucial to characterize the dynamical properties of the samba.

### 2.1 Preprocess and phase space reconstruction

In the analysis, after down-sampling the original data to 46.05 Hz, the first 250 samples (5.81 second long from the beginning of the recording) was excluded as initial setup of the actions, and 3250 samples (75.5 second long) of velocities were analyzed for each subject. In order to reduce measurement noise, for each movement of the markers, the local linear projective method was performed after phase space reconstruction of each time series on the 31 dimensional time delay space with 46 msec (i.e., $\{ t, t + \Delta t, t + 2\Delta t, \ldots, t + 30\Delta t \}$ where $\Delta t =$ 46 msec) (Takens, 1981). This technique is a nonparametric and unsupervised method which, in principle, reduces observation noise independent of the time series intrinsically generated from a nonlinear dynamical system. Figure 1 (b-4) shows the original data (open circles) and its noise-reduced data (filled circles) after applying the local linear projective method. Due to digitalization in the motion capture system, the original data only takes certain discrete values which may be potential source of observational noise in the measurement. As the result of the noise reduction, we obtained the 31 dimensional phase space of 3220 points for each coordinate of three dimensional positions of each marker movement in each subject and trial. An example of the reconstructed phase space is shown in Figure 1 (c-4).

### 2.2 Results

In order to see the rhythmic properties as phase shifts in repeating actions, we analyzed the temporal profiles of the velocities in right arm and wrist. Figure 2 shows the histogram of phase differences between body parts with the right shoulder as the reference point. Since the right elbow and wrist are the major body parts playing the shaker, their temporal structure was expected to reflect the rhythmic charateristics. However, not as expected, the five musicians showed quite different distributions in terms of phase shifts among body parts, even for the right wrist and elbow moved to play the shakers to the same tempo. The peaks found for the elbow and wrist are sharp for the musician A and D, compared with those found in the other musician. As for A, the peaks of the elbow and wrist come to the same phase, but the peak is less visible for the wrist. As for D, the peaks of the elbow and wrist come later than that of the shoulder. For the other musicians, no obvious feature is found. The frequency uniformly varying over the phase angle shows large fluctuations in arm movements for each musician. The histograms revealed both within-musician fluctuations and individual differences rather than similarity among actions. The results suggest that charaterization of the "same" action (i.e., playing to the samba rhythm) on the levels of physical properties may lead quite different patterns across subjects. Needless to say, changing physical properties such as tempos also directly changes phase differences. The level of physical properties is not sufficient for characterizing actions even if the major parameter of the actions (i.e., tempo) is well controlled.

Next, we analyzed the properties of actions by looking into the dynamical systems underneath body movements. A basic technique to characterize dynamical properties from an empirical time series is phase space reconstruction. A phase space reconstructed by time-delay embedding is visualied as a three dimensional subspace projection (Figure 1c-4). The phase space is originally a set of velocity vectors of the four markers including two sides of the shakers, right wrist and elbow. The trajectory on the reconstructed phase space shows an attractor or the state space the system may take. The phase space is 124 dimensional space consisting of 31 time-delay copies of the four dimensional time series. First we analyzed the dimensionality of the attractors as one of invariance for the dynamical system. It is formally measured by correlation dimensions (Kantz & Schreiber, 1997), and we found the correlation dimensions varying from 1.8 to 2.4 across five musicians and five conditions. These results suggested the state space of the samba rhythm is rather restricted on a low dimensional space.

Since the dimensions of the attractors are lower than three, it

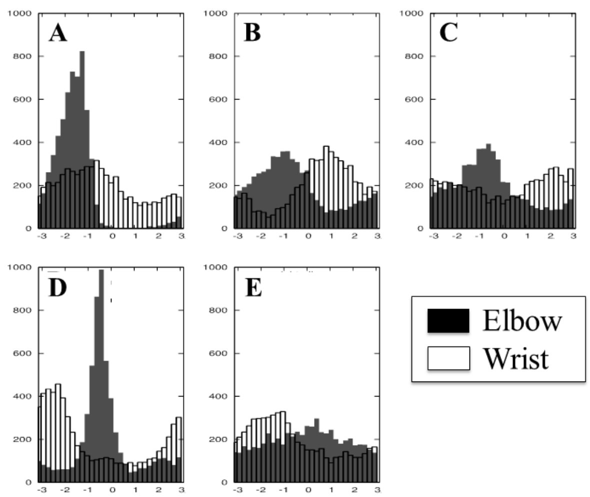


Figure 2: The distribution of phase shift of right elbow (gray) and right wrist (white).

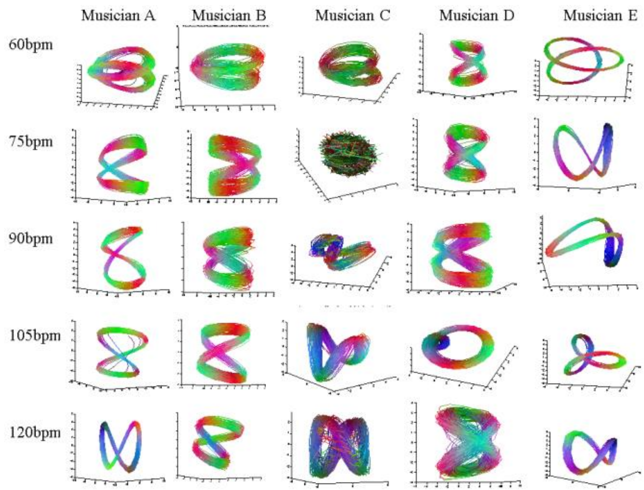


Figure 3: The reconstructed phase space embedding in three dimensional time delay space in Musician A-E playing at tempo 60, 75, 90, 105, and 120 BPM. The velocity of the trajectory on the three dimensional space is shown as RGB color code for its visibility.

allows us to visualize them on the three dimensional space without losing much information. Figure 3 shows the attractors estimated for all the five musicans on the five conditions. Visual inspection of the samba attractors grasps the gist of commonalities among the attractors. Consistently across most of the attractors, they share a similar shape of trajectories – a twisted double circle (which may appear different due to a specific visual angle of each attractor). These similar “shapes” of trajectories indicate that the topological nature of the attractors is similar. The result -- higher similarity between topolgoical properties of the state space -- is quite surprising with consideration to the individual differences on the physical level charateristics (Figure 2). The results suggest that the topological properties of the attractors were quite similar across different musicians and tempos, while their physical realizations of the actions were different person to person.

## 3. Discussion

One of challenges to the theory of action recognition is formalizing the possible attributes of characteristic actions. In the present study, we hypothesize that an intrinsic topological nature of actions as dynamical systems characterizes a similarity between actions. This is meant to describe actions on the basis of *invariances* under nonlinear transformations, rather than the specific features (coordinate systems) the actions have. In other words, this is to abstract the actions from their physical properties. In order to test the hypothesis, we investigated the samba playing action, which is repetitive rhythmic movement. The samba rhythm fluctuates within a certain range, showing a complex accent pattern, even if an auditory cue is given to keep the tempo constant (Figure 2). Therefore, no common property was found in movements across musicians in physical movements of the right arm.

Subsequently, we analyzed the same data from a perspective of actions as dynamical systems. In the analysis, we define a higher dimensional space in which action is mapped as a trajectory. We analyzed time-delay vectors, which embed the time series of movements, i.e., lower dimensional data, into higher dimensional space.

The analyses revealed topological similarities in the reconstructed phase space among the musicians and among the different playing conditions. The analyses using the symbolic dynamics quantified the similarities in terms of their topological structures. In sum, these results supported our hypothesis that human actions can be characterized on the basis of invariances as dynamical systems. This invariant nature of the dynamical property can serve as a possible basis for our perception of actions, and offers an explanation of why we perceive them as "the same actions." Interestingly, the patterns revealed in the current analysis (Figure 3) are not just abstract-level depiction, but they also correspond with the introspective view of the samba rhythm (Figure 4: obtained from the most experienced musician A in the post-experiment interview). His drawing represents a general periodic motion with accents at a particular part of the trajectory. The geometric shape of the trajectory closely corresponds with the reconstructed phase space (Figure 3).

The present study proposes the dynamical perspective of actions in which it is essential to characterize topological similarities of actions as attractors. It is viewed as a paradigm shift from cognition as inverse computation for an ill-posed problem to the computation of invariances under smooth transformations.

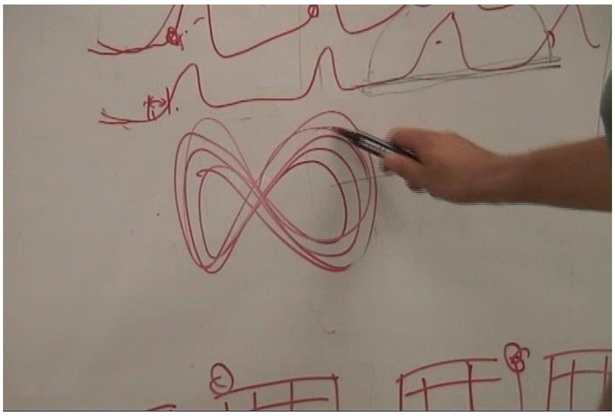

Figure 4: A drawing by the expert in his introspective explanation of the samba rhythm.

## Acknowledgments

This study was supported by Artificial Intelligence Research Promotion Foundation, a grant of the NeuroCreative Lab (NPO), and Grant-in-Aid for Scientific Research B KAKENHI No. 23300099 and Grant-in-Aid for Challenging Exploratory Research No. 25560297.

## References

Atkinson, A. P. (2009). Impaired recognition of emotions from body movements is associated with elevated motion coherence thresholds in autism spectrum disorders. *Neuropsychologia*, *47*, 3023–3029.

Bingham, G. P. (1987). Kinematic form and scaling: Further investigations on the visual perception of lifted weight. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 2, 155-177,

Blake, R. & Shiffrar, M., (2007). Perception of Human Motion. *Annual Review of Psychology, 58*, 47–73.

Chang, D. H. F., & Troje, N. F. (2008). Perception of animacy and direction from local biological motion signals. *Journal of Vision, 8*, (5):3, 1–10.

Chang, D. H. F., & Troje, N. F. (2009). Acceleration carries the local inversion effect in biological motion perception. *Journal of Vision*, *9*, (1):19, 1–17

Cook, J., Saygin, A. P., Swain, R., & Blakemore, S-H., (2009). Reduced sensitivity to minimum-jerk biological motion in autism spectrum conditions. *Neuropsychologia*, *47*, 14, 3275-3278.

DeMeijer, M, (1989). The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal Behavior*, *13*, 4, 247-268.

Giese, M. A. & Poggio, T. (2003). Neural Mechanisms for the recognition of biological movements, *Nature Reviews Neuroscience, 4*, 179-192.

Hobson, R. P. & Lee, A. (1999). Imitation and Identification in Autism, *Journal of Child Psychological Psychiatry, 40*, 4, 649-659.

Hubert, B., Wicker, B., Moore, D. G., Monfardini, E., Duverger, H., Fonse´ca, D. Da, Deruelle, C. (2006). Recognition of Emotional and Non-emotional Biological Motion in Individuals with Autistic Spectrum Disorders. *Journal of Autism Developmental Disorders*, *37*, 7, 1386-1392.

Johhanson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*, 2, 201-211.

Johnson, M. H., Bolhuis, J. J., & Horn, G. (1985). Interaction between acquired preferences and developing predispositions during imprinting. *Animal Behaviour, 33*, 1000–1006.

Kantz, H., & Schreiber, T. (1997). Nonlinear time series analysis. Cambridge, UK: Cambridge University Press.

Lange, J., & Lappe, M. (2006). A model of biological motion perception from configural form cues. Journal of *Neuroscience*, *26*, 11, 2894–2906.

Moore, Hobson, & Lee (1997). Components of person perception: An investigation with autistic, non-autistic retarded and typically developing children and adolescents., *British Journal of Developmental Psychology*, *15*, 401-423.

Pollick, F. E., Paterson, H. M., Bruderlin, A., Sanford, A. J., (2001). Perceiving affect from arm movement. *Cognition, 82*, B51–B61.

Pollick, F. E., Lestou, V., Ryu J. Cho, S-B. (2002) Estimating the efficiency of recognizing gender and affect from biological motion., *Vision Research*, *42*, 2345-2355.

Pollick, F. E., Paterson, H., Bruderlin, A. & Sanford, A. J. (2001) Perceiving affect from arm movement. *Cognition*, *82*, B51-B61.

Pollick F. E., Paterson, E. (2008). Movement style, Movement features, and the recognition of affect from human motion, In Shipley, T. F. & Zacks, J. M., *Understanding Events from Perception to Action*, New York: Oxford University Press, 286-307.

Takens, F. (1981). Detecting strange attractors in turbulence., In D. A. Rand and L.-S. Young. *Dynamical Systems and Turbulence, Lecture Notes in Mathematics*, vol. 898. Springer-Verlag. pp. 366–381.

Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision*, *2*, 371-387.

Troje, N. F. (2008). *Biological motion perception.* In Basbaum, A. et al. (Eds.), The senses: A comprehensive reference (pp. 231–238). Oxford: Elsevier.

Troje, N. F., Westhoff, C., & Lavrov, M. (2005). Person identification from biological motion: effects of structural and kinematic cues. *Perception & Psychophysics*, *67* (4), 667-675.

Turvey, M. T. (1998). Dynamics of effortful touch and interlimb coordination. Journal of Biomechanics, 31(10), 873-882.

Yamamoto, Y., Ishikawa, K., & Fujinami, T. (2006). Developmental stages of musical skill of samba. Journal of biomechanics, 39, S555.