The Beauty of Breaking Rhythms: Affective Robot Motion Design Using Jo-Ha-Kyū of Bunraku Puppet

Yang Chen
University of Tsukuba
cy@cavelab.cs.tsukuba.ac.jp

Ran Dong
University of Tsukuba
dong@cavelab.cs.tsukuba.ac.jp

Dongsheng Cai
University of Tsukuba
cai@cs.tsukuba.ac.jp

Shinobu Nakagawa
Osaka University of Arts
shinobu@osaka-geidai.ac.jp

Tomonari Higaki
Osaka University of Arts
thigaki@musicircus.net

Nobuyoshi Asai
University of Aizu
nasai@u-aizu.ac.jp

Figure 1: A) Manipulations of Bunraku Puppet in the theater. One puppet master and two other sub-masters (total three puppeteers) manipulate different part of Bunraku puppet to create sophisticated affective motions. B) The Tayū (chanter and narrator) and the shamisen player. They create Jo-ha-kyū (change of speed or tempo with breaking rhythms) to lead puppet master. C) Bunraku puppet motions are recorded by both the optical and magnetic motion capture systems. D) A robot with affective motions that are designed based on Bunraku puppet motions using deep learning and Jo-ha-kyū.

ABSTRACT

One of the UNESCO intangible cultural heritages Bunraku puppets can play one of the most beautiful puppet motions in the world. The sophisticated motions of the Bunraku puppet can express emotions interactively with a fixed facial expression overcoming the so-called "Uncanny Valley", simultaneously. In the present paper, we study the Bunraku motions using the famous concept so-called "Jo(Introduction)-Ha(Breaking)-Kyū(Rapid)". These emotional motions are synchronized with the Jo-ha-kyū. As a result, an android robot can express different affective motion synchronized to the different type of emotional chant or narration with a story that we call Jōruri.

CCS CONCEPTS

• **Applied computing** → **Performing arts**; • **Computing methodologies** → **Cognitive robotics**; *Motion capture*; *Neural networks.*

KEYWORDS

Motion analysis, Motion design, Bunraku puppet, Jo-ha-kyū

SIGGRAPH '19 Talks, July 28 - August 01, 2019, Los Angeles, CA, USA

ACM ISBN 978-1-4503-6317-4/19/07.
https://doi.org/10.1145/3306307.3328196

1 BUNRAKU PUPPETS

Bunraku, also known as Ningyō jōruri is the form of traditional Japanese puppet theatre (see Fig. 1 AB). Three kinds of performers play in a Bunraku performance. They are the Ningyōzukai (three puppeteers manipulate one puppet), the Tayū (narrator) and the shamisen (three string musical instrument) musicians. Bunraku puppet's motions and movements manipulated by one master and two sub-masters leading by Tayū's chants based on Jo-ha-kyū, which is a change of tempo with breaking rhythms. When human express our affections, we use facial expressions, eye movements, and gestures at the same time. On the other hand, Bunraku puppet masters express abundant emotions using mainly gestures without facial expressions. Unfortunately, these affective puppet motions have not been studied at all except for a few. As displayed in Fig. 1 C, we capture these sophisticated puppet motions and try to componentize these motions and use them to express affective motions.

2 JO-HA-KYŪ IN BUNRAKU PUPPETS

Jo-ha-kyū is a concept of modulation or breaking of rhythms or tempo in Japanese music and dances etc. Jo-ha-kyū can be roughly translated to "beginning, break, rapid" as shown in Fig. 2, it means that all actions or efforts should begin slowly, speed up, and then end swiftly[KONPARU 1980]. In Bunraku, all these puppet motions are synchronized and desynchronized with chanting and shamisen music based on Jo-ha-kyū. Tayū or chanter has him or her own tonal center and changes tempo based on Jo-ha-kyū according to the Jōruri (narration story), to leads the whole performance or rhythms as shown in Fig. 2.

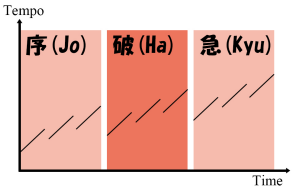

Figure 2: Change and breaking of rhythms or tempo (y-axis), Jo-ha-kyū in time (x-axis).

Figure 3 shows an example from our captured data, the typical change and break of speed in Jo-ha-kyū both in the speed of chant and puppet motions. As shown in Fig. 3(Left), the tempo of the chant, and puppet joint speeds that are Head, Hand, and Hip, respectively), are all correlated, and change speed or tempo according to the story. As shown in Fig. 3(Right), the chanting speed and head velocity are correlated and changing with the story to create the emotions.

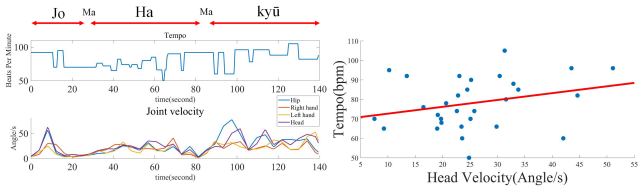

Figure 3: A part of jōruri play entitled Imoseyama On-na Teikin. The change and break of rhythms or tempo in both chant and puppet joints speed. They follow Jo-ha-kyū rule and joints (in ANOVA F=1.92, p=0.1282>0.05). The linear regression of tempo change with the median velocity of the head in measure unit (correlation coefficient r=0.3).

3 CHANGE OF ROBOTIC MOTION USING JO-HA-KYŪ

In this research, we proposed a framework to generate robot affective motion in Fig. 4. First, to reduce the degree of freedom (DOF) of Bunraku puppet to the robot, we directly use low dimensional Bunraku puppet motion data to retarget the high dimensional robot motion data and train the neural network. In Fig. 4, "B" and "R" represent DOF of Bunraku puppet (m) and robot (n), where m>n. We use both Bunraku puppet motion capture data and retargeted training data created by hand to train a deep neural network for retargeting affective motion.

Second, we use a convolutional auto-encoder proposed by [Holden et al. 2016] to extract affective robot motion manifold. The convolutional auto-encoder performs a one-dimensional convolution over the temporal domain, independently for each filter. The network provides a forward operation ϕ (encoding) and a backward operation $\phi^\dagger$ (decoding). The forward operation is:

$$\phi(X) = ReLU(\psi(X * W + b)) \tag{1}$$

The backward operation:

$$\phi^\dagger(X) = \psi^\dagger((H) - b) * \widetilde{W} \tag{2}$$

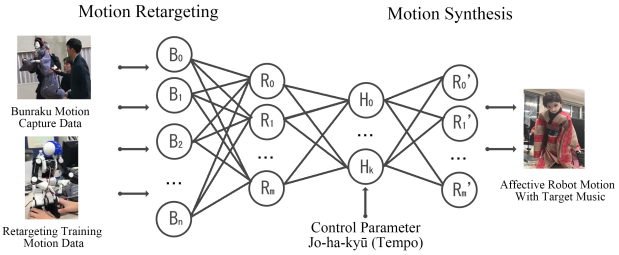

Figure 4: Affective Robot Motion Design framework.

where ψ is the max pooling operation and $\psi^\dagger$ is an inverse-pooling operation. The forward operation receives the vector X as inputs and output encoded values H in the hidden unit space. H means the hidden units. The cost function is given and is minimized with respect to the network parameters:

$$Cost(X) = \|X - \phi^\dagger(\phi(X))\|_2^2 \tag{3}$$

After training, a robot affective motion manifold is found. Finally, we use Jo-ha-kyū cost function discussed in session 3, to create motion corresponding to the targeted music from the hidden unit.

$$Johakyu(H) = \|\omega^H - T\|_2^2 \tag{4}$$

where T is the target tempo extract from input audio, ω^H is the angular velocity of the joint in the hidden unit.

$$H' = \arg\min_H \ Head(H) + Hand(H) + Hip(H). \tag{5}$$

By minimizing Eq. (5), we can generate affective robot motion from hidden space into visible space which corresponds to the input music or chant based on Jo-ha-kyū. The final output robot motions still have to be adjusted manually due to robot inertia balance.

4 CONCLUSIONS

In the present study, we use both optical and magnetic motion capture systems to capture the affective puppet motions. We also (1) use deep learning neural network to retarget affective robot motion from Bunraku puppet using direct teaching; (2) extract affective robot motion manifold using auto-encoder; (3) use a cost function to generate affective robot motion based on Jo-ha-kyū. Different affective robot motions are generated automatically according to the music or chant.

ACKNOWLEDGMENTS

The authors would like to thank the Bunraku performers: Tayu: Tsukoma Takemoto, Shamisen: Sosuke Takezawa, and Puppeteers team Kanjuro Kiritake and 2 others, for providing us the sound and motion data. The present work is supported by the 2nd term ACT-I by JST.

REFERENCES

Daniel Holden, Jun Saito, and Taku Komura. 2016. A deep learning framework for character motion synthesis and editing. *ACM Transactions on Graphics (TOG)* 35, 4 (2016), 138.

KUNIO KONPARU. 1980. *INVITATION TO NOH - SCIENCE BETWEEN JO-HA-KYU AND MA (JAPANESE).* TANKOSHA.