Bio-Inspired Design and Control of the Waseda Saxophonist Robot

J. Solis*
Waseda University
Tokyo, Japan

A. Takanishi†
Waseda University
Tokyo, Japan

K. Hashimoto
Toyota Motor Corporation
Toyota, Japan

Abstract—*This research is focused in developing an anthropomorphic saxophonist robot designed to reproduce the human organs involved during the saxophone playing. In this paper, we present the Waseda Saxophonist Robot No. 2 Refined (WAS-2R). In particular the shape of the oral cavity has been re-designed to increase the sound pressure range and potentiometers were embedded on the fingers to reduce the dynamic delay response of the wire-driven mechanism. In addition, a Pressure-Pitch Controller has been implemented to reduce the deviation of the sound pitch by implementing a feedback error learning algorithm for a Multiple-Input Multiple-Output system. A set of experiments were proposed to verify the effectiveness of the re-designed mechanisms and the improved control strategy. From the experimental results, we could confirm the improvements to extend the sound pressure rage to reproduce the decrescendo effect, to reduce the response delay from the finger mechanism as well as the deviations on the sound pitch.*

Keywords: Humanoid Robots, Bio-inspired Mechanism, Feedback Error Learning

I Introduction

The development of anthropomorphic robots is inspired by the ancient dream of humans replicating themselves. However, human behaviors are difficult to explain and model. Owing to the evolution of computers, electronics, and signal processing, this ancient dream is becoming a reality. In fact, current humanoid robots are able to perform activities such as dancing and playing musical instruments. However, these mechanical devices are still far from understanding and processing emotional states as humans do. Research on musical performance robots seems like a particularly promising path toward helping to overcome this limitation, because music is a universal communication medium, at least within a given cultural context. Furthermore, research into robotic musical performance can shed light on aspects of expression that traditionally have been hidden behind the rubric of “musical intuition”.

In 1984, at Waseda University, the WABOT-2 was the first attempt of developing an anthropomorphic music robot capable of playing a concert organ. Then, in 1985, the WASUBOT built also by Waseda, could read a musical score and play a repertoire of 16 tunes on a keyboard instrument [1]. The late Prof. Ichiro Kato argued that the artistic activity such as playing a keyboard instrument would require human-like intelligence and dexterity [2].

More recently; thanks to the technological advances on power computation, Musical Information Retrieval (MIR) and Robot Technology, several researchers have been focusing on developing anthropomorphic robots and interactive automated instruments capable of interacting with musical partners. As a result, different kinds of wind playing-instrument automated machines and humanoid robots have been developed for playing wind instruments [3-10]. Other researchers have been focusing in analyzing wind-instrument playing from a musical engineering approach by performing experiments with simplified mechanisms [12-13] and from a physiological point of view by analyzing medical imaging data of professional players [14-15]. In this research, we particularly deal with the development of an anthropomorphic saxophone-playing robot designed to mechanically emulate the required organs during the saxophone playing. Due to the interdisciplinary nature of this research, our collaboration with musicians, musical engineers and medical doctors will certainly contributes to better reproduce and understanding the required skills to play the saxophone instrument.

Since 1990, at Waseda University; we have been doing developing an anthropomorphic flutist robot ([5-6]) as a mean for understanding the human control, introducing novel ways of interaction between musical partners and robots and proposing applications for humanoid robots [16]. As a result of our research (Figure 1), we have succeeded in enabling the flutist robot to perform basic technical skills which are usually required by flutist beginners in order to produce a sound. Furthermore, the flutist robot is also able of performing some extended technical skills; which are typically practiced by intermediate level flutists. The research on the flutist robot has been particularly focused on understanding the human motor control from an engineering point of view. Therefore; we are aiming at developing an anthropomorphic flutist robot from two research approaches: reproducing the required motor dexterity to play the flute and displaying cognitive functions to coordinate the motion of simulated organs to express

* solis@ieee.org
† contact@takanishi.mech.waseda.ac.jp

emotions in musical terms.

Fig. 1. Waseda Flutist Robot's series

At Waseda University we have proposed; as a long-term goal, the development of MPRs designed to be able of interacting with musical partners at the emotional level of perception. Here, we refer as musical partners both human players as well as musical performance robots. For this purpose, we are focusing our research in enhancing the perceptual capabilities of the WF-4RIV as well as developing a new musical performance robot. The preliminary results of such a research are detailed in [17]. On the other hand, as it was previously introduced, we have also focused on developing a new musical performance robot which in the future will be able of interacting with the WF-4RIV. For this purpose, we have decided to develop a MPR which is able of performing other kind of wind instrument. Wind instruments fall into one of the following categories: brass instruments and woodwind instruments. One important difference between woodwind and brass instruments is that woodwind instruments are non-directional. This means that the sound produced propagates in all directions with approximately equal volume. Brass instruments, on the other hand, are highly directional, with most of the sound produced traveling straight outward from the bell. Thus, the wood instruments represent a higher challenge to reproduce the human motor control from an engineering point of view. Basically, there are three types of wood instruments: single reed (i.e. clarinet, saxophone), double reed (i.e. oboe, bagpipes, etc.) and flutes. Therefore, we have selected to develop a single-reed musical performance robot.

As a result, the authors proposed in [18] the development of an anthropomorphic saxophonist robot as an approach to enable the interaction with musical partners (i.e. with the flutist robot). Therefore; as a long-term goal, we expect that the proposed saxophonist robot is able not only of performing a melody, but also to dynamically interact with the musical partner (i.e. walking while playing the instrument, etc.).

As a result of our research (Figure 2), in [18], we have presented the Waseda Saxophonist Robot No.1 (WAS-1), which it was composed by 15 degrees of freedom (DOF) required to play an alto saxophone. In particular, lower lip (1-DOF), tongue (1-DOF), oral cavity, artificial lungs (air pump: 1-DOF and air flow valve: 1-DOF) and fingers (11-DOF) were developed. Both lips and oral cavity were made of a thermoplastic rubber (named Septon and produced by Kuraray Co. [19]). An improved version, the Waseda Saxophonist Robot No.2 (WAS-2) was presented in [20]; where the design of the artificial lips was improved and a human-like hand was designed. Furthermore, in [21], an Overblowing Correction Controller was implemented in order to assure the steady tone during the performance by using the pitch feedback signal to detect the overblowing condition and by defining a recovery position (off-line) to correct it.

However, still the range of sound pressure was still too limited to reproduce the dynamic effects of the sound (i.e. decrescendo) and deviations on the pitch were detected. Therefore; in this paper, the design of the oral cavity shape has been improved to expand the range of sound pressure and potentiometers were attached to each finger for implementing a dead-time compensation controller. From the control system point of view, a Pressure-Pitch Controller has been proposed to ensure the accurate control of the pitch during the steady phase of the sound produced by the saxophone.

Thus, in this paper, we describe the mechanical improvements on the oral cavity and finger mechanisms. In addition, the implementation of a finger dead-time compensation controller and Multiple-Input Multiple-Output controller to assure the accurate control of both air pressure and sound pitch

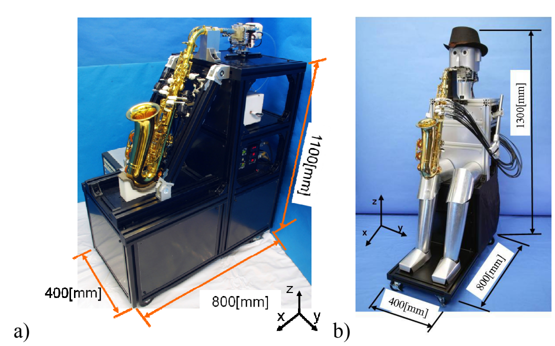

Fig. 2. Waseda Saxophonist Robot's series: a) Waseda Saxophonist Robot No. 1 (WAS-1); b) Waseda Saxophonist Robot No. 2 (WAS-2).

II. Waseda Saxophonist Robot No.2 Refined

In this year, we have developed the Waseda Saxophonist Robot No. 2 Refined (WAS-2R) which it has improved the shape of the oral cavity for increasing the sound range volume and added sensors to each finger for reducing the response delay. In particular, the WAS-2R is composed by 22-DOF that reproduce the physiology and anatomy of the organs involved during the saxophone playing as follows (Figure 3): 3-DOF to control the shape of the artificial lips, 16-DOF for the human-like hand, 1-DOF for the tonguing mechanism and 2-DOF for the lung system. In addition, to improve the stability of the pitch of the sound produced, a pressure-pitch controller system has been implemented.

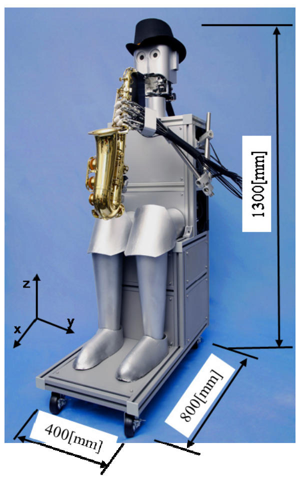

Fig. 3. Waseda Saxophonist Robot No. 2 Refined (WAS-2R).

- **Oral Cavity:** In the previous mechanism, it was possible to confirm the enhancement of the sound range produced by WAS-2 [20]. However; we detected that the note C3 was not possible to be produced. Therefore, we considered to analyze in more detail the oral cavity (in particular, the gap between the palate and the tongue) of professional saxophonist while playing the instrument. For this purpose, we have used an ultrasonic sound probe (ALOKA ProSound II, SSD-6500SV) to obtain images of the oral cavity from professional players while producing the sound of the note C4 (Figure 2a). As we may observe in Fig. 2b, when a higher volume sound is produced, a large gap between the palate and the tongue is observed. In contrast, while producing lower volume sounds, the gab is considerably narrowed. As a result from these measurements, a new oral cavity for the WAS-2R has been designed (Figure 4). Basically, based on the measurements obtained from images obtained from the professional player; the sectional area has been designed with 156 mm^2 (previous one was 523 mm^2).

- **Finger Mechanism:** In the previous mechanism, a human-like hand (actuated by a wire-driven mechanism) has been designed to enable the WAS-2 to push all the keys of the alto saxophone [20]. However; due to the use of the wire-driven mechanism, a dynamic response delay (approximately 110ms) has been observed. Therefore, in order to reduce such a delay time, we proposed to embed sensors for measuring the rotational angle of each finger. For this purpose, a rotary sensor (RDC506002A from Alps Co.) has been embedded into the each finger mechanism (Figure 5). In particular, each sensor was placed on a fixing mount device produced by a rapid prototyping device (CONNEX 500). As a result, we were able of attaching the sensing system without increasing the size of the whole mechanism. RC servo motors have been used to control the wire-driven mechanism designed for each finger. As end-effector, an artificial finger made of silicon has been designed. In order to control the sixteen RC motors, the RS-485 serial communication protocol has been used.

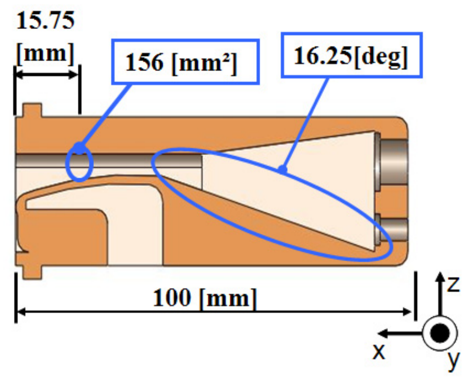

Fig. 4. Detail of the re-designed oral cavity of WAS-2R.

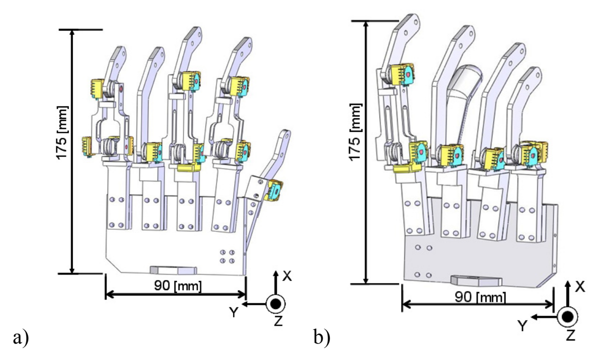

Fig. 5. Detail of the finger mechanism of the WAS-2R with the embedded sensors.

- **Control System:** In our previous research, a feed-forward air pressure controller with dead-time compensation has been implemented to ensure the accurate control of the air pressure during the attack time [20]. Moreover, for the control of the finger mechanism, a simple ON/OFF controller has been implemented. In particular, the feedback error learning during the attack phase of the sound has been used to create the inverse dynamics model of the Multiple-Input Single-Output (MISO) controlled system based on Artificial Neural

Networks (ANN). In addition, an Overblowing Correction Controller (OCC) has been proposed and implemented in order to ensure the steady tone during the performance by using the pitch feedback signal to detect the overblowing condition and by defining a recovery position (off-line) to correct it [21]. However, we still detect deviations on the pitch while playing the saxophone.

Therefore, in this paper, we proposed the implementation of the control system shown in Fig. 6. In particular, the improved control system includes a dead-time compensation controller for the finger mechanism (to reduce the effect of response delay due to the wire-driven mechanism) and a Pressure-Pitch Controller (PPC) for the control of the valve and lip mechanism (to assure the accurate control of the pitch).

Regarding the implementation of the dead-time compensation control (which it uses an element to predict how changes made now by the controller will affect the controlled variable in the future [22]); for each finger of WAS-2R, the pressing time of the saxophone's key is measured by means of the embedded potentiometer sensor (defined as L_N; where N represents the total number of DOF designed for the finger mechanism). By including the dead-time factor (referred as e^{sL}); it is possible to compensate the finger's response delay during the saxophone playing.

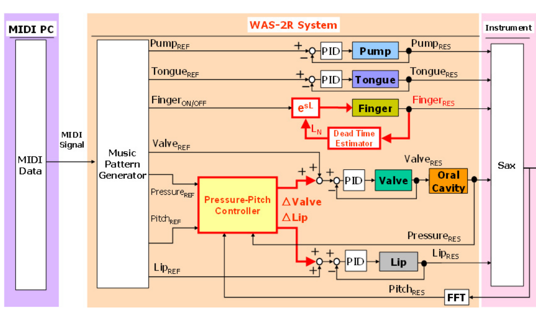

Fig. 6. Block diagram of the improved control system implemented for the WAS-2R (in red color are indicated the improved modules done).

On the other hand; regarding the pressure-pitch controller during the sustain phase of the sound, a Multiple-Input Multiple-Output (MIMO) system has been proposed not only to ensure the accurate control of the air pressure during the attack phase of the sound, but also to ensure the accurate control of both air pressure and sound pitch during the sustain phase of the sound. For this purpose, we implemented a feed-forward error learning method [23] to create the inverse model of the proposed MIMO system which is computed by means of an ANN. During the training process, the inputs of the ANN are defined as follows (during the attack phase, no modification was done to the previous system [20]): Pressure reference ($Pressure_{REF}$), pitch reference ($Pitch_{REF}$). In this case, a total of six hidden units were used (experimentally determined while varying the number of hidden units). As an output, the position of the air valve ($\Delta Valve$) and lower lip (ΔLip) are controller to ensure the accurate control of the required air pressure and pitch to produce the saxophone sound. Moreover, during the training phase, the air pressure ($Pressure_{RES}$) and sound pitch ($Pitch_{RES}$) are used as feedback signals and both outputs from the feedback controller are used as teaching signals for the effectively training the ANN. As a result from the training phase, during a saxophone playing performance, the created inverse model is used.

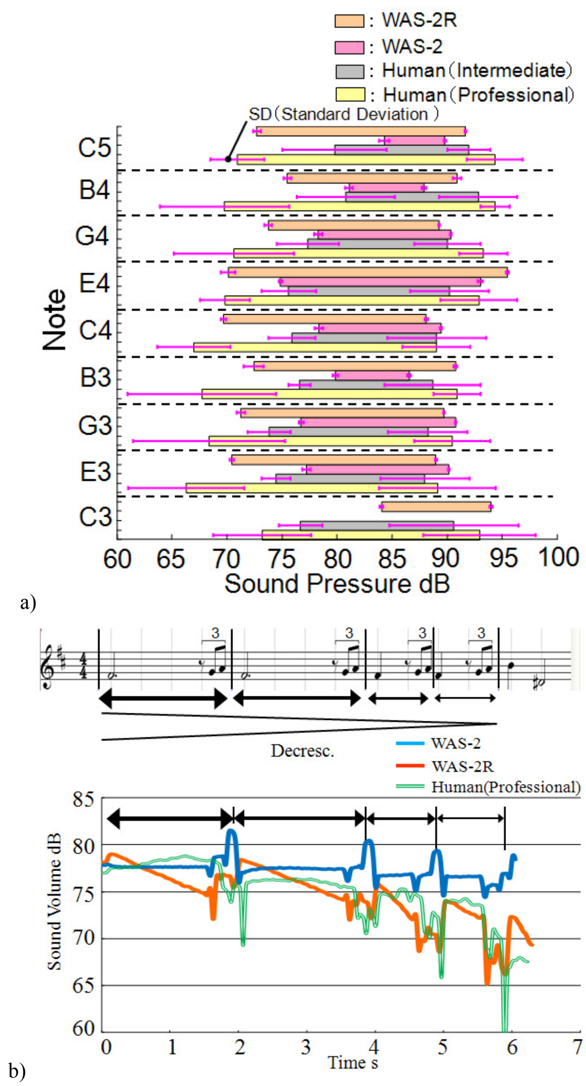

Fig. 7. Experimental results: a) Range of sound pressure; b) Reproduction of decrescendo.

IV. Experiments and results

- **Range of Sound Pressure:** In order to verify if the re-designed shape of the oral cavity contributes to extend the range of sound pressure; we have compared the previous mechanism with the new one while playing the notes from C3 to C. The experiment results are shown in Fig. 7a. The average sound pressure range for WAS-2R and WAS-2 are 17.7 dB and 9.69 dB respectively. Moreover, an intermediate player and professional are 13.2 and 22.6 respectively. From this result, we confirmed an increment of 83% thanks to the new shape of the oral cavity. Therefore, we could conclude that the shape of the gap between the palate and tongue has a big influence on the sound pressure range.

Thanks to this considerable improvement on the range of sound pressure, we proposed to compare the reproduction of the decrescendo; which is a dynamic sound effect that gradually reduces the loudness of the sound. For this purpose, we programmed the WAS-2 and WAS-2R to play the principal theme of the "Moonlight Serenade" composed by Glenn Miller. The experimental results are shown in shown in Fig. 7b. As we may observe, the WAS-2R was able of reproducing nearly similar to the performance of the professional one.

- **Finger Dead-Time Compensation:** In order to determine the effectiveness of the dead-time compensation to reduce the delay response of the finger mechanism, we programmed the WAS-2R to play the main theme of the "What a wonderful world" composed by Bob Thiele and George David Weiss" with and without dead-time compensation. The experimental results are shown in Fig. 8. As we could observe, when the melody is played without dead-time compensation, we may find considerable deviations on the pitch during the transitions from one note to other (D4-D#4 and C#4 and D4). However; when the dead-time compensation is used, no deviations on the pitch are observed. From this experiment, we could consider that the proposed control system effectively reduce the effect of the dynamic response delay due to the use of the wire-driven mechanisms. However, we may still find some small deviations of the pitch.

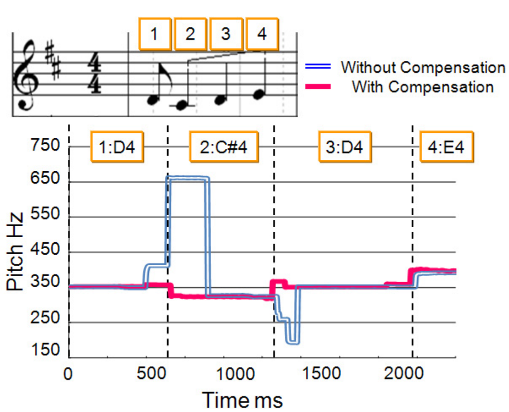

Fig. 8. Results of experiments to compensate dead time of finger with WAS-2R

- **Sound Pitch Control:** In order to determine the effectiveness of the proposed pressure-pitch controller to reduce the pitch deviations while playing the saxophone, we programmed the WAS-2R to play the main theme of the "Moonlight Serenade" composed by Glenn Miller before and after training the inverse model. In particular, as for the neural network parameters, a total of 6 hidden units were used. For the training process, a total of 144 steps were done. The experimental results are shown in Fig. 9; where 1[cent] is defined as (equi-tempered semitone/100). As we could observe, the deviations of the pitch after the training (Standard Error is 41.7) are considerable less than before training (Standard Error is 2372.8).

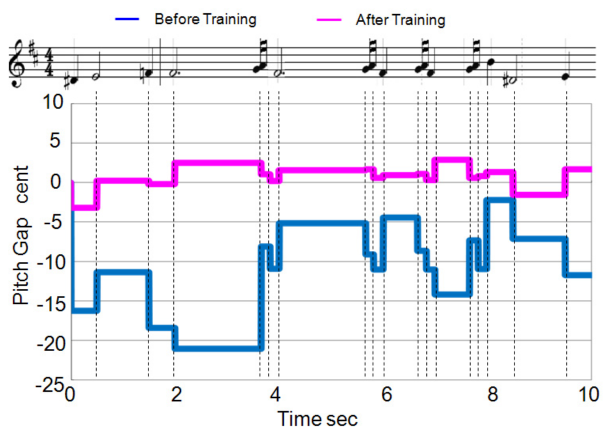

Fig. 9. Experimental results while comparing the deviations of the pitch before and after training the inverse model of the proposed MIMO system with the WAS-2R.

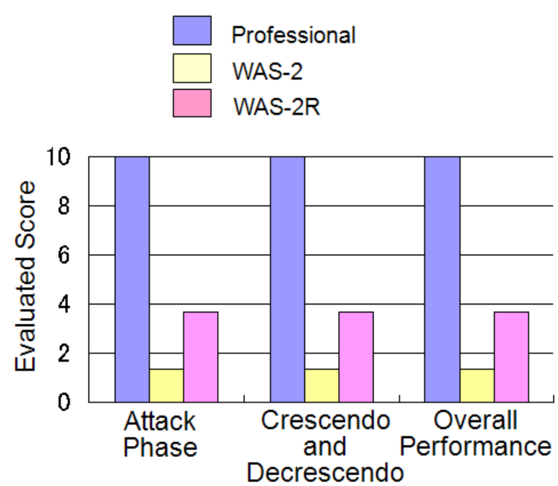

Fig. 10. Experimental results while comparing qualitatively the performance of the WAS-2R and WAS-2 by 3 professional saxophone players.

- **Performance Qualitative Evaluation:** Finally, in order to determine the improvements on the performance level of the WAS-2R, we have proposed to carry out a qualitative performance evaluation. For this purpose, we have asked 3 professional saxophone players to evaluate and compare the performance of the WAS-2 and WAS-2R. In particular, the recordings from both robots as well as a professional one were done while playing the "Moonlight Serenade". For this purpose, each subject was requested to evaluate each performance while considering the following parameters: attach phase, crescendo and decrescendo effects and overall performance. Each of the evaluation parameters were evaluated from a range of 0 to 10 (where the highest value belongs to recording from the professional one). The experimental results are shown in Fig. 10. As we may observe, all the subjects evaluated with higher scoring the performance of the WAS-2R compared with the WAS-2.

V. Conclusions

In this paper, we have presented the details of the mechanism design improvements on the Waseda Saxophonist Robot No. 2R. In particular, the shape of the oral cavity has been improved to extend the sound pressure range so that the decrescendo effect could be effectively produced. In addition, potentiometers were attached to each finger of the robot to reduce the response delay by implementing the dead-time compensation. On the other hand, a pressure-pitch controller has been implemented to ensure the accurate control of both air pressure and sound pitch during the sustain phase. For this purpose, the inverse model of a MIMO system has been created by means of the feedback error learning method.

As a future work, we are planning to design an actuated mechanism to modify the internal shape of the oral cavity. Moreover, different fingering techniques will be studied and mechanically reproducing by re-designing the finger mechanism. Finally, an evaluation function will be proposed to evaluate the quality sound during the attack, sustain, etc. phases and include it as a teaching signal to the feedback error leaning

References

[1] Sugano, S., Kato, I., WABOT-2: autonomous robot with dexterous finger-arm coordination control in keyboard performance, in Proc. of the International Conference on Robotics and Automation, New York, pp. 90-97, 1987.

[2] Kato, I., Ohteru, S., Kobayashi, H., Shirai, K., Uchiyama, A., Information-power machine with senses and limbs, in Proc. of the CISM-IFToMM Symposium on Theory and Practice of Robots and Manipulators, Udine, pp. 12-24, 1973.

[3] Doyon, A., Liaigre,L. Jacques Vaucanson: Mecanicien de Genie, PUF, 1966

[4] Klaedefabrik, K.B., Martin Riches - Maskinerne / The Machines. Kehrer Verlag, pp. 10-13, 2005.

[5] Solis, J., Chida, K., Suefuji, K, Takanishi, A., The development of the anthropomorphic flutist robot at Waseda University, International Journal of Humanoid Robots, Vol. 3(2), pp. 127-151, 2006.

[6] Solis, J., Taniguchi, K., Ninomiya, T., Takanishi, A. Understanding the Mechanisms of the Human Motor Control by Imitating Flute Playing with the Waseda Flutist Robot WF-4RIV," Mechanism and Machine Theory, Vol. 44 (3), pp. 527-540, 2008.

[7] Takashima S., Miyawaki, T. Control of an automatic performance robot of saxophone: performance control using standard MIDI files, Proc. of the IEEE/RSJ Int. Conference on Intelligent Robots and Systems - Workshop: Musical Performance Robots and Its Applications, pp. 30-35, 2006.

[8] Solis J., Petersen K., Ninomiya T., Takeuchi M., Takanishi A. Development of Anthropomorphic Musical Performance Robots: From Understanding the Nature of Music Performance to Its Application in Entertainment Robotics, in Proc. of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2309-2314, 2009.

[9] Dannenberg, R.B., Brown, B., Zeglin, G., Lupish, R. McBlare: A Robotic Bagpipe Player, Proc. of the International Conference on New Interfaces for Musical Expression, pp. 80-84, 2005.

[10] http://www2.toyota.co.jp/en/tech/robot/p_robot/index.html

[11] Degallier,S., Santos, C.P., Righetti, L., Ijspeert, A. Movement generation using dynamical systems: a humanoid robot performing a drumming task, in Proceedings of the IEEE-RAS International Conference on Humanoid Robots, pp. 512-517, 2006.

[12] Ando, Y. Drive conditions of the flute and their influence upon harmonic structure of generated tone, Journal of the Acoustical Society of Japan, pp. 297-305, 1970.

[13] Guillemain P., Vergez C., Ferrand D., Farcy A., An instrumented saxophone mouthpiece and its use to understand how an experienced musician plays, Acta Acoustica, 96(4):622-634, 2010

[14] Mukai S., Laryngeal movement while playing wind instruments, in: Proceedings of International Symposium on Musical Acoustics, pp. 239–241, 1992.

[15] Fletcher, N., Hollenberg, L., Smith, J. and Wolfe, J. The didjeridu and the vocal tract, in Proc. International Symposium on Musical Acoustics, pp 87-90, 2001.

[16] Solis, J., Suefuji, K. Taniguchi, K., Takanishi, A., Towards and autonomous musical teaching system from the Waseda Flutist Robot to Flutist Beginners, in Proc. of the IROS Workshop on Musical Performance Robots and Its Applications, Beijing, pp. 24-29, 2006.

[17] Petersen, K., Solis, J., Takanishi, A., Musical-Based Interaction System for the Waseda Flutist Robot: Implementation of the Visual Tracking Interaction Module," Autonomous Robots Journal, Vol. 28(4), pp. 439-455, 2010.

[18] Solis, J., Takeshi, N., Petersen, K., Takeuchi, M., Takanishi, A., Development of the Anthropomorphic Saxophonist Robot WAS-1: Mechanical Design of the Simulated Organs and Implementation of Air Pressure," Advanced Robotics Journal, Vol. 24, pp. 629-650, 2009.

[19] http://www.kuraray.co.jp/en/

[20] Solis J., Petersen, K., Yamamoto, T., Takeuchi, M., Ishikawa, S., Takanishi, A., Hashimoto, K. Design of New Mouth and Hand Mechanisms of the Anthropomorphic Saxophonist Robot and Implementation of an Air Pressure Feed-Forward Control with Dead-Time Compensation, Proceedings of the International Conference on Robotics and Automation, pp. 42-47, 2010

[21] Solis J., Petersen, K., Yamamoto, T., Takeuchi, M., Ishikawa, S., Takanishi, A., Hashimoto, K. Implementation of an Overblowing Correction Controller and the Proposal of a Quantitative Assessment of the Sound's Pitch for the Anthropomorphic Saxophonist Robot WAS-2, in Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems, 2010

[22] Kim, H., Kim, K., Young, M. On-Line Dead-Time Compensation Method Based on Time Delay Control. IEEE Transactions on Control Systems Technology, Vol. 11(2), pp. 279-286, 2003.

[23] Kawato, M., Gomi, H. A computational model of four regions of the cerebellum based on feedback-error-learning," Biological Cybernetics, vol. 68, pp. 95-103, 1992.