Enabling Humanoid Musical Interaction and Performance

Youngmoo E. Kim, David K. Grunberg, Alyssa M. Batula, Daniel M. Lofaro, JunHo Oh[1], and Paul Y. Oh[*]
Department of Electrical and Computer Engineering
Drexel University
Philadelphia, PA, USA
{ykim, dgrunberg, batulaa, dlofaro, pyo22}@drexel.edu
[1]Department of Mechanical Engineering
Korea Advanced Institute of Science and Technology
Daejeon, Korea
jhoh@kaist.ac.kr

ABSTRACT

Many people incorporate music into their daily lives, and the development of robots with musical awareness provides an opportunity for rich forms of human-robot interaction. Robots must, however, acquire a variety of skills before being able to participate in musical activities. In order to dance or play an instrument, for example, a robot be able to utilize substantial auditory and visual information. Our work focuses on providing such capabilities (audio and visual beat detection, pitch detection, motion control for producing musical notes) with the goal of enabling an adult-sized humanoid to be an interactive participant in a live musical ensemble. Miniature humanoids are used to prototype and refine many of these systems before deploying them on Hubo, an adult- sized humanoid developed by the Korean Advanced Institute of Science and Technology.

KEYWORDS: Dance, piano, robot audition, music

1. INTRODUCTION

Playing and enjoying music are integral activities in the daily lives of many people, and robots have been incorporated into several high-profile musical performances. Honda's ASIMO for example, conducted the Detroit Symphony Orchestra. Reviewers were impressed by the event, but noted that ASIMO could not respond to the musicians and was essentially a metronome.[1] The robot could not interact with the music or allow for any improvisation.

Developing robot systems that can truly understand music is still a challenging problem. It is not enough for the robot to render notes or melodies exactly as written on a sheet of paper. If the robot does not vary its performance in an expressive manner, its production will only appear flat and lifeless. A robot that is unable to interpret audio and visual information (such as the notes it plays), cannot account for mistakes, improvisations, or any deviation from the score. Humans can account for all of these factors, and robots that interact with music must be able to do the same.

Humanoids are desirable platforms for musical systems because they can potentially play instruments or produce motions similar to those used in human performances. We are developing our systems for Hubo, an adult-sized humanoid designed by the Korea Advanced Institute for Science and Technology (KAIST). Our goal is for Hubo to perform alongside human musicians. Many of our systems are first tested on a miniature humanoid, the Hitec Robonova-1. This robot provides a relatively robust and inexpensive prototyping platform before we apply our algorithms to Hubo. Both Robonova and Hubo have been used effectively in outreach events with the public.

2. RELATED WORK

One existing area of research is developing beat-sensitive robots for human interaction. One such robot, Keepon, is a small cartoon-like robot designed for studying

* Paul Oh is a member of the Department of Mechanical Engineering

This work is sponsored by a National Science Foundation Graduate Research Fellowship and by NSF Award 0730206.

[1] M. Stryker, ``DSO led by robot maestro,'' *Detroit Free Press*, 14 May 2008.

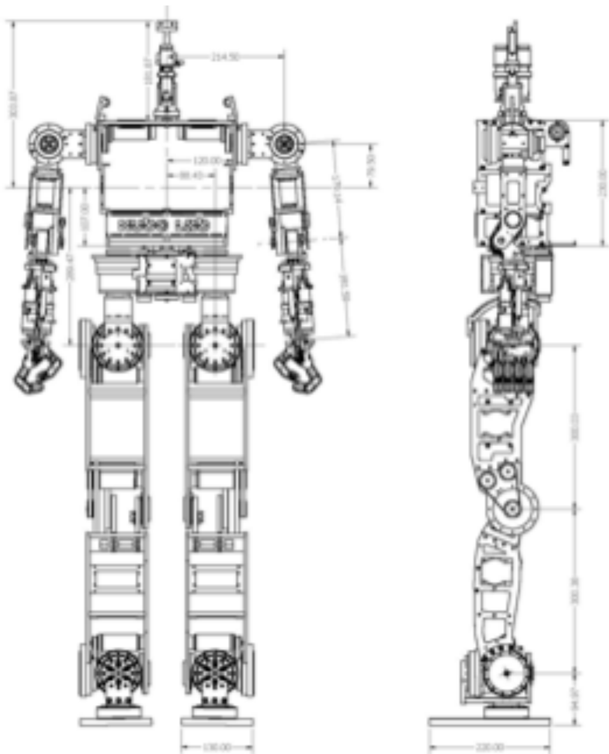

Figure 1: Hubo Schematic

interaction with autistic children. Keepon can detect and move in time with musical beats [1]. Another robot, called Haile, can listen to human drum sequences and synthesize its own accompaniments [2]. Furthermore, some robots sense beats visually rather than aurally. The robot Nico has demonstrated the ability to follow a conductor to determine beat positions [3].

Another branch of research has developed robots to play standard (human-playable) musical instruments, such as the Toyota Musical Robots [4]. These humanoids include a flutist, a violinist, and a trumpeter. Of particular interest to us are piano playing robots, such as Waseda's Wabot-2, which can play a variety of pieces for piano and organ [5]. The Wabot-2, however, cannot adjust its technique or position in the middle of a piece if it makes a mistake.

Researchers have also developed robots with the ability to dance. The HRP-2 has been enabled to reproduce traditional folk dances [6]. A few such robots already incorporate beat information, such as Honda's ASIMO, which can step in place with music based upon the beats that it "hears" [7]. Our group has enabled a Robonova to perform motions in synchrony to beats in music [8], [9].

3. HUMANOID RESEARCH PLATFORMS

Our work thus far has used two different robot platforms: the Robonova mini-humanoid robot and the Hubo KHR-4 adult-sized humanoid. Previous efforts have demonstrated that rapid prototyping algorithms with miniature humanoids has substantial benefits [8]. Development on large robots is more time-consuming and poses much higher risks in terms of cost and safety. The relatively low cost of mini-humanoids also broadens the audience for musically-aware robots.

Table 1. Robonova and Hubo degrees of freedom

Limb	Robonova	Hubo
Each arm	3	6
Each leg	5	6
Head and waist	0	2
Total	16	26

The Robonova-1 is a small (36 cm tall) humanoid with 16 degrees of freedom (DOF). The robot's motors are controlled with an ATmega 128 MPU processor. This controller had limited processing ability, so we reprogrammed it to allow an external computer to specify commands. This has improved the refresh rate to 50Hz [8]. We have also constructed new robot hands with a 3D printer. These hands are 1.27 cm longer than the original hands, and curve to a .64 cm square tip so that it can hit keys without striking neighboring notes.

The Hubo KHR-4 series humanoid robot is designed and built by the Hubo Lab at KAIST in Daejeon, South Korea. Drexel's Hubo ("Jaemi") is currently the only Hubo in America, and is currently located at the Drexel Autonomous Systems Lab (DASL) in Philadelphia, PA, USA. Jaemi Hubo is the primary focus of the Drexel-KAIST collaboration, which is supported through the US National Science Foundation's Partnership for International Research and Education (PIRE) program. This grant includes participation and cooperation between five universities in the United States and three in Korea.

Hubo is a 37 kg, 130 cm humanoid robot. It has 42 DOF, 26 of which can be used for dancing. A schematic of Hubo is shown in Fig. 1. Table 1 shows a comparison of Robonova and Hubo's degrees of freedom used for musical interaction.

Hubo and Robonova are often presented to the public as part of our outreach efforts. In the past two years, Jaemi Hubo performed at the Philadelphia Please Touch Museum (PTM), and was demonstrated to a Student-Professional Awareness Conference (S-PAC) sponsored by the IEEE. Jaemi interacts with visitors during demonstrations, such as by playing "Simon Says" at PTM, which pleased children in the audience. The musical abilities of both robots are also used in demonstrations.

Figure 2: Robonova Playing a Keyboard

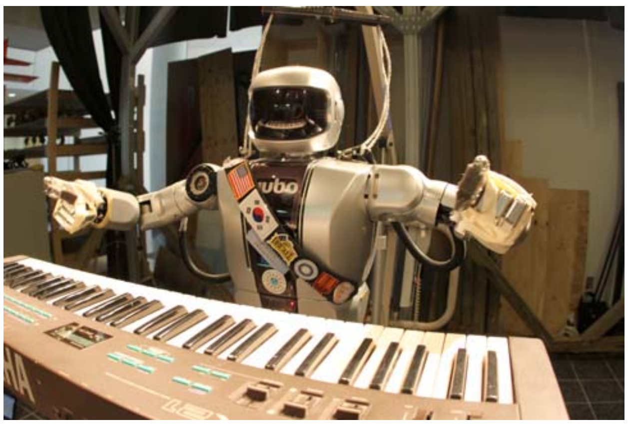

Figure 3: Hubo Playing a Keyboard

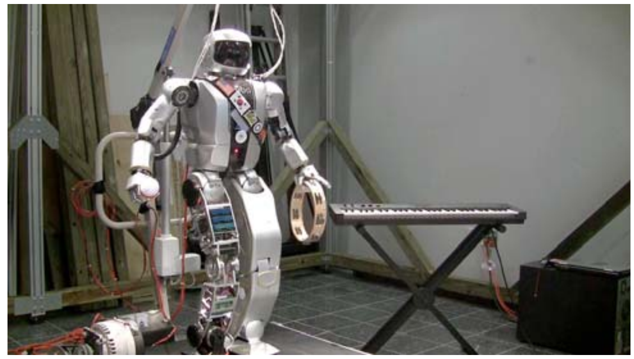

Figure 4: Hubo Playing Handheld Instruments

Figures 2 and 3 show both robots playing the piano, and Figure 4 shows the Hubo playing handheld instruments.

4. AUDIO AND VISUAL SENSING

4.1. Audio

Audio feedback is useful for maintaining correct pitch and timing during a performance. By identifying beat locations, the robot can synchronize its motions to the music, and by analyzing the frequency content of the music it produces, the robot can determine if it is playing the correct notes.

For beat tracking, our system exploits the fact that beats generally occur at a regular period throughout sections of a song. Our system filters audio into subbands, finds the energies of those subbands, and then uses autocorrelation to look for periodicity in the subband energy signals. The strongest periodicity usually corresponds to the tempo of the music [10]. The frame energies are used to find sequences of high- energy frames spaced according to the tempo. These frames are marked as probable beat locations.

As for pitch detection, we use the relationship between a note's fundamental frequency and pitch. This system assumes that the desired notes are known and looks for their fundamental frequencies in the frequency spectrum of the audio signal. If a peak is found at the note's fundamental frequency, the system determines that the note is present.

4.2. Vision

In order to implement visual beat tracking, we use cameras mounted on Jaemi Hubo's head. The live video input is equalized to compensate for different lighting conditions. The relative motion between consecutive frames is then found using the computer vision Horn-Schunck Optical Flow method [11]. The spectrum of this motion's mean magnitude and angle is analyzed to determine the tempo. The possible tempos are weighted using a triangular filter with a peak at 128 beats per minute (BPM), the average tempo of popular music [12]. This helps ensure the correct tempo is chosen. Using this method, Hubo is able to accurately track the beat when watching a trained musician conduct a steady tempo.

5. RESULTS

5.1. Piano Performance

The robot can currently play a one octave range with each hand. After a single note is played during a calibration step, the system can determine the note played and thus the location of the robot relative to the keys. The robot reads Musical Instrument Digital Interface (MIDI) files and uses the pitch detection algorithm to verify that it is correctly playing notes. Lastly, the robot can automatically calibrate itself by playing each accessible note and determining how far down to press its hand.

The pitch algorithm was evaluated on 230 correctly played notes, successfully identifying 229 (99.56%). It was then tested on 60 incorrect notes, identifying 54 (90%) as being incorrect.

5.2. Dance Performance

The dancing system allows a robot to dance in time to music. The system uses audio beat detection to find the beat in live music, then synchronizes its motions to the beats. This algorithm runs in real-time. Currently, both the Robonova and the Hubo can perform smooth, accurate dance gestures in response to audio.

The beat tracker has been tested on twenty songs in the pop genre and performs above 98% accuracy, using the F-Measure as a metric [13] The system can operate in real-time. Videos of dance routines are available at our website.[2]

6. CONCLUSIONS AND FUTURE WORK

Using an external computer, our humanoids can obtain and interpret sensory information from the environment to determine pitches and beat locations. The system is able to utilize this information in performances. In the near future, all of the computation will take place within the Hubo body, making the system fully self-contained.

We are interested in enabling the robot to make more decisions based on feedback, for example, by using a camera to locate piano key locations. Also, now that many of the technical challenges have been overcome, we may begin to explore the psychology of HRI more. Understanding what makes a robot seem more trustworthy or intelligent to a human musician could be a valuable branch of research.

REFERENCES

[1] M. Michalowski, R. Simmons, and H. Kozima, “Rhythmic attention in child-robot dance play,” in Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication, Toyama, Japan, September-October 2009.

[2] G. Weinberg and S. Driscoll, “The interactive robotic percussionist: New developments in form, mechanics, perception, and interaction design,” in Proceedings of the ACM/IEEE International Conference on Human Robot Interaction (HRI). New York: ACM, March 2007, pp. 97–104.

[3] C. Crick, M. Munz, T. Nad, and B. Scassellati, “Robotic drumming: synchronization in social tasks,” in Proc. of the 15th International Symposium on Robot and Human Interactive Communication, 2006.

[4] A. Kapur et al., “Collaborative composition for musical robots,” Journal of Science and Technology of the Arts, no. 1, pp. 48–52, 2009.

[5] C. Roads, “The tsukuba musical robot,” Computer Music Journal, vol. 10, no. 2, pp. 39–43, 1986.

[6] K. Shinozaki, A. Iwatani, and R. Nakatsu, “Concept and construction of a robot dance system,” The International Journal of Virtual Reality, vol. 6, no. 3, pp. 29–34, 2007.

[7] K. Yoshii et al., “A biped robot that keeps steps in time with musical beats while listening to music with its own ears,” in Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), San Diego, CA, October-November 2007.

[8] D. Grunberg, R. Ellenberg, Y. Kim, and P. Oh, “Creating an autonomous dancing robot,” in Proceedings of the International Conference on Hybrid Information Technology, 2009.

[9] Y. E. Kim, A. M. Batula, D. K. Grunberg, D. Lofaro, J. H. Oh, and P. Y. Oh, “Developing humanoids for musical interaction,” in Proceedings of the International Conference on Intelligent Robots and Systems, 2010.

[10] A. J. Eronen and A. P. Klapuri, “Music tempo estimation with k-nn regression,” IEEE Transactions on Audio, Speech, and Language Processing, vol. 18, no. 1, pp. 50–57, January 2010.

[11] B. K. P. Horn and B. G. Schunck, “Determining optical flow,” Artificial Intelligence, vol. 17, pp. 185–204, 1981.

[12] D. Moelants, “Dance music, movement and tempo preferences,” in Proc. 5th Triennal ESCOM Conference, Ghent University, Belgium, 2003.

[13] M. E. P. Davies, N. Degara, and M. D. Plumbley, “Evaluation methods for musical audio beat tracking algorithms,” Queen Mary University of London, Tech. Rep. C4DM-09-06, October 2009.

[2] http://www.music.ece.drexel.edu/research/robots