2010 IEEE-RAS International Conference on Humanoid Robots
Nashville, TN, USA, December 6-8, 2010

A Study on Musical Conducting Robots and Their Users

Brain Geuther, Autumn Breese, and Yunfeng Wang

Abstract— **Robots have started their adventure in the music conducting world recently, though music conducting is an art that takes years of experience and practice for human conductors. This paper presents our primary study on robotic conductors. This study includes the development of simple and affordable conducting robots, and the investigation of human-robot interaction for the specialized group of orchestra musicians. Our study has shown both encouraging and promising results for the application of conducting robots and robots' interaction with musicians.**

I. INTRODUCTION

WITH the increasing usages of robotics-related technology in human world, research on human-robot interaction (HRI) has gained great attention and importance in the last decade [1]. Many of the work have been conducted on technical advances in robotics, behavioral, cognitive, and social aspects of HRI, and the social contexts surrounding HRI. HRI issues have been applied to different groups of people including children, adults, elderly people, robots' operators and collaborators [2]-[5]. In this paper, we exam HRI to a more specialized group of people – orchestra musicians.

In May 2008, Honda's ASIMO robot walked onstage at the Max M. Fisher Music Centre, waved to the audience, and raised its hands to conduct the Detroit Symphony Orchestra. For the next three minutes, ASIMO gestured in an approximation of conducting technique, while the orchestra performed "The Impossible Dream." [6] The event made national and international headlines in the following weeks, as word of the robotic conductor hit the arts world and blogosphere. The event was clearly a success. However, little work has been done on HRI in the context of conducting robots and their conducted orchestra musicians. The work presented in this paper is a contribution to this topic. It first introduces the conducting robot we developed. The results of field trials are then presented and discussed.

II. BACKGROUND ON ORCHESTRAL CONDUCTING

The Conductor's role in a rehearsal setting is extremely important. During a performance, it is the job of the conductor to keep the ensemble together, prepare the next musical moment, and, in some cases, even give the audience something to watch. The job of a conductor during a rehearsal is much more complex.

Prior to rehearsals the conductor chooses a repertoire that is appropriate for the caliber of the ensemble. The conductor spends a great deal of time rehearsing the orchestra. With no prior explanation the conductor uses gestures and beat patterns to express to the ensemble what he would like to hear. If the ensemble is not performing to his wishes, the conductor vocalizes what he would like. The conductor changes the way he or she conducts if the conducting proves to be ineffective. A conductor should be able to communicate everything he would like to get out of the music from the ensemble through motion without verbalizing.

The levels of communication between the conductor and the orchestra can be described as a supervisory control system [7]. The conductor must achieve an authoritative supervision to maintain control of the ensemble. However, the ensemble has the ability to communicate with other musicians within the ensemble. If the ensemble's tempo begins to fall apart, section leaders generally communicate with their section to bring the ensemble back together.

An orchestral conductor conducts the beat pattern with his right hand or with both hands. Dynamics are indicated by the size of the beat pattern, by raising or lowering the left hand, and through facial and body expressions. Cueing is indicated with either hand. The conductor makes eye contact with the section he is cueing and the gives them the preparatory beat and their entrance.

III. CONDUCTING ROBOT

Though a few institutions/organizations could have developed the humanoids like ASIMO, it should not be the obstacles to prevent the study on the interaction between conducting robots and orchestra musicians. To perform this primary study, we developed a simple conducting robot using a commonly used robotics kit.

A. Robotic Platform

The platform selected to develop this robotic system should not have the cost, technology and time constraints for building a robot such as ASIMO. We picked the Lego Mindstorm NXT. This platform is very affordable and allowed for rapid prototyping. Importantly, the

B. Geuther and Y. Wang are with the Department of Mechanical Engineering, The College of New Jersey, Ewing, NJ 08628. Y. Wang is the corresponding author (e-mail:jwang@tcnj.edu)

A. Breese is with the Department of Computer Science, The College of New Jersey, Ewing, NJ 08628.

microprocessor used in the NXT brick is multithreaded. This capability is essential to control the musical conducting behavior of robots.

B. Design Consideration

The robotic conductor does not attempt to idealize and become identical to a human conductor. The design is focused to have some key functions of human conductors. The goal of the robot is to conduct, not idealize a human conductor. The robot's aesthetics should bridge the gap between the human and robotic worlds by combining human and robotic traits.

Based on the above considerations, the physical model is to have two arms. The right arm conducts beat patterns and dynamics. The left arm portrays cueing. With the limitations of three motors with the NXT, it is decided that the right arm has two degrees of freedom using two motors. This gave it a larger workspace to contribute to the wider range of motion required for this arm. The left arm has one degree of freedom using one motor.

C. Robot Construction

The NXT platform has a few limiting factors that must have been taken into consideration in the design and construction of the prototype. The greatest factors include number of outputs, size of program, synchronization, battery life, vibration analysis, and achievable speeds for the motors.

The number of motor outputs was limited to three. In an attempt to create a familiar environment a humanoid design was an ideal situation. Realistic workspace designs were created based upon the number of revolute joints and other transformation parameters. However with a maximum of three revolute joints, the ideal situation could not be achieved. The design decisions ended up with the right arm becoming a 2R manipulator, and the left arm being a 1R manipulator as shown in Figure 1. A gear system was generated for one motor to control two revolute joints at the same time for the right arm. This attempted to get a more unique workspace for the baton arm to achieve more realistic movement and conducting styles.

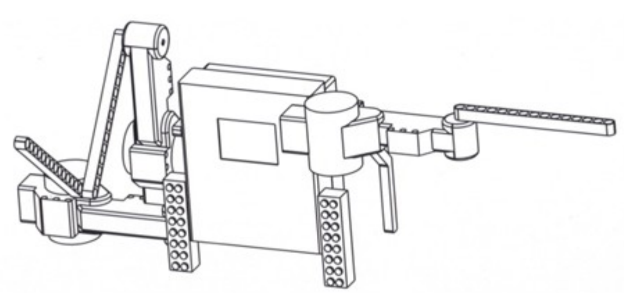

Fig. 1: CAD model of the Proposed Design

Another physical restriction upon the NXT was the size of memory that was on the NXT microprocessor. Early on in the design, there was an attempt to minimize and optimize code so that size was not an issue. This is when it was decided to create a midi-interpreter in which a command binary file was created to place on the limited space microprocessor. The command binary file consisted of a series of steps for the motor controller to interpret. Each step consisted of a beat pattern, time, volume, and cueing value. The beat pattern was used to allow the robot to distinguish which beat it was to conduct. The time value defined the quantity of time to conduct that beat in milliseconds. A relative size of motion was stored in the volume value. The cueing value presented an annotated value for where the right arm is supposed to be pointing.

Synchronization is a difficult objective to achieve in most robotic systems. The NXT microprocessor is multi-threaded, which made synchronization easy. The software that the NXT came with would make the synchronization very difficult. Early on with this project, this difficulty was noticed, and alternative languages were researched. A language called Not eXactly C was found to be compatible with the NXT platform in controlling motors and other functions via the microprocessor [8]. This language is compiled by a compiler called Next Byte Code, which generates a byte-code that is usable by the NXT [9]. The synchronization design that was implemented can be described as a stack based operation. During each time step, significant down to 1ms, a function would send out the commands for the 3 motors at the same time. If the processor could not handle multi-threaded processing, this would have been much more difficult to implement. Each of the commands divided up the time step to execute each beat's respected movements.

Battery life was another limiting factor in the design of the robot. The piece that the robot could conduct could be no longer than an estimated 45 minutes. For testing case, multiple variations of the same piece were performed to attempt to achieve an ideal situation of how to communicate with the orchestra. In order to achieve this, the portion of the piece conducted was restricted to being less than 4 minutes. The original goal was to successfully conduct 50 measures of a specific music score, and this fell well within the power capabilities of the NXT. For the best precision, a newly charged battery was used for the performance.

Vibration analysis is always important when generating a robotic system. If the robot is required to move and stop in small amounts of time, care must be put into the design to minimize the vibrations caused by the conservation of momentum. Small and low mass moment arms were used to minimize this factor. Also, foam fittings were used as a damper for the tolerances in the gears. The foam fittings also doubled as an alteration in appearance, which will be discussed later. As it can be seen in Figure 1, the right arm's left-right rotational motor was placed at the bottom to reduce the quantity of mass that the robot would have on the top part of the frame. The left arm's motor, however, could not have been moved for visibility reasons. This caused some unwanted vibrations while cueing, and could have been solved with a correctional power being applied during programming, but was ignored to keep the precision of 1ms during program execution.

The achievable range of speeds for the motor given a specific power was analyzed. Figure 2 displays the calibration line of the motor, to achieve an estimated velocity at a given power percentage of the motor.

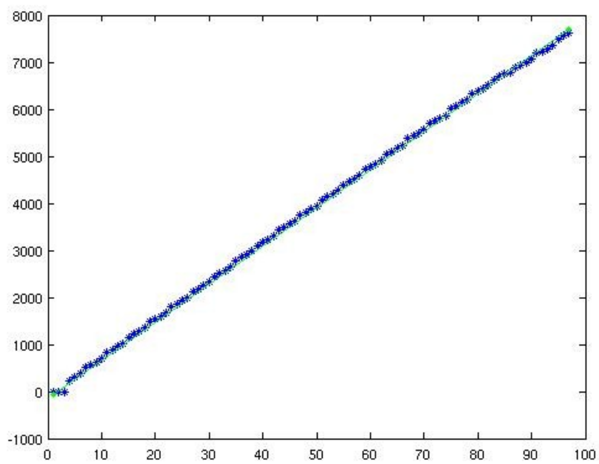

Fig. 1: Power Percentage Vs. Degrees per 10 seconds

The effective range of these motors is from about 10 to 100 percent power. The least square analysis for a linear line to portray the calibration data was used to estimate velocities that could be achieved. One gear reduction was used in the left-right movement of the right arm to increase the precision of the movement. A gear ratio of 5 to 1 worked effectively for this joint. This gearbox also increased the load that could be on the motor, which is why a second motor could be mounted on it. The controlling algorithm for the right arm was designed to estimate motor powers to create a relative size of movement. This could be implemented for any beat pattern, but 2/4 was focused upon perfecting. The algorithm design estimated a power to apply on each motor that would achieve a goal position after the given time step.

D. Appearance

Two types of appearances were created. One is the pure Lego bricks style robots as shown in Figure 3. The other type was enhanced with foam armor. The purpose of this appearance was two-fold: to make the robot more humanoid and make the motions more visible to the orchestra as shown in Figure 4. The center faceplate of the robot was designed to cover the NXT unit. Two arms were constructed and shaped to vaguely model human anatomy. All of the foam pieces were painted a bright yellow to make the arms more visible against a dark background. A piece of carbon fiber was placed in the robot's right arm to mimic a baton.

Both appearances were tested to see if there was a significant difference in the orchestra's reaction.

IV. DATA STORAGE SYSTEM

The data storage system is used to convert a piece of music score to the data that can be used by the developed conducting robot. It was selected to be the Midi file type [10]. This is a relatively known and used music score data storage system. The Midi file designed for this implementation could be viewed as a command file, in which someone generates a file that defines what they would like the robot to do. The important factors for the file were dynamics, beat patterns, and cueing. In the file design, a track was created to supplement each piece of data, as it changed over time. This storage system increased the modularity of the entire robotic system, so that it could perform different pieces depending upon which command Midi file was interpreted.

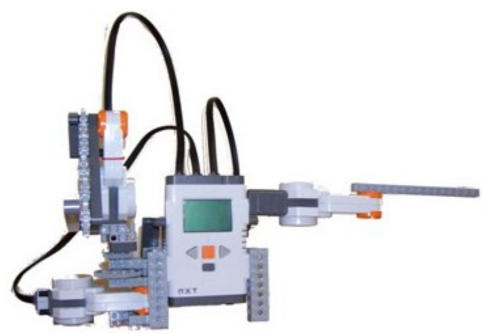

Fig. 3: Pure Lego-brick appearance

Fig. 4: Humanoid feature enhanced appearance

V. FIELD TRIALS

Two field trials were arranged with our college's Orchestra in a separation of one week. Two types of investigations were designed. One is to test the effectiveness of this NXT-based conducting robot. The other is to test the general feelings of musicians on robotic conductors.

The piece of music selected is Beethoven's VII Symphony, movement 2. The robot conducted for about 1 minute. The robot was required to communicate beats, volume, and cues.

Due to the small size of the robotic conductors and nature of the musical score, only the string section (violins, violas, and celli) of the orchestra attended the trials (25 musicians). The movements of the robotic conductors were also projected on a big screen for visibility. Figure 5 shows two video clips of the field trials.

Fig. 5: Video clips of the field trials

VI. FIELD TRIALS RESULTS

A. Robot's effectiveness and musicians' responses

The effectiveness of the developed conducting robots and musicians' responses were obtained by a specially designed questionnaire filled by members of the ensemble. The questionnaire provided scores in five different areas: tempo, dynamics, section cuing, conducting style/beat pattern, and level of interaction (as follows):

1. How effective was the system's portrayal of the piece's tempo?
2. How effective was the system's portrayal of the piece's dynamics?
3. How effective was the system's portrayal of the piece's section cueing?
4. How effective was the system's conducting style and beat pattern?
5. What was the level of interaction between the conducting system and you, the orchestra?
6. Comment upon the appearance of the robot
7. Any other comments.

For the questions 1 to 6, they were rated on a scale from 1 to 10 with 1 meaning "not at all" and 10 "very".

i. Results from Field Trial 1

Table 1 shows the average, maximum, and minimum ratings for each comment.

Table 1 Results from field trial 1

Trial 1 Strings			
	Average	Minimum	Maximum
Tempo	7.64	5	10
Dynamics	4	1	8
Cueing	4.6	1	9
Beat Pattern	6.92	4	10
Interaction	4.6	1	10

Some of the poorer ratings were due to the musicians disagreeing with the chosen tempo. On average, musicians either did not comment on the tempo or they felt that it was clear. One musician felt that the motions needed to be bigger.

Many musicians felt that they could see the robot trying to indicate dynamics, but the dynamics needed to be exaggerated to be clearer. Others felt they simply could not see the difference in the size of the conducting pattern. A suggestion received was to make the piano and pianissimo gestures even smaller. One musician wrote that they did not understand what the robot was communicating, in spite of the directions provided by the group.

Most musicians felt that cueing was too abrupt; one musician felt they were too slow. Several musicians focused on the fact that the second cue was pointing at the second violins instead of the violas. One musician suggested that cues would be easier to understand if the robot could make eye contact with the section of the ensemble it was cueing.

Musicians indicated that they thought the beat pattern was clear but too systematic or metronomic. One commented on how the upbeats were much larger and clearer than the downbeats.

The musicians rated the level of interaction between the robot and the ensemble. However, the robot was not receiving any input and was pre-programmed. Some musicians noted that the robot was not reacting with the orchestra while others felt that the orchestra was interacting with the robot.

Many musicians preferred the enhanced appearance and felt that it made it more "human" and easier to follow. A few musicians commented that the yellow made it easier to follow. The musicians who elaborated on why they preferred the robot without the foam seemed to feel that the foam was nothing more than a distraction.

ii. Results From Field Trial 2

Based on the musicians' responses from the first field Trial, several improvements have made on the robot. A week later, the second trial was performed for the improved robot. An identical questionnaire was distributed after the second field trial. Table 2 depicts a summary of these results.

Table 2 Results from field trial 2

Trial 2 Strings			
	Average	Minimum	Maximum
Tempo	8.2	5	10
Dynamics	7.34	3	10
Cueing	8.39	2	10
Beat Pattern	7.04	3	10
Interaction	5.7	1	10

Most musicians wrote that the beat pattern was clear, but there were problems with understanding the first few measures. While a few individuals wrote that they did not notice dynamics, a majority of the ensemble wrote that they were clear and/or better than last week. Comments were still received about the fact that the 2nd violins were cued instead of the violas. The majority of the musicians agreed that they interacted with the robot, but that the robot was not interacting with them. All of the ratings are higher than the previous week.

iii. Summary of effectiveness testing

The testing of the robotic conductor was a success. The orchestra was able to understand and follow the beat pattern to play at the intended tempo. Most of the cues were correct and aided the musicians in entering at the correct time. Even though dynamics were somewhat misunderstood the difference in volume was still noticeable. Meaningful feedback was received from the ensemble that facilitated improvements for subsequent tests. At one point, the human conductor had to step in to give the robot authority over the ensemble. The section leaders were very effective in assisting the rest of the musicians in accurately following the robotic conductor. The robot was able to successfully conduct 101 measures of Beethoven's VII Symphony movement 2.

B. General impression on robotic conductor

After both field trials, a feedback form on the acceptance and interesting level of robotic (non-human) conductors were given to members of ensemble. They were rated on a scale from 1 to 10 with 1 meaning "not at all" and 10 "very".

Orchestra members found being conducted by non-human conductors only moderately acceptable overall with responses ranging from Unacceptable to Completely Acceptable with Average=5.45, Mode=6, Range =1-10, and Standard Deviation =2.37.

In spite of that, they found the experience fairly interesting overall with responses ranging from Boring to Exciting with Average=7.31, Mode=10, Range=1-10 and Standard Deviation=2.75.

VII. Conclusion

This paper presented an empirical study on the interaction between robotic conductors and orchestra musicians. It also introduced a design solution for constructing a Lego Mindstorm NXT based conducting robot. Though the developed robot for this study is like a robotic metronome without the emotions of human conductors, the HRI results using this robot provided constructive information for the development direction of conducting robots, and the musicians' perception on being conducted by non-human conductors.

Acknowledgments

The authors would like to thank the orchestra of The College of New Jersey and its conductor Dr. Philip Tate for their collaboration and support. The authors would also appreciate for the useful suggestions from Dr. Andrea Salgian, Dr. Christopher Ault and Dr.Teresa Nakra. This work was supported by the National Science Foundation under Grant No. 0855973.

References

[1] T. Fong, I. Nourbakhsh, K. Dautenhahn, "A survey of socially interactive robots," *Robotics and Autonomous Systems*, Vol 42, pp. 143-166, 2003.

[2] T. Kanda, T. Hirano, D. Eaton, "Interactive robots as social partners and peer tutors for children: a field trail," *Human Computer Interaction*, Vol. 19, pp.61-84, 2004.

[3] P.J. Hinds, T.L. Roberts, H. Jones, "Whose job is it anyway? A study of human-robot interaction in a collaborative task," Human Computer Interaction, Vol. 19, pp.151-181, 2004.

[4] C.D. Kidd, W. Taggart, S. Turkle, "A sociable robot to encourage social interaction among the elderly," *IEEE International Conference on Robotics and Automation*, Orland, Florida, pp. 3972-3976, 2006.

[5] Kanda, Takayuki, " An affective guide robot in a shopping mall.," *ACM/IEEE International Conference on Human-Robot Interaction*, La Jolla, California, pp. 173-180, 2009.

[6] "Honda ASIMO Robot conducting the Detroit Symphony Orchestra", May 13, 2008,

http://world.honda.com/news/2008/c080423ASIMO/,
http://www.youtube.com/watch?v=Cf5szwz6Qzc

[7] G. Johannsen, "Human Supervision and Control in Engineering and Music," *Journal of New Music Research*, Vol. 31, pp.179-190, 2002.

[8] J. Hansen (2007) Not eXactly C (NXC) Programmer's Guide Version 1.1.1 b33 Available: http://bricxcc.sourceforge.net/nbc/nxcdoc/NXC_Guide.pdf

[9] J. Hansen (2007) NeXT Byte Codes (NBC) Programmer's Guide Version 1.1.1 b33 Available: http://bricxcc.sourceforge.net/nbc/doc/NBC_Guide.pdf

[10] Standard MIDI File Format Available: http://faydoc.tripod.com/formats/mid.htm