Music Playing Robot

A. Alford, S. Northrup, K. Kawamura, K-W. Chan
Center for Intelligent Systems
Vanderbilt University
Nashville, TN 37232

J. Barile
Microsoft Corp.
Redmond, WA

Abstract

In our lab at the Center for Intelligent Systems, we are currently developing a dual-armed humanoid service robot called ISAC. One of our focus areas in this research is the use of humanoid robots for entertainment. One application we have developed is using ISAC to play a musical instrument called the Theremin.

In this paper we present our Theremin-playing robot and describe the software system used to control the robot. The software runs on commercial PC equipment under the Windows NT operating system. The robot, being a rubbertuator-based humanoid, is quite well suited for entertaining with the Theremin. The motion is natural and vibrato is achieved by oscillating the muscles. The robot achieves perfect pitch and thus can play notes more accurately than novice Theremin players.

1 Introduction

In our lab at the Center for Intelligent Systems, we have developed service robots for aiding the disabled [3, 6, 10]. We are currently developing a dual-armed humanoid service robot called ISAC. One of our focus areas in this research is the use of humanoid robots for entertainment. One application we have developed is using ISAC as a drawing robot [12]. We are currently developing a system using ISAC to play a musical instrument called the Theremin [8, 11].

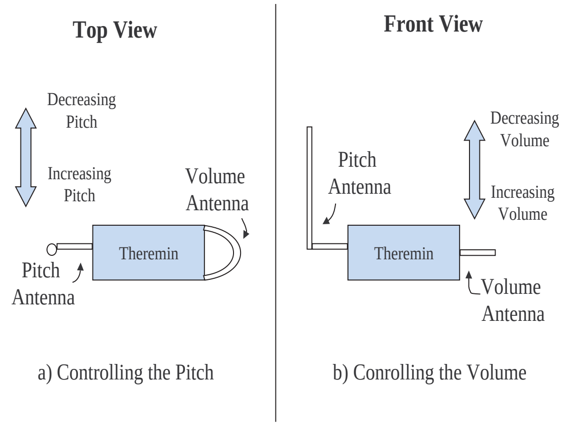


Figure 1: Controlling the Sound from the Theremin

The Theremin is an electronic musical instrument which is played without physical contact. The Theremin was invented in 1919 by Lev Sergeyevich Termen, who later changed his name to Leon Theremin. A typical Theremin resembles a box with two protruding metal antennae, one on either side. A single musical tone is controlled through interactions with the antennae. One antenna controls pitch, the other volume. As the musician's hand approaches the pitch antenna, the pitch of the sound increases (Fig. 1(a)). As the other hand approaches the volume antenna, the volume of the sound decreases(Fig. 1(b)). Most Theremins are capable of producing a 5 octave range.

The Theremin is difficult to play for several reasons. Since the player never touches the instrument, there is no physical point of reference to associate with specific tones, and perfect pitch is required. In addition, the response of the capacitive antennae can vary with temperature and humidity. Only a handful of people have ever mastered the Theremin, most notably Clara Rockmore [4].

Other musical robots have been developed in the past. For example, Wabot-2 [2], a piano-playing robot developed at Waseda University in Japan. MUBOT, the Musician Robot, was a project at the University of Electro-Communications in Japan [1, 5]. Several MUBOTs were developed, which could play the recorder, violin, or cello. Kurzweil [7] describes several examples of robots and Artificial Intelligence systems used in music as well as other arts.

2 System Description

Overview

Fig. 2 shows the system overview. A human plays notes on the keyboard, which sends information to the

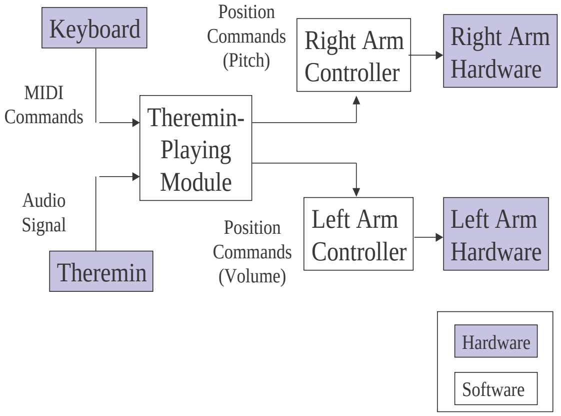


Figure 2: Theremin-Playing Robot System

Theremin-Playing Module (TPM) using the General MIDI format. This information indicates which note has been played. The TPM uses this code to determine the desired pitch, or frequency, in Hertz. The keyboard also sends messages indicating the volume of the note, as well as control messages that affect the motion of the robot's arms. The TPM samples the audio signal from the Theremin, and using the desired pitch and volume information, computes position commands to send to the Arm Controller modules. These modules move the arms end-effectors to the desired positions.

Except for the robot manipulators, all equipment used in the system is available commercially. The Theremin is a Big Briar EtherWave Theremin. The keyboard is a Yamaha CS1x Control Synthesizer. However, the system will work with any Theremin with a similar antenna arrangement, and with any keyboard which provides General MIDI output. The software runs under Windows NT on a PC which has a standard sound card.

The robot manipulators are SoftArms, which are powered by artificial pneumatic muscles called McKibben actuators. These artificial muscles contract when inflated. A pair of muscles is used to move a single joint; this is analogous to the operation of human flexor/extensor muscles. The SoftArms are naturally flexible and safer to use in close proximity to humans than are typical electro-mechanical robots. The disadvantage, however, is that the SoftArm response is much slower and less accurate than electro-mechanical systems.

Arm Controller Modules

The position of each of the robots arms is controlled by a separate software module. This module computes the inverse kinematics which converts Cartesian position commands to angular position commands. The resulting desired angular positions are realized using a nonlinear control loop [9].

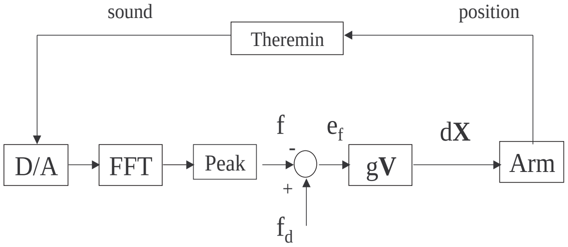


Figure 3: Audio Feedback Loop

Theremin-Playing Module

The core of the TPM is an audio-servo loop (Fig. 3). This loop moves the robot's right arm (the Pitch arm) such that the Theremin emits a note of a desired pitch. The audio signal from the Theremin is sampled by the PCs sound card at 8kHz. The sampled audio data is then transformed to frequency information using a Fast Fourier Transform (FFT). 512 data-points, zero padded to 2048, are transformed each time. This provides a spectrum with frequency resolution of 3.90625 Hz. The frequency component with the largest magnitude is then taken to be the fundamental frequency of the Theremin's audio signal, represented as f in Fig. 3.

The difference e_f between the desired frequency f_d and the actual frequency f is calculated. The change in arm position, $\mathbf{dX}$ is computed using the following equations:

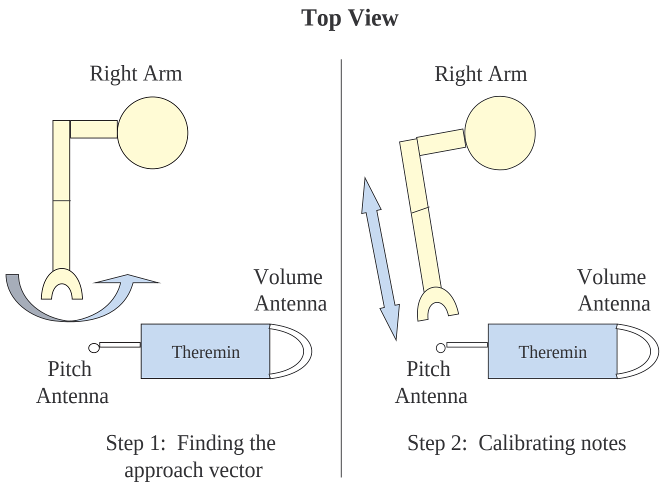
$$\mu = \log(1.0 + |e_f|) \tag{1}$$

$$g = \frac{K_p \ \mu}{f_d} \tag{2}$$

$$\mathbf{dX} = e_f \ g \ \mathbf{V} \tag{3}$$

The incremental arm movement, $\mathbf{dX}$, is always along the approach vector $\mathbf{V}$. This distance, $e_f \ g$, is inversely proportional to the desired frequency, since the Theremin is more sensitive to position changes at higher frequencies. A non-linear factor, μ, is included to damp the motion as the error e_f decreases; this is to avoid overshoot.

The approach vector $\mathbf{V}$ is found using a calibration which is performed when the TPM is first activated (Fig 4). The purpose of the calibration is to find the arm positions that correspond to the various notes of the scale. Since there is no unique position for each



Step 1: Finding the approach vector

Step 2: Calibrating notes

Figure 4: Feedforward Calibration Procedure

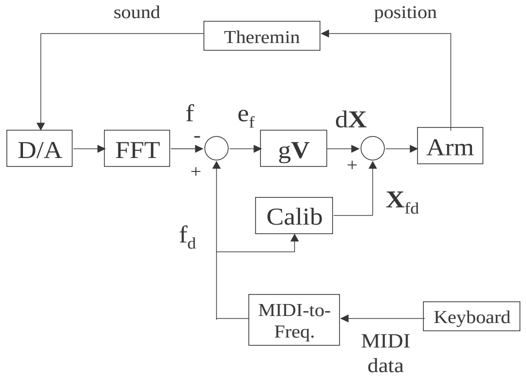


Figure 5: Audio Feedback Loop with Calibrated Feed-forward

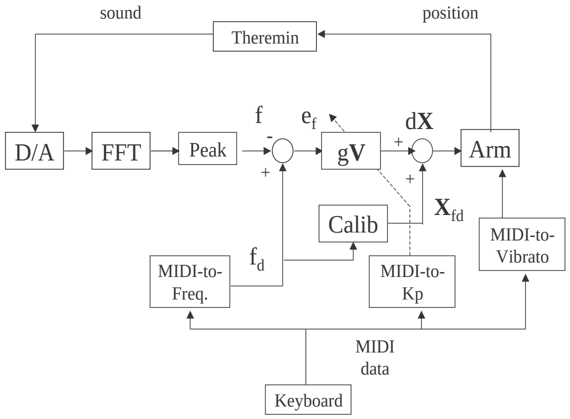


Figure 6: Complete Pitch Control System

note, the TPM first finds an approach vector $\mathbf{V}$. This vector is found by rotating the robot's base joint. This has the effect of panning the robots end-effector in a horizontal arc near the Theremin's pitch antenna. As the end-effector moves, the TPM monitors the fundamental frequency f of the Theremin audio signal, maintaining a running maximum measured value and the corresponding base joint angle. Since the f increases as the arm moves closer to the pitch antenna, the angle of highest audio frequency is the angle of closest approach. When the arc is complete, the base joint is returned to the angle of closest approach. This angle is denoted θ_m . The approach vector is then defined as $\mathbf{V} = (\cos\theta_m, \sin\theta_m, 0)$.

To calibrate to positions along the approach vector which correspond to notes of the musical scale, the desired frequency f_d is set to the lowest note of the scale. The arm is moved along $\mathbf{V}$ until the frequency f of the Theremin audio signal equals f_d. The position of the arm is then stored with f in a table. The calibration continues with the successive notes of the scale until the entire scale has been calibrated.

Fig. 5 shows the audio servo loop with feed-forward information using the lookup table generated by the calibration. When a human plays notes on the keyboard, the keyboard sends MIDI Note-On messages to the TPM. Each message contains a code which indicates the key. This code is used by the TPM as an index into a table of frequencies to determine the desired frequency.

In addition to piano keys, the MIDI keyboard used in the system also has several control knobs. These also generate MIDI messages which the TPM uses to vary certain aspects of the robot's play. One of these control knobs controls the proportional constant K_p. Two other control knobs are used to add vibrato to the music. In this mode, a sinusoidal signal is added to the robot's joint control signals. The control knobs vary the amplitude and frequency of the sinusoidal signal. Fig. 6 shows the complete pitch control system.

The volume of the Theremin audio signal is increased or decreased by moving the robot's left arm (Volume arm) away from or closer to the volume loop antenna (Fig. 7). Two positions, one close to the antenna and another approximately 70cm higher, are predefined for Note Off and Note On, respectively. During normal play, when a new note command is received, the Volume arm is first moved to the Note Off position, then the desired frequency f_d is sent to the pitch control system. When the Pitch arm has moved to the calibrated feedforward position, the Volume arm is returned to the Note On position and the

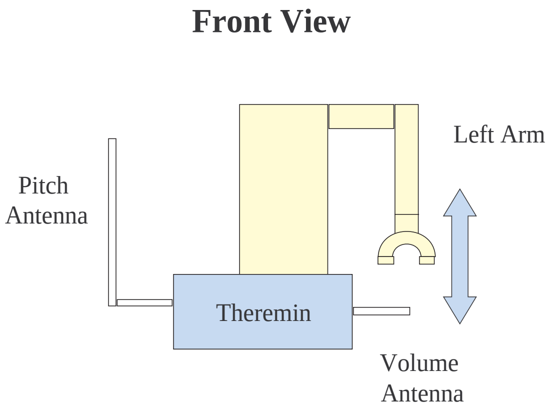


Figure 7: Volume Control

audio feedback loop ensures that the proper pitch is attained. Besides the pre-programmed Note On and Note Off positions, the MIDI keyboard can also send volume control messages which move the volume to positions between the Note On and Note Off positions.

3 Performance

Although capable of playing a pre-recorded musical score, ISAC is intended to be an extension of the human musician—much like a musical instrument. In contrast to highly accurate electro-mechanical music-playing robot systems such as Wabot-2 [2], ISAC is not intended reproduce exactly a written score. Instead, the combination of the human at the keyboard, ISAC, and the Theremin create a unique musical instrument.

As mentioned previously, the Theremin is a difficult instrument to play. The Theremin's sensitivity to environmental changes requires the player to have "perfect pitch"—that is, the ability to identify exactly the note being played by an instrument. The TPM effectively gives ISAC "perfect pitch." The performance of the robot then becomes function of its servo control bandwidth. ISAC's softarms are not ideal in this respect; their control bandwidth is quite limited. However, the arms are human-like in appearance and therefore may be more appealing to watch from an entertainment perspective. Additionally, the flexing of the muscles during the oscillatory vibrato is quite natural appearing and easy to obtain with a rubbertuator-based humanoid robot. Electro-mechanical arms can vibrate (although it is not conducive to longevity) but they are very stiff and "robot-like" in high frequency movement.

A video of ISAC playing the Theremin was shown at a Theremin Festival in Portland, Maine, in 1997. The conference attendees were mainly Theremin musicians with an appreciation for the difficulties in playing the Theremin. They enjoyed the robot's perfect pitch and were amused by the vibrato—vibrato is often over-used by beginning Thereminists to cover up "sour" notes.

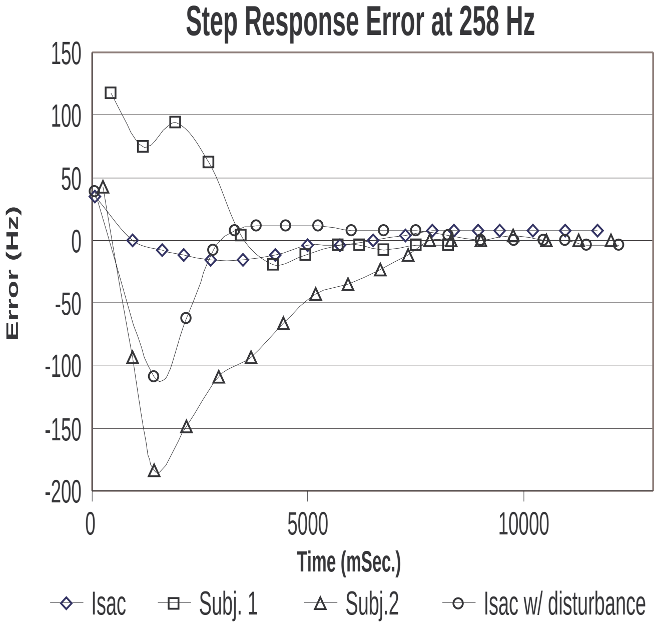


Figure 8: Error for ISAC and Musically Experienced Humans

To illustrate the performance of ISAC's Theremin-playing ability, we first asked several people to try playing the Theremin. These test subjects were classified in two groups: those having some experience with musical instruments and those with none. Using the keyboard, the experimenter played a sequence of reference notes, while the test subject tried to achieve the same notes on the Theremin. The experimenter would hold each note until the test subject reached the reference note. We recorded the frequency of the reference note along with the frequency played by the Theremin. While the subject was playing the Theremin, the experimenter could see the FFT results on the monitor to know when to terminate the test. The test subject, however, could only use the audio signals from the Theremin and the keyboard for comparison. The sequence consisted of the following notes: C, 258Hz; F, 348Hz; C, 520Hz; G, 391Hz; E, 328Hz; A, 219Hz.

Before ISAC played the same sequence of notes, the TPM calibration was performed. ISAC then played the sequence of notes and the same data was recorded

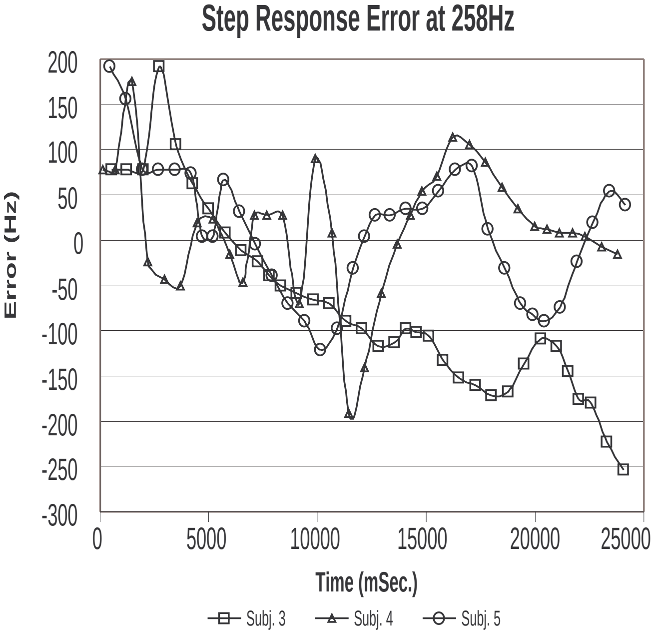


Figure 9: Error for Musically Inexperienced Humans

as for the humans. Then, the TPM calibration was disturbed by moving the Theremin, and the experiment was repeated.

Fig. 8 shows ISAC and the two best musically experienced humans playing the same note. The humans and ISAC had comparable settling times, but the mean squared error for the humans was much greater than for ISAC. However, when ISAC's calibration was disturbed by moving the Theremin, ISAC's mean squared error approached that of the best human subject tested.

For completeness, we include Fig. 9 that shows the attempts of three musically inexperienced humans at playing the Theremin. One of the subjects converged to the correct frequency at about 25 seconds, but the other two never converged without external coaching.

4 Conclusion

In this paper we presented a Theremin-playing robot and described the software system used to control the robot. The software runs on commercial PC equipment under the Windows NT operating system. The robot, being a rubbertuator-based humanoid, is quite well suited for entertaining with the Theremin. The motion is natural and vibrato is achieved by oscillating the muscles. The robot achieves perfect pitch and thus can play notes more accurately than novice Theremin players.

References

[1] http://www.kajitani.mce.uec.ac.jp/mubot/.

[2] Special issue on wabot-2. *Bulletin of Science and Engineering Research Laboratory*, (112), 1985.

[3] Sugato Bagchi and Kazuhiko Kawamura. Isac: A robotic aid system for feeding the disabled. In *AAAI Spring Symposium on Physical Interaction and Manipulation*, Stanford, CA, 1994.

[4] Albert Glinksy. *The Theremin In The Emergence Of Electronic Music.* PhD thesis, School of Education, Health, Nursing, and Arts Professions, New York University, 1992.

[5] Makoto Kajitani. Development of musician robots. *Journal of Robotics and Mechatronics*, pages 254–255, 1989.

[6] Kazuhiko Kawamura, D. M. Wilkes, Todd Pack and Magued Bishay, and Jason Barile. Humanoids: Future robots for home and factory. In *International Symposium on Humanoid Robots*, pages 472–475, Waseda University, Tokyo, Japan, October 1996.

[7] Raymond Kurzweil. *The Age of Intelligent Machines.* MIT Press, Cambridge, MA, 1990.

[8] L. Lange. The theremin strikes a chord with engineers. *EE Times*, pages 1,156, July 1997. Issue 964.

[9] R. T. Pack, M. Iskarous, and K. Kawamura. Comparison of fuzzy and nonlinear control techniques for a flexible rubbertuator-based robot joint. In *Proceedings of the 1994 International Fuzzy Systems and Intelligent Control Conference, Institute of Fuzzy Systems and Intelligent Control*, pages 361–370, March 1994.

[10] Robert Pack and Moenes Iskarous. The use of the soft arm for rehabilitation and prosthetics. In *RESNA 1994 Annual Conference*, pages 472–475, Nashville, TN, 1994.

[11] Motoko Rich. This musician could also play air guitar. *Wall Street Journal*, March 1998.

[12] A. Srikaew, M. Cambron, S. Northrup, R. A. Peters, D. M. Wilkes, and K. Kawamura. Humanoid drawing robot. In *Proceedings of the IASTED International Conference on Robotics and Manufacturing*, Banff, Canada, July 1998.