Constructing a Trumpet Playing Robot with a Simplified Lip Model

Andrei Faitas
Institute of Informatics
University of Oslo
Oslo, Norway
andreif@ifi.uio.no

***Abstract*—The physics behind the way humans play brass wind instruments are both delicate and complex, and constructing a robot that can produce similar results poses an interesting challenge. In this body of work I mean to explore the implementation of a simplified model of human lips in order to discover if the most basic requirements for making a trumpet produce a tone can be met with common resources and a simple Machine Learning algorithm. This paper presents the construction of a prototype robot from 3D-printed plastic parts and artificial silicon lips, controlled by a Multi-Layer Perceptron algorithm.**

I. INTRODUCTION

The human mouth is a complicated array of muscles which all interact in a complex manner that is hard to model and artificially reproduce. As such, creating an artificial trumpet player becomes comparatively hard. However, in [Rod95] authors explored with success a simplified physical model of how human lips interact with brass wind instruments to produce tones. In it they identified the most essential features and mechanisms required. Using this model as a basis I have designed and constructed a prototype that implements said features and mechanisms in an attempt to artificially recreate a basic trumpet player. The construction consists of 3D-printed plastic parts, silicon to act as artificial lips and plastic tubing connecting the mouth and lips to an air compressor.

The goal for the robot is also relatively simple as I initially only want to explore whether or not pursuing further improvements is at all worth while. The robot is only meant to be able to play a single note; if this cannot be done with the resources available then I must conclude that I am not able to implement the proposed lip model. As such, control over the valves of the instrument will not be necessary.

I attempt to achieve this by supplying constant air pressure to the robotic mouth and then use a Multi-Layered Perceptron (MLP) neural network to adjust the lip tension with a servo until the robot identifies a B^b (the musical note) via its microphone, i.e learning to play that particular note. In [RV96] it was shown that the amplitude and oscillation of the dynamic system the simplified model forms could be precisely predicted, but the contribution of this paper is an exploration of whether an MLP can learn to make the same predictions.

Discovering whether the simplified model is sufficient for successfully playing a trumpet holds interesting potential as it can give strong implications for a general method for creating artificial brass musicians, able to play not only trumpets, but any manner of brass wind instruments. Equally valuable is the prospect of discovering the limitations of this model; what pitches are not attainable and what quality of tone can be produced.

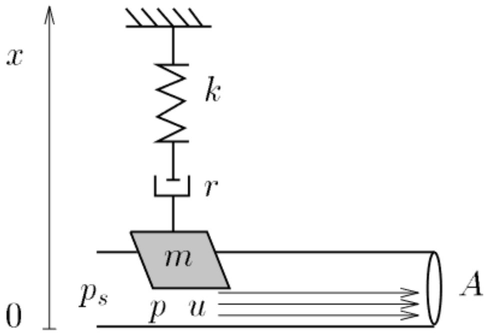

Fig. 1: Schematic of the basic single mass model as proposed in [Rod95]. Image is from their follow-up paper [citation]

II. BACKGROUND AND RELATED WORK

A. The Simplified Lip Model

The flow of air through a pair of lips can, and is typically, modelled as a pair of spring loaded masses opposite one another, with the air flowing between the masses being modulated by their relative distance. In [Rod95] authors propose a single mass model for playing wind instruments, explaining that in [Mar42] it was shown that the displacement of the lower lip is negligible and so only the oscillation of the upper lip is of concern.

Figure 1 shows a schematic of the model, where m is a spring loaded mass displaced by a spring with stiffness k and damping coefficient r. As stated in [citation] the mouth pressure p_s is considered to be steady, whereas the air flow in the mouth is neglected. The upper lip displacement, the pressure and the air flow under the upper lip are designated by x, p and u respectively. Even though the simplifications in this model are substantial, the authors have shown that it captures the most essential sonic characteristics of playing brass wind instruments.

B. Some physics of the Trumpet

The trumpet is a long, conical tube which flares out sharply at the end. This tube acts as a resonator for the oscillating air flow produced by the lips of the player ”buzzing” their lips, essentially rapidly opening and closing them. The oscillating air creates a standing wave through the instrument with nodes at the ends, which means the longer the instrument is the further apart these nodes will be and the longer the wavelength, producing a deeper tone. The standard B^b trumpet will have a length of 1.48 meters which will produce a standing wave with its fundamental harmonic frequency at approx. 233 Hertz yielding the note B^b (B flat). Doubling the frequency of lip oscillation creates a third node in the middle of the standing wave. This is called the 2nd harmonic frequency which the human listener perceives as an octave higher in pitch. In this way the player can change the tone by adjusting the tension in their lips, allowing them to play the 3rd and 4th harmonic frequency and so on. The introduction of the three-valved-trumpet in the 1600’s allowed the player to mechanically redirect airflow through alternate tubes in the instrument, effectively modulating the length and allowing for a complete range of notes to be played by a single horn.

III. PROTOTYPE DESIGN

In this section I present and explain the design of the prototype, i.e my idea for solving the problem. It will detail the process and the components used in the construction.

A. System Overview

My proposed solution requires the following components in order to be realized:

- A rigid body to both fasten the artificial lips and secure the instrument
- Artificial lips that can mimic the flexible texture of human lips
- A motorized linear translation joint to press the lips together
- A microphone to pick up the sound it might produce
- An MLP machine learning algorithm to learn how to play a B^b
- A compressor to supply an air flow through the lips

In an ideal sense, these components should be enough. This means that it is the details of the system design that will ultimately decide the success of these experiments. The artificial lips, for instance, poses a great challenge in that the qualities of the chosen material (silicon) is beyond the author to identify. The stiffness of the material is directly related to the oscillation of the vibrating upper lip; without knowing anything about optimal stiffness or even how to measure it, it must be discovered through experimentation. It is also unclear if a microphone will be too affected by environmental noise for it be used for effective training of the MLP.

An overview of the system design is shown in figure 2. The MLP algorithm first receives a one-hot encoded vector $x \in \{0, 1\}^{12}$, where each of the 12 elements corresponds to one of the 12 notes in the western musical system. Currently I will restrict the input to only being B^b which corresponds to the second element in the vector. In other words it will only receive the vector $x = [0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]$. This will then translate via the MLP to a certain linear force by the servo pressing the upper lip against the bottom lip and modulating lip tension. The artificial lip then begins to oscillate according to the lip tension and the air flowing between them, causing sound to be produced in the instrument. This sound is then captured by the microphone. The system then performs a frequency analysis to identify the fundamental frequency of the sound that was played. The difference between this recorded frequency and the target $B^b \approx 233$ Hertz is then used in the back propagation part of the MLP algorithm to adjust the force the servo delivers down to the lips. This is repeated until the system produces the target note B^b. The supplied air flow from the compressor is constant.

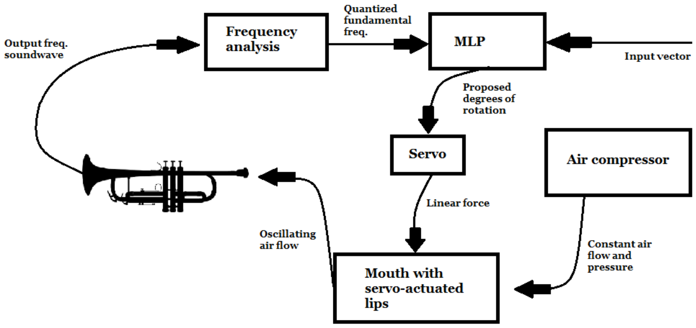

Fig. 2: A schematic showing the system design.

B. 3D model

The rigid mouth fixture and silicon mold for the lips were both designed with Autodesk Fusion360 CAD software. The mouth fixture (see fig. 3 and 4) consists of three main parts; the lower jaw which features a connector between the mouth and air compressor, the upper jaw which features braces for attaching a servo motor to drive the lip closing, and a conical tube channeling the air from the back of the lower jaw to the lips like a mouth cavity.

The silicon mold (fig. 5 and 6)is designed to yield lips where the opening is smaller than the mouthpiece of the instrument, ensuring that air will not leak out from the corners of the mouth. It is also made to ensure the lips angle slightly outward from the mouth so as to more closely resemble the parallelepipedic shape of the mass described in the simplified model. (See fig. 1)

Finally, a shaft and end effector (fig. 7) was designed to create a specialized piston in order to translate the rotational force of the servo to a linear one.

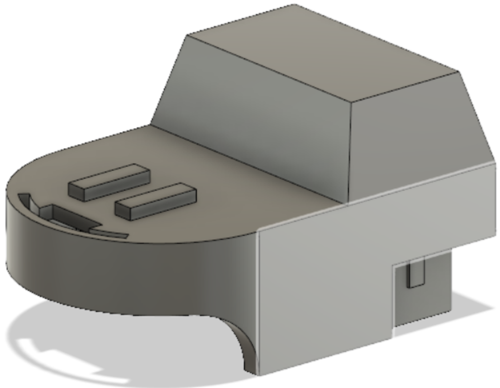

Fig. 3: Upper jaw with braces for servo attachment and opening for piston end effector.

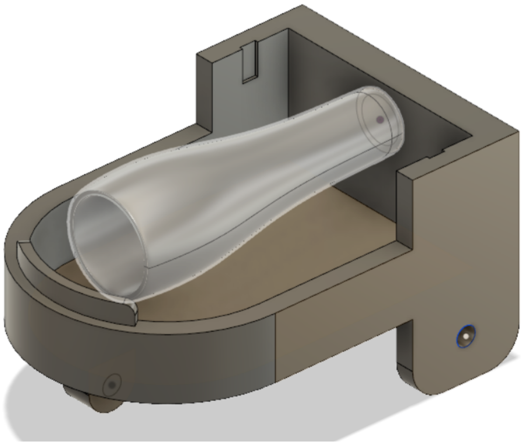

Fig. 4: Lower jaw with conical tube.

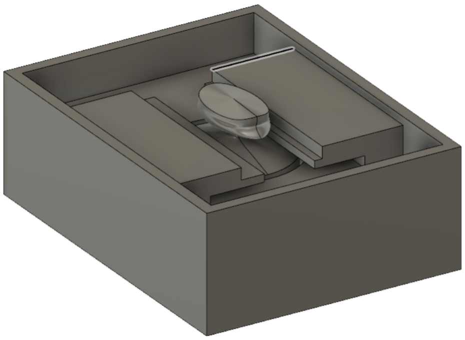

Fig. 5: Mold for the artificial lips.

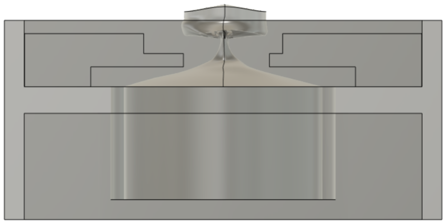

Fig. 6: Side view of the lip mold with a lowered opacity for cross section like look.

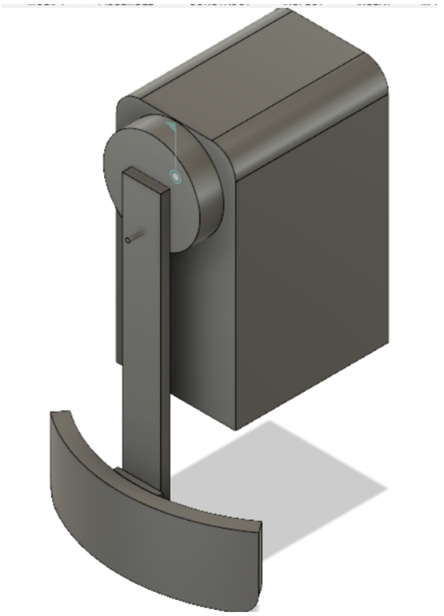

Fig. 7: Shaft and end effector, combined to create linear action piston. For reference it is here shown attached to a model of the actuating servo.

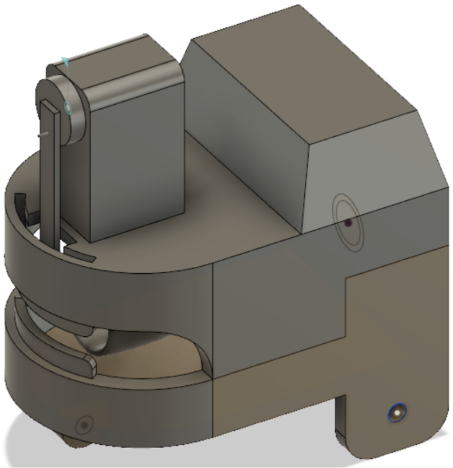

Fig. 8: Assembled model.

Fig. 9: The realized design.

C. *Realized Prototype*

The realized prototype was constructed using 3D-printed parts of PLA plastic, a Dynamixel AX-12 servo motor, and silicon. The trumpet is in fact a flugelhorn which strongly resembles a trumpet, the difference being a slightly wider bore which should not impose any noticable difference in how the system should be designed; if it works with a flugelhorn it should work with a trumpet.

IV. CONCLUSION AND FURTHER WORK

Experiments remain inconclusive as none have been performed. Certain assumptions can however be made based on the literature referenced in this paper. Firstly, the basic model proposed in [citation] is simple enough that its implementation is entirely feasible, and the results from their own experiments are very promising. Secondly, looking at how humans increase lip tension to play higher notes, it seems safe to assume there is some strictly increasing, exponential relationship between tension and frequency - such a relationship is well within the capabilities of an MLP to learn.

Further work includes improving lip mold design to create artificial lips that perform better and fit the rigid mouth fixture better, designing an adjustable fixture for the instrument that allows one to control how hard the instrument is pressed against the lips, and measure results from experiments. Furthermore, the MLP should be expanded to allow for other training inputs besides the fundamental harmonic frequency, and to allow for autonomous adjustment of air pressure supplied by the compressor as higher notes require higher air velocities, though only if the prototype first succeeds at learning to play the fundamental.