

A Real Time Visual Feedback Voice Training System for Teachers

Leung, Fung Yee Elaine¹ and Rossiter, David²

¹ fyleung@ust.hk

² rossiter@ust.hk

Department of Computer Science,
The Hong Kong University of Science of Technology

ABSTRACT

A teacher's professional role is dependent on the effective and competent usage of his or her voice. However teachers in Hong Kong are often deficient in their ability to project their voices clearly and confidently. There is an obvious need for some form of voice training to be introduced to enhance their proficiency in their roles as educators. To address this, we have introduced a voice training system to help solve these problems. The ultimate goal of this project is to promote awareness and provide training for improving enunciation and voice projection among teachers. A real time visual feedback voice training system (RV) has been built for teachers to help them acquire effective vocal techniques. Following the simple instructions from the system, they can visualize their voices and make improvements accordingly. This project can be extended to students as well as those who want to be skilled and confident public speakers and even those who are training as singers.

Keywords

Real time, visual feedback, voice training, voice projection, singing

APPROPRIATE TRAINING IN CORRECT VOCAL USAGE FOR TEACHERS

Prevalent Problems of Vocal Abuse and Misuse

Teachers are among those professionals who make great demands on their voices; they talk all day long in the classroom and have to project loud enough so that all their students can hear them clearly. Unfortunately, this stress on the voice can lead to health problems.

The two common health problems are vocal abuse and misuse. Vocal abuse is any behavior or occurrence that strains or injures the vocal cords. This may include excessive talking, throat clearing, coughing, screaming, or yelling. Vocal misuse is improper voice usage such as speaking too loudly or at an abnormally high or low pitch.

Frequent vocal abuse and misuse can damage the vocal folds and cause temporary or permanent changes in vocal function, voice quality, and possible loss of voice.

To address these issues, we have developed a system to facilitate voice training for teachers. Teachers will be guided to perform various speaking tests and analysis with the assistance of RV. They may then develop a better concept of correct voice usage.

RV, the Voice Training System

RV is designed to assist teachers in visualizing their voices. The results shown by the system serve as an indicator of how well or badly their voice is being generated. For example, in the simple mode analysis shown below, teachers are encouraged to use their voice to make a bigger circle display. With this visual assistance, they can grasp the abstract techniques of voice usage easily and then deliver a clearer and better lecture.

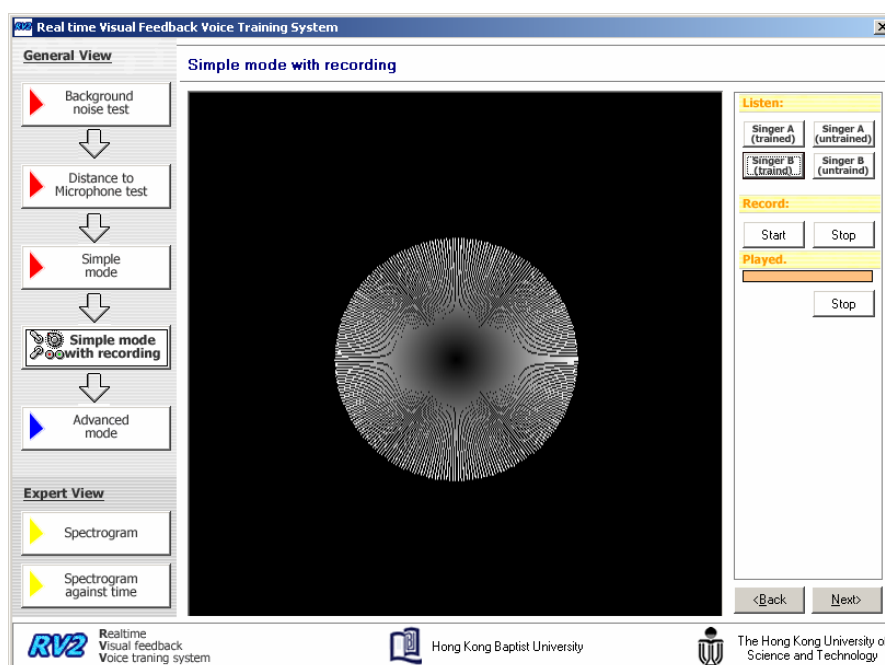


Figure 1. Screen shot from RV, simple mode with recording

Background Theory

When we sing or speak, sound is produced. Sound is comprised of various frequency bands. The typical range of frequency that human beings can hear is 20 Hz to 20,000 Hz. However, our ears are not equally sensitive to all frequencies generated. Some can be heard more distinctly than others.

The following spectrogram is collected from the singing of a trained singer. It shows the signal amplitude against frequency. The vertical bands represent various frequencies from 0 Hz to 6000 Hz. The taller the band, the stronger the signal perceived at that specific frequency (refer to Figure 2).

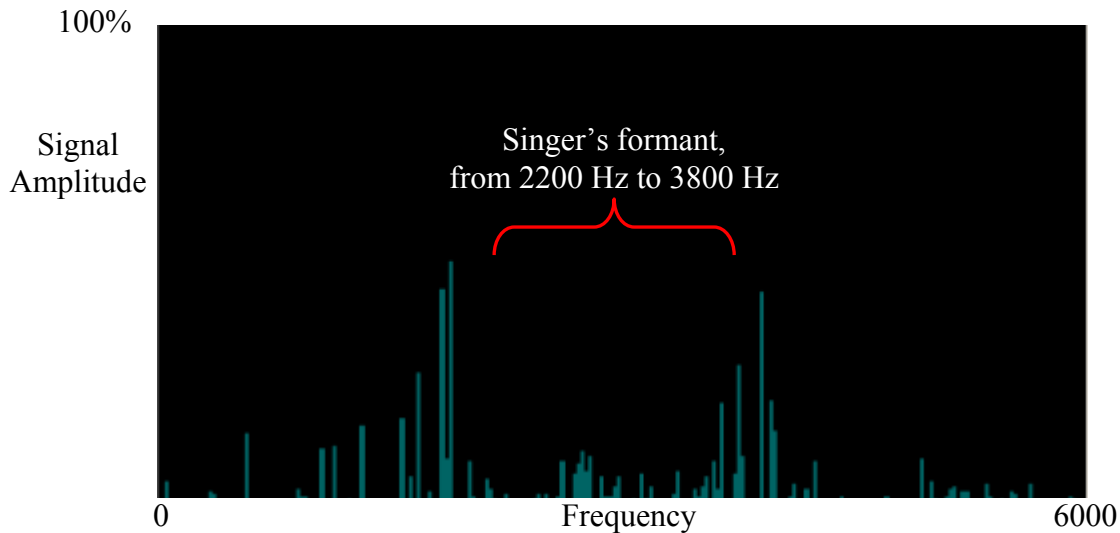


Figure 2. Sound signal represented in frequency domain, signal amplitude vs. frequency

The ‘singer’s formant’ is a peak resonance in the 2200 Hz to 3800Hz range produced by a voice trained to create this effect. A singer’s formant is necessary if a singer is to be heard over an orchestra, especially in a large opera orchestra, in a large hall. It has been a noticed effect for hundreds of years and is often referred to in older literature as the ‘ring’ in the voice.

RV measures the singer’s formant of the user’s voice, and finds the ratio of its specific frequency range to the total frequency range. This formant ratio is then used for the radius of the displaying circle. As a result, the higher the level of energy in the singer’s formant range attained, the bigger the circle obtained from the system. It indicates that a more effective and better voice is made.

System Design

RV consists of two main parts: environment tests and frequency analysis. In order to minimize unwanted background noise or distraction during the input, RV suggests users go over two simple tests before analysis proceeds. At this stage, the first two tests simply measure the loudness/amplitude of the voice.

The first test is the **Background noise test**. It helps reduce possible error and noise. The user should check that the sound level does not cause the green light to move past the threshold arrows.

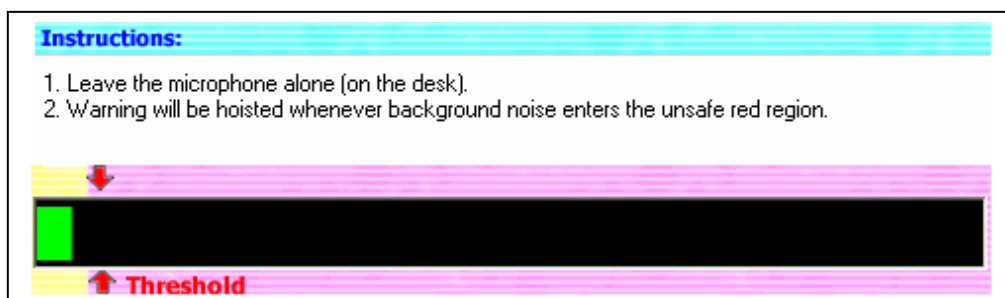


Figure 3. Background noise test

Next comes the **Distance to microphone test**. This helps the user find the appropriate microphone speaking distance. This prevents too loud a voice from being recorded and thus avoids occurrences of ‘clipping’. (Clipping occurs when the sound values get too high resulting in bad or distorted sound.)



Figure 4. Distance to microphone test

For the analysis part, RV provides three modes of operation for different users depending on their comfort level in signal processing and sound-related experience. The **Simple mode**, with or without the recording features, displays the formant circle when users speak with the microphone. In the **Advanced mode**, the software it will plot the ratio values against time for better comparison. And the **Expert mode** will draw the spectrogram so that more technical information like energy distribution can be observed. Let us go through the three modes in more detail.

Simple mode measures the effective frequency that is the singer’s formant. A bigger circle displayed indicates that a more effective voice is being produced. Users can listen to trained and untrained singers’ singing and compare their circles made on the display. Also, users may record their own voice and do a comparison as well.

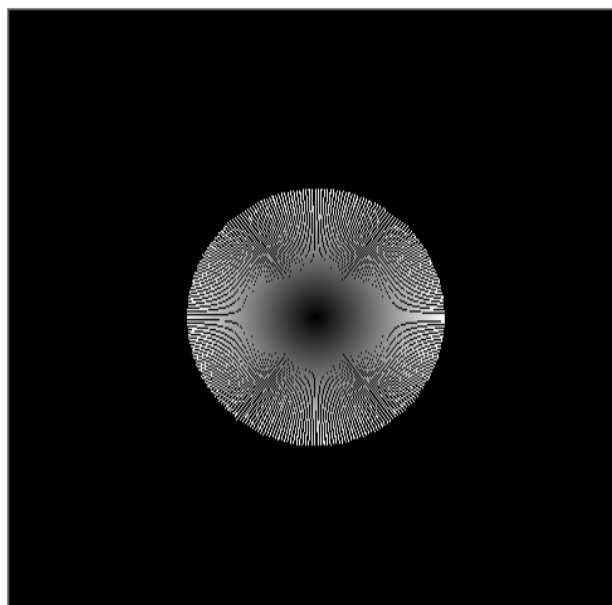


Figure 5. Simple mode, showing the circle with radius calculated from the formant ratio

Advanced mode provides a better visualization for comparison. Users can observe the ratio made within a specific duration. They can reset the time duration and select the range of vocal efficiency to be displayed.

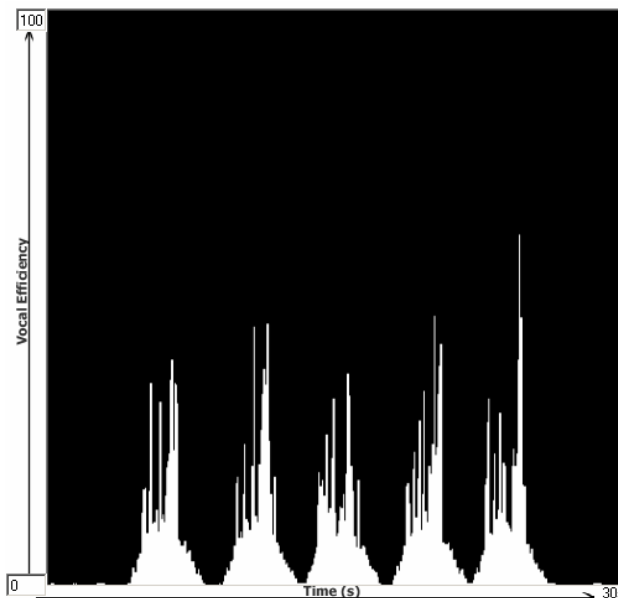


Figure 6. Advanced mode, showing a plot of the formant ratio against time

Expert mode is designed for more advanced users with more background knowledge. They can view the spectrogram and find out more useful information from it.

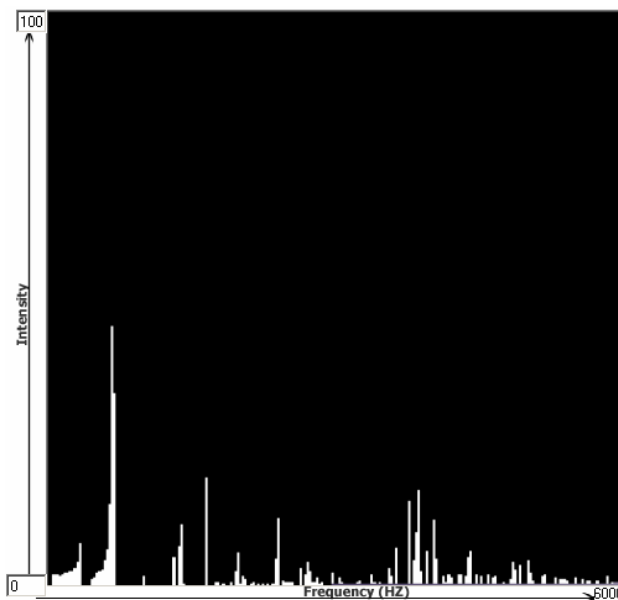


Figure 7. Expert mode, showing the spectrogram

Using all the above visual assistance, teachers can find out the weak or poor words spoken in their recording. By continually keeping on practicing with RV, they can improve their voice usage and get a higher ratio (for the circle display in simple mode).

Tutorials

Educating teachers in correct voice usage and delivering attractive speaking is important. Tutorials on proper gesture, correct mouth shape, air usage and so on, are all essential in contributing to proper speaking. We will provide users with interesting pictures or perhaps video to facilitate better interpretation. Others features, such as definitions of technical terms and system software problems, will be shown in the Help or FAQ section of the system.

CONCLUSION

Promoting awareness of healthy voice usage among teachers is of paramount importance. Teachers should not stretch their voices to project louder, but utilize the formant frequency for effective voice usage. We have developed RV (though currently not fully implemented), to serve as an indicator for finding the effective voice. RV can be extended to those students who would like to be skilled and confident public speakers and even those who are training as singers.

REFERENCES (Retrieved: May 6, 2004)

Hanson, Lloyd W., DMA Professor of Voice Pedagogy. Discussion forum about singer's formant
<http://www.vocalist.org/group/vocalist-temporary/message/9812.html>

Kim, Youngmoo Edmund, Ph.D. Thesis. 'Singing Voice Analysis/Synthesis'
<http://web.media.mit.edu/~moo/thesis/>

National Institutes of Health. 'National Institute on Deafness and Other Communication Disorders 2004'
<http://www.nidcd.nih.gov/health/education/news/teachers.asp>

Oppenheim, Willsky and Nawab, *Signals and systems*, 2nd edition, Prentice Hall

Robison, Dr. Clayne, Professor of Voice at Brigham Young University. 'Beautiful singing dot com'
<http://beautifulsinging.com/singing/spectra.html>

Tay, Verena. National University of Singapore. 'Voice Awareness – Nurture (tips for NUS Teachers)'
<http://www.cdtl.nus.edu.sg/verena/voice/care.htm>

WIKIPEDIA – the free encyclopedia
http://en.wikipedia.org/wiki/Discrete_Fourier_transform