A Study of the Gábor Transform

Peter Schultz - Feb 8 2019

I. Introduction and Overview

Although the Fourier transform is an extremely powerful tool in data analysis, it has a major drawback. Its inability to localize a signal's frequencies in time or space makes it less usable in many applications. To remedy this, Gábor Dénes proposed a method to localize both time and frequency. In this report, we will analyze two applications of the Gabor transform, and how different parameters can be altered to achieve different goals. The first part analyzes a popular piece of music written by the composer Handel. In the second part of the assignment, two recordings of "Mary Had a Little Lamb" (MHALL) played on the piano and recorder are analyzed.

II. Theoretical Background

This method, denoted the Gabor transform or the short-time Fourier transform, allows for the extraction of frequency data from a signal in specific locations. This works by multiplying the original signal by a Gabor window, which isolates a certain portion of the signal, and Fourier transforming that isolated portion. The window is subsequently shifted, and the process is repeated until the entire signal is covered. This information is presented in a spectrogram, which represents the power of different frequencies with a colormap, and time or space on the horizontal axis. Furthermore, in order to get different information out of the signal, the width, shape, and translation distance of the window can be varied. The definition of the Gabor transform is given below.

$$\mathcal{G}[f](t,\omega) = \tilde{f}_g(t,\omega) = \int_{-\infty}^{\infty} f(\tau)\bar{g}(\tau - t)e^{-i\omega\tau}d\tau = (f, \bar{g}_{t,\omega})$$

III. Algorithm Implementation and Development

Part I

In general, the spectrograms generated in the first part of the assignment use a "super gaussian" window with this form.

$$g = e^{-(width)*(t-\tau)^{10}}$$

The width of the window will be varied, as will how the τ vector (or t-slide) is "slid" along the time window. The t-slide vector is normally moved 0.1 seconds each iteration, but this will be changed to show the effects of different translation values. To create the spectrogram, the window above is multiplied by the original signal and fourier transformed at each iteration, and the resulting row vector spectrum is added to the spectrogram in the vertical direction at the corresponding time.

Part II

The algorithm for the second section is very similar to that of the first section, except that the parameters are kept constant throughout. The t-slide vector is shifted by 0.4 seconds each iteration, and the width of the window is 10^9. This is a very narrow window, which is used so that only one note is present within the window each iteration. This makes it easy to extract the dominant frequency and match it with its corresponding musical note. Each iteration of the Gabor transform, if the max value in the spectrum is above a certain cutoff, which is empirically chosen to avoid choosing noise in the recording during a rest in the piece, then it is added to a list of dominant frequencies and scatter-plotted along with its corresponding time point.

IV. Computational Results

Part I

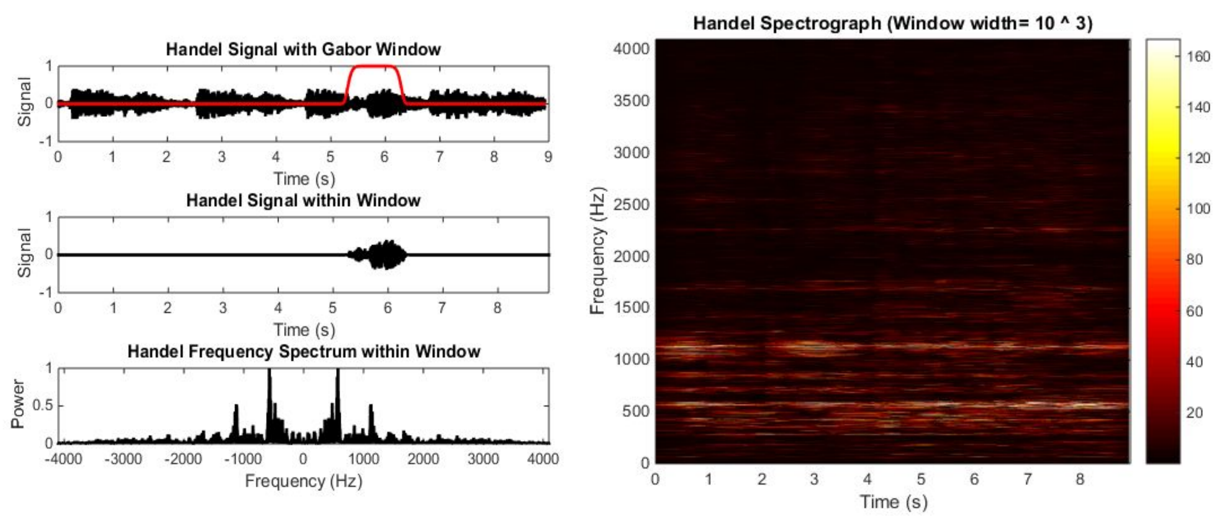

Figure 1: a) Plots of signal with window, and spectrum within window (left). b) Resulting spectrogram (right).

This is a result of a typical Gabor transform, with a window width of 10^3.

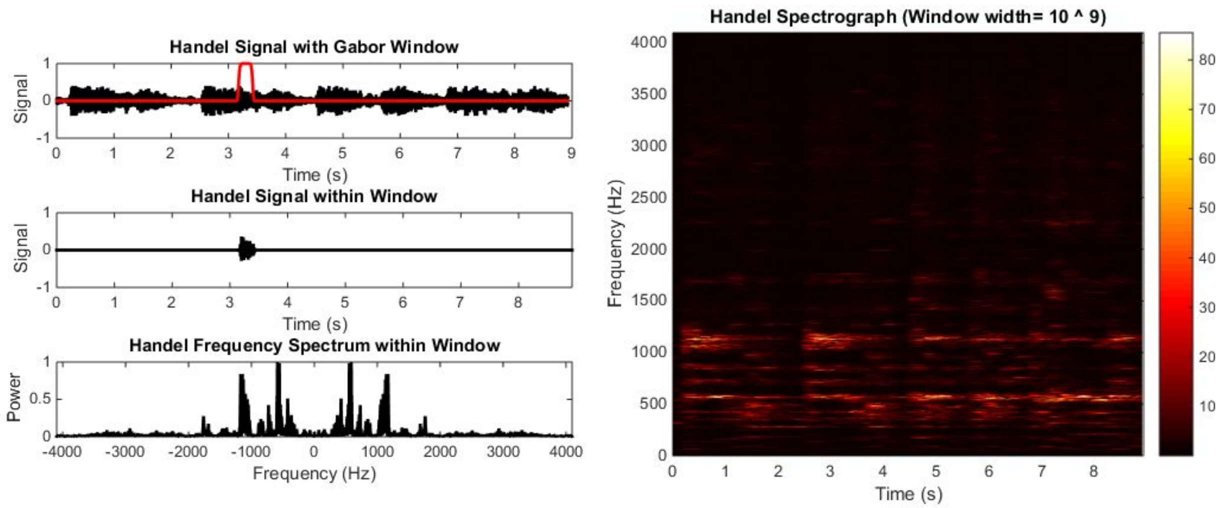

Figure 2: a) Narrow window (left). b) spectrogram with narrow window (right).

Figure 2 above demonstrates that with a very narrow window, there is virtually no signal power in the low frequencies. This is because low frequencies cannot be detected when the window is too small to contain them. However, there is substantial temporal information in the spectrogram. A narrower window results in increased time resolution, but decreased frequency resolution.

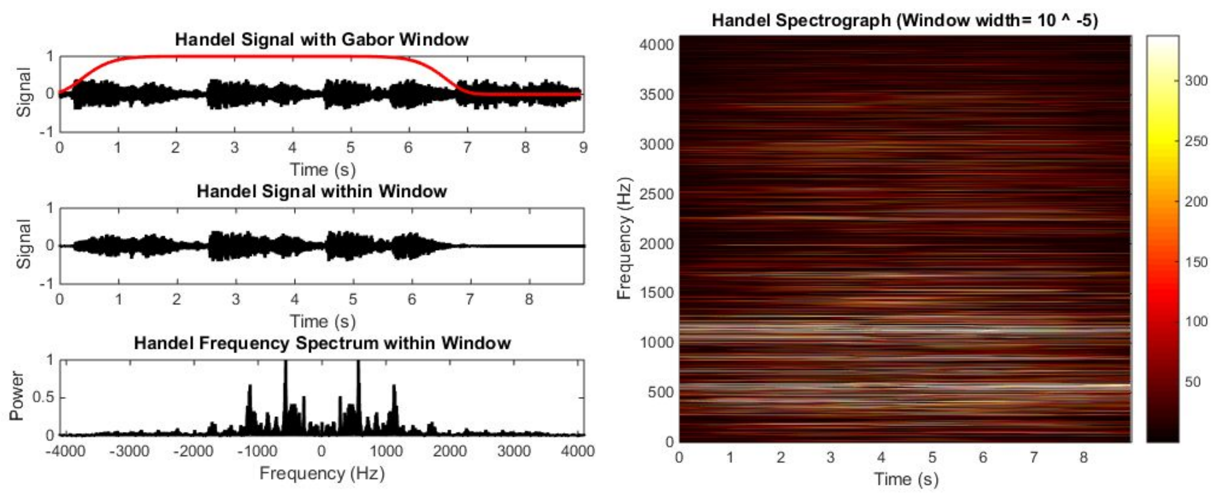

Figure 3: a) Handel signal, wide window (left). b) spectrogram, wide window (right).

In figure 3 above, it is apparent that with a very wide window, most of the temporal information is lost. Most of the signal powers appear consistent throughout the time series, which is incorrect, but all of these frequencies are detected across the whole time window when the window is so large. One positive attribute however, is that the lower frequencies are more easily detected with a large window. A wider window results in decreased time resolution, but increased frequency resolution.

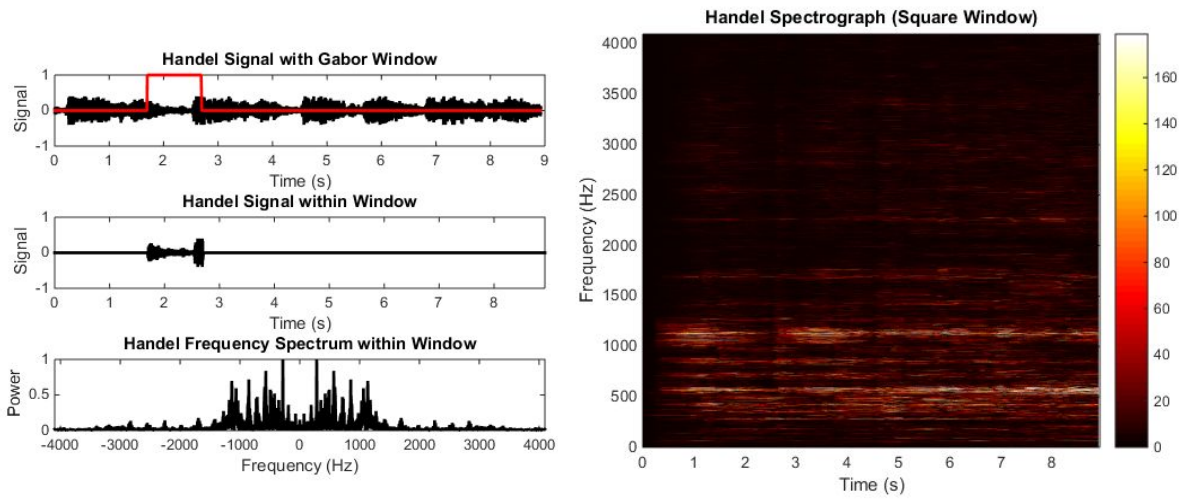

Figure 4: a) Handel signal, step window (left). b) spectrogram, step window (right)

Figure 4 demonstrates the spectrogram result when a square step window is used for the Gabor window. This spectrogram is very similar to the original, narrow Gaussian window, but may prove more useful for some specific applications.

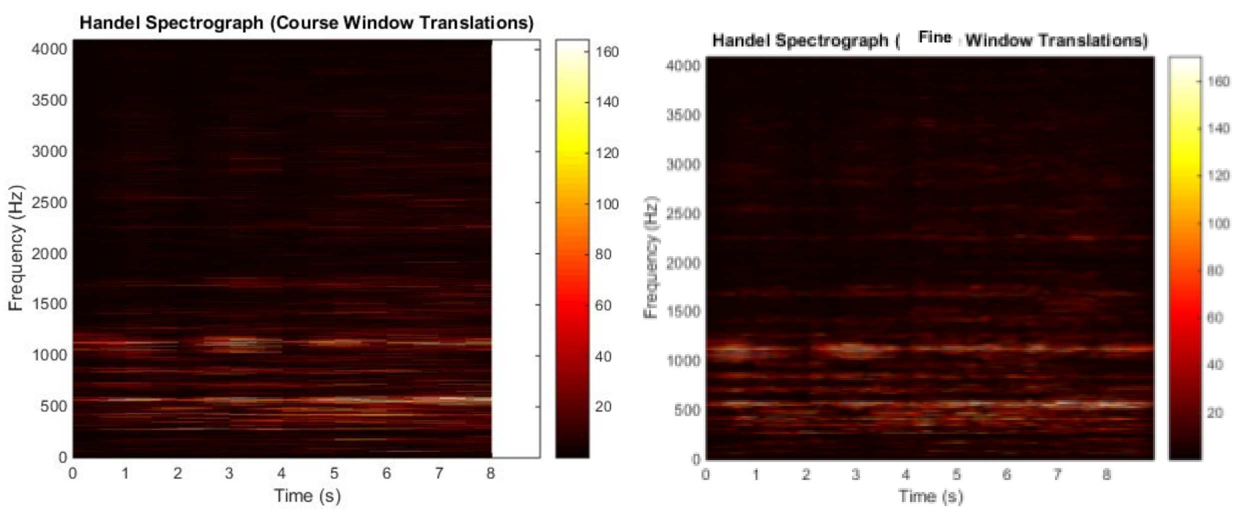

Figure 5: spectrogram with course window translations (left, Δt = 1 second) and (right) with very small window translations (Δt = 0.02 second)

In the left spectrogram, it is apparent that with very course translations for the Gabor window, the spectrogram becomes pixelated with respect to the horizontal axis. Specifically, there are very wide bands representing the frequency power at each timepoint. The signal powers in different frequencies are highly segmented. This spectrogram, however, took less time to generate than other ones.

The spectrogram on the right is highly detailed with respect to the time axis, however, it took more than 60 seconds to generate. This is due to the high amount of computation, which is not worth the extra time since it appears much the same as other spectrograms which somewhat courser window translations.

Part II

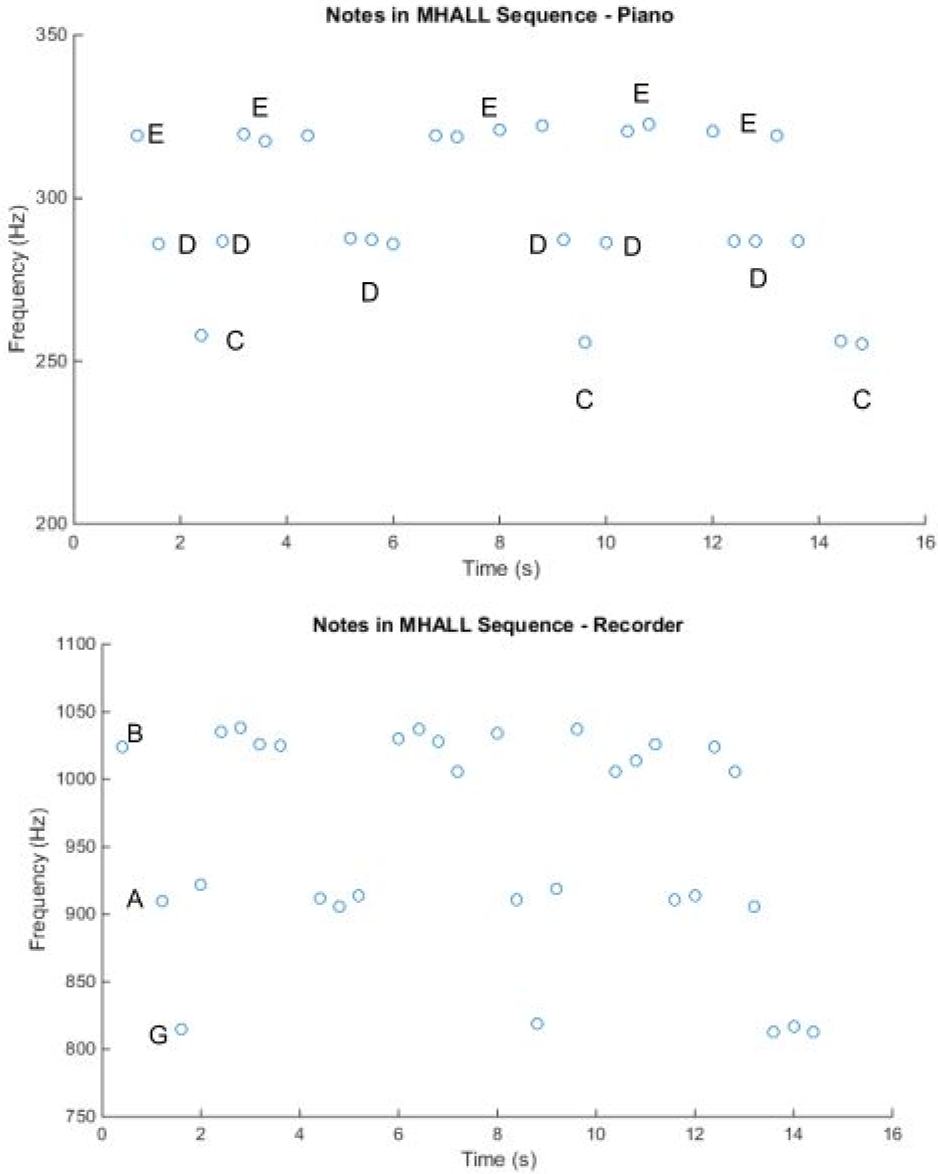

Figure 6: Scatterplot of notes present in MHALL played on the piano (top), and recorder (bottom).

Figure 6 shows the notes present in the MHALL sequence played on the piano and recorder, respectively. In the case of the recorder, the notes are around 30Hz sharp compared to the chart provided in the assignment specification. There is some slight variation in the exact frequency of each note, this is likely because of slight variations in the way the note was played, and how it was recorded. The variations are more apparent in the recorder rendition because the tone of the note being played is more likely to change based on how it is played.

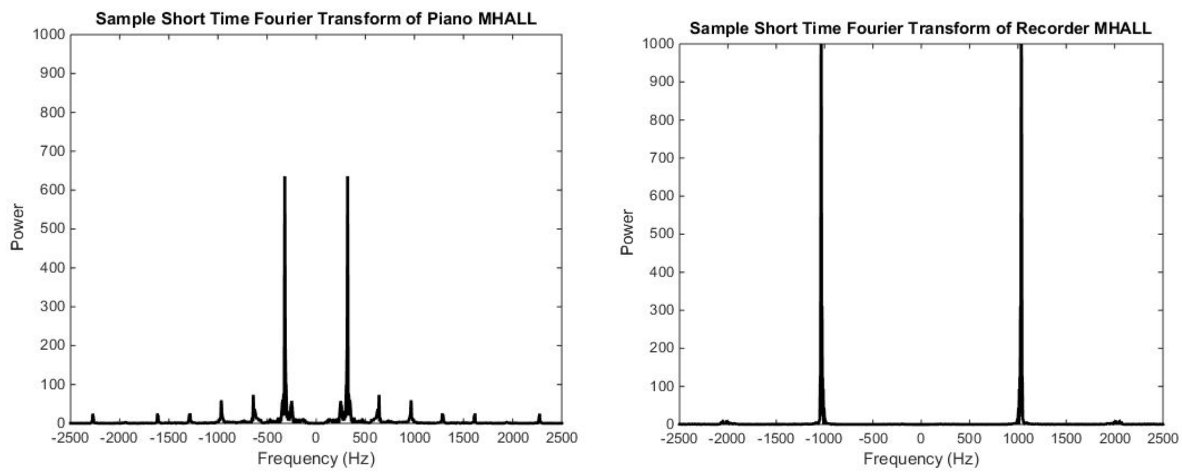

Figure 7: An example of a short time Fourier transform of MHALL played on the piano (top), and recorder (bottom).

One interesting observation about the spectrums present for the two instruments demonstrated in figure 7 is that the piano spectrum contains many overtones, which are secondary frequencies that are present in the signal, higher than the primary frequency. These overtones contribute to the "fullness" of the instrument, a characteristic of its timbre. These overtones are virtually nonexistent in the recorder spectrum.

V. Summary and Conclusions

In conclusion, the Gabor transform, along with the Fourier transform, also proves to be a highly useful tool. However, it does have some drawbacks of its own. For instance, it is impossible to increase time resolution while keeping frequency resolution constant, and vice versa. Additionally, decreasing the time distance that is shifted at each iteration can improve the resulting spectrogram's time resolution, but also greatly increases computation time.

Appendix A MATLAB functions used

audioread: reads audio files
audioplayer: creates audioplayer object
playblocking: synchronous playback of audio samples in audioplayer object
scatter: creates scatter plot
fft: performs the fast fourier transform
fftshift: shifts values in a vector about center
heaviside: step function
pcolor: creates a pseudocolor or checkerboard plot