Assignment 4

Di Riccio Tommaso
18/03/2024

Choice of the samples

To deal with this assignment I decided to select two wind instruments, **flute** and **saxophone**, and two string instruments, **guitar** and **mandolin**.

To have a fair comparison of the instruments' spectrograms I used the **C4 note** for all of them. I chose this note as it is the first note of the widely used C major scale. With a fundamental frequency of 262 Hz, it falls within the medium frequency range.

The tests were also conducted on the D5 note, confirming that the observations reported for C4 are applicable to it as well. For the sake of clarity, the corresponding plots have been omitted.

Code

```
# Function to plot the spectrogram of an audio file
def plot_spectrogram(time_series, sampling_rate, savepath):
    # Short-time Fourier transform
    fourier = librosa.stft(time_series)
    # Convert the amplitude to decibels
    spectrogram_db = librosa.amplitude_to_db(np.abs(fourier), ref=np.max)
    # Compute spectral centroid
    spectral_centroid = librosa.feature.spectral_centroid(y=time_series, sr=sampling_rate)

    # Plot the spectrogram with the spectral centroid
    fig, ax = plt.subplots()
    img = librosa.display.specshow(spectrogram_db, x_axis='time', y_axis='linear', ax=ax)
    ax.set(title='Spectrogram')
    times = librosa.times_like(spectral_centroid)
    ax.plot(times, spectral_centroid.T, label='Spectral centroid', color='w')
    ax.legend(loc='upper right')
    fig.colorbar(img, ax=ax, format="%+2.f dB")
    fig.savefig(savepath)
```

This function computes and plots the spectrogram and the spectral centroid of the time series which represents the audio file of an instrument.

It uses the LibRosa library to perform a Short-Time Fourier Transform, then converting the resulting amplitudes to decibels (a logarithmic scale). Using ***ref=np.max*** the max amplitude values will correspond to 0 dB and all other amplitudes will be expressed relative to this maximum value.

For computing the spectral centroid, the function ***librosa.feature.spectral_centroid(y, sr)*** is used, passing the audio time series and its sampling rate.

Then, the Matplotlib library is used to plot the spectrogram, with the spectral centroid in overlay.

Code

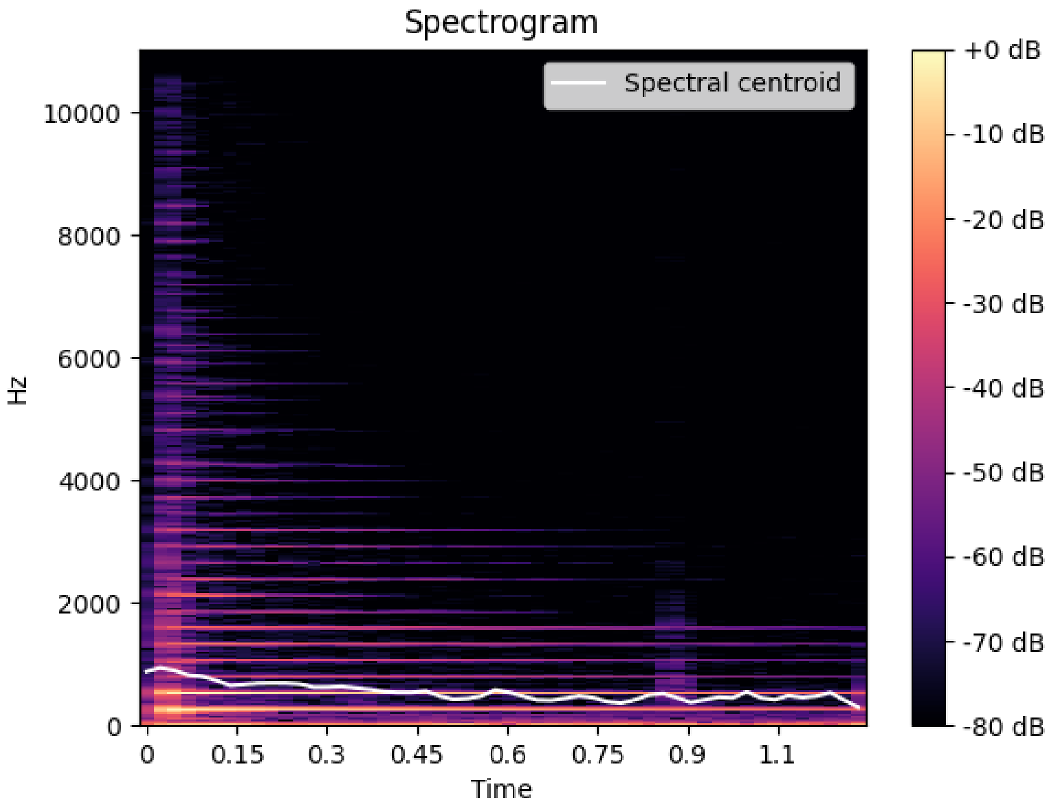
```
# Path to the audio file
guitar_audio = 'all-samples/guitar/guitar_C4_very-long_forte_normal.mp3'

# Path to save the spectrograms
guitar_savepath = 'spectrograms/guitar_spectrogram.png'

# Load the audio file
guitar, sr_guitar = librosa.load(guitar_audio, offset=0.17, duration=1.2)

# Plot the spectrograms
plot_spectrogram(guitar, sr_guitar, guitar_savepath)
```

Given the path, the function *librosa.load(path, offset, duration)* load the audio file as a floating-point time series. In this case, since the sample rate isn't specified, the default rate of 22050 Hz is utilized. The offset and duration parameters are used to omit silence at the beginning or end of the audio.

The function from the previous slide is then called to plot the spectrogram and the spectral centroid.

This process is repeated for all four instruments.

Results - Guitar



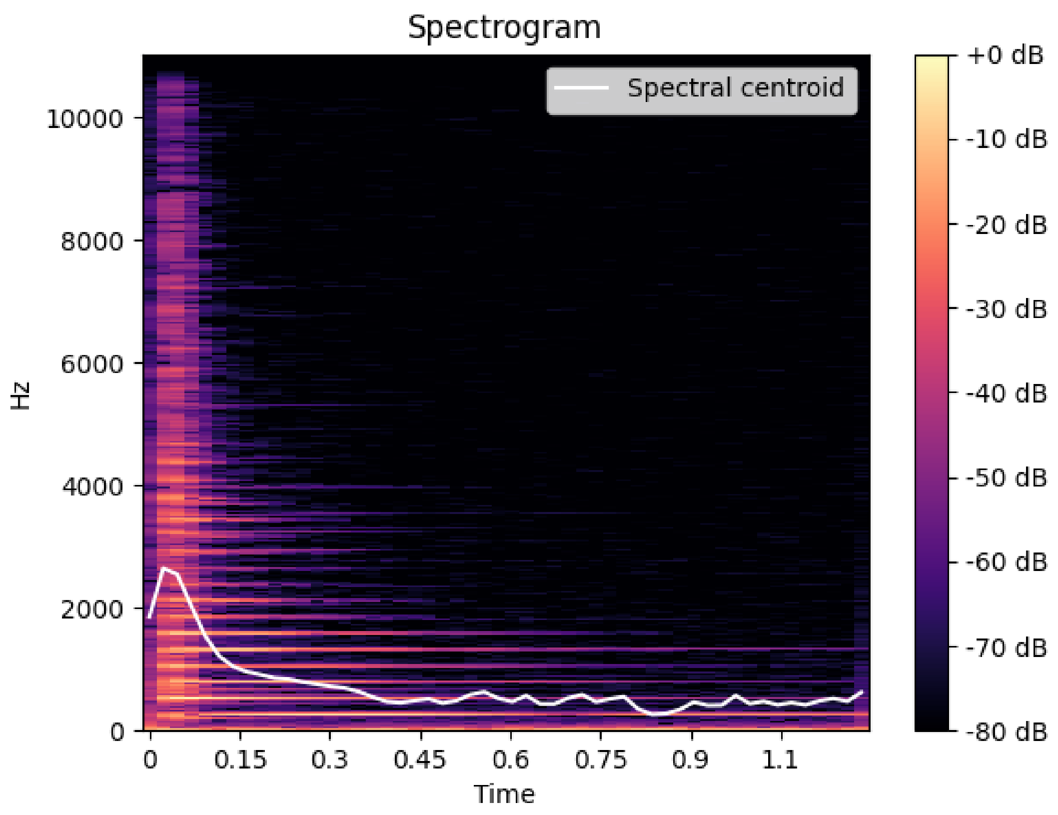
Guitar – C4 note

The most notable feature visible in the guitar spectrogram is the rapid decay of frequencies intensity above 2000 Hz over time.

Furthermore, is noticeable that the high frequencies exhibit lower intensity even when the note is just been played.

This is supported by the behavior of the spectral centroid, which remains relatively stable over time and not so distance from the fundamental frequency of the played note (262 Hz).

Results - Mandolin



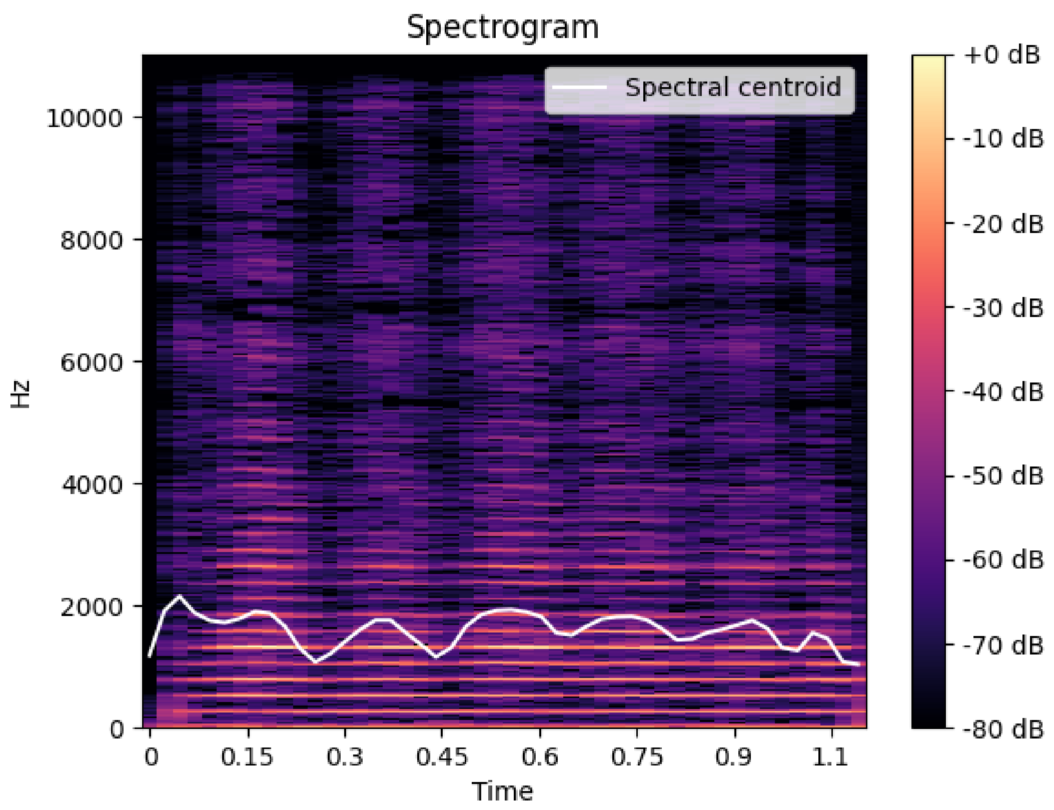
Mandolin – C4 note

The mandolin is played similarly to a guitar, as both instruments involve plucking the strings and have a resonating chamber. Indeed, the spectrogram reveals a comparable decay of high frequencies intensity over time.

Upon closer inspection of the spectrogram's initial segment, it becomes apparent that high frequencies are more pronounced in the mandolin compared to the guitar. Consequently, the spectral centroid is higher, exceeding 2000 Hz when the note is plucked.

After 0.6 seconds, when all the frequency above 2000 Hz have vanished, it is very difficult to distinguish the two instruments.

Results - Flute



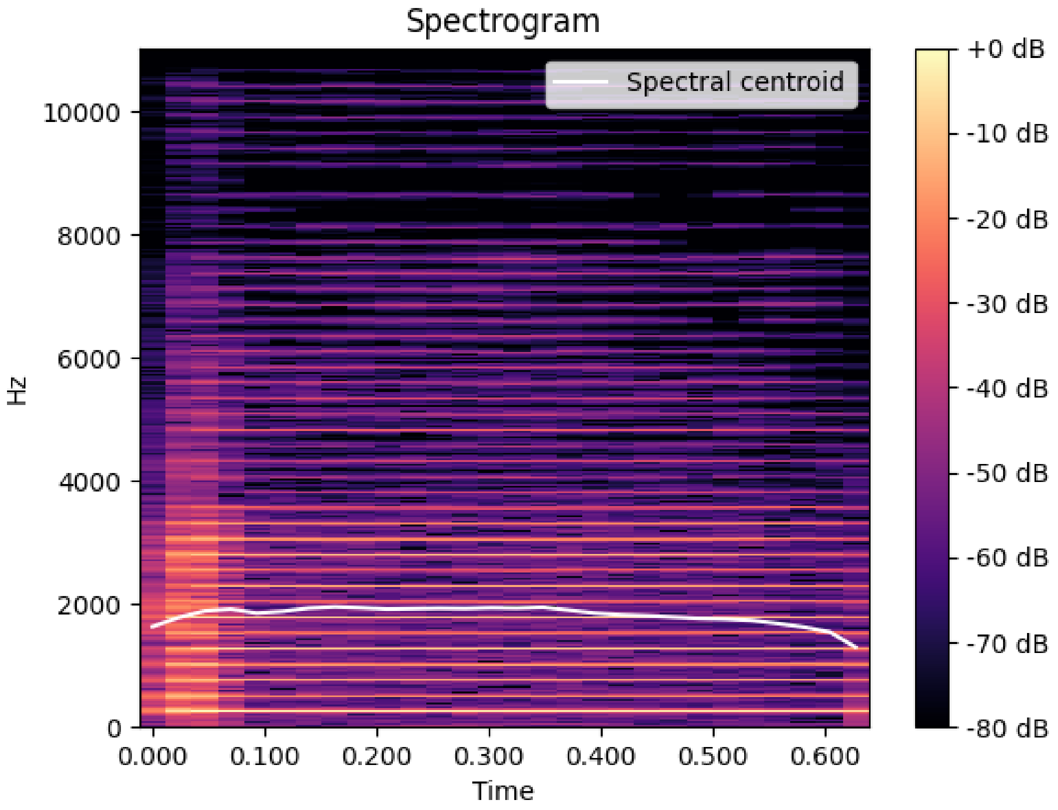
Flute – C4 note

Moving to the wind instruments, we analyze the flute spectrogram.

The most recognizable behavior arises from frequencies above 4000Hz. We can see that their intensities increase and decrease moving along time.
This behavior is clearly depicted by the fluctuation of the spectral centroid, which follows a wave pattern.

Comparing this with the guitar and mandolin, it's noticeable that, despite the oscillation, the high frequencies persist from the beginning to the end of the audio file during note playback.

Results - Saxophone



Saxophone – C4 note

The final instrument we analyze is another wind instrument, the Saxophone.

Like the flute, it stands out distinctly from the guitar and mandolin as frequencies above 2000Hz exhibit no intensity decay in over time.

Another observation is the strong intensity of frequencies below 4000 Hz, particularly noticeable when the note begins playing.

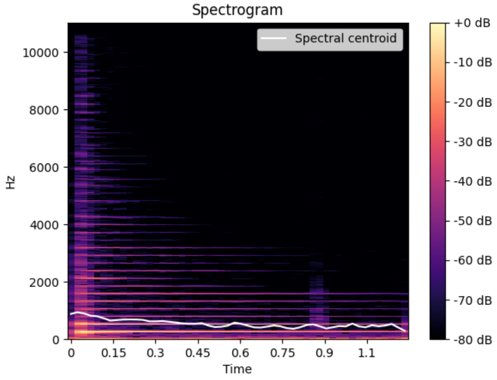
Moreover, we don't observe oscillations in the frequencies dB. In fact, the spectral centroid remains relatively steady at 2000Hz. This characteristic enables us to differentiate the saxophone from the flute.

Results - Recap

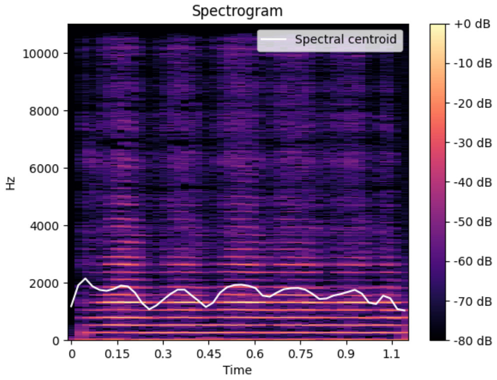
After the analysis, we may consider using the spectral centroid as a distinguishing feature as follows:

- **Guitar**: exhibits a low spectral centroid (<1000 Hz) that remains stable over time.
- **Mandolin**: present a higher spectral centroid (>2000 Hz) when the note is plucked, which then decreases over time.
- **Flute**: shows a spectral centroid that fluctuates over time between 1000 Hz and 2000 Hz.
- **Saxophone**: demonstrates a high spectral centroid (≈2000 Hz) that remains stable over time.

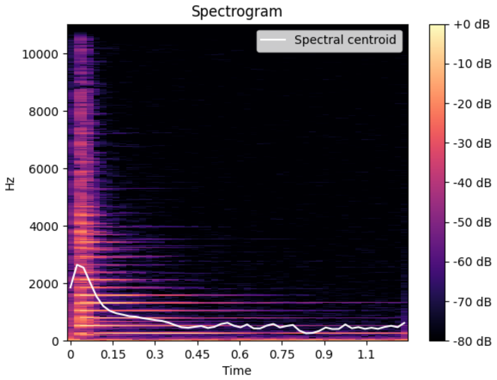
N.B. frequencies numbers are valid only for the C4 note



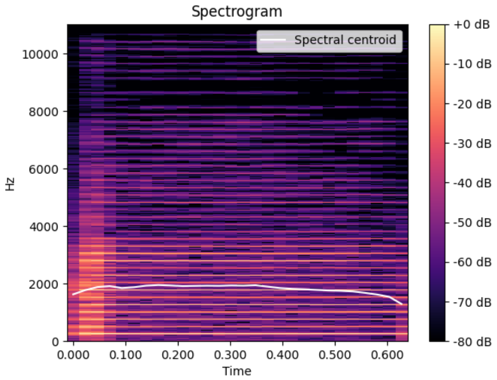
Guitar – C4 note



Flute – C4 note



Mandolin – C4 note



Saxophone – C4 note

Weak aspects

Some weak aspects of the proposed approach to distinguish the instruments are:

- Different playing techniques, such as strumming the guitar instead of plucking a single string, can significantly influence the spectral characteristics of an instrument. This aspect has been ignored in the analysis.

- The analysis was performed on the four instruments playing the same note. When a higher note is played on the guitar, the spectral centroid may exhibit a behavior similar to the mandolin playing a lower note.

- Considering other instruments with similar characteristics to those analyzed might reveal overlap of spectral centroid behavior, potentially leading to misclassification between instruments.

To address these weaknesses, a more comprehensive approach could involve incorporating additional spectral features.