Question 2 Task A

Report

1 Introduction

Spectrograms are visual representations of the frequency content of audio signals over time. In signal processing, the choice of windowing function plays an essential role in determining the resolution of the spectrogram. The primary goal of this report is to compare and analyze the spectrograms generated using three different windowing techniques: Hann, Hamming, and Rectangular windows. These techniques differ in how they smooth the frequency content, impacting the trade-off between frequency and time resolution.

This study uses the UrbanSound8K dataset, consisting of urban sound recordings, to compute the spectrograms and assess the performance of the three windowing techniques.

2 Methodology

To generate the spectrograms for comparison, the following steps were followed:

1. Audio File Selection: A random audio file from the UrbanSound8K dataset was chosen for analysis. The dataset is publicly available and consists of recordings of various urban sounds.

2. Windowing Techniques:

   - Hann Window: A cosine-shaped window that reduces the side lobes of the frequency response.
   - Hamming Window: Similar to the Hann window but with a slightly higher side lobe.
   - Rectangular Window: A simple boxcar window that provides the least smoothing but may result in greater frequency leakage.

3. Short-Time Fourier Transform (STFT): The STFT was computed using a window size of 2048 samples and a hop length of 512 samples for all three windows. The spectrograms were then converted to decibels (dB) for easier visualization.

4. Visualization: The spectrograms were plotted side by side to allow for direct comparison between the three windowing techniques.

5. Classification : A Support Vector Machine (SVM) with an RBF kernel was trained for each windowing technique. The dataset was split into 80% training and 20% testing sets.

3 Results

3.1 Spectrograms

The following spectrograms were generated using the three windowing techniques for a single audio file. Each spectrogram is a visual representation of the signal's frequency content over time, with the horizontal axis representing time and the vertical axis representing frequency.

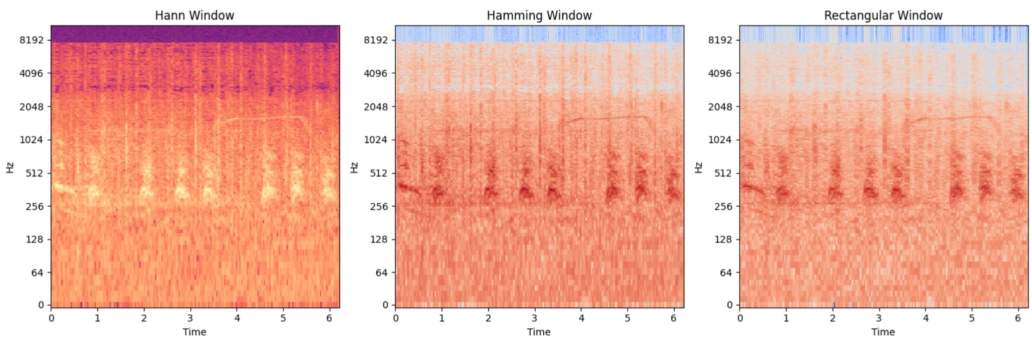

Figure 1: Spectograms

3.2 Analysis of Spectrograms

The visual analysis of the spectrograms reveals the following observations:

- Hann Window: The Hann window provides a good balance between frequency resolution and time resolution. It effectively reduces spectral leakage (side lobes) and provides a smooth transition between adjacent time frames, making it suitable for signals that require good frequency resolution with minimal leakage.

- Hamming Window: The Hamming window is similar to the Hann window but offers slightly higher side lobes. As a result, it has a more significant trade-off between time and frequency resolution, and its spectral leakage is higher than that of the Hann window. This may result in less accurate frequency information at higher frequencies.

- Rectangular Window: The Rectangular window exhibits the highest spectral leakage. It has a sharp transition at the edges of the window, leading to a significant leakage effect and a relatively poor frequency resolution. While it provides high time resolution, the lack of smoothing makes it less suitable for tasks where frequency accuracy is critical.

4 Results

The classification accuracy for each windowing technique is presented in Table 1.

Window Type	**Accuracy**
Hann Window	51.46%
Hamming Window	50.26%
Rectangular Window	50.83%

Table 1: Classification Accuracy for Different Windowing Techniques

The results indicate that the Hann window performed slightly better than the other techniques, achieving the highest classification accuracy of 51.46%. The differences in performance suggest that the choice of windowing function can impact feature extraction but may not significantly affect classification accuracy in this case. The relatively low accuracy values indicate that further feature engineering and model tuning may be required.

5 Conclusion

The three windowing techniques—Hann, Hamming, and Rectangular—each have their strengths and weaknesses in generating spectrograms.

This study compared three windowing techniques in a speech understanding task. While the Hann window yielded the best results, the overall accuracy was relatively low, suggesting the need for alternative feature extraction methods or classifier improvements.

The Hann window strikes a good balance between frequency and time resolution, making it ideal for general-purpose spectrogram generation. The Hamming window, though similar to the Hann window, exhibits slightly higher leakage, which may affect the accuracy of the frequency content representation. Finally, the Rectangular window, while offering high time resolution, suffers from significant spectral leakage, which diminishes the quality of the spectrogram for tasks requiring precise frequency analysis.

In practical applications, the choice of windowing technique depends on the specific requirements of the task at hand. For example, if accurate frequency information is crucial, the Hann window is likely the best choice, while the Rectangular window might be better suited for real-time analysis where time resolution is prioritized.