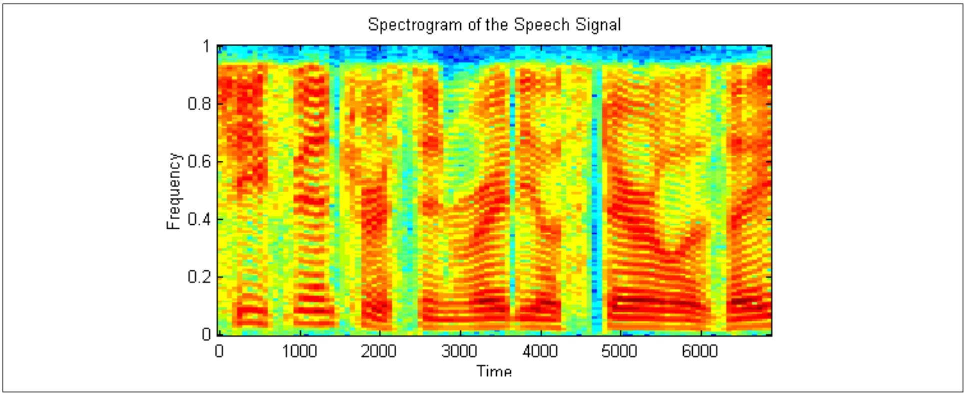

Figure 7-3. A speech spectrogram representing the frequencies found in a speech sample.

Recurrent Cells

Gradient Instability

Recurrent networks tend to degrade signal over time. Think of it as attenuating a signal by a multiplicative factor at each timestep. As a result, after 50 timesteps, the signal is quite degraded.

As a result of this instability, it has been challenging to train recurrent neural networks on longer time-series. A number of methods have arisen to combat this instability, which we will discuss in the remainder of this section.

There are a number of elaborations on the concept of a simple recurrent neural network that have proven significantly more successful in practical applications. In this section, we will briefly review some of these variations.

Long Short-Term Memory (LSTM)

Part of the challenge with the standard recurrent cell is that signals from the distant past attenuate rapidly. As a result, RNNs can fail to learn models of complex dependencies. This failure becomes particularly notable in applications such as language modeling, where words can have complex dependencies on earlier phrases.

One potential solution to this issue is to allow states from the past to pass through unmodified. The long short-term memory (LSTM) architecture proposes a mechanism to allow past state to pass through to the present with minimal modifications. Empirically using an LSTM "cell" (shown in Figure 7-4) seems to offer superior learning performance when compared to simple recurrent neural networks using fully connected layers.