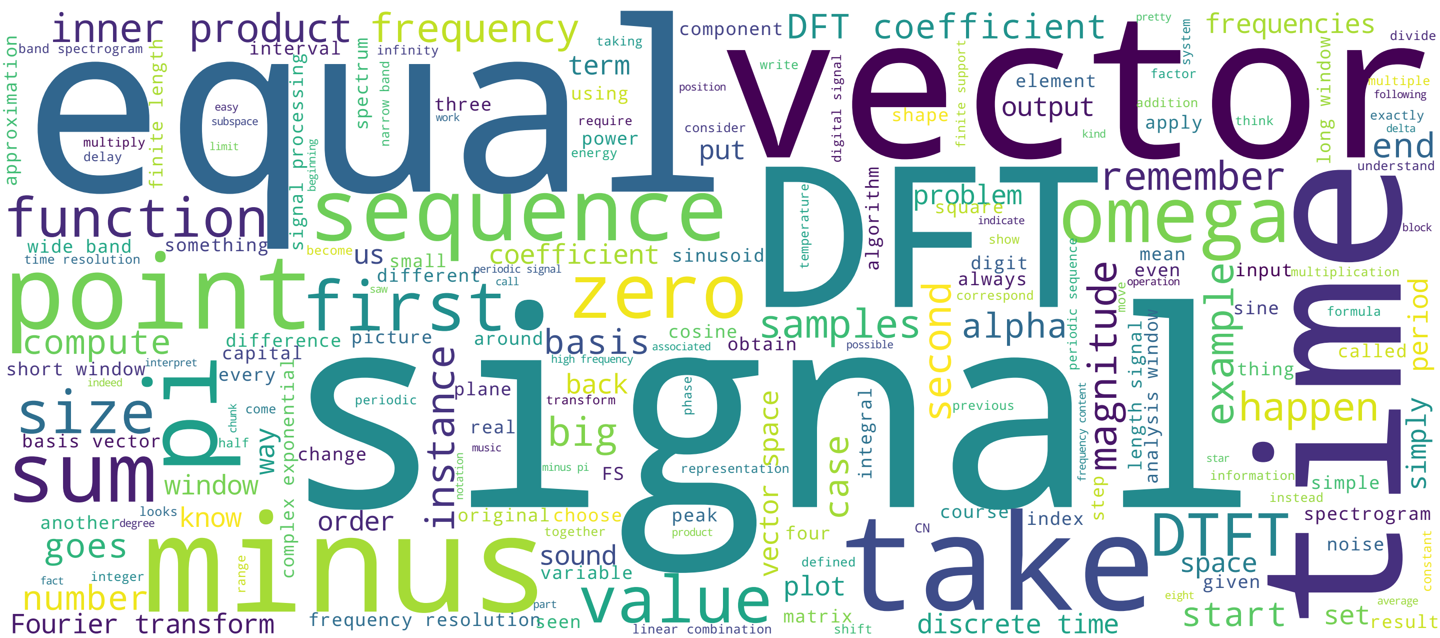
The Spectrogram

Idea:

- color-code the magnitude: dark is small, white is large
- use $10 \log_{10}(|X[m;k]|)$ to see better (power in dBs)
- plot spectral slices one after another

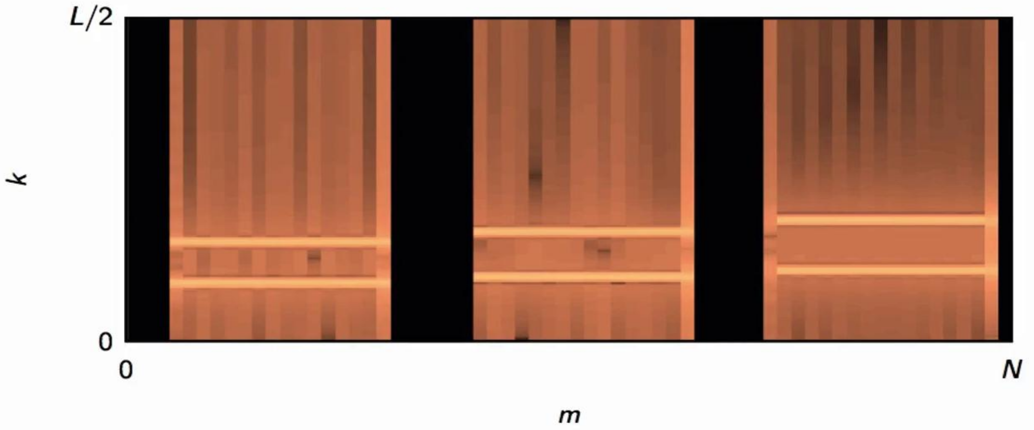


Search MOOC

Video

EPFL

DTMF spectrogram



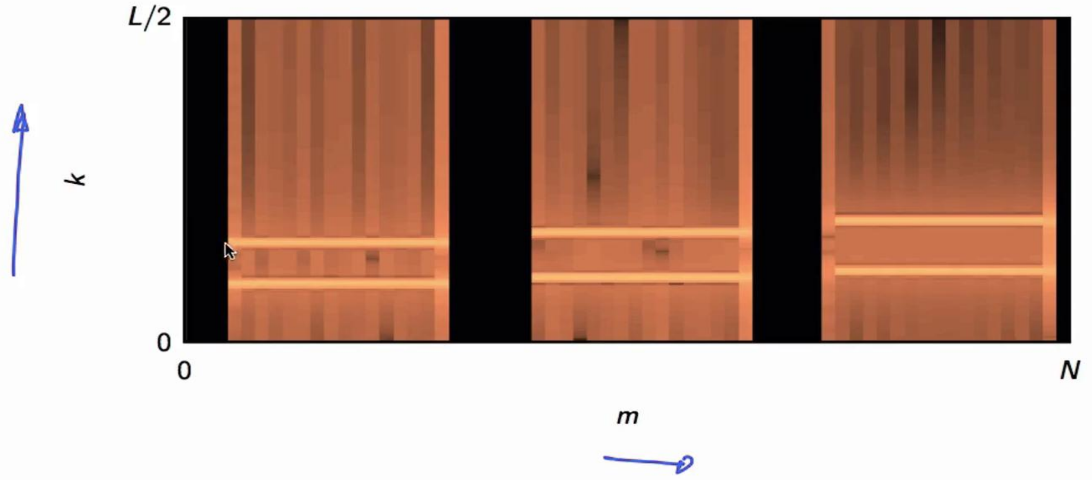
The spectrogram is a clever way of showing this time varying spectral information in one single plot. If you think about it, the short time before you transform is a complex valued function of two variables. And so to plot it properly, we would need a four dimensional plot, which of course is not possible. We can restrict ourselves to the magnitude of the DFT, at which point the SDFT becomes a real valued function of two variables, which requires a three dimensional plot. Now, this is not only quite hard to do, but also rather difficult to interpret. To make it easier to understand what we do instead is colour code the magnitude of the full year transform. And we use dark colour or dark hues for small values and whitish or brilliant hues for large values. We also take the logarithm of the magnitude in order to compress the range of values that are associated to magnitude and to better map them over a colour scale. And we put the spectral slices one after another to obtain an image like picture of the time variance spectrum. So this is the spectrogram of the TTMF signal. On the horizontal axis we put the variable M. So the starting point for each spectral slice.

Notes

Summary

0m 00s

DTMF spectrogram



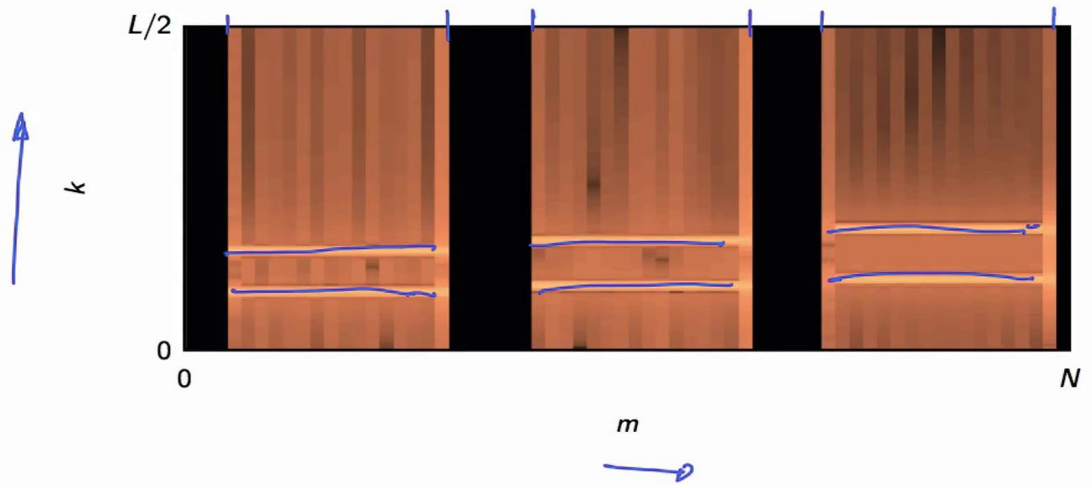
Here on the vertical axis, we put the DFT coefficient. We have a real signal, so we just go from zero to L over two where L is the size of our DFT window. You can see here that the black pixels in the picture correspond to very small values for the DFT. So this black areas indicate the silence regions in the TTMF signal. At the same time, the bright bands here correspond to high values of the DFT coefficients.

Notes

Summary

1m 20s

DTMF spectrogram

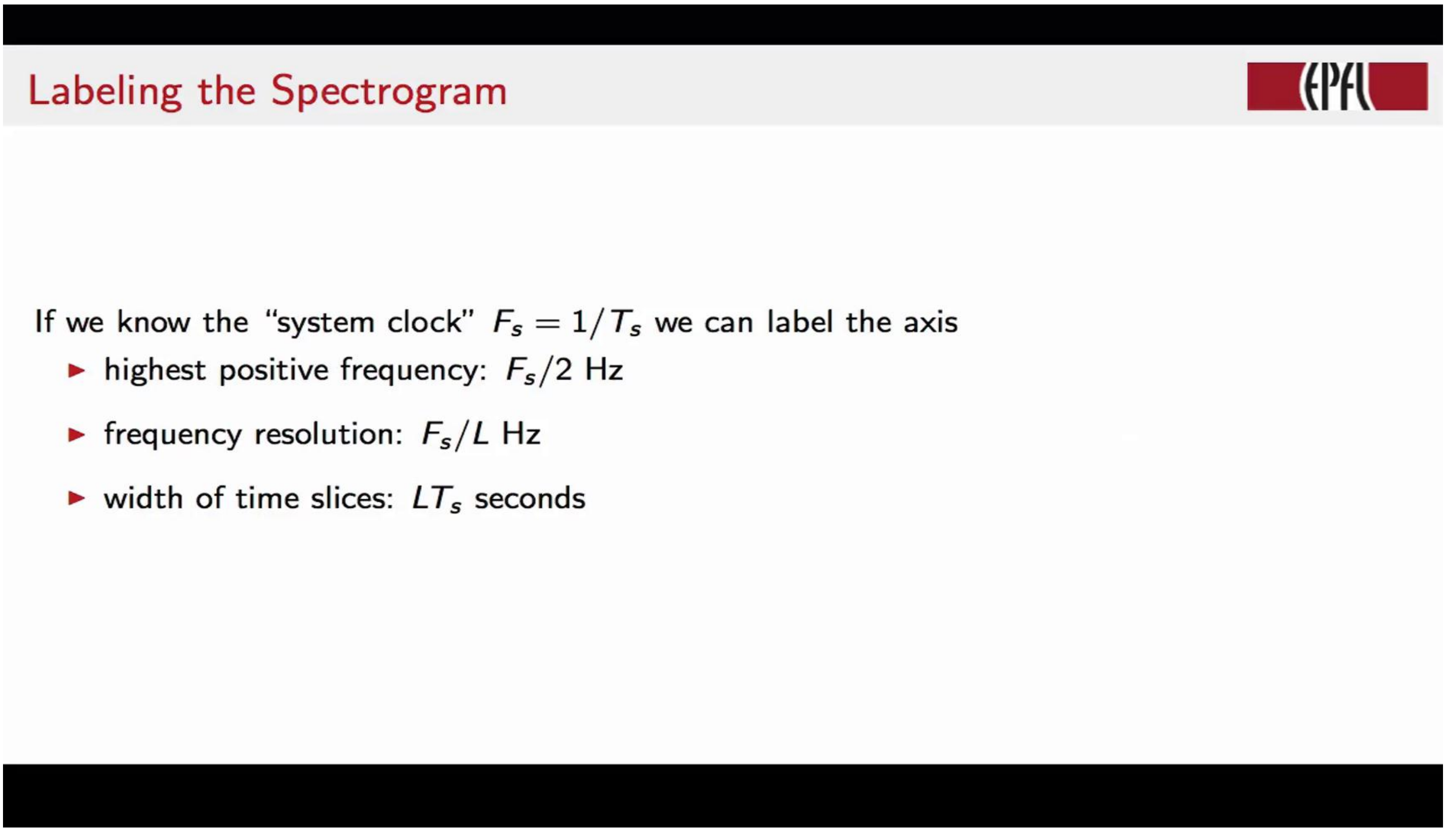


So these are actually the frequencies in each digit being dialled. And so in the plot we have. Shown at the same time, both the time information, we have a good estimation of where each digit begins and ends and of the frequency content that is associated to each digit. So we can read this picture and find out that the digits were 1, 5, 9 in sequence.

Notes

Summary

1m 50s

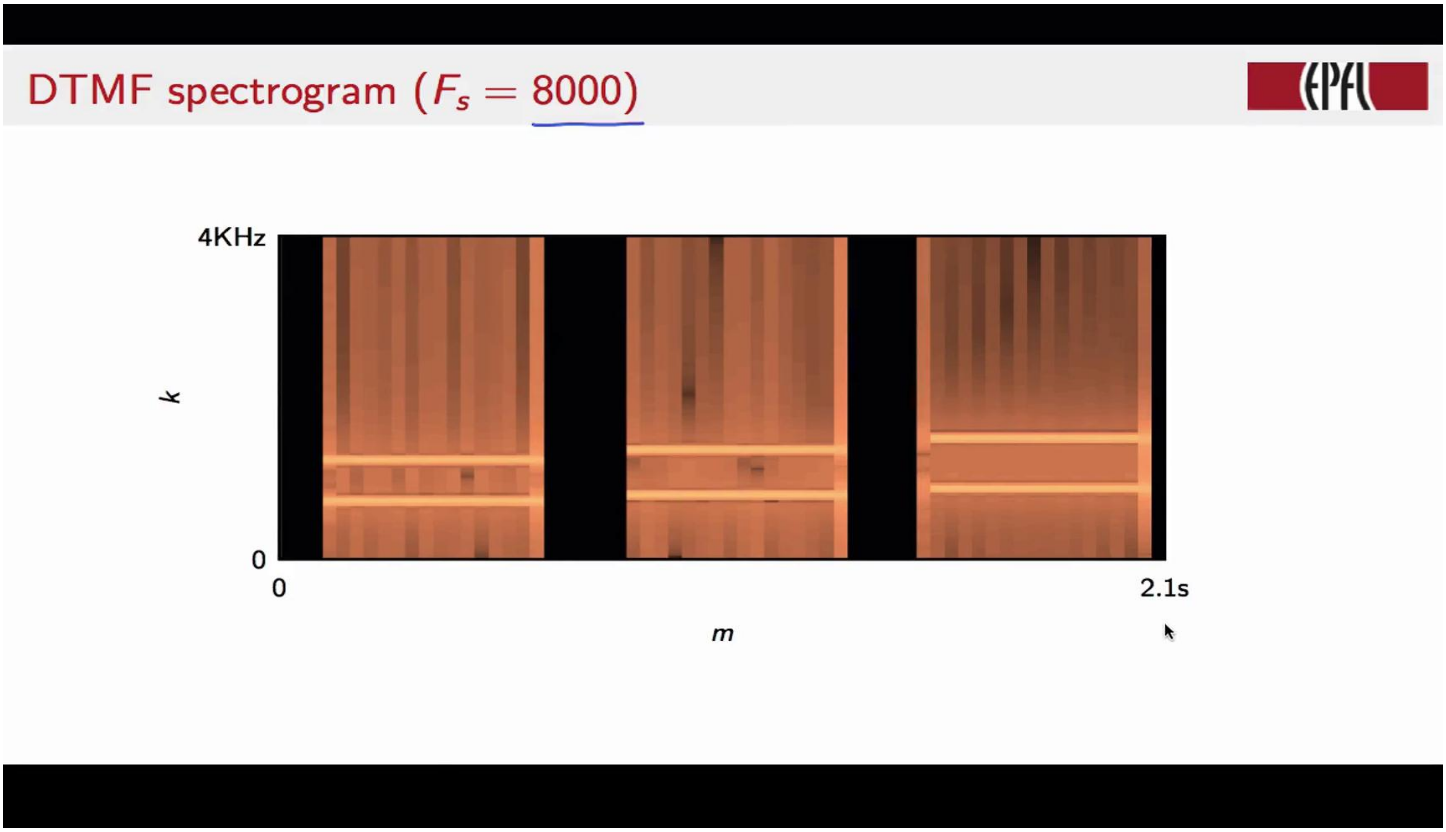


If we know the system clock for the signal or the sampling rate, we can label the axes just in the same way we did for the DFT. So remember, the highest positive frequency is FS over two where FS is the sampling rate of the signal. The frequency resolution. How fine a frequency we can resolve in the DFT will be given by FS over L and the width of the times lies. So the time resolution is l times TS seconds where TS is one over FS.

Notes

Summary

2m 17s

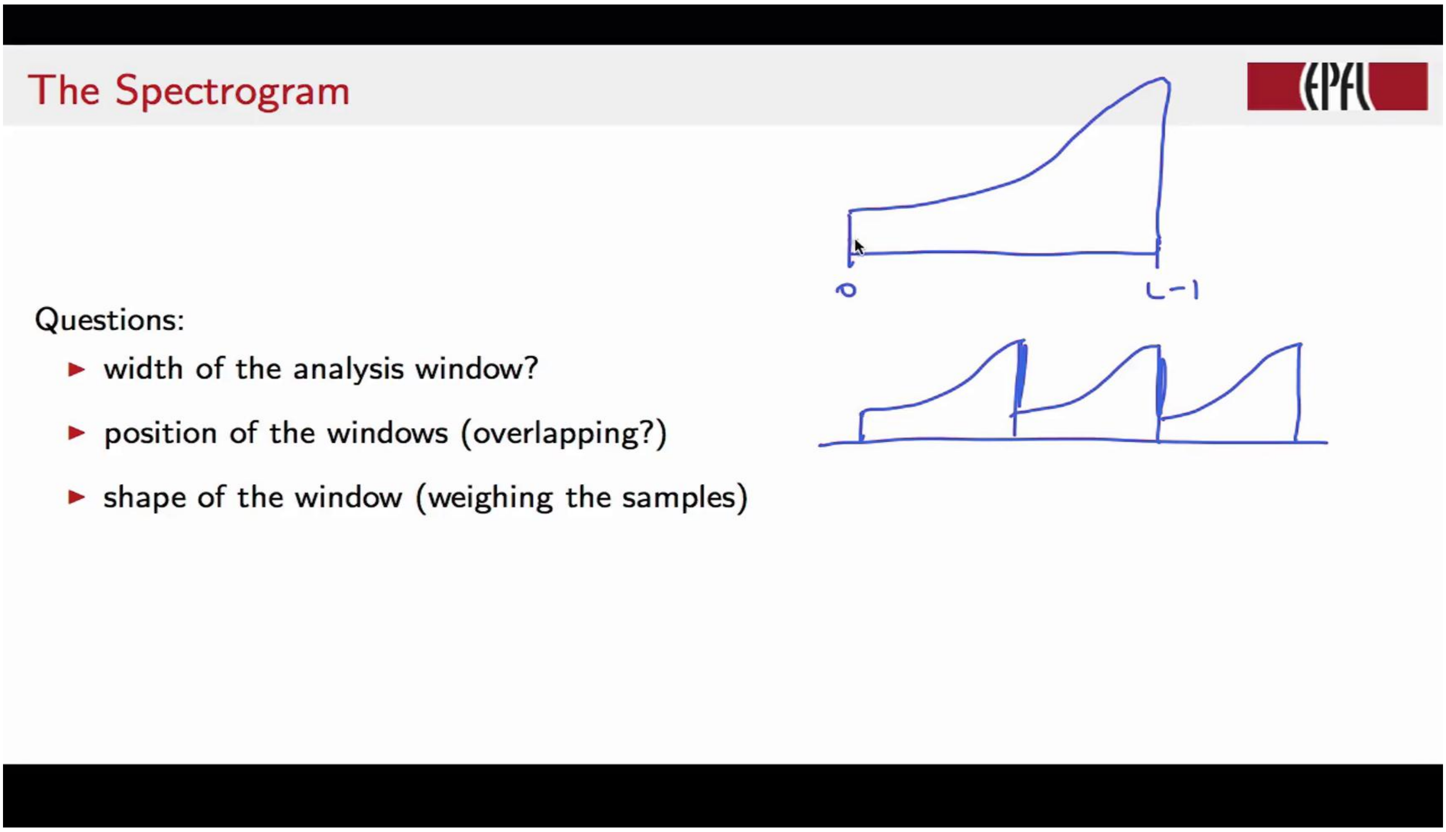


So if we apply this to the DTMF signal, which was sampled at eight kilohertz, we can label the axis like. So we have a maximum frequency of four kilohertz and a total duration of the signal of 2.1 seconds.

Notes

Summary

2m 49s



The natural questions that should come to mind of this point are. What about the width of the analysis window, which was 256 points? Why is it the optimal size? What happens if we take a larger window? What happens if we take a smaller window? How do we position this windows along the signal? Do they overlap? And if so, by how much? And what is the shape of the window that we should use? But shape here I mean the following here, we're taking chunks of L samples and just taking a DFT of the raw data. Now, suppose my signal is a very smooth signal over this window that goes like this. So this is a smooth signal and my DFT should have basically just low frequency coefficient. But now remember that to the DFT, everything is a periodic sequence. So what the DFT really sees is something like this. Now here we have all of a sudden a big jump at the discontinuity. And this will create spurious high frequency content in the DFT coefficients. We can counteract this side effect by taking the raw data in each chunk and using the tapering window. So for instance, suppose we have a tapering window shaped like a triangle.

Notes

Summary

3m 04s

The Spectrogram

Questions:

- width of the analysis window?
- position of the windows (overlapping?)
- shape of the window (weighing the samples)

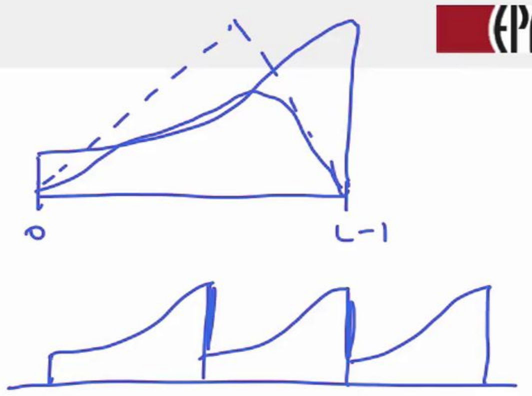
We multiply each sample by the value of the window and we will have a signal that is pretty much identical to the original data in the middle of the window. But then it tapers to zero at the extremities. And so in the end you would get something without jumps in the periods version like so.

Notes

Summary

4m 26s

The Spectrogram



Questions:

- width of the analysis window?
- position of the windows (overlapping?)
- shape of the window (weighing the samples)

So the whole story is that we could really spend weeks talking about all the tricks and tweaks that we can apply to a spectrogram in order to extract some kind of information from a real world signal. But since we don't have that kind of time here, we will just talk about the main trade-off, which is related to the size of the analysis window.

Notes

Summary

4m 43s

Wideband vs Narrowband

F_s/L $2\pi/L$

Long window: narrowband spectrogram

- long window ⇒ more DFT points ⇒ more frequency resolution
- long window ⇒ more “things can happen” ⇒ less precision in time

Short window: wideband spectrogram

- short window ⇒ many time slices ⇒ precise location of transitions
- short window ⇒ fewer DFT points ⇒ poor frequency resolution

Spectrograms can be either wide band or narrowband according to the frequency resolution of the associated DFT. So if we choose a long window, if our L is big. In that case, we have a narrowband spectrogram. Why is that? Because a long window will give us more DFT points and therefore more frequency resolution. Remember, the frequency resolution in the end is equal to the sampling frequency divided by l two Pi divided by alpha. If we remain in the complete abstract, discrete time world.

Notes

Summary

5m 04s

Wideband vs Narrowband

Long window: narrowband spectrogram

- long window $\Rightarrow$ more DFT points $\Rightarrow$ more frequency resolution
- long window $\Rightarrow$ more “things can happen” $\Rightarrow$ less precision in time

Short window: wideband spectrogram

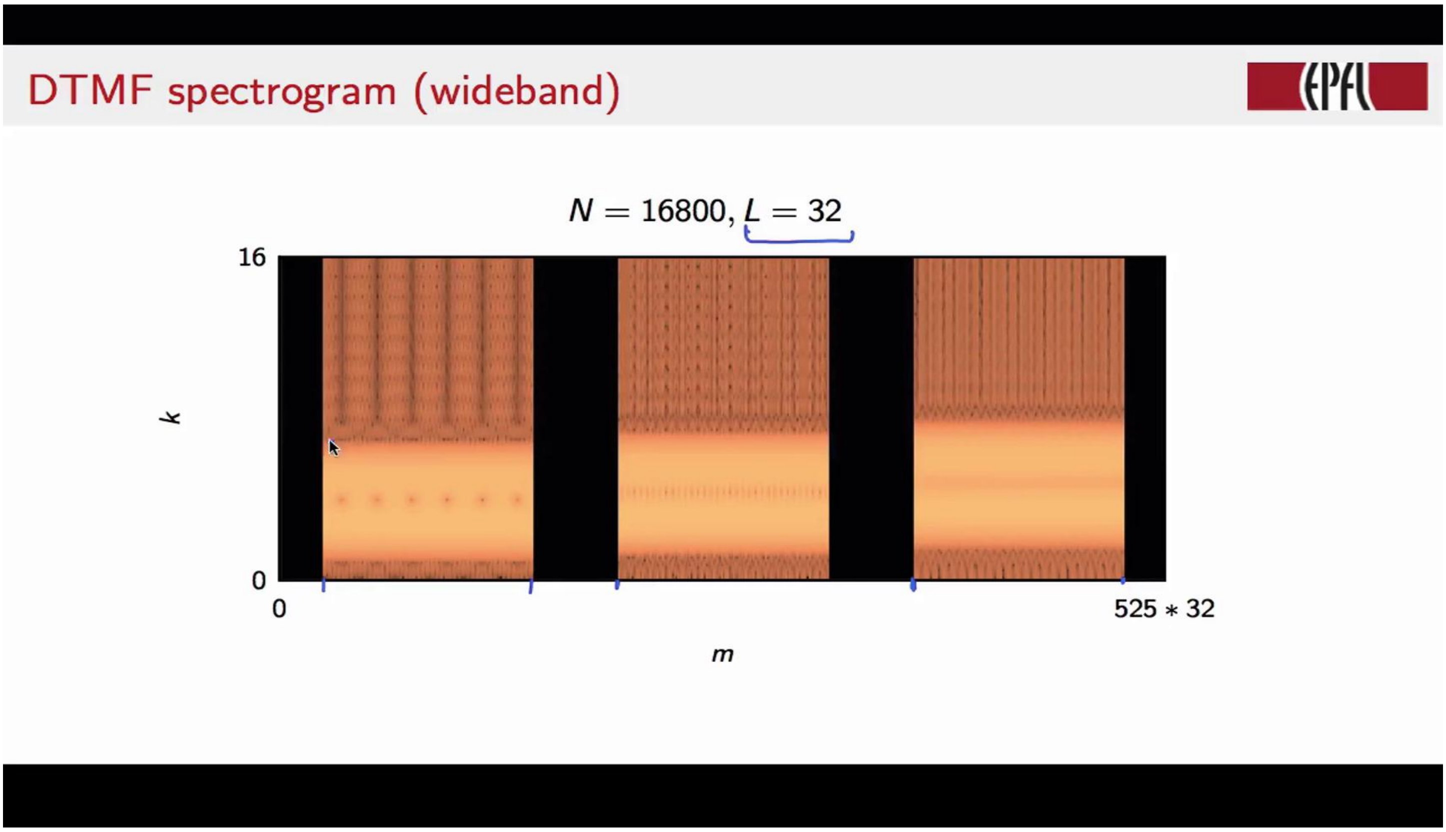
- short window $\Rightarrow$ many time slices $\Rightarrow$ precise location of transitions
- short window $\Rightarrow$ fewer DFT points $\Rightarrow$ poor frequency resolution

However, in a long window, more things can happen. And so we have less precision in the time resolution. In the limit a long window is the DFT of the entire signal, and we have seen that that completely obliterates the time information. Conversely, if we choose a short window, then we have a wide band spectrum. A short window will create many times lysis because you will divide the whole support of the signal into more chunks. And so we have a much more precise location of the transitions. But a short window will give us fewer DFT points. And so the frequency resolution will be poor.

Notes

Summary

5m 41s



So let's use our DTNF signal once again and look at the difference between a wide band and narrow band spectrogram. Here is an example of wide band spectrogram where the analysis window is just 32 samples. With such a short analysis window, we have a very good localisation in time of the start and stop point for each burst of sound.

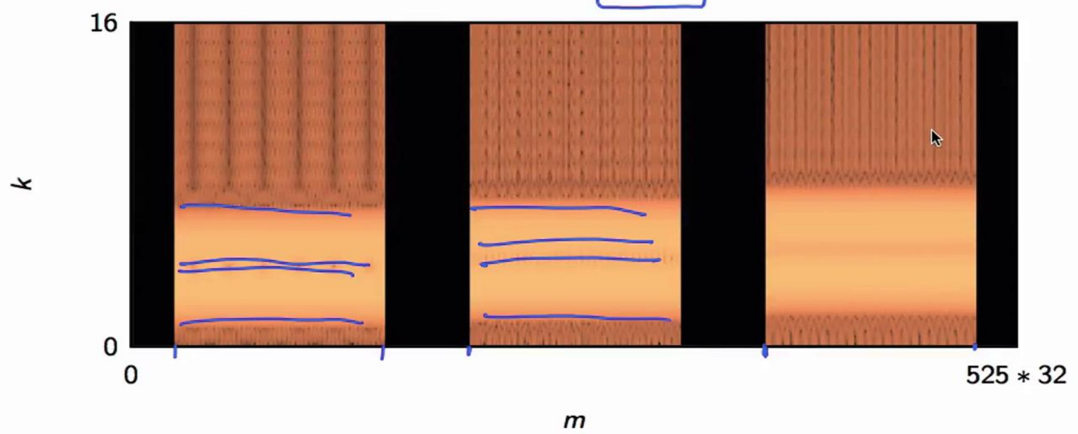
Notes

Summary

6m 20s

DTMF spectrogram (wideband)

$N = 16800, L = 32$



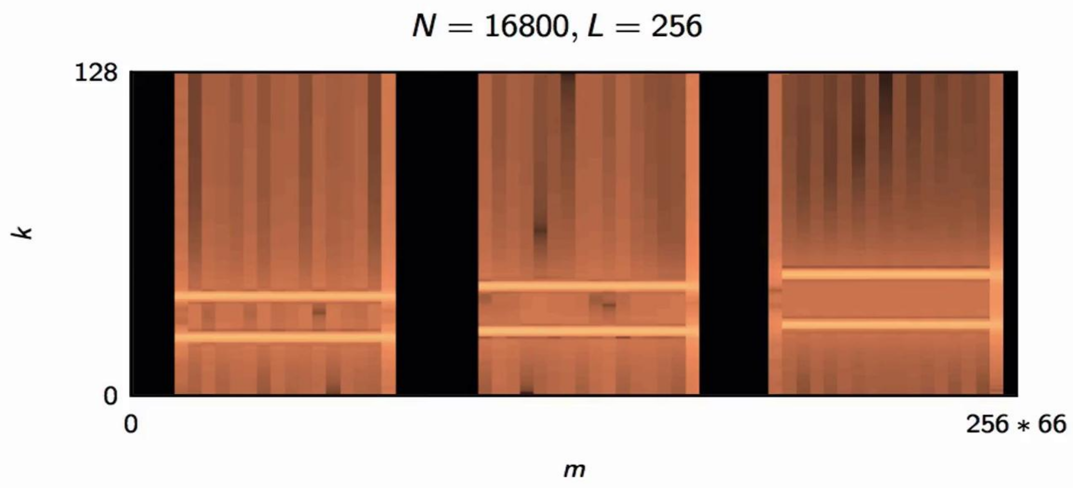
But you can see that the frequency bands are extremely wide. Also having a very short window creates artefacts in the high frequency range. Because as we sweep the window over the signal, we will be encompassing uneven numbers of periods for the underline sound as always.

Notes

Summary

6m 41s

DTMF spectrogram



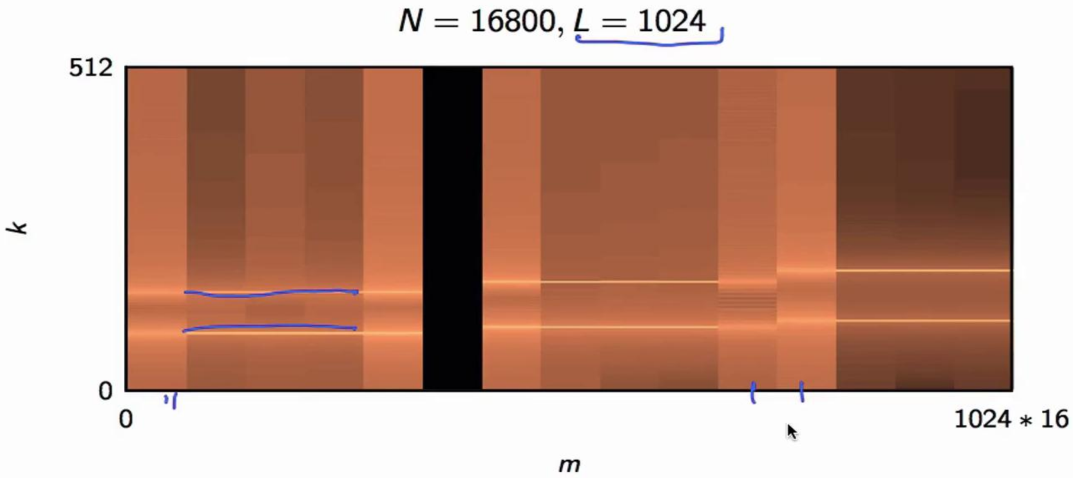
This is a spectrogram that we saw in the beginning. The window now is 256 point. So let's say that the spectrogram is in between an extremely narrow band and an extremely wide band. This is a very good compromise to interpret what's going on in the signal.

Notes

Summary

7m 01s

DTMF spectrogram (narrowband)



If we increase the size of the window to 1024, so four times larger, then we have an extremely narrow band spectrogram. You see now that the frequencies are localised extremely precisely. But on the other hand, the time resolution is very poor. We're completely missing the silence here in the beginning, for instance, and the silence between these two digits.

Notes

Summary

7m 19s