E9 205 Machine Learning for Signal Processing

Introduction to Machine Learning of Sensory Signals

12-08-2019

Instructor - Sriram Ganapathy (sriramg@iisc.ac.in)

Teaching Assistant - Prachi Singh (prachisingh@iisc.ac.in)

http://leap.ee.iisc.ac.in/sriram/teaching/MLSP_19/

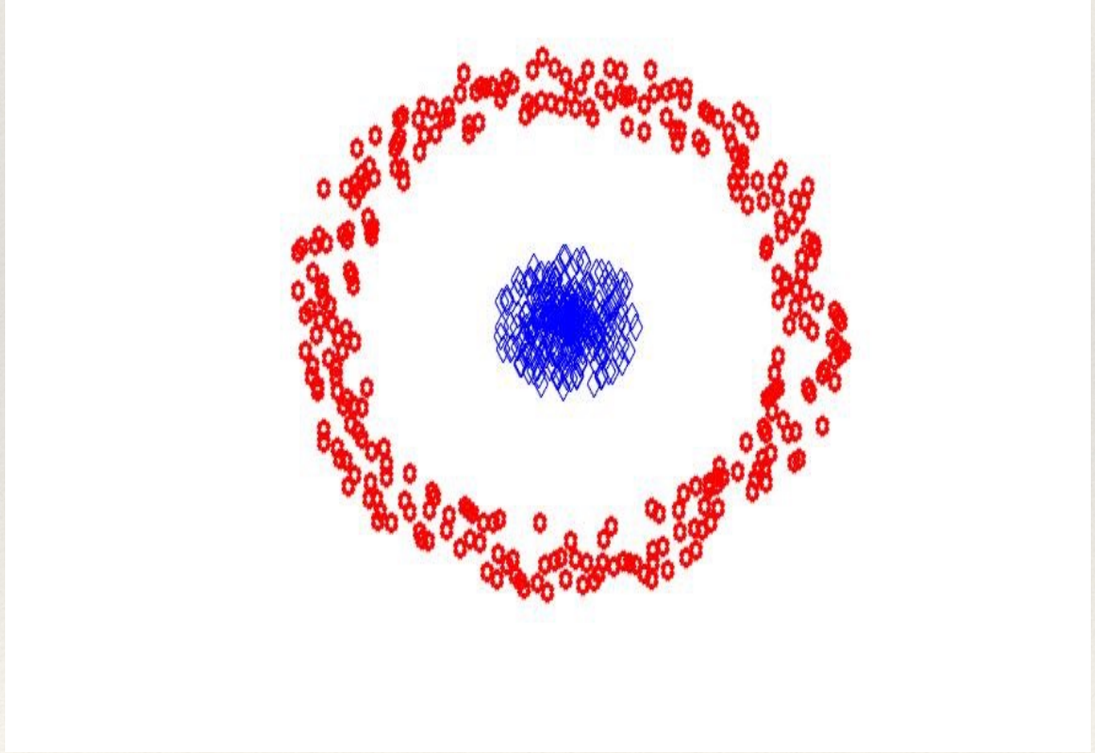
Feature Extraction

- Feature Extraction
  - Using measured data to build desirable values.
    - Attributes of the data that are informative and non-redundant.
    - Resilience to noise / artifacts.
    - Facilitating subsequent learning algorithm.

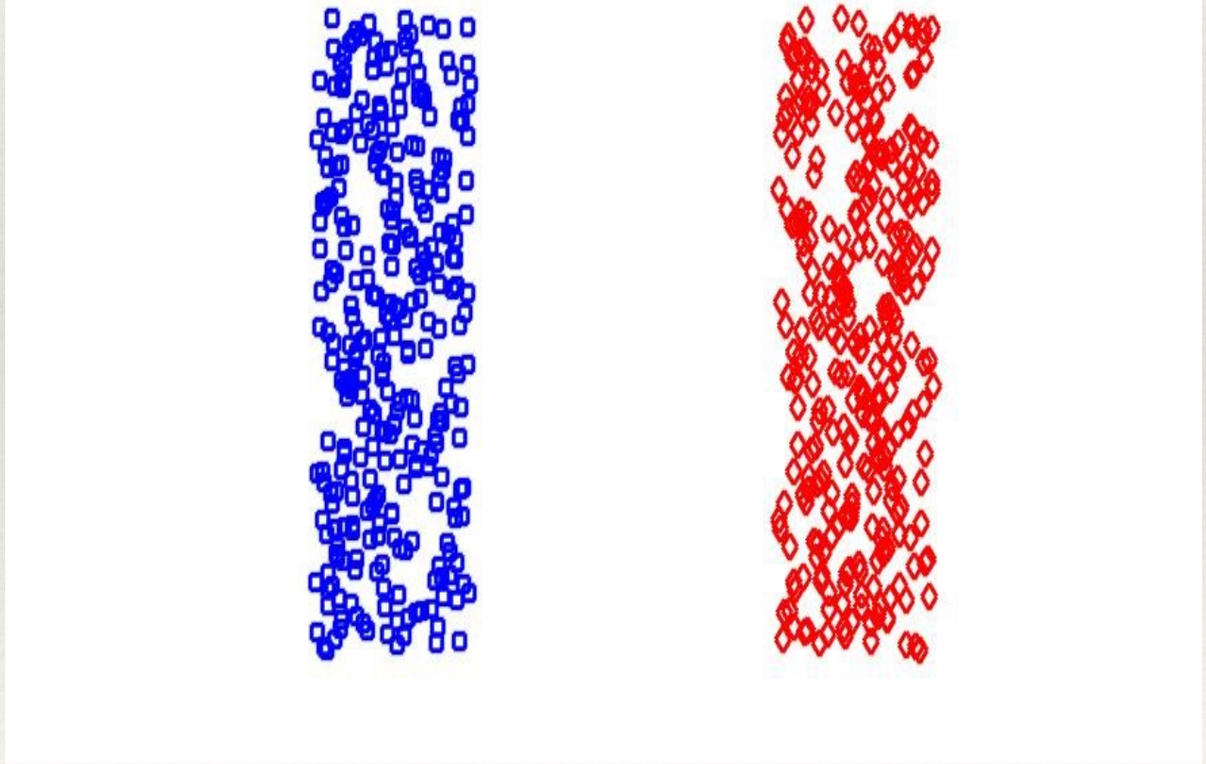
Feature Extraction

- Representation Problem

Cartesian Coordinates



Polar Coordinates



Feature Extraction

Scope for this course

I. Feature Extraction in Text.

II. Feature Extraction in Speech and Audio signals.

III. Processing of Images.

Established 1911

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Text Processing

Text Modeling - Introduction to NLP

- Definitions
  - Documents, Corpora, Tokens (Terms)
  - Term Frequency (TF)
  - Collection Frequency (CF)
  - Document Frequency (DF)
  - TF-IDF
  - Bag of words model

Example [Manning and Schutze, 2006]

Word	cf	df
try	10422	8760
insurance	10440	3997

► **Figure 6.7** Collection frequency (cf) and document frequency (df) behave differently, as in this example from the Reuters collection.

term	df_t	idf_t
car	18,165	1.65
auto	6723	2.08
insurance	19,241	1.62
best	25,235	1.5

► **Figure 6.8** Example of idf values. Here we give the idf's of terms with various frequencies in the Reuters collection of 806,791 documents.

https://nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf

Perplexity

- Measuring the goodness of language modeling
-

$$\begin{aligned} \mathrm{PP}(W) &= P(w_1 w_2 \ldots w_N)^{-\frac{1}{N}} \\ &= \sqrt[N]{\frac{1}{P(w_1 w_2 \ldots w_N)}} \end{aligned}$$

On a Wall-street Journal Corpus

	Unigram	Bigram	Trigram
Perplexity	962	170	109

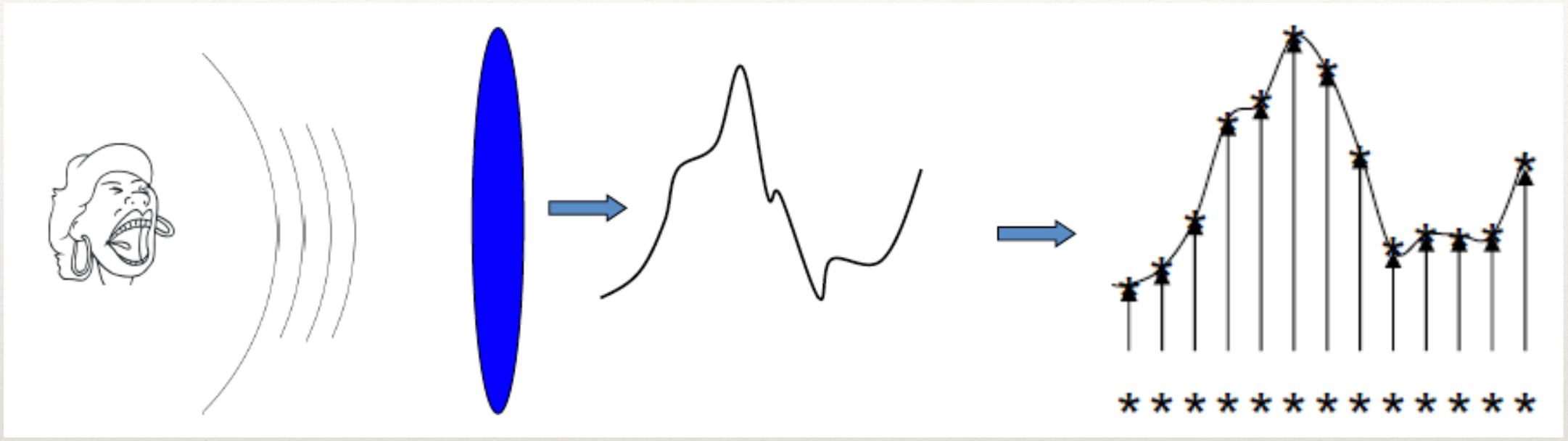
https://web.stanford.edu/~jurafsky/slp3/4.pdf

Speech and Audio Processing

Speech and Audio

- Speech / Audio - 1D signals
  - Generated by pressure variations producing regions of high pressure and low pressure.
  - Travels through a medium of propagation (like air, water etc).
  - Human sensory organ - eardrum.
    - Converting pressure variations to electrical signals.
    - Action mimicked by a microphone.

Sound waves in a computer



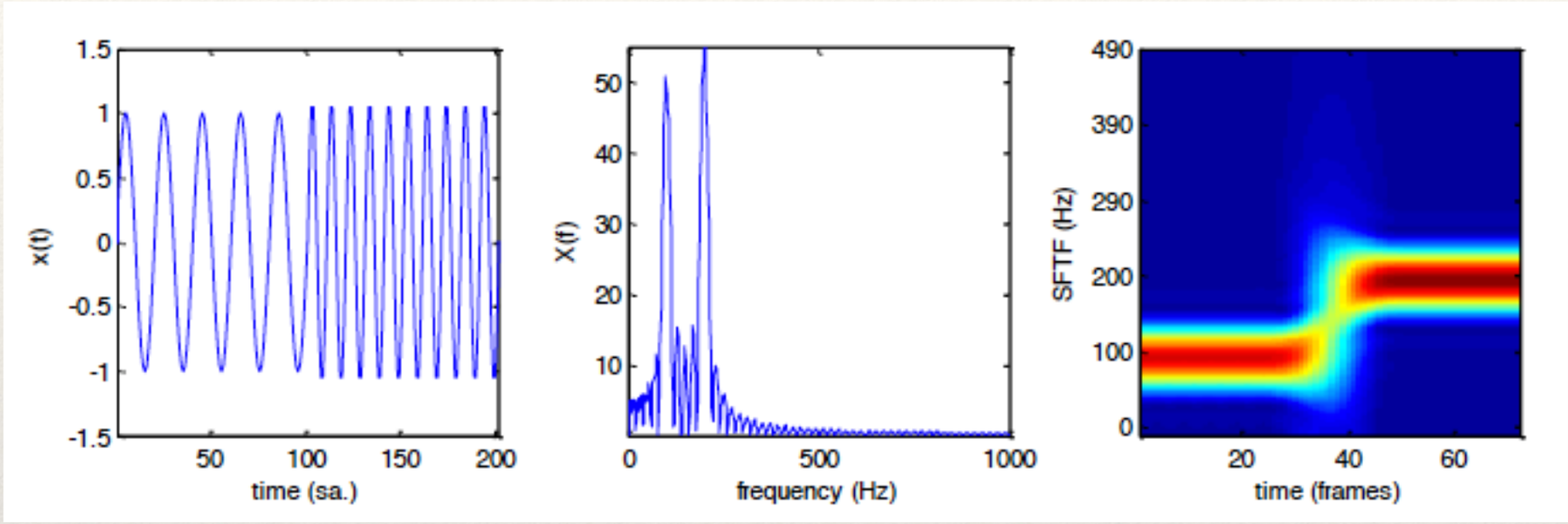
- Analog continuous signal from the microphone
  - Discretized in time - sampling.
  - Digitized in values - quantization.

http://mlsp.cs.cmu.edu/courses/fall2014/lectures/slides/Class1.Introduction.pdf

Established 1911

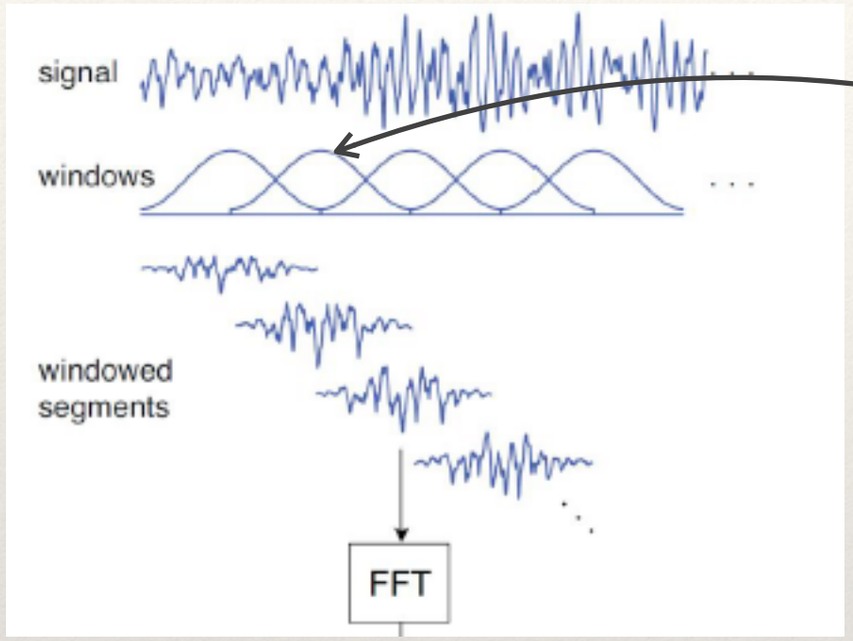
INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

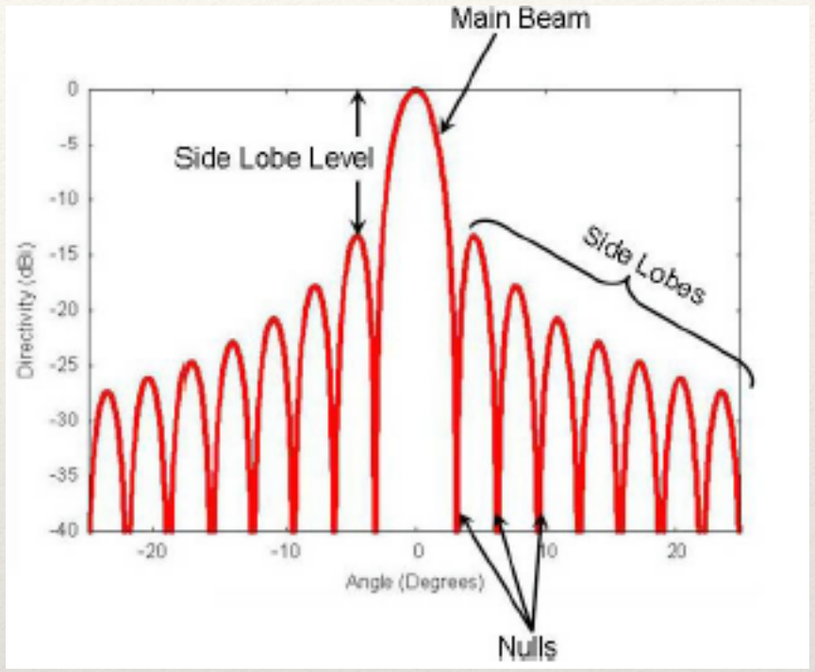
Why do we need time varying Fourier Transform



- When the signal properties change in time
  - DFT will only capture the average spectral character
  - Short-window analysis can indicate the change in spectrum.

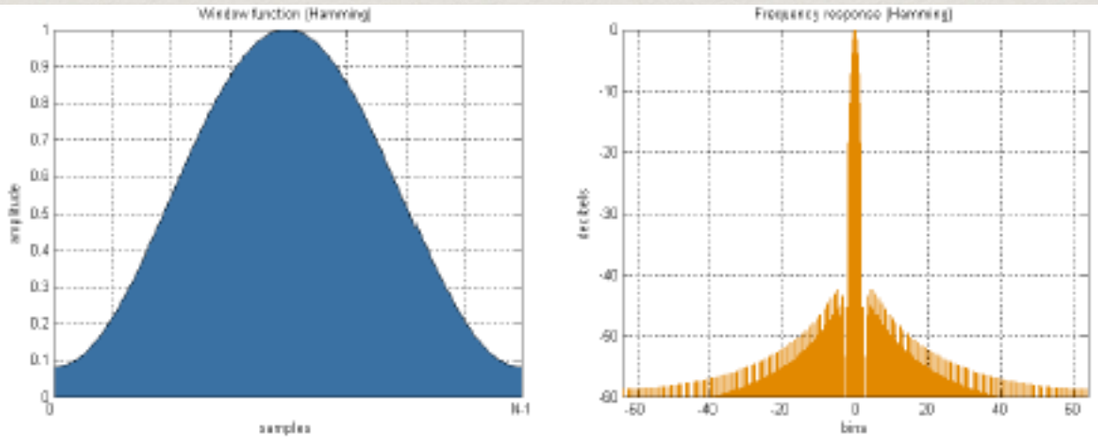
Summary of STFT Properties





$$X[k, n_0]$$

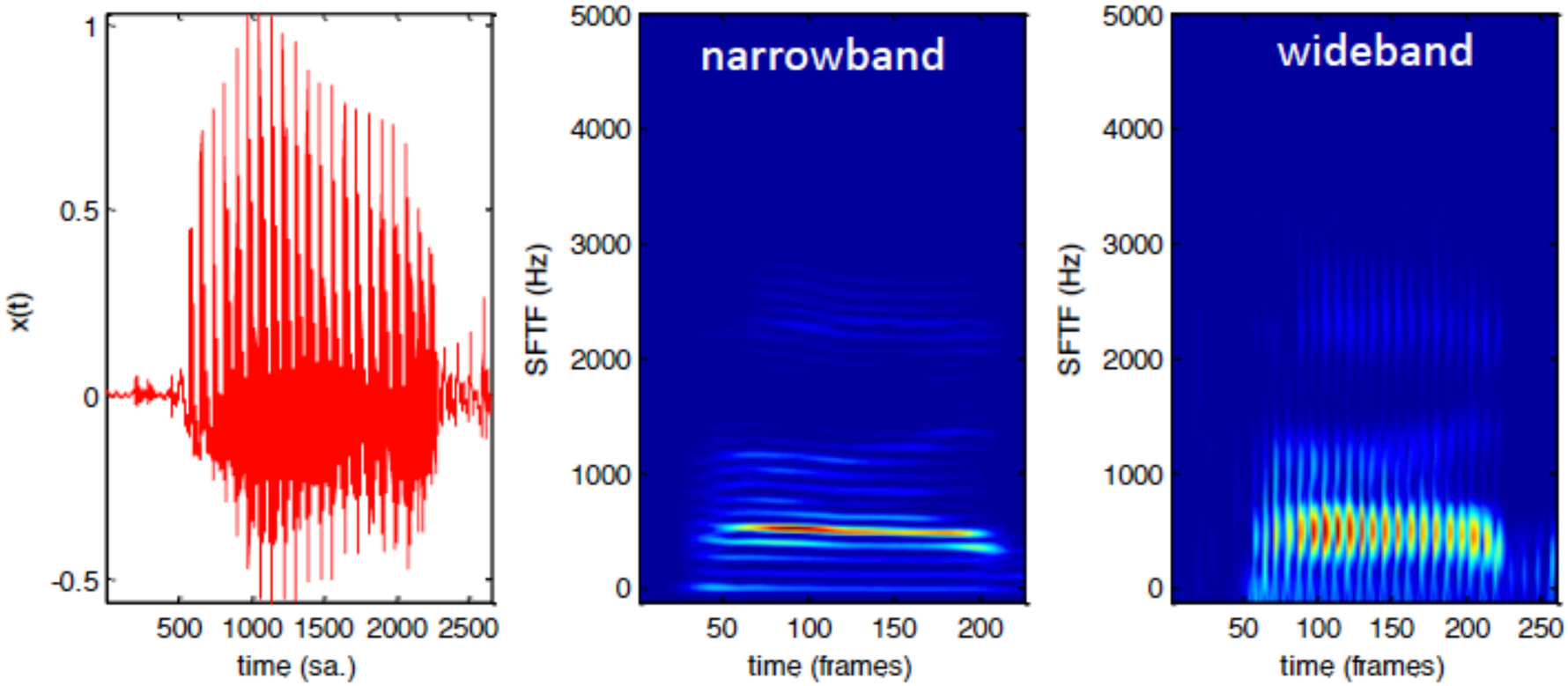
Hamming



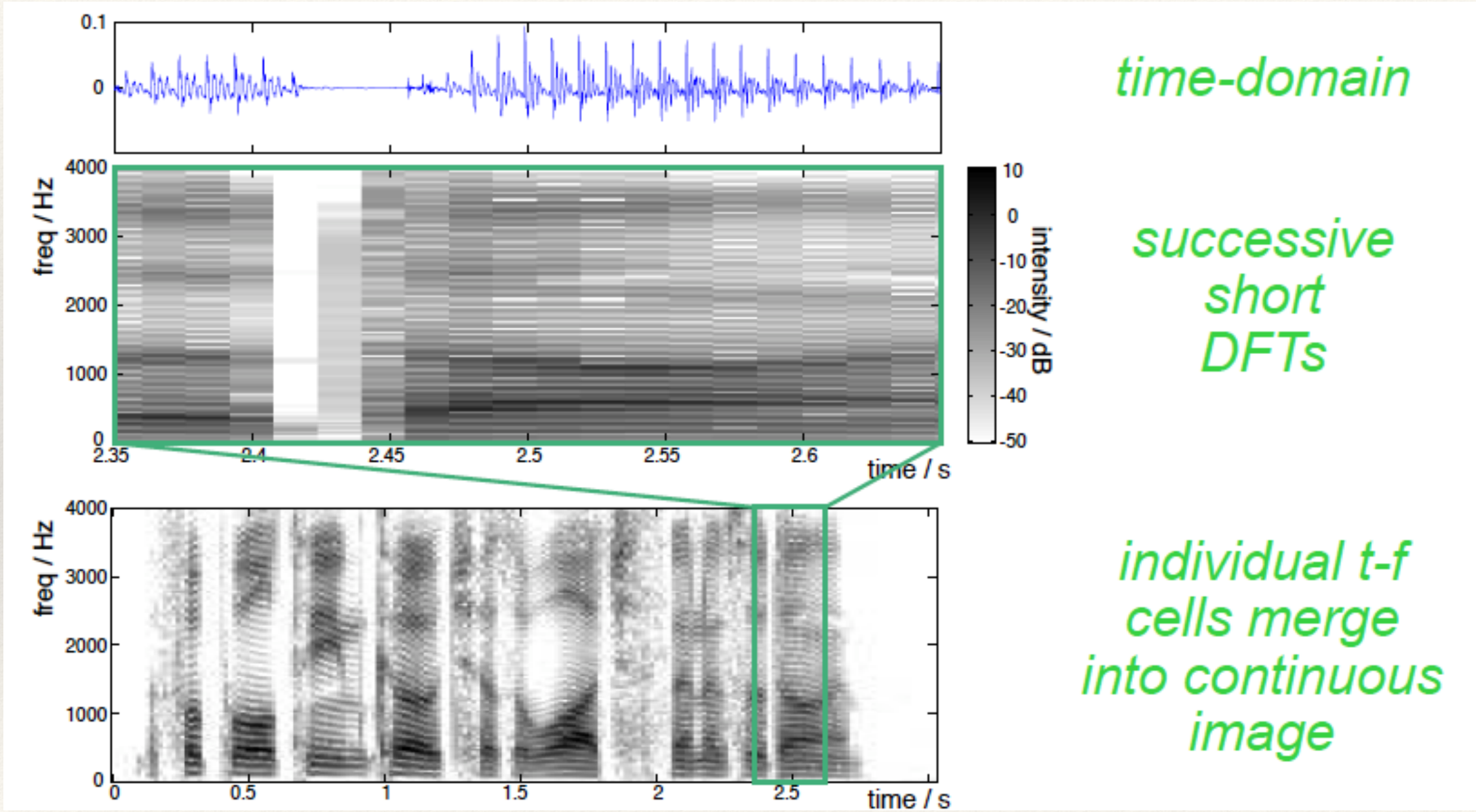
http://en.wikipedia.org/wiki/Window_function

Narrowband versus Wideband

- Short windows - poor frequency resolution - wideband spectrogram
- Long windows - poor time resolution - narrowband spectrogram



Spectrogram of Real Sounds



Dan Ellis, “STFT Tutorial”