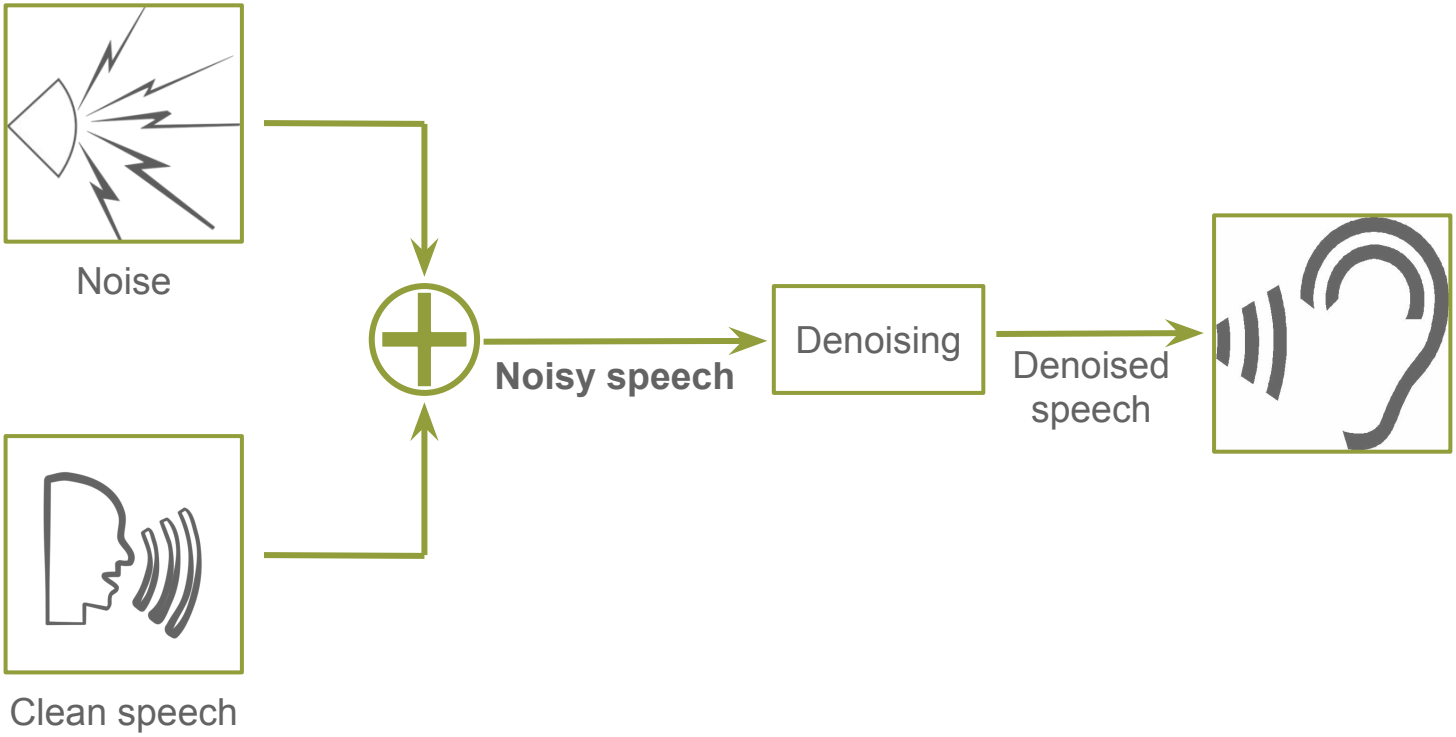
Speech Denoising via Nonnegative Matrix Factorization

Oleg Alenkin
Artyom Chashchin
Vladimir Chernykh
German Novikov

December 2016

Problem


Noise
Clean speech
Noisy speech
Denoising
Denoised speech

Applications

Automatic Speech Recognition

Telephone conversations

Hearing aids

Data

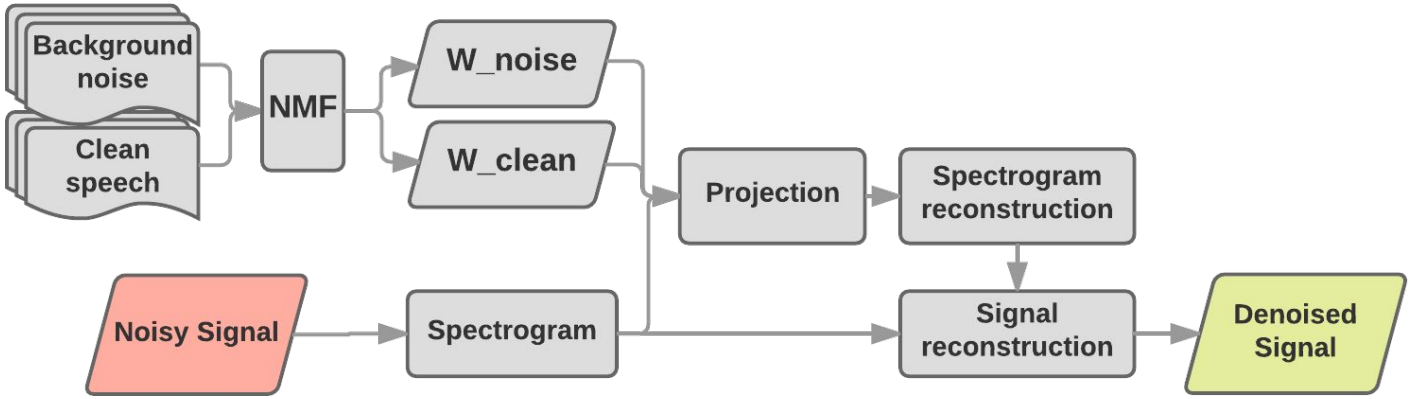
- CHiME Speech Separation and Recognition Challenge
  http://spandh.dcs.shef.ac.uk/chime_challenge/chime_download.html
  Recordings of WSJ utterances + 8 hours of noise

- Berlin Database of Emotional Speech
  http://www.emodb.bilderbar.info/download/
  Clean utterances

- Aurora noising
  http://aurora.hsnr.de/download.html

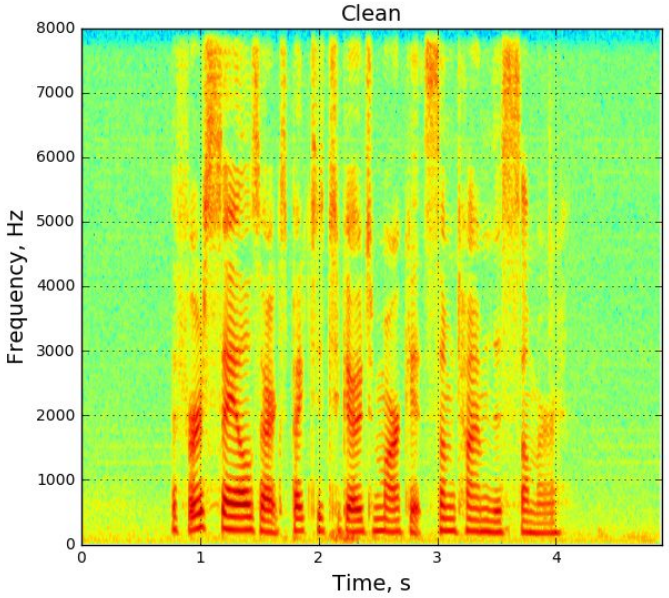
Solution

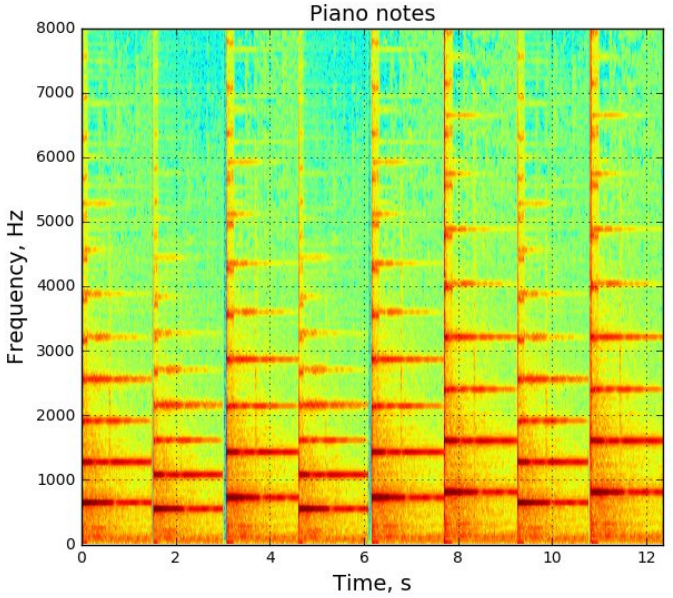
- Learn frequency patterns from speech and noise via NMF
- Decompose new signal by joint “dictionary” of patterns
- Take only projection corresponding to “clean speech”



Spectrogram

- Represent signal in Time-Frequency domain
- Built by means of Short-Term Fourier Transform - FFT in sliding window
- STFT - Complex spectrogram S ⇒ Amplitudes V

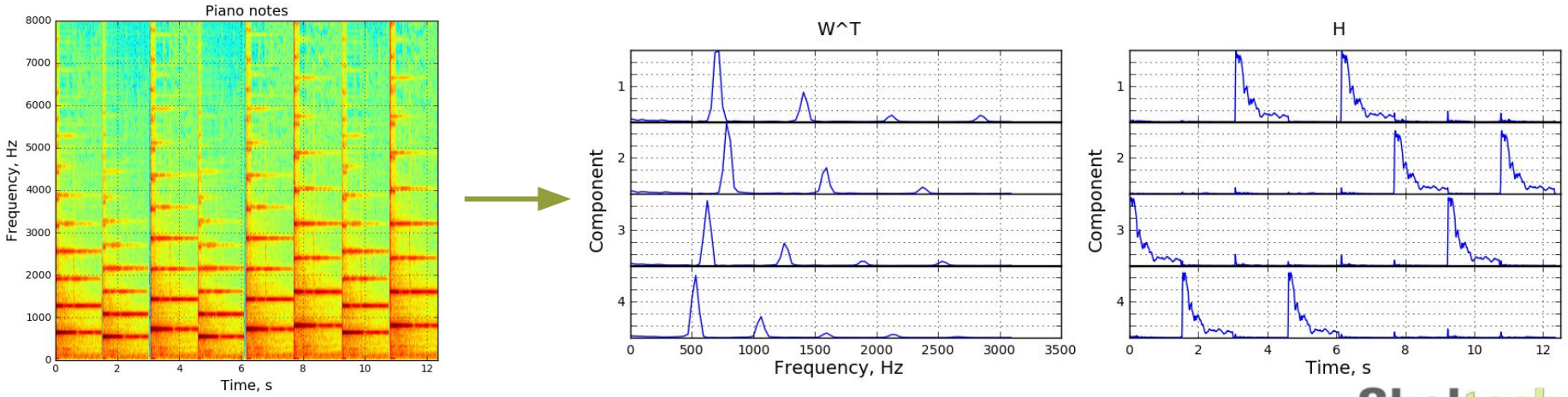




Skoltech
Skolkovo Institute of Science and Technology

NMF

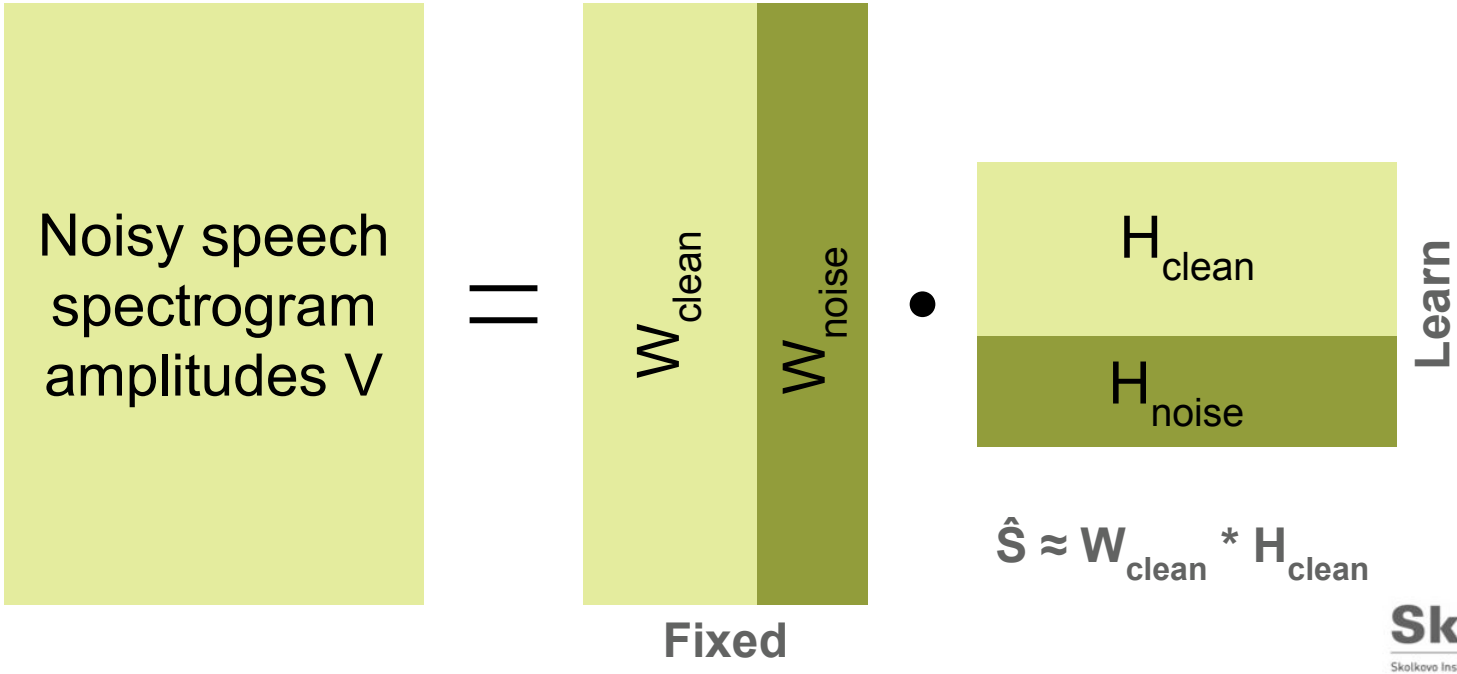
- Factorization V ≈ W * H, where V, W, H - real nonnegative
- Interpretation: W - frequency patterns, H - time-activation matrix
- Hidden dimension ≈ number of phonemes ≈ 40
- Learn speech and noise “building blocks”



Skoltech
Skolkovo Institute of Science and Technology

Projection

- Join “dictionaries” - concatenate matrices W_{noise} and W_{clean}
- Project signal onto them



Skoltech
Skolkovo Institute of Science and Technology

NMF: how to compute

Optimization problem:

$$(W^*, H^*) = \operatorname{argmin}_{W \geq 0, H \geq 0} D\left(V, WH\right)$$

$$D\left(P, Q\right) = \sum_{i=1}^{m} \sum_{j=1}^{n} d(p_{ij}, q_{ij})$$

The most popular metrics:

$$d(p, q) = (p - q)^2 \quad \text{Frobenius norm}$$

$$d(p, q) = p \ln(\frac{p}{q}) - p + q \quad \text{KL divergence}$$

Multiplicative Update Method:

$$[\nabla_H]_{kj} = \frac{\partial D\left(V, WH\right)}{\partial h_{kj}} = [\nabla_H^+]_{kj} - [\nabla_H^-]_{kj}$$

$$h_{kj} \leftarrow h_{kj} - \frac{h_{kj}}{[\nabla_H^+]_{kj}} \left([\nabla_H^+]_{kj} - [\nabla_H^-]_{kj}\right)$$

Skoltech
Skolkovo Institute of Science and Technology

NMF: how to compute

Alternating Nonnegative Least Squares:

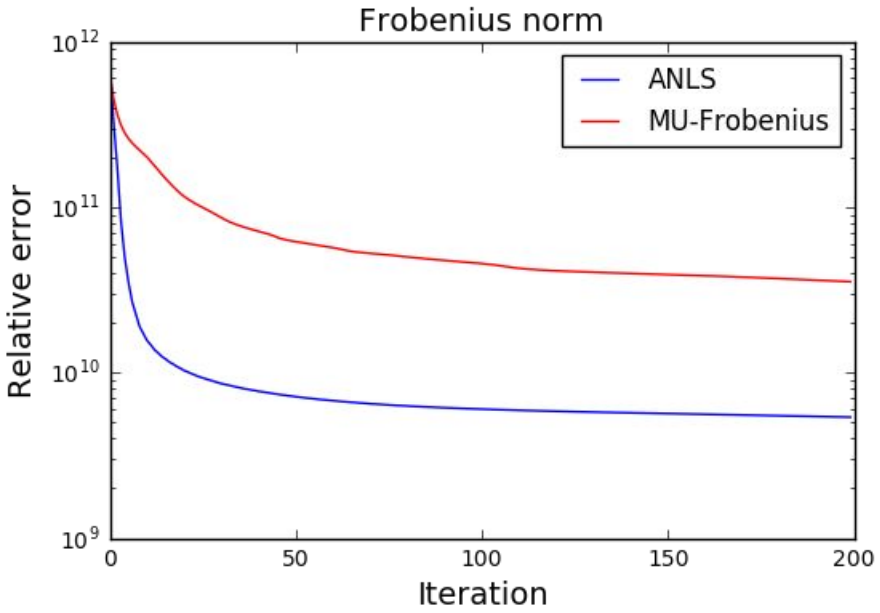
1) Initialize $W_{ia}^1 \geq 0, H_{bj}^1 \geq 0, \forall a, i, b, j.$

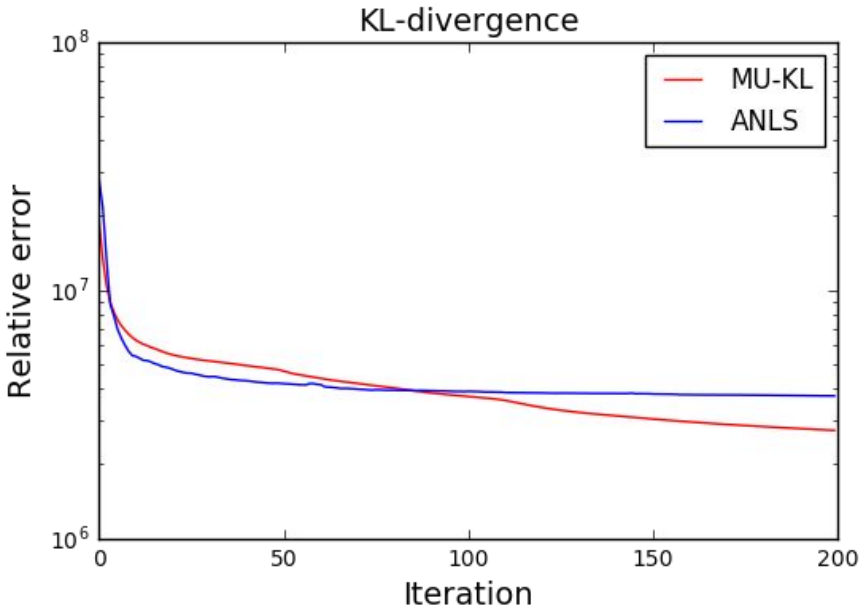
2) For k=1,2,...

$$W^{k+1} = \text{argmin}_{W \geq 0} D\left(V, WH^k\right),$$

$$H^{k+1} = \text{argmin}_{H \geq 0} D\left(V, W^{k+1}H\right).$$

Methods Convergence





Skoltech
Skolkovo Institute of Science and Technology

Signal reconstruction

- Naive method with zero phase:

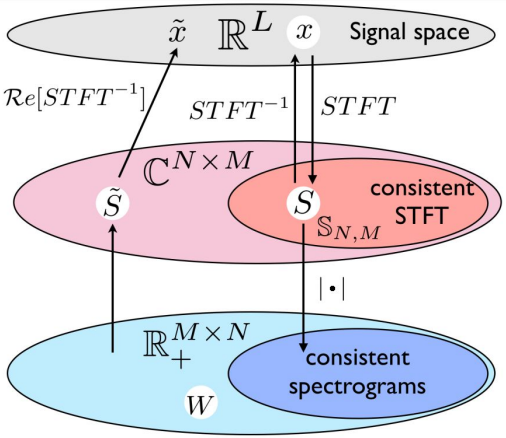
$$\hat{X} = STFT^{-1}(\hat{S})$$

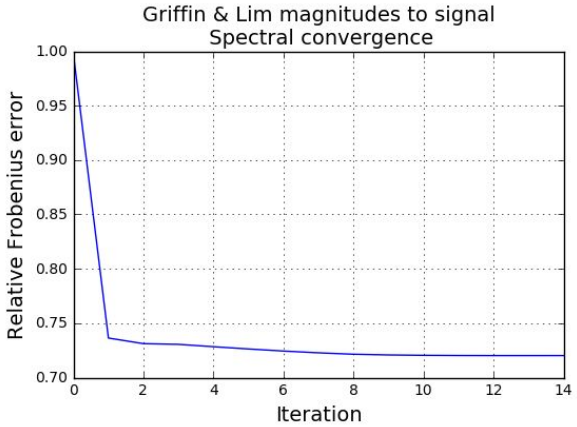
- Noisy signal phases:

$$\hat{X} = STFT^{-1}(\hat{S} \times \exp(i \angle STFT(X_{noisy})))$$

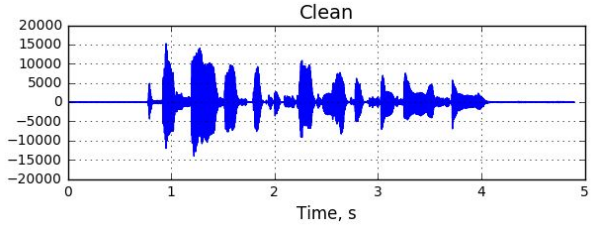
- Griffin & Lim iterative method:

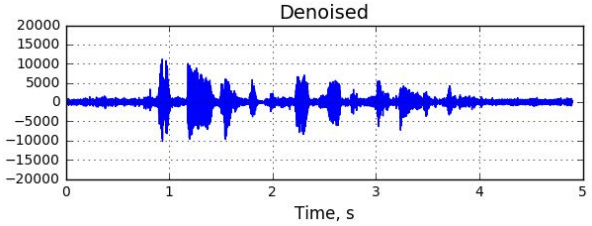
$$\hat{X}_n = STFT^{-1}(\hat{S} \times \exp(i \angle STFT(\hat{X}_{n-1})))$$

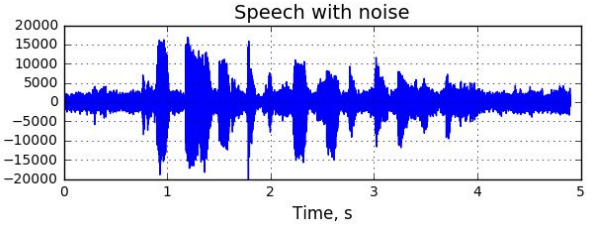


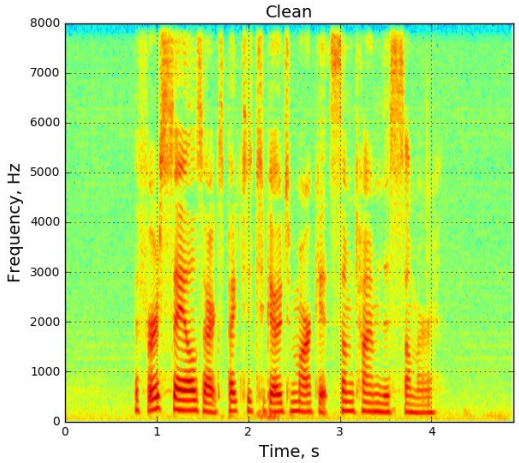


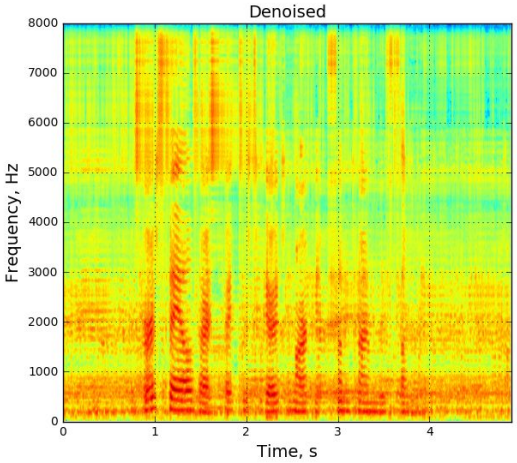
Demo

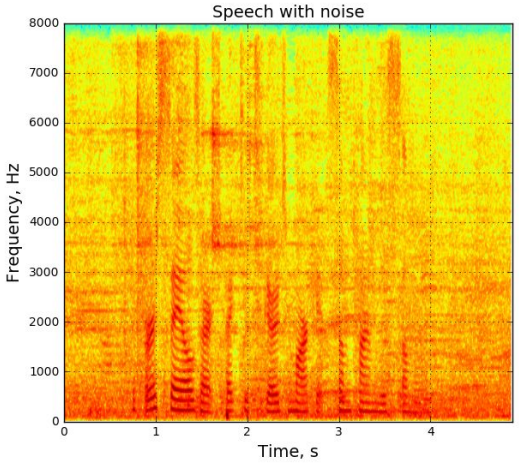












Skoltech
Skolkovo Institute of Science and Technology