Signal Reconstruction from Sampled Data Using Neural Network

Poster presentation by
Ron Teichner and Asya Leikin

Supervisor: Dorit Baras
Sampling and Reconstruction

For the following signal:

- Sampling above Nyquist sampling rate:

- Sampling below Nyquist sampling rate – aliasing:
Neural Network

A Neural Network generally maps a set of inputs to a set of outputs after a training procedure on a finite set of input-output examples.

Training is the act of presenting the network with some sampled data and modifying the weights to better approximate the desired output.
The Human auditory system

- A very complex system with the phenomena of:
  - Sensitivity to log-spectrum - Octaves
  - Critical bands
  - Threshold of hearing
  - Masking
The Paper:
“SIGNAL RECONSTRUCTION FROM SAMPLED DATA USING NEURAL NETWORK”
Akihito Sudou, Pitoyo Hartono, Ryo Saegusa, Shuji Hashimoto

- A signal sampled with Nyquist rate underwent data reduction.
- Reconstruction was made both with a linear perceptron and a multilayered perceptron (with one hidden layer).
- For comparison a reconstruction using FIR filter with sinc was made.
- The NN did the training by minimizing the MSE.
- Two experiments were made:
  - Comparing the errors between the different reconstructions and the original signal for different data reduction rates.
  - Comparing the results of the different reconstructions while changing the number of inputs to the NN and the data reduction rate.
Results from the paper:

Note: the criteria for a good reconstruction was the average MSE not the sound quality.

- The experiment showed that the perceptron achieved better reconstruction then the FIR reconstruction. Also it was shown that for larger data reduction the mistake in the reconstruction is larger.
- They didn't see a meaningful difference in the reconstructions by the linear perceptron and the multi-layer perceptron with one hidden layer.
- The second experiment showed that using a perceptron with larger number of inputs improves the reconstruction
Time Domain

- We are using 2 signals: speech and musical.

- Down sampling N → Up sampling N → Sinc/ NN reconstruction

- Up sampling does zero interpolation and is probably not optimal for the reconstruction. We tried other types of interpolation - linear:

- LP: \( w_s = f/2*N \) → Down sampling N → Up sampling N → Regular sinc reconstruction

- ...
Frequency Domain

- Down sampling N
- Up sampling N/interpolation
- DFT transform
- NN reconstruction
- 10log

- Zero padding x7/ x10
- DFT transform
- Phase and Amp. separate
- NN reconstruction for each
- NN reconstruction for phase only

Gave the best results
Phase and amplitude reconstruction vs. only phase reconstruction:

- Red - the reconstructed signal
- Phase and amplitude reconstruction and then combination
- Using the unreconstructed amplitude and the reconstructed phase.
- We can see that the amplitude reconstruction adds a lot of noise.
Zero padding

- Zero padding is actually a way to increase the frequency resolution.
- We did only phase reconstruction and used either the original (non-reconstructed amplitude) or the average amplitude.
- The average amplitude showed better results.
- We used PESQ and our hearing as an estimate for the quality of reconstruction here are the best results that we got for the speech signal.

<table>
<thead>
<tr>
<th></th>
<th>PESQ</th>
<th>Comp. rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>L=30 seg l=0.1 pad=7</td>
<td>1.679</td>
<td>1.276</td>
</tr>
<tr>
<td>L=30 seg l=0.1 pad=6</td>
<td>1.677</td>
<td>1.413</td>
</tr>
<tr>
<td>L=10 seg l=0.1 pad=7</td>
<td>1.670</td>
<td>2.305</td>
</tr>
<tr>
<td>L=20 seg l=0.1 pad=7</td>
<td>1.685</td>
<td>1.643</td>
</tr>
<tr>
<td>L=30 seg l=1 pad=7</td>
<td>1.716</td>
<td>3.297</td>
</tr>
<tr>
<td>L=30 seg l=0.5 pad=7</td>
<td>1.689</td>
<td>2.803</td>
</tr>
</tbody>
</table>
Sending and receiving:

- **Sending:**
  - Down sampling, interpolation, Zero padding
  - FFT transform, 10log, phase and amplitude separation
  - NN training on the phase and saving the weights

  Sending: l- length of the signal after down sampling, N- down sampling rate, type of the interpolation, zero padding rate, L- window length, segment length, down sampled signal, weights in the correct order.

- **Receiving:**
  - Interpolation, Zero padding
  - FFT transform, 10log, phase and amplitude separation
  - Phase reconstruction using the weights received, finding the average amplitude
  - Composing the signal from the reconstructed phase and the average amp.
  - Noise reduction in time domain
Conclusions

- The reconstructed signals are very noisy and need a lot of noise reduction processing if a good quality of sound is required but if the goal is only to understand the sentence spoken the quality is good enough.
- This method proved to be not useful as a compression method because there are much better ways of compression that deliver a much better sound quality after decompression.