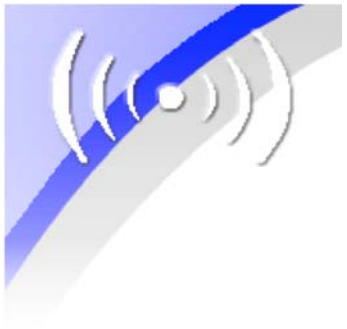


Master’s thesis presentation

Parameter Estimation of a Plucked String Synthesis Model via the Genetic Algorithm

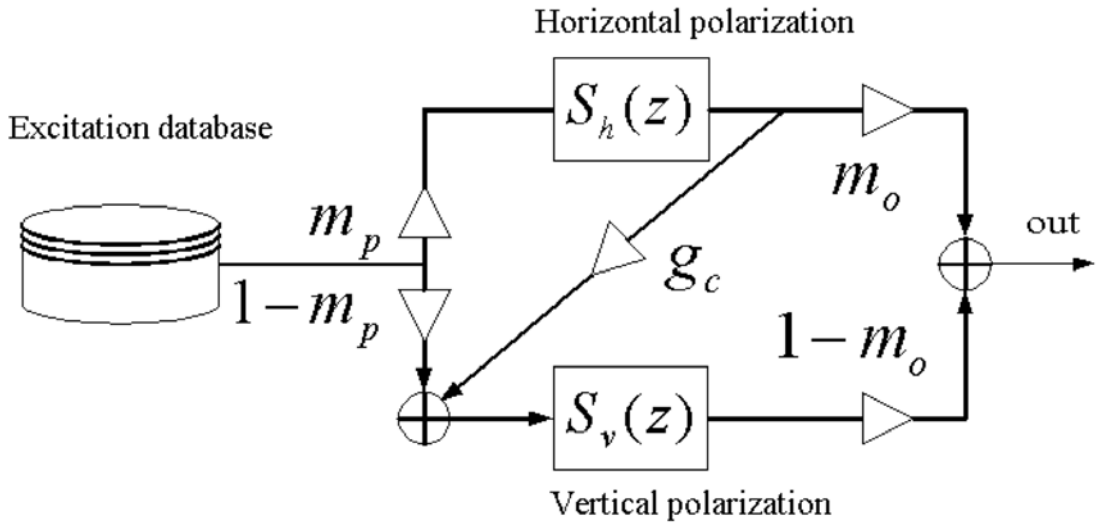
Janne Riionheimo
Helsinki University of Technology
Laboratory of Acoustics and Audio Signal Processing

Outline

- Plucked string synthesis model
- Estimation procedure
  - Genetic algorithm (GA)
  - Fitness calculation
  - Implementation
- Results
- Conclusions

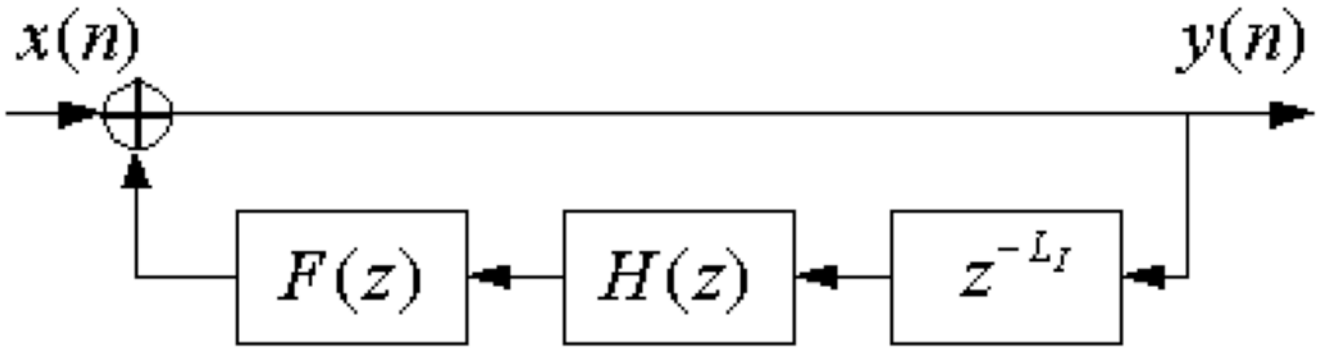
Plucked String Synthesis Model

- Proposed by Karjalainen et al. (1998) and based on digital waveguide synthesis theory (Smith, 1992)
- Dual polarization and coupling
- Used intensively to synthesize various plucked string instruments
- Excitation signal is obtained by subtracting the harmonics from recorded tones and equalizing (Välimäki and Tolonen 1998)



Plucked String Synthesis Model (2)

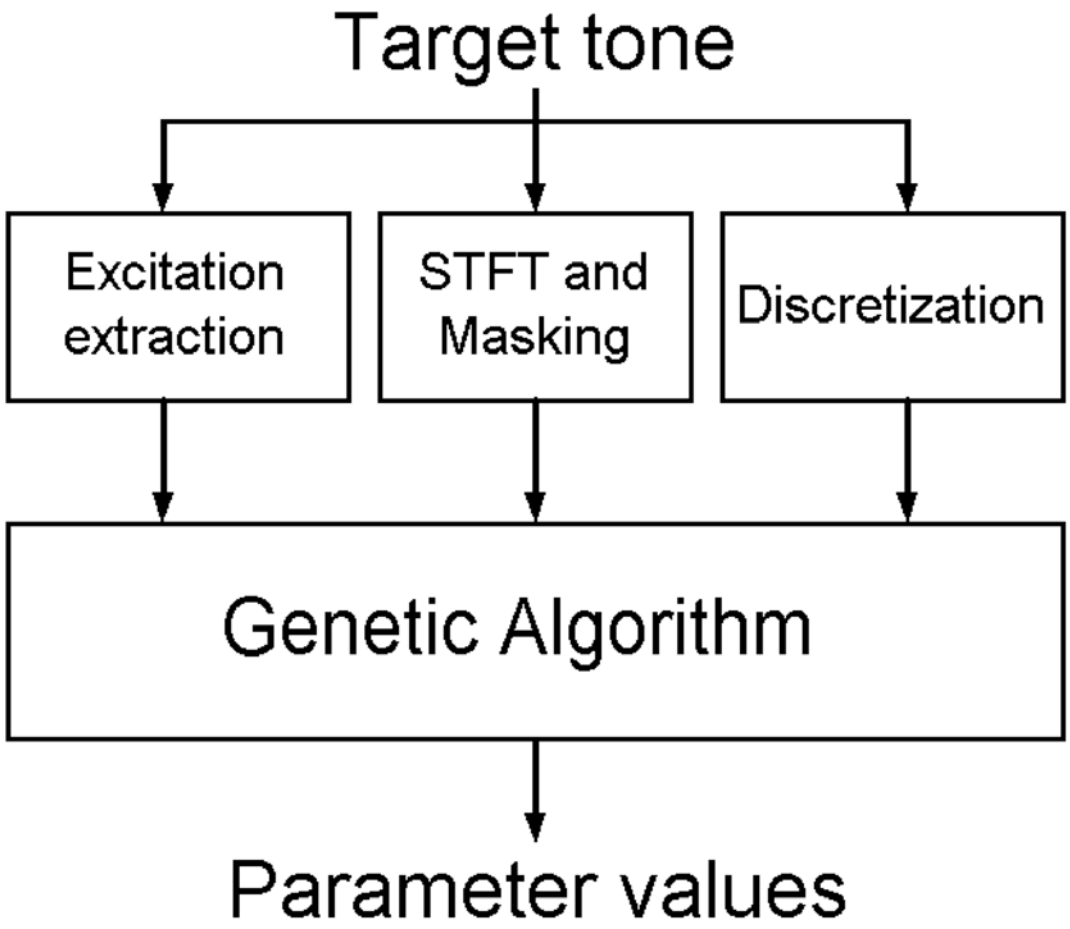
- Proposed by Jaffe and Smith (1983).
- Delay line and fractional delay filter *F(z)* control fundamental frequency
- Filter *H(z)* is controlled by loop gain parameter *g* and frequency dependent gain *a*
- Parameters estimation earlier: Algorithm + hand tuning



Parameters to estimate

parameter	control
$f_{0,h}$	fundamental frequency of the horizontal string model
$f_{0,v}$	fundamental frequency of the vertical string model
g_h	loop gain of the horizontal string model
a_h	frequency dependent gain of the horizontal string model
g_v	loop gain of the vertical string model
a_v	frequency dependent gain of the vertical string model
m_p	input mixing coefficient
m_o	output mixing coefficient
g_c	coupling gain of the two polarizations

Parameter Estimation Procedure



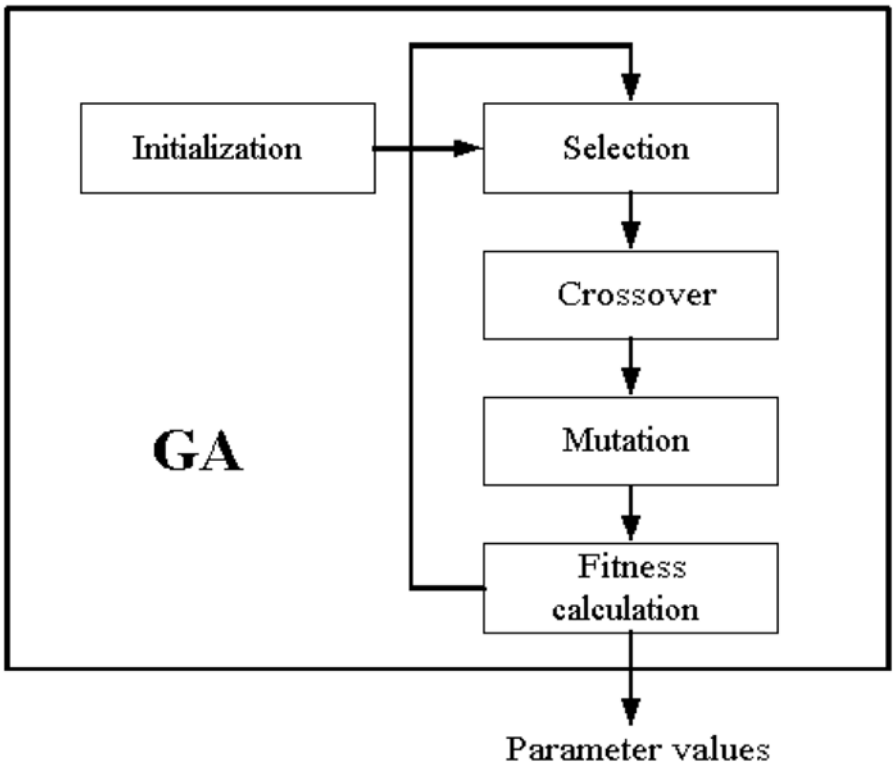
Genetic Algorithm

- Mimics the evolution of nature → fittest individuals survive
- Operate on a population of chromosomes

f_0	d_f	g_h	a_h	g_v	a_v	m_p	m_o	g_c
330	0.9	0.991	-0.21	0.922	-0.64	0.11	0.78	0.23

- In every generation, a new set of chromosomes is created by crossing and mutating the fittest individuals
- Desired characteristics are passed into the next generation while imperfect individuals become extinct

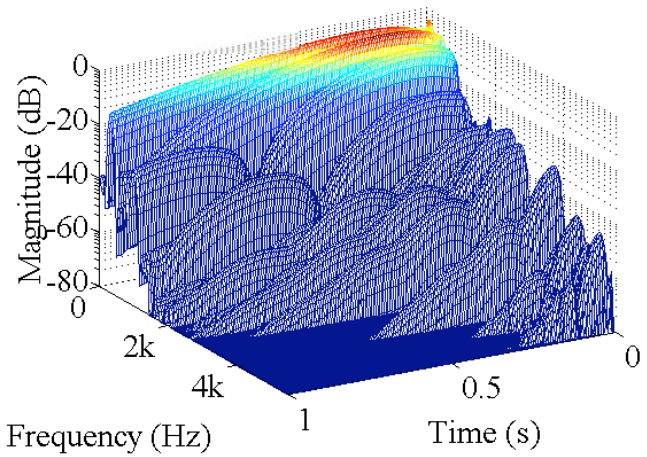
Genetic Algorithm (2)



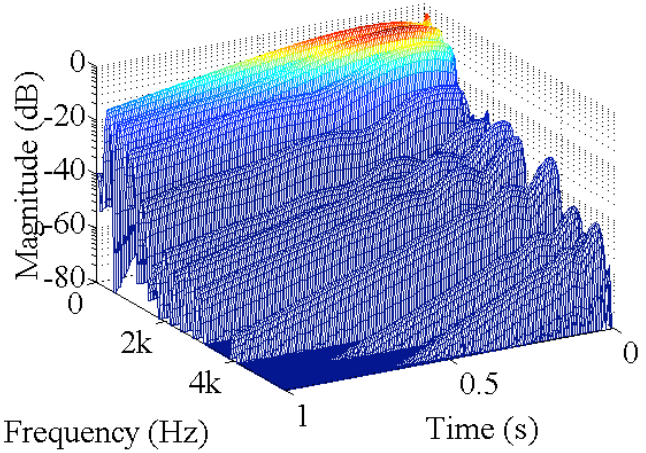
How to rank sounds?

- Listening tests
- Psychoacoustic model
  - Frequency dependence
  - Frequency masking

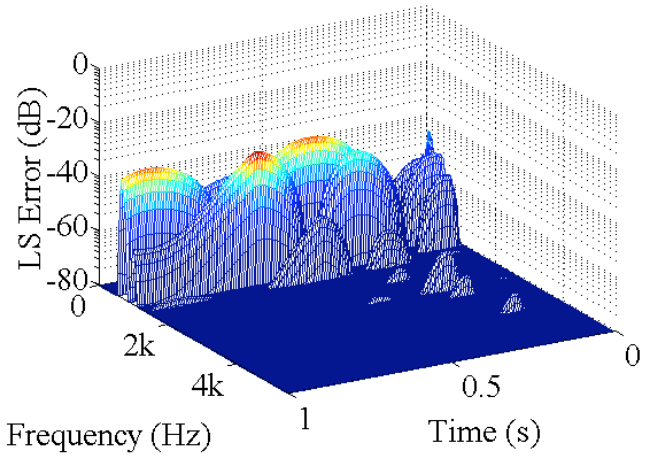
Fitness/error calculation

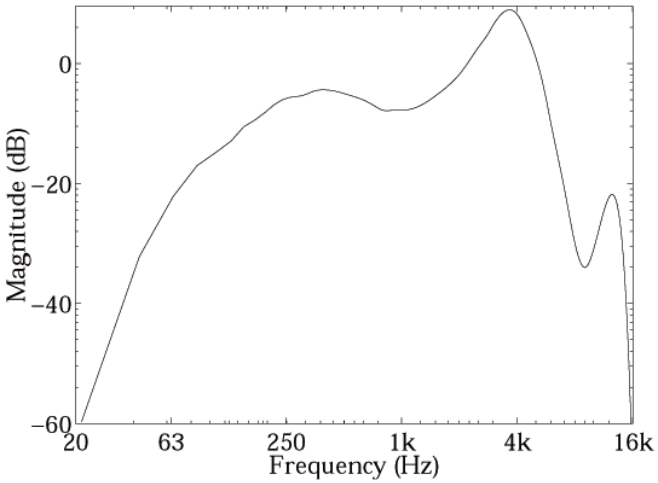


-



$()^2=$

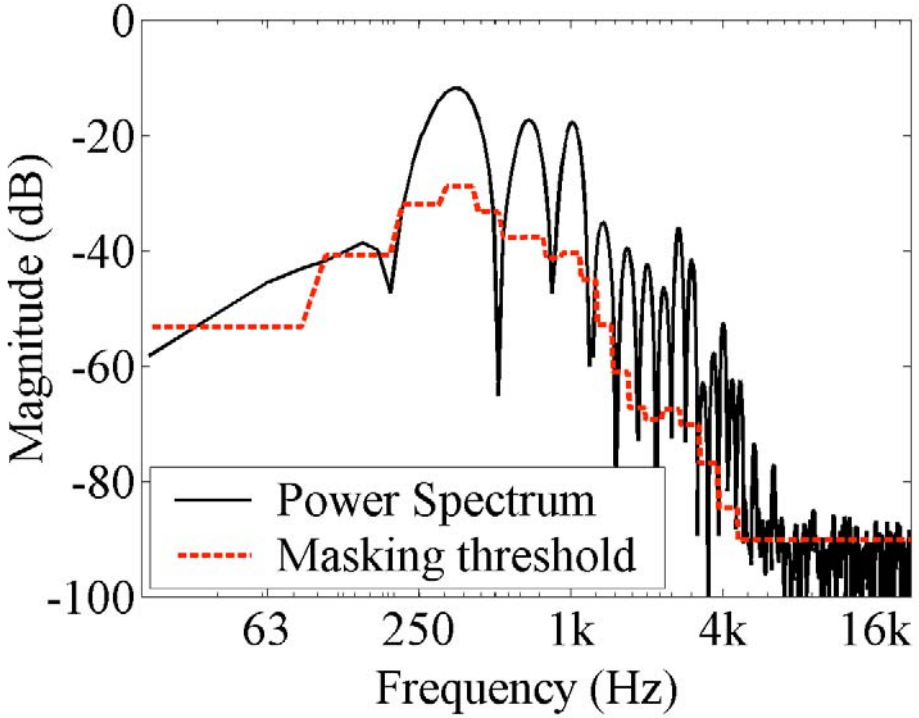




Error= 21.2

Calculation of Threshold of Masking

- Based on Johnston's model (Johnston 1988)
  - Energy per critical band
  - Spreading function
  - Final masking threshold
- Frequency components below the threshold are inaudible and therefore unnecessary



Discretizing the parameters

- Perceptually reasonable discretation:

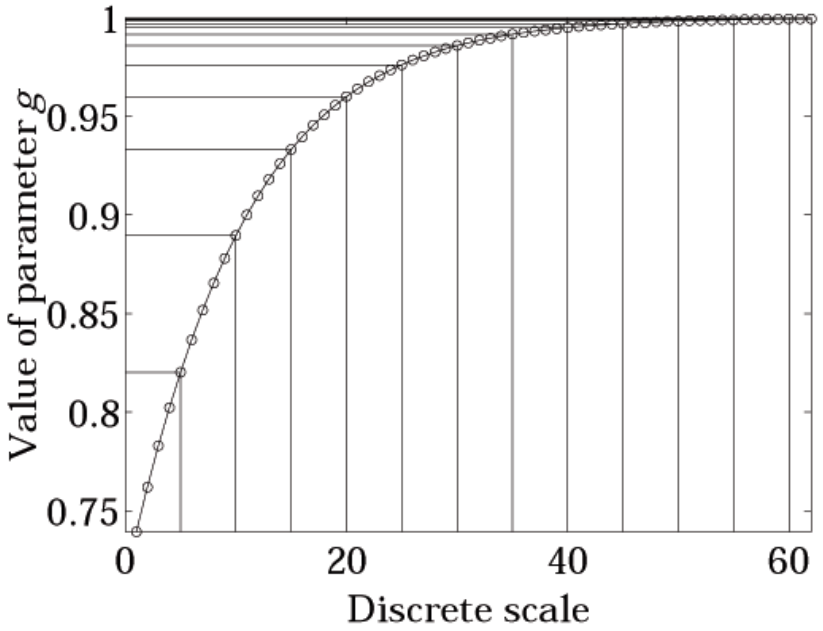
Use only parameter values that produce different sounding tones!

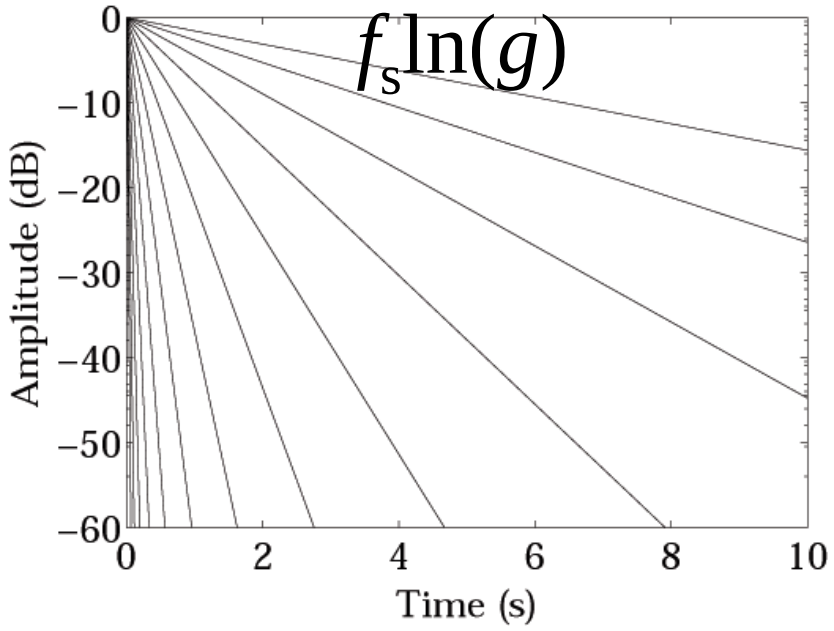
- Based on previous studies, informal listening and sensitivity of the fitness function
- Research is in progress in Aku lab?

Discretizing the decay parameters

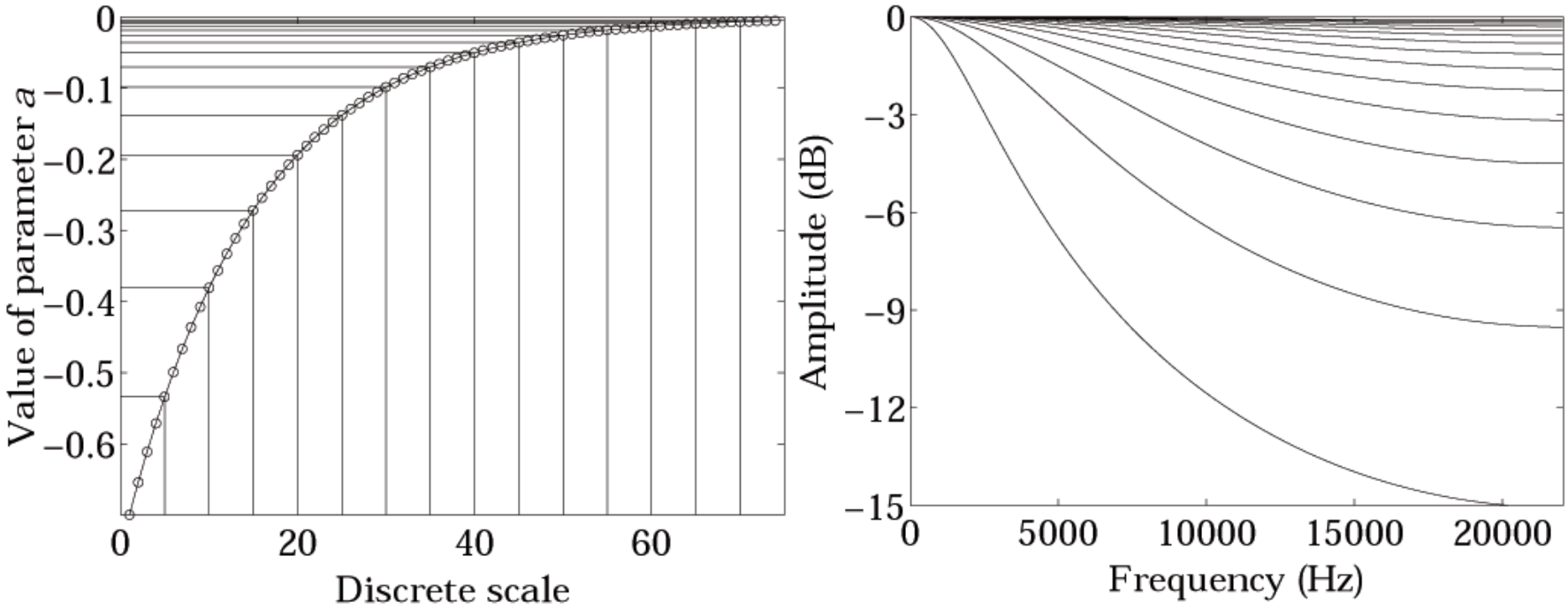
- Inaudible variation of the time constant τ is 75% - 140% and parameter *a* between 83% - 116% (Järveläinen and Tolonen 2001)
- We have ± 10% for the τ and ± 7% for the *a*

$$t = -L / f_s \ln(g)$$



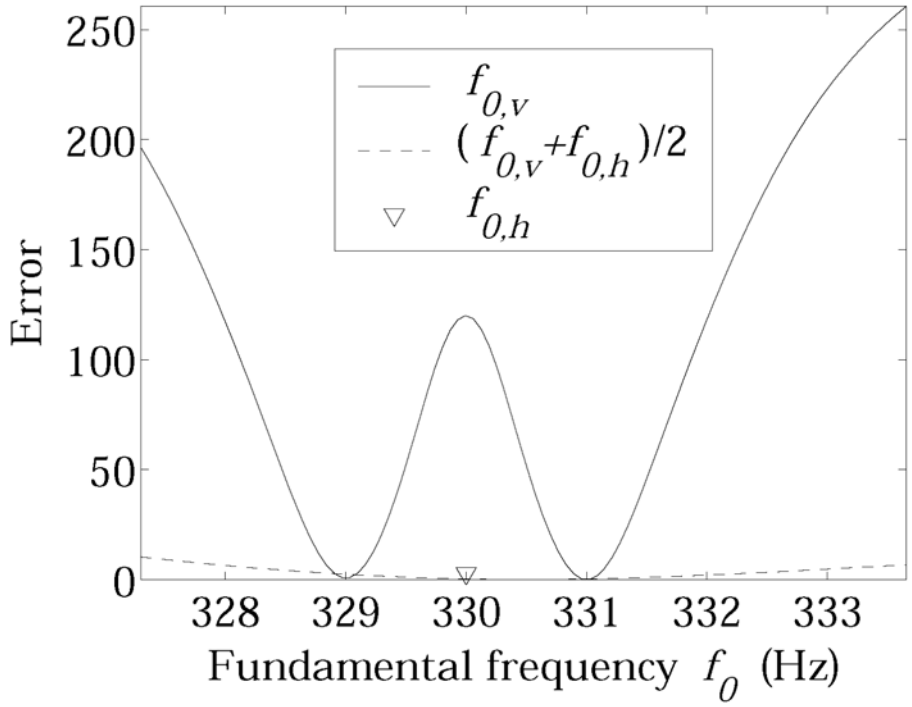


Discretizing the decay parameters (2)



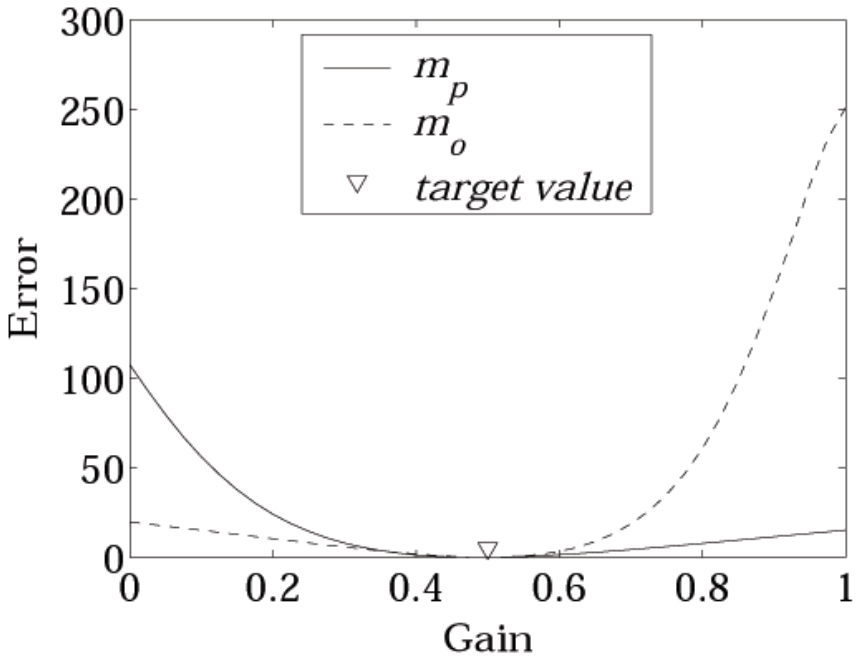
Discretizing the fundamental frequency

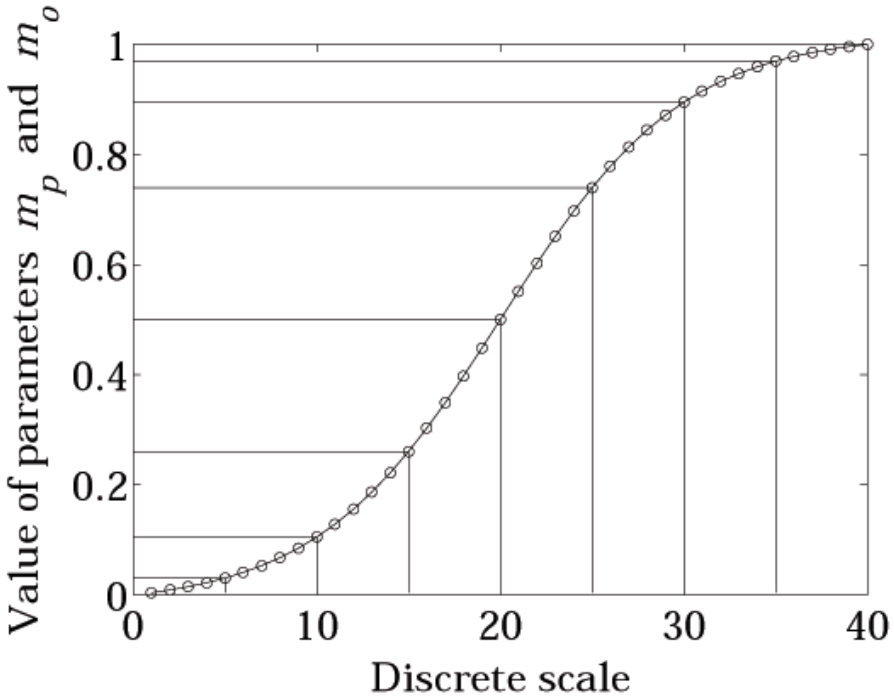
- Sensitivity of the fundamental frequency parameters is examined by varying the parameters and calculating the error
- Beating is significant. Estimate mean frequency f_0' and difference d_f



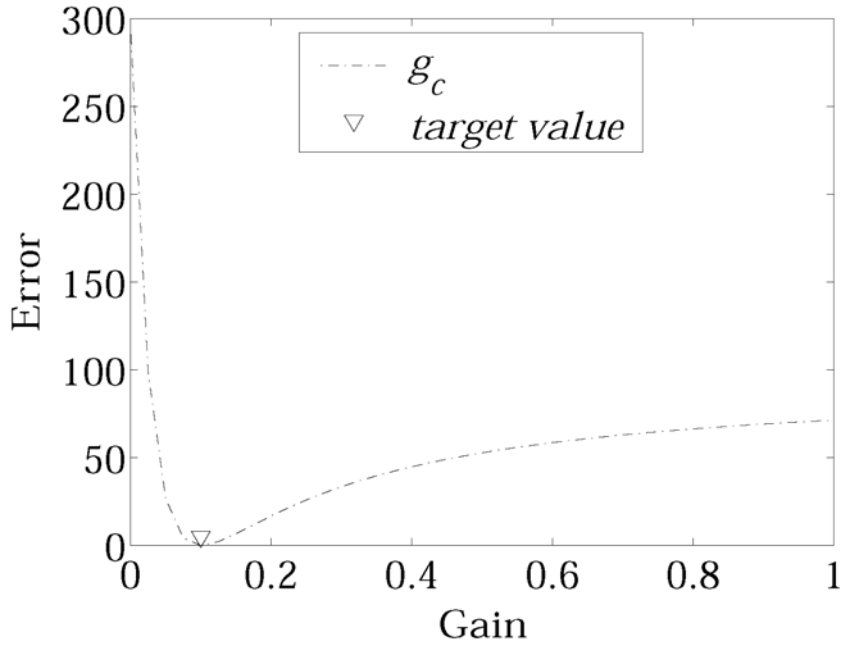
Discretizing the mixing parameters

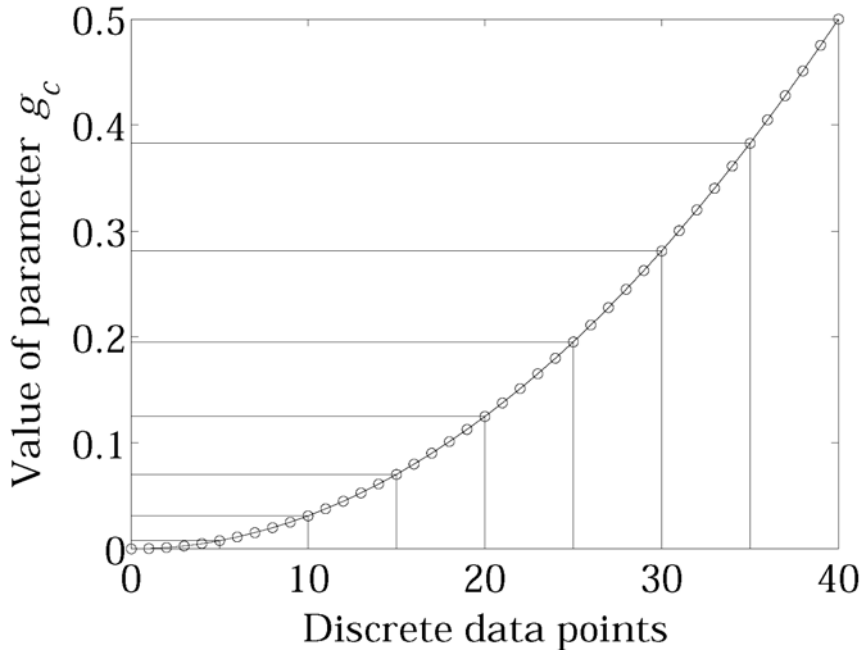
- Sensitivity of the mixing parameters is examined by varying the parameters and calculating the error
- Most sensitive at the boundaries which accords with listening





Discretizing the mixing parameters (2)

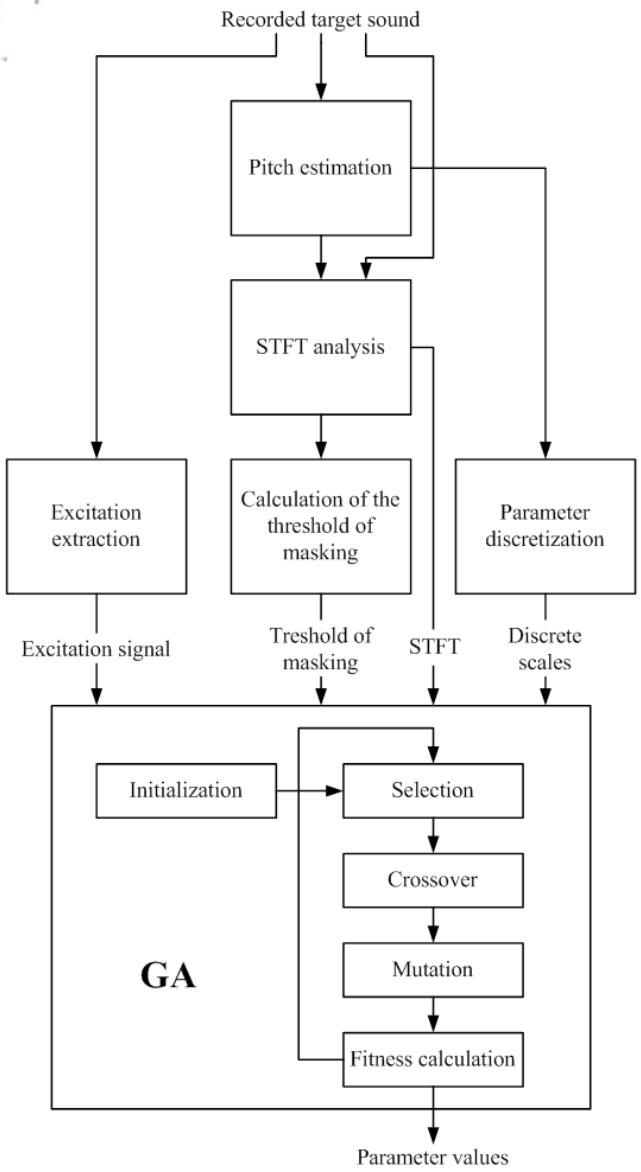




Discretizing the parameters

- 62 values for g
- 75 values for a
- 20 values for f_0'
- 100 values for d_f
- 40 values for m_o, m_p and g_c
- Total 2.77×10^{15} combinations

Implementation



Results

- Procedure was first tested with synthetic tones - target values are known
- Mixing and coupling parameters were noticed not to be orthogonal, implying that similar tones can be synthesized with totally different parameter values
- If m_o and m_p are fixed the precise parameter values are estimated when using similar excitation signals

Synthesized target tone

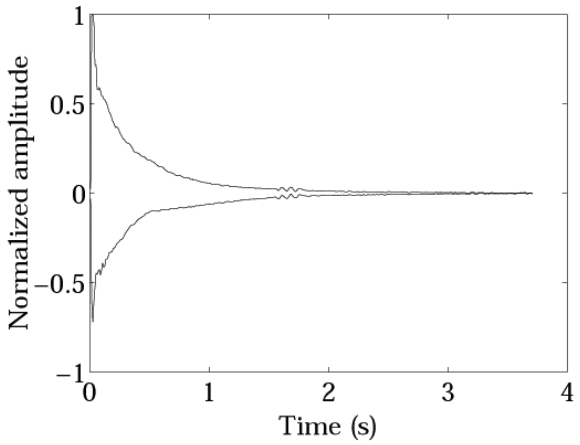
Estimated tone with error = 0.0112

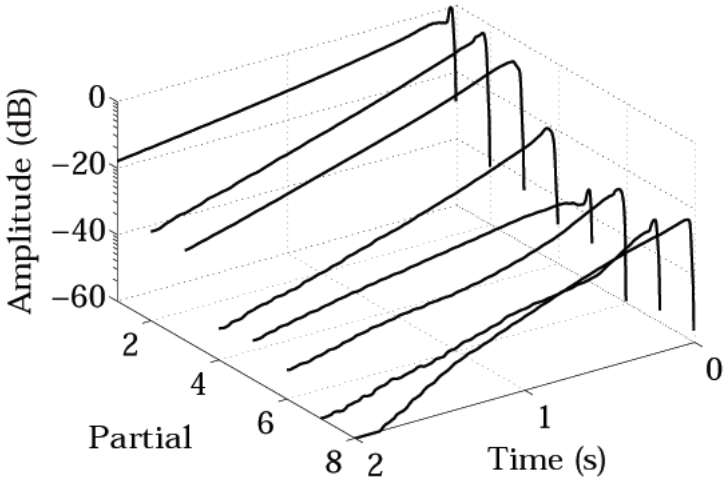
Estimated tone with error = 0.464

Estimated tone with extracted excitation signal

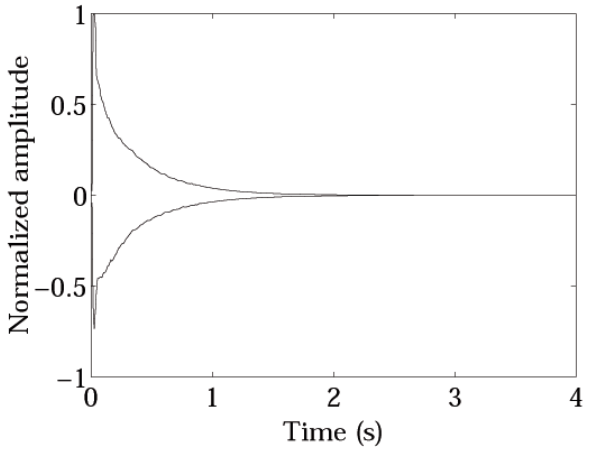
Results (2)

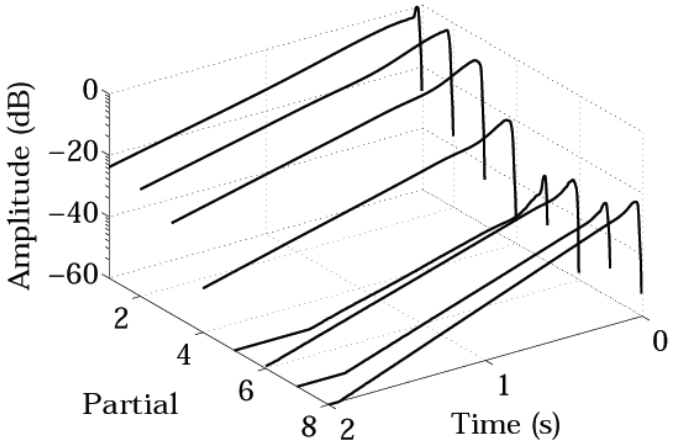
Real tone





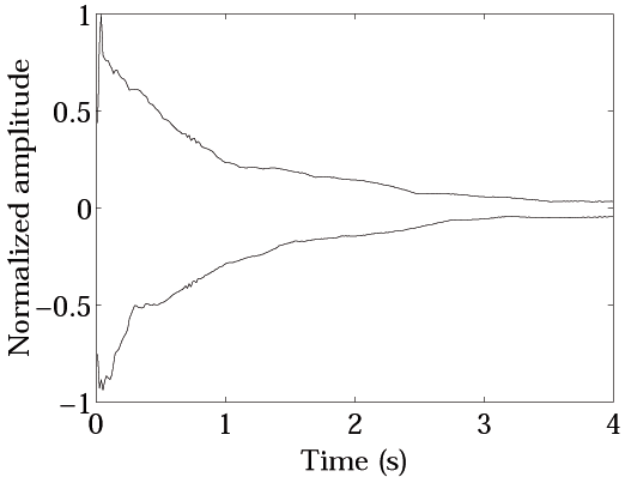
Synthesized tone

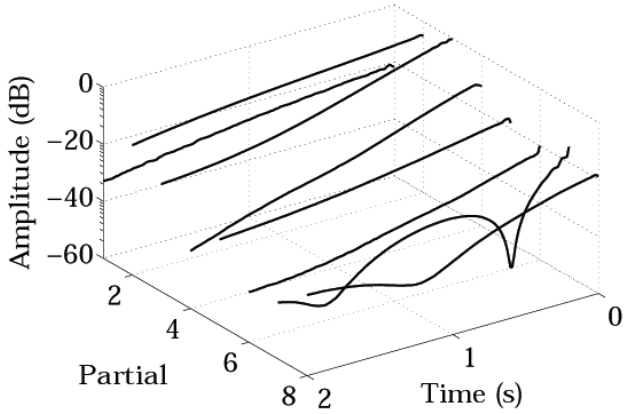




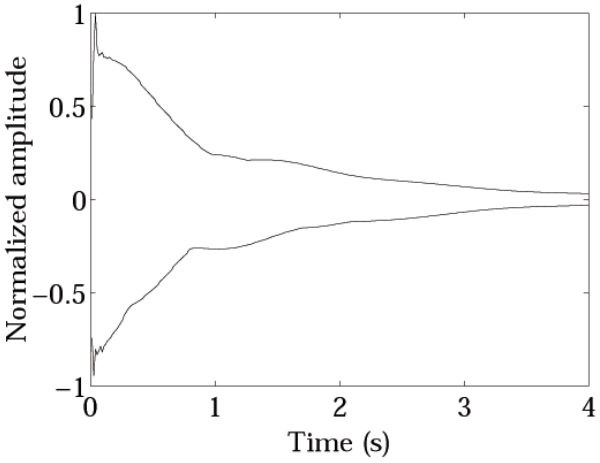
Results (3)

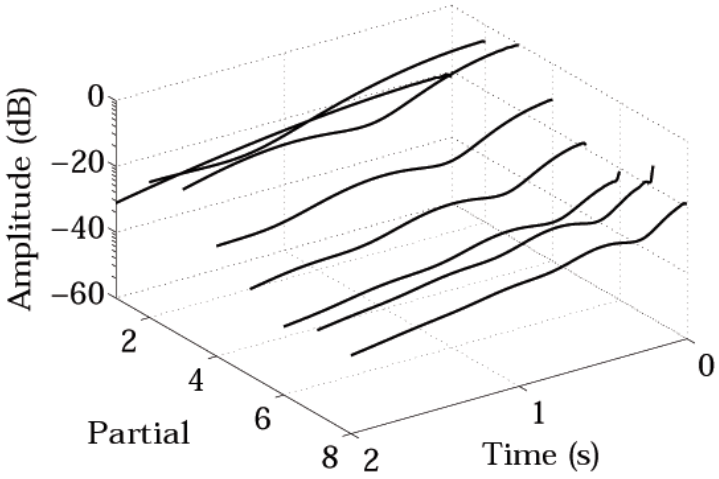
Real tone





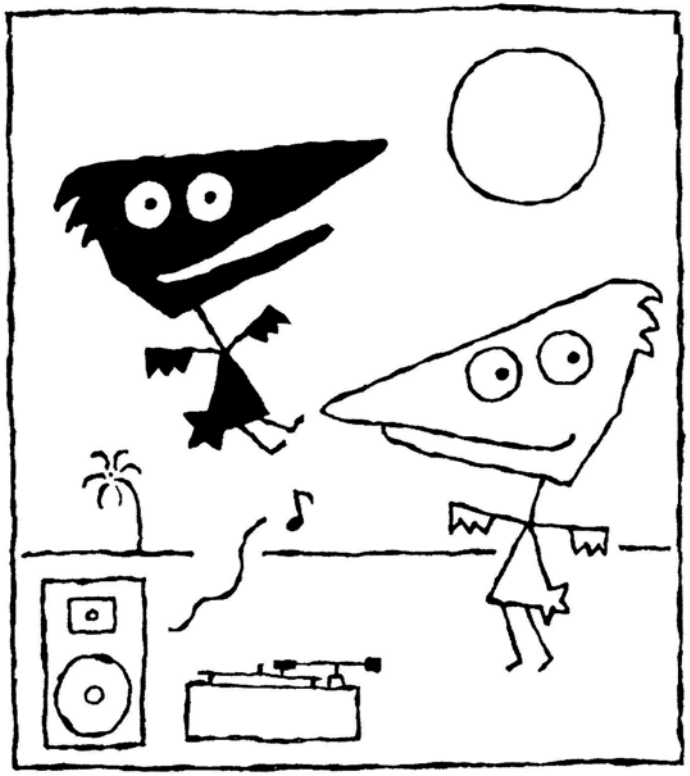
Synthesized tone





Conclusions

- Genetic algorithm is a universal optimizing tool, which is well suited to our problem
- The most important question is the fitness calculation. How to rank sounds?
- Our procedure works great with "simple" tones
- Our procedure works with "complex" tones, but the result is not identical with the target tone
- Do we want exactly similar tones or just tones that sound realistic?
- Is the expression power of the model enough for "complex" tones?
- m_o, m_p and g_c not orthogonal
- More perceptual studies and better auditory models



http://acoustics.hut.fi/publications/papers/jasp-ga/