On the pitch strength of harmonic complex tones and comb-filter noises

Markus Fruhmann
AG Technische Akustik, MMK, TU-München, 80290 München, Germany, Email: fruhmann@mmk.ei.tum.de

Introduction

In an early investigation on pitch strength, Fastl and Stoll [1] found that pitch strength is strongest for sinusoids. All other kinds of sounds produce weaker pitches at comparable perceived pitches. For example, for harmonic complex tones, pitch strength reaches about 80 percent of the pitch strength for a sinusoid at same basic frequency.

Some more noise-like sounds have also been investigated like narrowband and bandpass noises, but also so called comb-filter noises that are created using a delay and add process with a white noise as input signal.

When using the delay and add process for several times, a so called 'Iterated-Rippled-Noise' (IRN) is generated, which, depending on the number of iterations, creates a much higher pitch strength than comb-filter noise [2].

In the following study, the spectrum of an IRN is interpreted as comparable to that of a harmonic complex tone. In an experiment, the pitch strength of harmonic complex tones is compared to that of IRNs with similar spectral properties. Additionally a simulation of the resulting data is shown and discussed.

Stimuli and Procedure

The procedure of magnitude estimation with an anchorsound was chosen for the experiments. Thereby pairs of stimuli with a duration of 0.5 seconds were presented with pauses of 0.4 seconds in-between. The first sound within a pair of sounds is the so called 'anchor' followed by the test stimulus. The subjects' task was to judge the pitch strength of the testsound relative to the pitch strength of the anchor. A constant value of 100 was given to the anchor so that the results are a percentual values giving the relative pitch strength of the testsound in comparison to the anchorsound used.

The stimuli used as testsounds were so called Iterated Rippled Noises. These sounds are created using multiple stages of a delay and add network with same values for delay and gain in each step. The FFT-spectrum of these sounds shows peaks at multiples of the inverse delay time marking the pitch evoked by these sounds.

As the purpose of the experiments was to compare harmonic complex tones and IRN, the delay time for the IRNs was set to the inverse value of the fundamental frequency. Additionally the IRNs were filtered using an FFT filter to achieve a comparable spectral composition, namely the same number of harmonics as the anchor. By this, four different numbers of harmonics ranging from 4 to 32 were used together with delay times from 2 to 8 msec. The overall stimulus level was set to 50 dB.

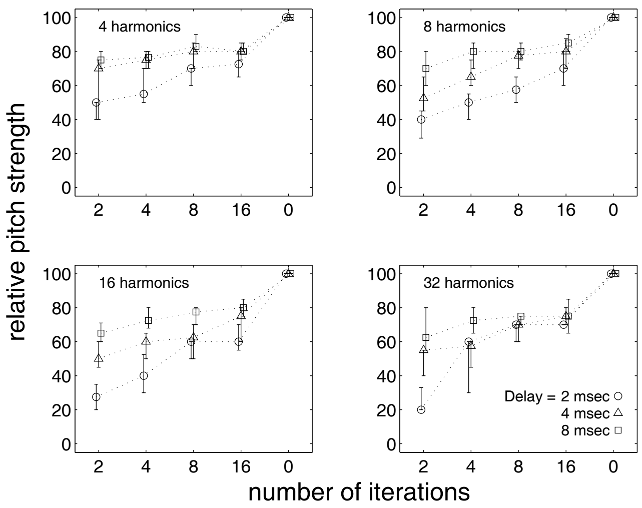

Figure 1: *Experimental results: Pitch strength of iterated rippled noise with varying number of iterations compared to a harmonic complex tone ('0' iterations). Basic frequency and number of harmonics are the same for testsound and anchor.*

All pairs of sounds were presented in random order as well as the order within a pair of sounds to avoid effects of adaption. By presenting pairs of same sounds, the subjects' concentration and reliability were also tested.

The experiments were carried out in two series with each of the eleven normal hearing subjects. Within each trial, every sound configuration was presented three times, giving an overall of six judgements per subject and sound.

Results and Discussion

The experimental results are shown in Figure 1 as medians and interquartile ranges. Thereby the median is calculated as median of the median values for each test subject and the interquartiles are calculated as the median of all subjects' interquartile values. The anchor sound (harmonic complex tone) is indicated as '0' iterations.

A stromg dependency of the results on the number of iterations can be seen. With a growing number of iterations, the pitch strength rises to about 80 percent the pitch strength of a harmonic complex tone. Depending on the delay time and thus the basic frequency, the pitch strength decreases more for higher pitches (2 msec $\simeq$ 500 Hz) than it does for lower pitches (8 msec $\simeq$ 125 Hz).

The data can be explained quite well regarding the human auditory system as a spectral analyzer, which was proposed by Feldtkeller and Zwicker ([3]) in the middle of the last century. Thereby the hearing system can be interpreted as a filterbank consisting of a set of filters called the critical bands.

For frequencies below 500 Hz, the width of the critical bands is almost constant at a value of 100 Hz. Looking at the composition of the test stimuli used, this is about the basic frequency of the sounds for a delay time of 8 msec. Thus these sounds can be resolved worse by the hearing system and the noise floor occuring in IRN play a minor role for the perception of pitch strength. Recall that the amount of noise within an IRN is inverse proportional to the number of iterations used in the creation process.

The opposite is the case for higher basic frequencies. Here the spectral components of a harmonic complex tone can be resolved very well, as, at least for lower frequencies, the critical bandwidth is well below the spacing of the harmonics. Thus the noise-like components in the IRN have a much bigger influence on the perception of sound leading to rising differences between anchor and testsound at low numbers of iterations.

With an increasing number of iterations, the spectrum of an IRN is getting more and more similar to that of a harmonic complex tone, which represents the spectrum of an IRN with an infinite number of iterations. So it seems plausible that the pitch strength of IRN is getting more and more the same as the pitch strength of a harmonic complex tone, which is reached asymptotically.

Modeling the experimental results

Terhardt et. al. [4] described an algorithm to predict the salience of tonal components included in a sound. This method takes into account spectral as well as virtual pitch phenomena.

The basic idea behind the algorithm is a spectral analysis of sounds according to the properties of the human auditory system. To achieve this, a transform like the short-time Fourier transform or the Fourier-t-Transform (see [5]) is used. From the output of the filterbank audible pitches are calculated taking into account effects of masking. Each perceivable pitch (spectral and virtual) is weighted due to its spectral location and its level excess over the neighbouring spectral components.

Based on this algorithm the salience of the spectral pitch closest to the fundamental frequency was calculated and regarded as the sound's pitch strength. The data predicted by Terhardt's modell are shown in Figure 2. The calculated values of pitch salience are normalized to the value for the corresponding harmonic complex tone. It can be seen that the basic behaviour of the experimental results is reproduced in a way that the dependency of pitch strength on the number of iterations and the delay time of the IRN-algorithm shows the same behaviour as in the experimental results. The differences for different delay times, on the other hand, can only be reproduced in a very rough way. Taking into account that this modell originally was intended to be used for sounds with line spectra only (e.g. harmonic complex tones), the calculated results seem reasonable and improvements are conceivable concerning the adaptation to more complicated signals.

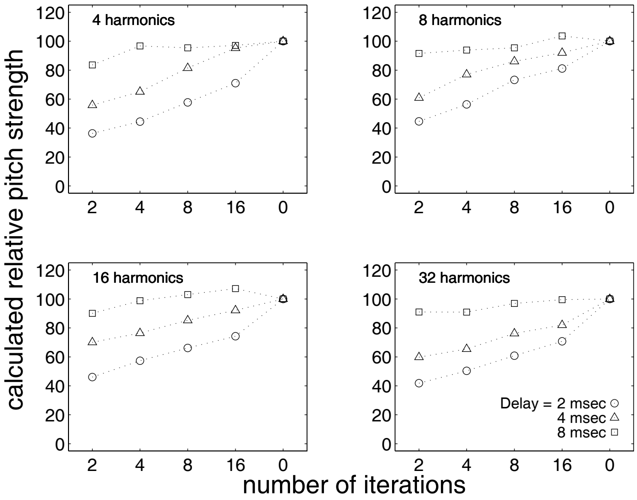

Figure 2: *Data calculated by the algorithm for pitch salience proposed by Terhardt in [4]. Data is normalized to the calculated values for the corresponding harmonic complex tone.*

Conclusion

In the present study, harmonic complex tones and iterated rippled noises have been compared regarding their pitch strength. Results showed that the pitch strength of IRN grows with the number of iterations. Additionally the basic frequency or better the delay time of the IRN also plays an important role.

A concept was presented explaining the data found when interpreting the human auditory system as a signal analyzer using a filter bank of critical band wide filters. This would explain that pitch strength differences are higher for higher basic frequencies, as the spectral components can be resolved in a better way.

Finally, data calculated with a model of pitch salience was compared to the experimental results showing some analogies but also some differences for the modelled data.

Acknowledgments

The author would like to thank Prof. Dr.-Ing. H. Fastl for supporting and encouraging the author. This work is supported by Deutsche Forschungsgemeinschaft FA 140/1.

References

[1] Fastl, H., Stoll, G.: Scaling of Pitch Strength. Hearing Research **1** (1979), 293-301

[2] Krumbholz, K., Patterson, R. D., Nobbe, A., Fastl, H.: Microsecond temporal resolution in monaural hearing without spectral cues? J. Acoust. Soc. Am., **113** (2003), 2790-2800

[3] Feldtkeller, R., Zwicker, E.: Das Ohr als Nachrichtenempfänger, Hirzel-Verlag (1956)

[4] Terhardt, E., Stoll, G., Seewann, M.: Algorithm for extraction of pitch and pitch salience from complex tonal signals. J. Acoust. Soc. Am. **71** (1982), 679-688

[5] Terhardt, E.: Fourier Transformation of time signals: conceptual revision. Acustica **57** (1985), 252-256