Perception and Adjustment of Pitch in Inharmonic String Instrument Tones

Hanna Järveläinen[1], Tony Verma[2], and Vesa Välimäki[3,1]

[1]Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, P.O. Box 3000, FIN-02015 HUT, Espoo, Finland

[2]Hutchinson mediator, 888 Villa St. Mountain View, CA 94043

[3]Pori School of Technology and Economics, P.O.Box 300, FIN-28101 Pori, Finland

Email: hanna.jarvelainen@hut.fi

Abbreviated title

Pitch of Inharmonic String Tones

Abstract

The effect of inharmonicity on pitch was measured by listening tests at five fundamental frequencies. Inharmonicity was defined in a way typical of string instruments, such as the piano, where all partials are elevated in a systematic way. It was found that the pitch judgment is usually dominated by some other partial than the fundamental; however, with a high degree of inharmonicity the fundamental became important as well. Guidelines are given for compensating for the pitch difference between harmonic and inharmonic tones in digital sound synthesis.

Keywords

Sound synthesis, perception, inharmonicity, string instruments

Perception and Adjustment of Pitch in Inharmonic String Instrument Tones

Introduction

The development of digital sound synthesis techniques (Jaffe and Smith 1983), (Karjalainen et al. 1998), (Serra and Smith 1989), (Verma and Meng 2000) along with methods for object-based sound source modeling (Tolonen 2000) and very low bit rate audio coding (Purnhagen et al. 1998) has created a need to isolate and control the perceptual features of sound, such as harmonicity, decay characteristics, vibrato, etc. This paper focuses on the perception of inharmonicity, which has many interesting aspects from a computational viewpoint. A model is presented for estimating the effect of inharmonicity on pitch in string instrument sounds.

It is well known that the stiffness of real strings causes wave dispersion by making the velocity dependent on frequency. If the string parameters are known, the inharmonic partial frequencies can be computed based on the wave equation. The partials are elevated according to the following formula (Fletcher 1962):

$$f_n = nf_0\sqrt{1 + Bn^2}$$

(1)

$$B = \frac{\pi^3 Q d^4}{64 l^2 T}$$

(2)

In these equations n is the partial number, Q is Young's modulus, d is the diameter, l is the length and T is the tension of the string, and f_0 is the fundamental frequency of the string without stiffness. B is the inharmonicity coefficient for an unwrapped string. Its value depends on the type of string and the string parameters. Completely harmonic partial frequencies are obtained with $B = 0$. One should note that the frequency f_1 of the first partial in the inharmonic complex tone is actually higher than the fundamental frequency f_0 of the ideal string without stiffness. In addition to stiffness, there can be other sources of inharmonicity. For example, strong inharmonicity can be observed in the attack transients of harpsichord tones, where the restoring force of string displacement is nonlinear (Weyer 1976). However, the present study focuses on the systematic stretching of partials which is mainly caused by string stiffness.

The effects of mistuning a single partial within a complex tone are relatively well understood (Moore et al. 1985, Cohen 1984). Mistuning can be detected in different ways. If the mistuned component is one of the 4-6 lowest harmonics, it either affects the residue pitch of the complex or it segregates from it and is heard as a separate tone. Higher harmonics than the sixth are no longer heard separately, but inharmonicity is detected by using either beating or roughness as a cue.

Moore et al. (1985) showed how mistuning a single partial affects the pitch of the whole complex. With mistuning up to about 2-3 %, the residue pitch of the complex changes linearly according to the amount of mistuning. With mistuning greater than that, the pitch effect gets weaker, and finally the mistuned component segregates from the complex. The change of the residue pitch due to mistuning one partial was then used to approximate the relative dominance of the partials. The perceived pitch was found to be a weighted average of the lowest five or six harmonics. The weights were assigned to each harmonic according to the pitch effect that it caused. Later on, Dai (2000) suggested that the dominance region of pitch is not only connected to the partial number but also to absolute frequency, the most dominant partials being around 600 Hz.

Besides pitch, Moore's group also studied other aspects of the perception of mistuning (Moore et al. 1985). They measured the thresholds for detecting a mistuned component in a complex tone and showed that mistuning is perceived in different ways depending on the harmonic number. Shortening the stimulus duration produced a large impairment in performance for the higher harmonics, while it had little effect on the performance for the lower harmonics. It was reasoned that particularly for long durations beats provide an effective cue, but for short durations many cycles of beats cannot be heard. For the lower harmonics, beats were generally inaudible, and the detection of mistuning appeared to be based on hearing the mistuned component stand out from the complex. The thresholds varied only weakly with duration.

Even though the effects of inharmonicity are generally audible at least in the lowest piano tones, the partials still tend to fuse together, causing a single pitch percept unlike the sounds of bells or chimes, for instance. However, it is expected that the pitch is shifted from that of an otherwise similar harmonic tone. If the inharmonicity of synthesized tones were gradually increased to greater levels than what is observed in natural instruments, ambiguity of pitch would result, and the listeners would start to hear out individual partials or detect several pitches.

The pitch shifts of inharmonic complex tones were observed already in the 1950's by de Boer, who studied the so called first pitch-shift effect, the change of pitch as a function of center frequency for complex tones with a fixed and equal frequency spacing (de Boer 1976). He modelled the effect mathematically by a sawtooth function. A pitch-shift was also detected, when the frequency spacing of a complex tone with fixed center frequency was varied. Vassilakis (1998) related this to the first pitch-shift effect and explained both by a single model. The explanations of the pitch-shift effects as well as the pitch ambiguity are based on the detection of periodicity in the signal waveform (Cariani and Delgutte 1996), (Schneider 2000). Compared to harmonic tones, the waveform of inharmonic tones is less uniform, and the detection of periodicity becomes harder.

There are few studies on sounds that exhibit systematic inharmonicity like string instruments. Slaymaker (1970) commented that the timbre of sounds with compressed or stretched octave relations is bell- or chimelike, and that chords played with such tones begin to sound out-of-tune. Mathews and Pierce (1980) found that stretching the partials affects tonal harmony by destroying the consonance of musically consonant chords. Thus the finality of cadences is reduced, for instance.

One of the most evident effects of inharmonicity in string instruments is the stretched tuning of the piano to maintain harmonic consonance between musical intervals. Because of inharmonicity, the higher partials of low tones become sharp with respect to corresponding higher tones, and unpleasant beats occur. To minimize the beats, the bass range is tuned slightly flat and the treble range slightly sharp compared to the equal temperament. In the middle range, the adjustment is only a few cents (1/100 of a semitone), but at both ends of the keyboard, differences of 80-120 cents are common (Martin and Ward 1961). Lattard (1993) has simulated the stretched tuning process computationally.

The effect of inharmonicity on timbre was recently studied in (Järveläinen et al. 2001). A model was presented for the thresholds for detecting a difference in timbre as a function of fundamental frequency. It was found that based on the timbre criterion, it is only occasionally necessary to implement inharmonicity in digital sound synthesis, while it might be otherwise neglected to achieve computational savings. It is therefore expected that some finetuning is needed in order to maintain the correct musical scale between harmonic and inharmonic synthetic tones.

In the present study, the change of pitch due to inharmonicity was measured as a function of the inharmonicity coefficient B with fundamental frequency as a parameter. Guidelines are given to approximate the pitch correction to compensate for the pitch increase in digital sound synthesis.

Listening Tests

The pitch of inharmonic sounds was measured for four notes, A1 (55 Hz), G3 (196 Hz), A4 (440 Hz), and C#6 (1108.7 Hz). The task was to match the pitch of an inharmonic tone to that of a harmonic tone. The pitch increase was measured as a function of B for each note.

Test Sounds

In order to generate realistic sounding tones for the experiment, sinusoidal modeling was used. Real piano tones were analyzed for both their inharmonic partials and soundboard response. To achieve this, the partials were modeled and subtracted from the original signal. This produced a residual signal that consisted of the characteristic `knock' that occurs during the onset of a piano tone. By varying the frequency of each partial according to the inharmonicity coefficient $B$ then adding in the pitchless `knock', realistic sounding piano tones with varying degrees of inharmonicity were generated for the experiment. The original piano tones were taken from the McGill University Masters Series recordings (http://www.music.mcgill.ca/resources/mums/html/mRecTech.html) and downsampled to 22.05 kHz at 16 bit resolution.

Seven inharmonic test tones were generated for each note. The first test tone was completely harmonic with $B = 0$. The value of B increased fairly logarithmically to B_{max} which was chosen for each note based on preliminary listening. With $B = B_{max}$ the pitch judgment was still possible but usually ambiguous.

Each test tone had a duration of 0.9 seconds and consisted of the lowest six partials with amplitudes corresponding to the analyzed piano sounds. Even though the higher partials still contain spectral energy, it is generally thought that the lowest six partials are responsible for the perceived pitch (Moore et al. 1985). A piano-like timbre was already created by the six partials, but to study the possible effect of the higher ones, corresponding test sounds with 12 partials were generated for A1 and G3. Otherwise, the number of partials was fixed in order to eliminate any possible effect on the test results. Table 1 summarizes the properties of the test sounds. Typical spectra of the sounds are presented in Figure 1, which shows the difference between the least and most inharmonic test tone of fundamental frequency 440 Hz (A4). <INSERT FIGURE 1 ABOUT HERE>

The adjustable tones were otherwise identical to the test tones but harmonic with $B = 0$. Sine tones were rejected for the practical reason that the aim is to correct the differences between harmonic and inharmonic tones, and because the pitch comparison was easier between sounds of similar timbre. <INSERT TABLE 1 ABOUT HERE>

Subjects and Test Method

Six subjects participated in the experiment. The listeners were personnel of Helsinki University of Technology, and most of them had a musical background as well as earlier experience of psychoacoustic tests and an interest for sound synthesis issues. Many had participated also in earlier studies of the present authors. Two of the listeners were the authors HJ and TV, but they gained no special benefit except hearing the sounds already in the synthesis stage. None of the subjects reported any hearing defects. Although a relatively small group, compromising quantity for quality was considered justified. The sounds were played through headphones in a silent listening room, and the subjects were allowed to practise before the test.

The method of adjustment was used in the experiment. The task was to adjust the pitch of a harmonic tone until it matched that of an inharmonic tone. The fundamental frequency of the harmonic tone was changed by a scroll bar in the graphical user interface. The quantizing intervals were 0.2 Hz, 0.3 Hz, 0.4 Hz, and 0.5 Hz for the notes A1, G3, A4, and C#6, respectively.

Test Results

Modeling the Pitch Increase

The test results are shown in Fig. 2 for sounds with six partials. The results are consistent when *B* is reasonably small, but with the highest values two different tendencies can be seen. Some of the subjects apparently chose the pitch of the fundamental that segregated from the complex, while others were still judging according to the overall impression.

The test results were analyzed by the Analysis of Variance (ANOVA) for significant differences in the mean results for different *B* values (Lehman 1991). The results were highly significant for all notes ($p < 0.01$), confirming that there is a true difference in perceived pitch for at least one *B* value for each note. The results obtained by using 12 or 6 partials were tested against each other. The number of partials had no significant effect on the results. Thus the pitch effects were mainly connected to the lowest six partials. <INSERT FIGURE 2 ABOUT HERE.>

The results suggest that the perceived pitch is given by some nonlinear function of *B*. The unknown function seemed to have the same general shape as any of the partials. Since each partial referred to a different fundamental frequency of an imaginary harmonic tone, it was tested whether the perceived pitch could be dominated by a single partial or if it was likely to be a combination of several. A candidate for the perceived pitch was found for each partial ($N = 1...6$) as follows:

$$f_0 = \frac{f_N(B)}{N} = f_0\sqrt{1+Bn^2}$$

(3)

The dominant partial number N was found for each note by nonlinear model fitting in the least-squares sense. The highest values of B were usually left out of the analysis, since the perceived pitch was obviously dropping towards the fundamental. The value of N that gave the optimal fit was very close to an integer in each case, being 6.0, 2.97, 3.18, and 2.03 for A1, G3, A4, and C#6, respectively. Thus it was reasoned that the perceived pitch as a function of B was dominated by a single partial. For A1, the mean pitch contour followed the sixth partial, i.e., the pitch was judged equal to the frequency of the sixth partial divided by six. With increasing fundamental frequency, the number of the dominating partial decreased. For G3 and A4, the third partial was dominant, and the second partial dominated for C#6 .

Interestingly enough, the dominant partial was not the one with the largest amplitude. An example is seen in Figure 1 for A4 . The third partial dominates the pitch judgment, even though the fundamental and the fourth partial are the most prominent.

The dominant partial pitch estimate as well as the fundamental frequency of the test tone are shown by dashed lines in Figure 3. The mean pitch judgments are shown by solid lines. It is seen that the perceived pitch separates from the estimate for the highest values of B and drops towards the fundamental. In this area the pitch is ambiguous, since individual partials start to segregate from the complex. <INSERT FIGURE 3 ABOUT HERE.>

The general thresholds for detecting a pitch difference between harmonic and inharmonic sounds could not be measured in this study. However, it was reasoned that even though the dominant partial gave a good pitch estimate for the lowest *B* values as well as for the average, the difference between the estimate and the nominal fundamental frequency might be insignificantly small. On the other hand, there was a need to approximate an upper limit for *B,* above which the estimate would no longer be valid. Relevant ranges were thus approximated for the dominant partial pitch estimates as functions of *B* by testing the significance of the difference between the mean pitch judgments and the nominal fundamental frequencies as well as the dominant partial pitch estimates. A significant t-test ($p < 0.05$) means that the mean pitch judgments differ significantly from a specified value; either the dominant partial pitch estimate or the nominal fundamental frequency. The basic idea of the relevant range is presented in Figure 4. <INSERT FIGURE 4 ABOUT HERE.>

For the lowest *B* values, the pitch increase was insignificant for each note. For the highest values, the dominant partial estimate began to fail and the pitch judgment was dominated more and more by the fundamental frequency. The lower and upper cutoff values of *B* are presented for each note in Table 2. Between those limits there is a region in which the dominant partial gives the best pitch estimate. Below the lower limit the pitch increase is insignificant, and above the upper limit pitch is ambiguous and is judged inconsistently.

Comparison of Pitch and Timbral Effects

Another objective of this study was to find the relation between the pitch and timbral effects of inharmonicity. In a previous study, the audibility of the timbral effect of inharmonicity was measured (Järveläinen et al. 2001). The possible pitch increase was compensated based on the results presented in (Järveläinen et al. 2000). Detection thresholds, expressed through the inharmonicity coefficient *B*, were measured for timbre differences as a function of fundamental frequency *f* and modeled as follows:

$$\ln(B_{\mathrm{timbre}}) = 2.57\ln(f) - 26.5 \qquad (4)$$

The lower limits were used for estimating the B_{pitch} required for a significant pitch increase. B_{timbre} and B_{pitch} are compared for different fundamental frequencies in Figure 5. One should bear in mind that while the detection thresholds for timbre were directly obtained from subjective tests, the thresholds for significant pitch increase are statistical approximations based on the pitch-matching tests. Furthermore, the test sounds in the previous and current experiments were generated in different ways, the current ones sounding more natural and piano-like. The previous test tones were generated by additive synthesis without considering the characteristic attack and initial amplitudes of the partials. However, they had a rich harmonic content to ensure that the timbre effects were reproduced. Unlike pitch perception, the perception of timbre depends strongly on the higher partials. The sounds in the timbre experiment contained all partials up to 10 kHz. <FIGURE 5 ABOUT HERE>

For low fundamental frequencies, the difference in timbre was perceived at much lower values of B than the pitch difference. For higher fundamental frequencies, the thresholds approach each other, and the slopes suggest that they might even cross at a still higher frequency. This would mean that at fundamental frequencies greater than about 2 kHz, the pitch increase could be detected even if there were no audible difference in timbre.

The relation of pitch and timbral effects becomes interesting, if we consider successive sounds in a musical context. Using the audibility of the timbral effects as a criterion, implementing inharmonicity might be necessary for some of the sounds, while others could be replaced by an otherwise identical but harmonic tone. Therefore, we should be able to correct the pitch of the inharmonic sounds relative to the harmonic ones in order to maintain the correct musical scale. Another question is whether, having skipped implementation of inharmonicity because of inaudible timbral difference, a pitch correction is still needed.

Additional Experiments

Since the original experiments raised a number of questions, two additional small-scale listening tests were organized. The test procedures were the same as before. The objective of the first test was to find the relation of detecting differences in pitch and timbre at high fundamental frequencies. The second additional experiment measured perceived pitch after removing the dominant partial that was found in the first experiment. Only five subjects could participate the additional experiments, and one of them had to be exluded from the final results of the timbre experiment because of a strong bias towards detecting a difference when there was none.

Pitch and Timbre Experiments for A7

The pitch-matching experiment was re-run using 3520 Hz (A7) as fundamental frequency, which represents the highest part of the piano keyboard. Seven inharmonic test tones were generated with inharmonicity coefficients logarithmically spaced between 0 and B_{max} = 0.3. However, the Nyquist limit at 11 kHz reduced the number of partials to no more than two. This of course destroyed the piano-like timbre of the sounds, and the second partial was sometimes heard as a separate tone. A solution would have been to increase the sampling frequency for the highest sound, but in that case it would not have been in line with the previous experiments, and no comparisons would have been possible between the results. For this reason, the test was carried through as planned. However, the previous results suggested that the pitch judgment would depend on the lowest two harmonics even if more partials were present, and the objective was merely to see wether the fundamental would become even more dominant for the highest sound.

The results of the pitch-matching test for A7 are presented in Figure 6. The subjects indeed based their judgments almost entirely on the fundamental

frequency. The nonlinear model fitting was performed again as described in the previous section, giving $N = 1.01$. This confirms the dominance of the fundamental frequency. The significance of the pitch difference was also tested as before. The value of B required for a significant pitch increase was 0.010.
<INSERT FIGURE 6 ABOUT HERE.>

The thresholds for detecting a difference in timbre were measured for A7 as reported earlier. The sounds contained three partials up to B=0.121 and only two thereafter. The average threshold was $B = 0.0113$. The model given in Eq. 4 and in (Järveläinen et al. 2001) predicts a slightly lower threshold of $B = 0.004$. Even though the timbre measurement for A7 can be considered only informal, it shows that a difference in timbre is indeed observed. The threshold is much lower than the point where the number of partials drops from three to two.

Figure 5 summarizes the results of both timbre and pitch experiments for all fundamental frequencies. We can now consider the relation of pitch and timbral effects in a frequency band from 55 Hz up to 3.5 kHz. The threshold for significant difference in pitch was in the order of 0.001 for the lowest five notes but increased to 0.01 for A7. The pitch and timbre thresholds are almost identical for the highest note. Thus the middle region, which is seen in Figure 6 for A1 through C#6, vanishes for A7. Inharmonicity is either undetected, or it has an effect on both timbre and pitch.

We can conclude that for the tones between A1 and C#6, there is no need for any pitch corrections whenever inharmonicity is left unimplemented according to the timbre criterion. Also when inharmonicity is implemented, its effect on pitch will not reach a significant level until the value of B is about 0.001. At least for values greater than that, the pitch increase should be compensated. For A7 the situation is different, since the threshold for detecting a difference in pitch is very close to the threshold for detecting a difference in timbre. It is thus possible that the pitch effect is detected for smaller values of B than the timbral effect. For this reason the pitch correction should be made even if the timbral effect were inaudible. In such cases the inharmonic tone could be replaced by a harmonic one whose fundamental frequency has been slightly increased according to the model given in Eq. 3 with $N = 1$.

Effect of Removing the Dominant Partial

It was found that for each note, the best pitch estimate was given by a single partial. The second additional experiment was organized to find out, whether only one of the partials is actually responsible for the pitch of the sound or if it could be a joint effect of all partials. For this experiment, the dominant third partial was removed from the original inharmonic test sounds for A4, and the pitch-matching experiment was repeated as before.

The pitch judgments were clearly different in the original pitch-matching test and the current test. They were quite similar until about $B = 0.001$, but for larger values of B the pitch dropped notably for tones with the missing partial. The results of the original and the additional test for A4 are compared in Figure 7. The pitch no more followed the third partial as clearly as in the first test. For one of the subjects, the perceived pitch followed the second partial almost exactly. However, the nonlinear model fitting gave $N = 2.54$. This suggests that no single partial could be considered dominant once the third partial was removed. <INSERT FIGURE 7 ABOUT HERE.>

Discussion

Pitch compensation becomes necessary in the synthesis of string instruments, if inharmonicity is partly ignored. Another field of application is the synthesis of very inharmonic, imaginary sounds. In both cases, pitch corrections are needed to maintain the musical scale.

Typical values of B for piano strings lie roughly between 0.00005 for low bass tones and 0.015 for the high treble tones (Conklin 1999). In the bass range, pitch corrections are surely unnecessary, even though it would be worthwhile to implement inharmonicity based on the timbre criterion. On the other hand, it is recommended that in the highest treble the pitch increase should always be compensated. If the number of partials is low for the highest tones, implementation of inharmonicity could be omitted in other respects.

According to the current results, compensating for the pitch differences between harmonic and inharmonic synthetic tones is not too complicated. We have presented lower limits for the B parameter, below which no significant pitch effect was observed, and shown how the average pitch judgments follow some of the higher partials.

Our results are consistent with the previous research that was discussed in the introduction. In the work of Moore et al. (1985b), Dai (2000), and the pitch models of Terhardt et al. (1982a) and Goldstein (1973), it is stated that the pitch percept is based on only the lowest partials. This is also seen from our tests, see Fig. 3. Furthermore, the test results were similar regardless of the number of partials (6 or 12).

However, previous research disagrees on the dominance of individual partials. Our results with systematically mistuned sounds are nearest to those of Dai (2000), who suggested that the dominant partials are around 600 Hz in frequency. We found that the dominant partial number decreased with increasing fundamental frequency. For the lowest fundamental, the pitch judgment followed an estimate given by the sixth partial, while for the highest fundamental, the second partial was dominant. However, we found no absolute dominance of the 600 Hz region.

Of course, the most interesting question is how to choose the dominant partial for any fundamental frequency. Unfortunately, based on these limited experiments, we cannot give a complete answer. The 600-Hz rule seems to work for fundamental frequencies up to 200 Hz, but for higher frequencies the dominance region is higher in absolute frequency. However, the dominant partial number drops from six to one with increasing fundamental frequency.

We have been considering two classes of inharmonicity effects: pitch and timbre. The pitch effects are mainly caused by the lowest five or six partials, while timbral effects, such as roughness and beats, concentrate on the higher partials. However, a great deal of inharmonicity in the lower partials could cause them to segregate from the complex. Even though the segregation of mistuned partials has drawn a lot of attention in earlier studies (Moore et al. 1986, Hartmann et al. 1990), it is not of primary importance to the perception of string instrument sounds. Based on the systematic mistuning, the partials of string instrument tones tend to stay together. If the sounds are inharmonic enough to cause segregation, the timbre turns unnatural and the pitch becomes ambiguous.

In the present experiments, the highest tone A7 was exceptional because of the low number of partials, and segregation might have affected the pitch and timbre judgments for relatively small values of B. Because the perceived pitch followed the fundamental exclusively, it is expected that the second partial was heard separately most of the time. This would undoubtedly mean that a great deal of the timbral differences were actually detected based on the frequency of the second partial. Now the threshold for detecting timbral differences covers roughness and beating as well as partial segregation, and whichever is detected first will set the detection threshold. From a synthesis point of view this causes no problem, because the model of inharmonicity presented in Eq. 1 and 2 makes no difference between the different kinds of timbral effects. Once inharmonicity is implemented, the synthetic sound will cause the same perceptual effects as the original.

The present study aims at giving practical guidelines for compensating for the pitch effects of inharmonicity in sound synthesis, which explains the use of synthesized musical instrument tones as test material. Practical knowledge of the perception of musical instrument sounds has become increasingly interesting with the development of object-based and structured representations of sound. Perceptual research can help to reduce the computational load of existing sound synthesis methods such as physical modeling, or it can be applied in the design of new coding schemes, which aim at parametric control of the perceptual features of sound.

Future Work

Partial segregation needs to be considered more carefully, if imaginary and very inharmonic sounds are synthesized. The behavior of string instrument tones also requires some further study in high frequencies, where partial segregation was observed. It would be useful to re-run the experiments for A7 with sampling frequency increased to 44.1 kHz to allow more partials.

One of the most important future goals is to give a more complete model for the dominant partial number as a function of fundamental frequency. It would require a large-scale listening experiment that could not be organized during this study.

It also remains a future task to measure the actual detection thresholds for pitch differences. It will be interesting to compare the difference limen to what is known about frequency discrimination of sine tones (Wier et al. 1977) and partials within complex tones (Moore et al. 1984). For instance, if the frequency difference between the dominant partial of the harmonic and inharmonic tones correlated to those measures, it would give more evidence for the dominance of a single partial.

Conclusions

The pitch effect caused by inharmonicity was measured for string instrument sounds for five fundamental frequencies. In each case, the perceived pitch followed an estimate given by one of the higher partials. The dominant partial number dropped from six to one as the fundamental frequency increased from 55 Hz to 3520 Hz. However, more detailed experiments are needed for modeling the pitch of inharmonic tones for the whole pitch range of the piano, for instance.

Acknowledgments

The financial support of Pythagoras graduate school, Nokia Research Center, and the Academy of Finland is gratefully acknowledged. The authors also wish to thank Dr. Aki Härmä for the listening test software.

References

de Boer, E.1976. On the “residue” and auditory pitch perception, in *Handbook of sensory physiology*. Keidel, W.D. and Neff, W.D (eds). Berlin: Springer-Verlag.

Cariani, P.A. and Delgutte, B. 1996. Neural correlates of the pitch of complex tones, I: Pitch and pitch saliance, II: Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region of pitch. *J. Neurophysiology*, 76, 1698-1716, 1717-1734.

Cohen, E. A.1984. Some effects of inharmonic partials on interval perception. *Music Perception* 1, 323-349.

Conklin, H. A. 1990. Generation of partials due to nonlinear mixing in stringed instruments. *J. Acoust. Soc. Am.,* 105(1), 536-545.

Dai, H. 2000. On the relative influence of individual harmonics on pitch judgment. *J. Acoust. Soc. Am.,* 107(2), 953-959.

Fletcher, H., Blackham, E. D., and Stratton, R. 1962. Quality of piano tones. *J. Acoust. Soc. Am.,* 34(6), 749-761.

Fletcher, H., and Rossing, T. 1991. *The Physics of Musical Instruments*. New York: Springer-Verlag.

Goldstein, J. L. 1973. An optimum processor theory fot the central formation of the pitch of complex tones. *J. Acoust. Soc. Am.,* 77(5), 1496-1516.

Jaffe, D.A., and Smith, J.O. 1983. Extensions of the Karplus-Strong plucked-string algorithm. *Computer Music J.*, 7(2), 56-69.

Järveläinen, H., Välimäki, V., and Karjalainen, M. 2001. Audibility of the timbral effects of inharmonicity in stringed instrument tones. *Acoustics Research Letters Online*, 2(3), 79-84.

Järveläinen, H., Verma, T., and Välimäki, V. 2000. The effect of inharmonicity on pitch in string instrument sounds. *Proc. Int. Computer Music Conf.*, (pp. 237-240). San Francisco: International Computer Music Association.

Karjalainen, M., Välimäki, V., and Tolonen, T. 1998. Plucked string models: from the Karplus-String algorithm to digital waveguides and beyond. *Computer Music J.*, 22(3), 17-32.

Lattard, J. 1993. Influence of inharmonicity on the tuning of a piano - measurements and mathematical simulation. *J. Acoust. Soc. Am.,* 94(1), 46-52.

Lehman, R. S. 1991. *Statistics and Research Design in Behavioral Sciences*. Belmont, CA: *Wadsworth Publishing Company.*

Martin, D., and Ward, W. 1961. Subjective evaluation of musical scale temperament in pianos. *J. Acoust. Soc.Am.,* 33(5), 582-585.

Mathews, M. V., and Pierce, J. R. 1980. Harmony and nonharmonic partials. *J. Acoust. Soc. Am.,* 68(5), 1252-1257.

Moore, B. C., Glasberg, B. R., and Shailer, M.J. 1984. Frequency and intensity difference limens for harmonics within complex tones. *J. Acoust. Soc. Am.* 75(2), 550-561.

Moore, B. C., Glasberg, B. R., and Peters, R. V. 1985. Relative dominance of individual partials in determining the pitch of complex tones. *J. Acoust. Soc. Am.,*77(5), 1853-1860.

Moore, B. C., Peters, R. V., and Glasberg, B. R. 1986. Thresholds for hearing mistuned partials as separate tones in harmonic complexes. *J. Acoust. Soc. Am.,* 80(2), 479-483.

Purnhagen, H., Edler, B., and Feredikis, C. 1998. Object-based analysis/synthesis audio coder for very low bit rates. *104th convention of the Audio Engineering Society*, preprint 4747.

Schneider, A. 2000. Inharmonic sounds: Implications as to “pitch”, “timbre”, and “consonance”. *J. New Music Research*, 29(4), 275-301.

Serra, X., and Smith, J.O. 1990. Spectral modeling synthesis: a sound analysis/synthesis system based on a deterministic plus stochastic decomposition. *Computer Music J.*, 14(4), 12-24.

Slaymaker, H. F. 1970. Chords from tones having stretched partials. *J. Acoust. Soc. Am.,* 47(6), 1569-1571.

Terhardt, T., Stoll, G., and Seewann, M. 1982. Pitch of complex signals according to virtual pitch theory: test, examples, and predictions. *J. Acoust. Soc. Am.,* 71(3), 671-678.

Tolonen, T. 2000. Object-based sound source modeling for musical signals. *109th convention of the Audio Engineering Society*, Preprint 5174.

Vassilakis, P. 1998. A single model explaining the first and second pitch-shift effects as alternative manifestations of a single phenomenon. *Proc. 5th ICMPC*, (pp. 107-114).

Verma, T., and Meng, T. 2000. Extending spectral modeling synthesis with transient modeling synthesis. *Computer Music J.* 27(2), 47-59.

Weyer, W.D. 1976. Time-frequency-structures in the attack transients of piano and harpsichord sounds I, II. *Acustica*, 35, 232-252, 1976. *Acustica*, 36, 241-258, 1976.

Wier, C.C., Jesteadt, W., and Green, D.M. 1977. Frequency discrimination as a function of frequency and sensation level. *J. Acoust. Soc. Am.*, 61(1), 178-184.

Tables

Table 1. Characteristics of the test tones.

Note	f_0	B_{max}	Partials
A1	55.0 Hz	0.005	1...6 and 1...12
G3	196.0 Hz	0.01	1...6 and 1...12
A4	440.0 Hz	0.01	1...6
C#6	1108.7 Hz	0.1	1...6

Table 2. Upper and lower boundaries for the relevant range of dominant partial pitch estimation.

Note	*B* at lower limit	*B* at upper limit
A1	0.0008	0.005
G3	0.001	0.01
A4	0.001	0.005
C#6	0.001	0.05

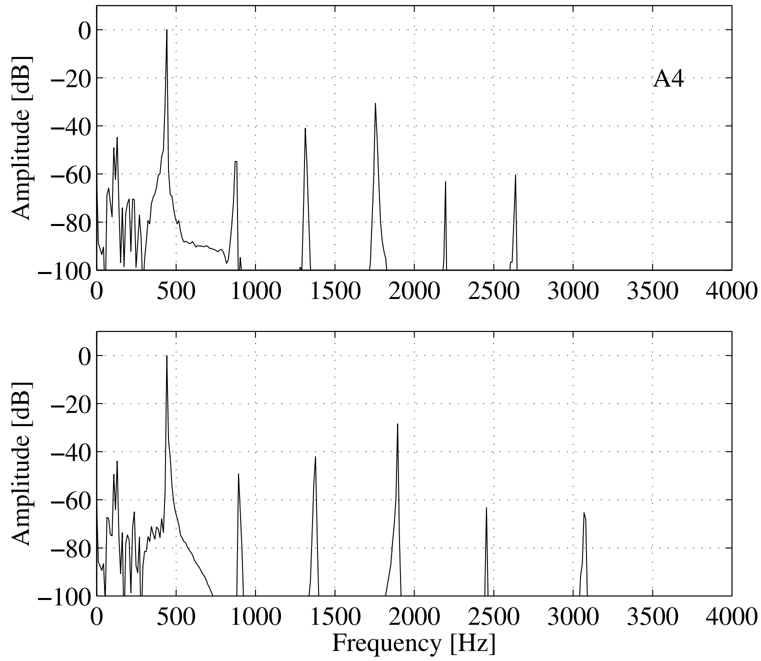

Figure 1: Spectrum of the least (top) and most (bottom) inharmonic test tone A4 (f0 = 440 Hz, B=0.01).

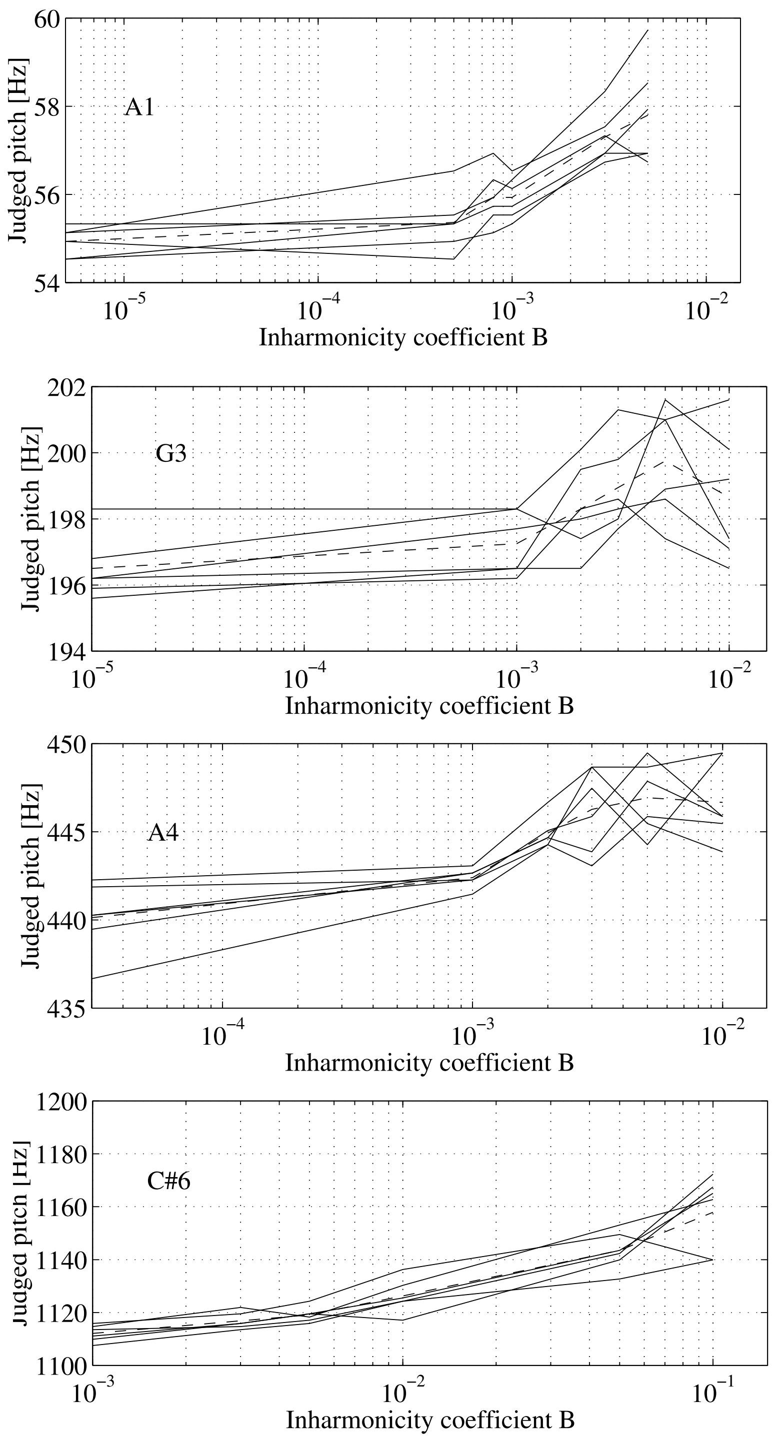

Figure 2: Top to bottom: Individual results for A_1, G_3, A_4, and $C\sharp_6$ with 6 partials (solid lines), and the corresponding mean results (dashed lines).

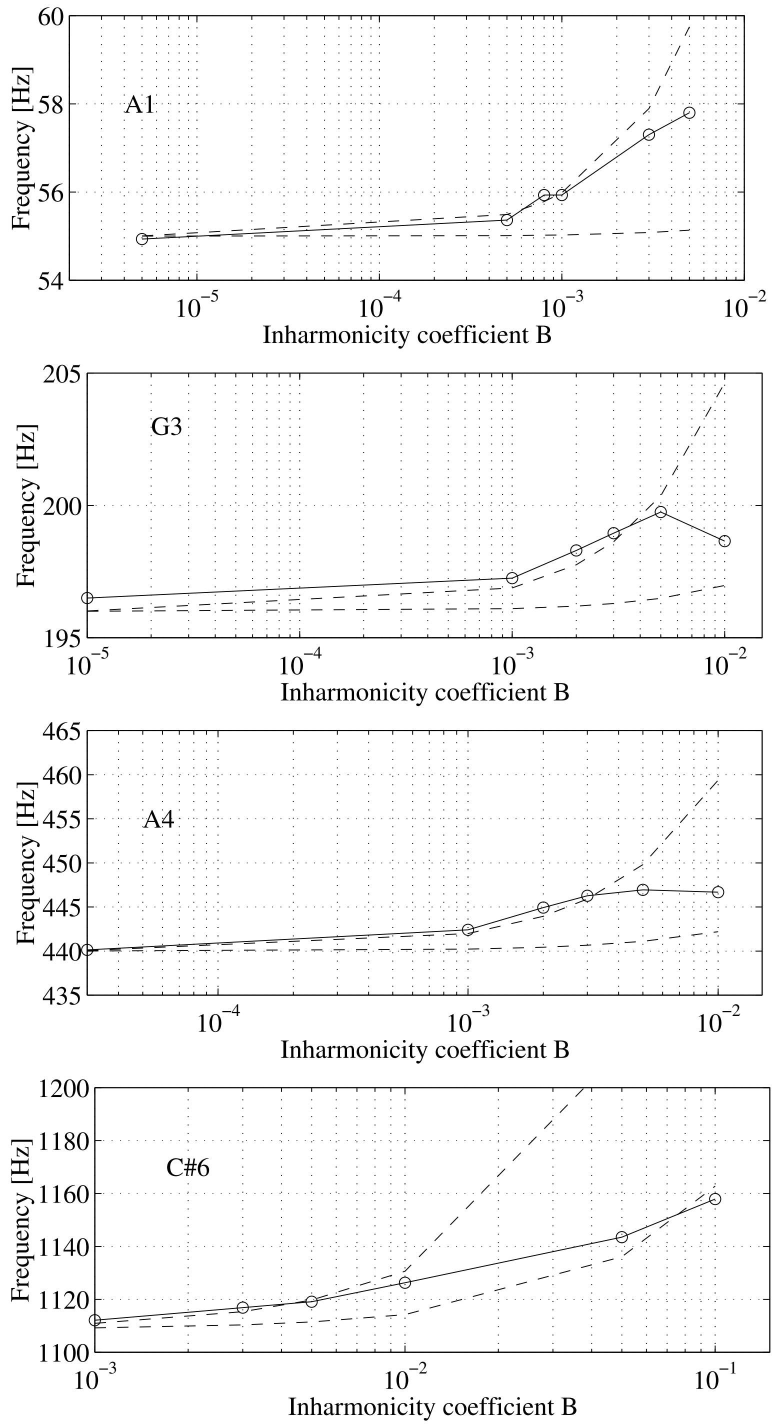

Figure 3: Estimation of the pitch according to the dominant partial. Mean of the pitch judgments (solid lines with 'o'), the fundamental frequency (lower dashed), and the dominant partial pitch estimate (upper dashed). Pitch estimates top to bottom: $f_6/6$, $f_3/3$, $f_3/3$, and $f_2/2$ for A_1, G_3, A_4, and $C\sharp_6$, respectively.

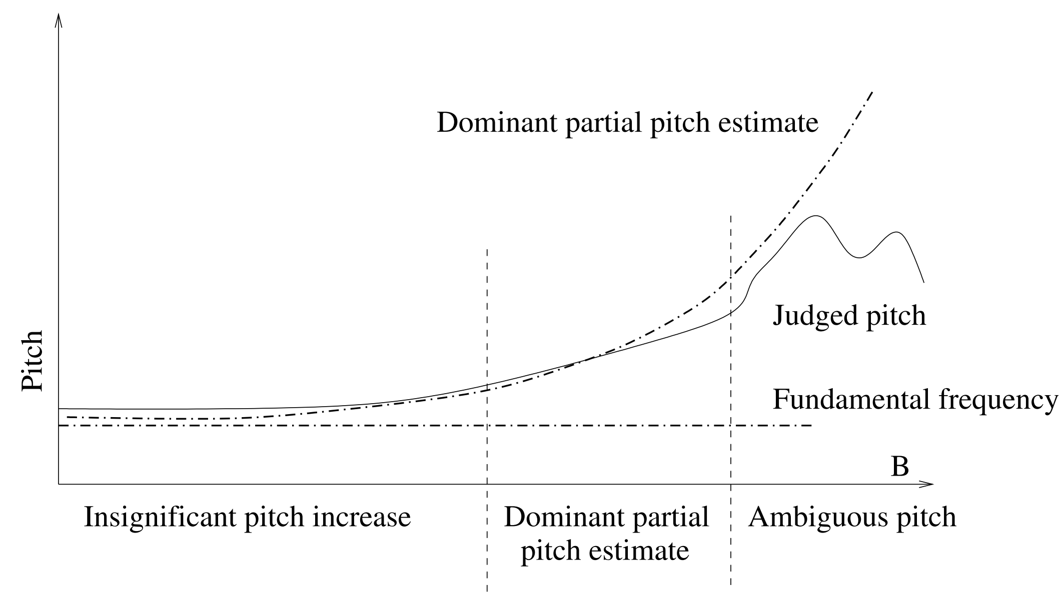

Figure 4: General idea of estimating the relevant range for dominant partial pitch estimation.

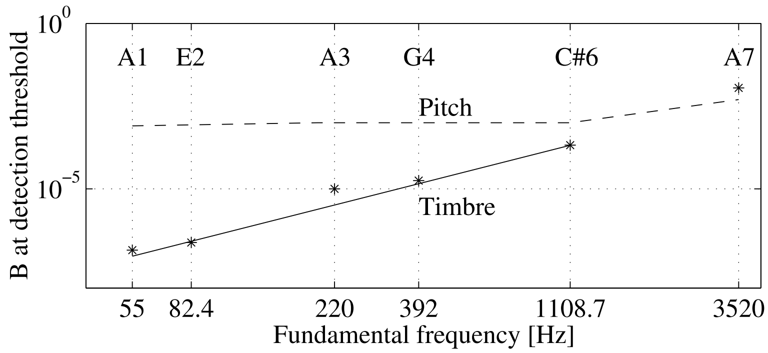

Figure 5: Thresholds for detecting differences in pitch and timbre for A_1 through A_7.

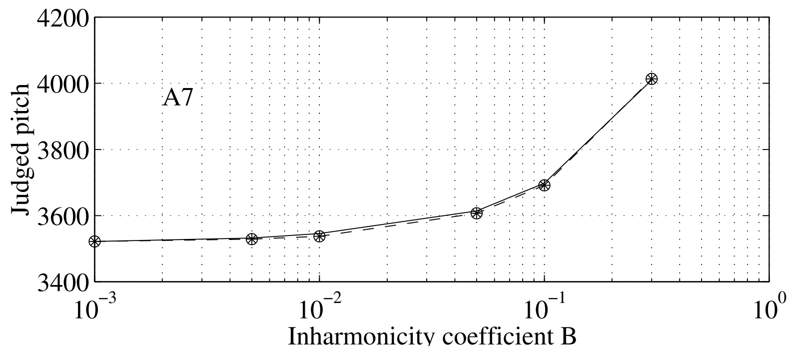

Figure 6: Mean results of the pitch-matching test for A_7 (solid line), and fundamental frequency (dominant partial) of the test tones (dashed line).

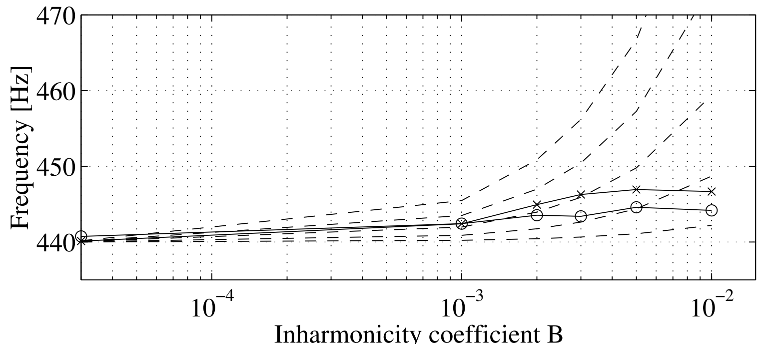

Figure 7: Mean results of the pitch-matching test for A4 (solid line with 'x') and for A_4 with missing third partial (solid line with 'o'), and candidates 1 through 5 for the dominant partial pitch estimate (dashed lines).