HELSINKI UNIVERSITY OF TECHNOLOGY
Department of Electrical and Communications Engineering
Laboratory of Acoustics and Audio Signal Processing

Heidi-Maria Lehtonen

Analysis and Parametric Synthesis of the Piano Sound

Master's Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Technology.

Espoo, November 29, 2005

Supervisor: Professor Vesa Välimäki
Instructor: M.Sc. Jukka Rauhala

HELSINKI UNIVERSITY OF TECHNOLOGY

ABSTRACT OF THE MASTER'S THESIS

Author:	Heidi-Maria Lehtonen		
Name of the thesis:	Analysis and Parametric Synthesis of the Piano Sound		
Date:	Nov 29, 2005	**Number of pages:**	55
Department:	Electrical and Communications Engineering		
Professorship:	S-89 Acoustics and Audio Signal Processing		
Supervisor:	Prof. Vesa Välimäki		
Instructors:	M.Sc. Jukka Rauhala		

In this thesis, an overview of the sound production mechanism of the piano is given. The acoustical properties of the instrument are studied in order to make a baseline for a physical and parametric model for the piano. In addition, the most important features of the piano sound, such as inharmonicity, the complicated decay process of the tones and the properties of the soundboard and the pedals, are investigated. The differences between the grand piano and the upright piano are considered in brief.

As the digital waveguide technique is the most feasible physics-based sound synthesis technique at the moment, the synthesis procedure that is followed in this thesis is based on this technique. An overview of the main aspects of this synthesis scheme is given, and the most important modeling issues are taken into account from the piano sound synthesis point of view. A novel filter design technique for modeling the losses occurring in the piano sound is presented with some practical design examples. In addition, the modeling of the sustain pedal is discussed and signal analysis is performed in order to gather information for the synthetic sustain pedal algorithm. The analyzed signals are obtained from two recording sessions which were carried out in two parts during the year 2005.

Keywords: digital filter design, digital signal processing, music, musical acoustics, piano, signal analysis, sound synthesis

TEKNILLINEN KORKEAKOULU		DIPLOMITYÖN TIIVISTELMÄ
Tekijä:	Heidi-Maria Lehtonen	
Työn nimi:	Pianon äänen analyysi ja parametrinen synteesi	
Päivämäärä:	29.11.2005	**Sivuja:** 55
Osasto:	Sähkö- ja tietoliikennetekniikka	
Professuuri:	S-89 Akustiikka ja äänenkäsittelytekniikka	
Työn valvoja:	Prof. Vesa Välimäki	
Työn ohjaaja:	DI Jukka Rauhala	

Tässä työssä tutkitaan pianon äänentuottomekanismia sekä akustisia ominaisuuksia. Tarkoituksena on luoda lähtökohdat pianon äänen parametriselle mallintamiselle. Lisäksi tutkitaan pianon äänen tärkeimpiä ominaisuuksia, kuten epäharmonisuutta, osaäänesten monimutkaista vaimenemisprosessia, kaikupohjan ja pedaalin ominaisuuksia sekä näiden tekijöiden vaikutuksia ääneen. Flyygelin ja pystypianon eroja tarkastellaan lyhyesti.

Koska digitaalinen aaltojohtomallinnus tarjoaa parhaat lähtökohdat fysikaaliseen soitinmallinnukseen, tämä työ pohjautuu tähän tekniikkaan. Digitaalisen aaltojohtomallinnuksen pääpiirteet esitellään, kuten myös pianon kannalta olennaisimmat mallinnukseen liittyvät asiat. Lisäksi esitellään uusi tekniikka häviösuotimen suunnittelua varten, sekä annetaan muutama esimerkki käytännön suodinsuunnittelusta tällä tekniikalla. Tämän lisäksi tarkastellaan kaikupedaalin mallintamista sekä suoritetaan signaalianalyysi tehokkaan mallinnusalgoritmin löytämiseksi. Analysoitavat signaalit on äänitetty kahdessa äänityssessiossa vuoden 2005 aikana.

Avainsanat: digitaalinen signaalinkäsittely, digitaalisuotimien suunnittelu, musiikki, musiikkiakustiikka, piano, signaalianalyysi, äänisynteesi

Acknowledgements

This Master's thesis has been carried out in the Laboratory of Acoustics and Audio Signal Processing at Helsinki University of Technology during the year 2005. The work has been funded by the Academy of Finland as a part of the CAPSAS project (project no. 104934).

My foremost thanks go to my supervisor, Professor Vesa Välimäki, whose encouragement and many helpful comments have contributed significantly to my whole working process. Also, I want to thank my instructor M.Sc. Jukka Rauhala, who has given me many valuable advises during this work.

I wish to thank my colleagues at the Akulab, who have created a pleasant working environment. Especially, I would like to thank M.Sc. Henri Penttinen for providing me information about recording techniques, Mr. Antti Huovilainen for helping me with coding-related problems and Mr. Jussi Pekonen for helping me to carry out the recordings. I would also like to thank Mrs. Lea Söderman for taking care of all practical issues.

Finally, I would like to thank my family for their love and prayers. My father, Professor Pentti Huovinen, has encouraged me to believe in myself and in my dreams. My mother, Dr. Saara Huovinen, has supported me in every aspects of my life during this whole process. My brother Ville and sister Elisa, thank you for your encouragement. I would also like to thank my parents-in-law for their support.

Last, but definitely not least, I would like to thank my dear husband Hannu. Thank you for your patience and love during the good and especially the hard times.

Otaniemi, November 29, 2005

Heidi-Maria Lehtonen

Contents

Abbreviations

DFT	Discrete Fourier Transform
FDN	Feedback Delay Network
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FM	Frequency Modulation
IFFT	Inverse Fast Fourier Transform
IFIR	Interpolated Finite Impulse Response
IIR	Infinite Impulse Response
KS	Karplus-Strong
LPC	Linear Predictive Coding
VCO	Voltage Controlled Oscillator

List of Figures

Chapter 1

Introduction

In this thesis, the aim is to analyze the piano sound and study the modeling of the instrument. The piano is one of the most widely used instrument in the western music. Moreover, the word piano covers actually two separate instruments, the grand piano and the upright piano. They both are extremely complex from the physical point of view. This fact makes the modeling task fascinating and challenging. The acoustics of the grand piano and the upright piano have been an interesting research topic over decades, but despite the dozens of studies the behavior of the instruments are anything but clear. An overview to the acoustics of the grand piano is given in Chapter 2. This analysis can be applied mostly to the upright piano as well.

An acoustical piano may not be suitable choice for all situations and for all pianists. For example, a digital piano may be a better solution for those who live in apartment houses, as the digital instrument does not cause structural noise, which travels from one apartment to another. At the moment, most of the digital pianos use the sampling technique, in which recorded piano tones are played back from the database. The advantage is that the sounds of high-quality digital pianos are authentic and the computational requirements are not significant. On the other hand, the need for memory is heavy, since all tones must be saved in the database with several different dynamic levels. Also, some features of the piano sound, such as the coupling between the played tones, cannot be modeled correctly, since the tones are simply played back separately from each other. These facts give a fruitful starting point to the development of a parametric piano synthesis model where the basis lies on the physics of the instrument. In addition, if the implementation of the model is done in a computationally efficient way, physics-based sound synthesis can provide a good alternative for digital pianos using the sampling technique. At the moment, the digital waveguide modeling has turned out to be the most feasible physical sound synthesis technique for plucked and struck string instrument modeling. It simulates the traveling wave solution of the wave equation and provides a computationally efficient solution for the real-time implementation. The digital waveguide modeling of the piano is studied in Chapter 3.

The piano has a very characteristic sound. The features of the piano sound must be taken into account in the modeling procedure. Parametric modeling provides a good baseline for a physical model, as the parameters can be extracted from recorded piano tones and studied by means of different signal analysis tools. These parameters are used to control the physical model, which imitates the sound production mechanism of the piano. The main features of the piano tones are inharmonicity, which comes from the stiffness of the piano strings, and complicated decay of the tones, which basically comes from the energy leakage into the soundboard and other energy losses. A novel loss filter design technique for modeling the decay of the tonesis presented in Chapter 4.

In addition to the string model, models for soundboard and sustain pedal are needed in a high-quality parametric piano. Despite their importance, these are the least-studied parts of the acoustic piano. Especially, only few models for the sustain pedal have been presented in literature. An overview of the sustain pedal effect is given in Chapter 5, and a reverberator algorithm is presented for modeling the effect.

Despite the fact that the piano has been studied by a number of researchers, the work is by no means completed. It is curious that this instrument is very popular and known widely around the world, but the physics of the instrument still remain partially unknown. This affects also the modeling work. In order to make reliable physical sound synthesis, the key assumption is that the physics of the instruments are known, at least at some level. Moreover, the operation of human auditory system in the piano sound perception should be known better in order to make the synthesis model more efficient. This is a part of the future work, which is discussed in Chapter 6. This chapter also concludes the work.

Chapter 2

Overview of the Piano

The piano is probably the most popular instrument used in the western music. It has a very characteristic sound and a wide dynamic range, and its playing range is more than seven octaves. Usually the piano is used as a solo instrument but often it is used to accompany other solo instruments or singing. Its popularity probably arises from its versatility and the fact that it is capable of producing both powerful and sensitive sounds. Also, the control of the sound production mechanism is relatively easy, as there are only three different control parameters: The number of the key, the velocity of the press, and the pedals.

The roots of the modern piano go back into the beginning of the 18th century (Fletcher and Rossing, 1991). In 1709, Bartolomeo Christofori of Florence modified the harpsichord in order to make the instrument capable of variations in tone. He replaced the jacks with hammers and called the new instrument the “gravicembalo col piano et forte”. During three centuries, the instrument has evolved into two distinct instruments, the grand piano and the upright piano. In Fig. 2.1 a Steinway & Sons grand piano is shown and in Fig. 2.2 a Yamaha upright piano is presented. The pictures were taken in Espoo Music School during the recording session in February 2005. For more information about the recordings, see Appendix A.

The grand piano consists of five main parts, the keyboard, the action, the strings, the soundboard, and the frame (Fletcher and Rossing, 1991). The keyboard consists of 88 keys of which 52 are white and 36 are black. From the keystroke the information message is transmitted to the action, which controls the hammer. The hammer hits the string and sets it into vibration. The kinetic energy of the hammer is transformed into vibrational energy, which is stored into the normal modes of the string. This energy is transmitted into the soundboard via the bridge. The soundboard is a thin, wooden plate positioned under the frame. The cast-iron frame, positioned to the upper part of the wooden case, keeps the instrument together, and it is designed to withstand the high tension of the strings. The

Figure 2.1: A grand piano.

Figure 2.2: An upright piano.

strings are attached to the tuning pins at the player end and to the hitch-pin rail at the other end. The general structure of the instrument is presented in Fig. 2.3.

2.1 Keyboard

The player controls the output sound mainly with the keyboard. The choice of the key determines the pitch of the tone, whereas the velocity of the finger press determines the loudness. This interaction between the keyboard and the player is of great importance in

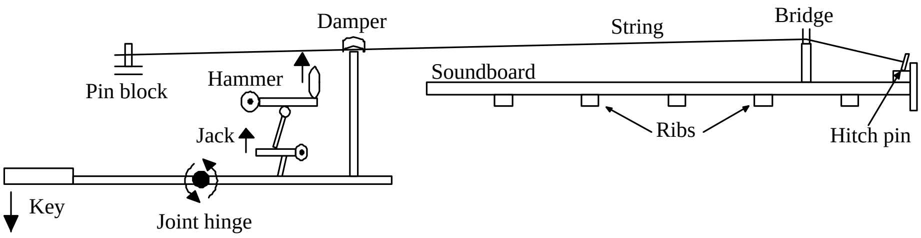

Figure 2.3: The general structure of a grand piano.

the piano performance. Obviously, it is important that the feedback from the instrument to the player is homogenous throughout the whole keyboard. The needed force for the key depress should be the same for all keys. At the same time the “touch weight” should not feel too light or too heavy to the pianist.

The “touch of the pianist” is a concept that divides the physicists’ and pianists’ opinions. Physicists claim that the resulting sound does not depend on anything else than the velocity of the finger press. On the other hand, many skilled pianists have the opinion that the “touch” is an important factor in the art of piano music. By applying different kinds of touches, the pianist can have major influence on the shadings of the music. It should be pointed out that Askenfelt and Jansson (1990) have found out that at least the key motion depends on the touch: The performance of the action is different depending on the playing style, e.g., staccato and legato.

2.2 Action and the Hammer

The highly complicated machinery of the piano is called the action. This true specimen of skill is responsible for transmitting the players intentions and accounts from the control mechanism of the instrument to the strings. The most important part of the action is the interaction between the hammer and the strings. Fig. 2.3 illustrates the principal idea of the action and in the following, its operation is presented in general level. More comprehensive studies can be found in (Askenfelt and Jansson, 1990) and in (Fletcher and Rossing, 1991).

Askenfelt and Jansson (1990) divided the piano action in four major parts: The key, the lever body, the hammer, and the damper. When the key is depressed, the damper lying on the string is lifted up. Right after this, the lever body is rotated upwards and the hammer is accelerated when the jack pushes the roller positioned under the hammer. Just before the hammer strikes the string, the jack reaches the escapement dolly and the top of the jack is moved away from the hammer roller. The hammer continues the rest of its journey to the string freely.

After the hammer strike, the rebounding hammer falls down and it is captured by the check. This way the hammer is prevented from striking the string a second time, and the action is now ready for repetition. If the key is held down after the hammer-string interaction, the action will remain in this state. If the key is released right after the hammer stroke, the action will return to its rest position.

The hammers have a strong effect on the tone quality. Their material, strength and size contribute to the loudness as well as to the playability. The hammers are covered with two layers of wool felt stretched over wooden moldings. The harder the felt material, the louder and brighter the sound. In order to keep the loudness uniform throughout the scale, the treble hammers must be harder than the bass hammers.

In modern instruments, the size and mass of the hammers decrease from bass to treble in order to optimize the relation between aforementioned properties. If the hammer is heavy, its contact time with the string increases. This, in turn, may increase losses in the case that the roundtrip travel time for the initial pulse becomes shorter than the hammer-string interaction time.

More information related to piano hammers and research dealing with hammers can be found e.g. in (Conklin, 1990) and in (Conklin, 1996a).

2.3 Piano Strings

The strings are often called the heart of the piano. A concert grand piano has 243 steel-made strings. The lowest strings are long and massive (length of even 2 m) while the highest strings are thin and short. At the treble end, the length of the strings is only about 5 cm. At the bass end, the lowest strings overlap the middle strings providing them to get nearer to the soundboard. The strings are terminated at the tuning pins at the player end and at the hitch pin at the far end. The pins are made of steel and have a diameter from 3 to 4 mm. The speaking length of the string, though, is restricted to the bridge, see Fig. 2.3.

The first eight strings are single strings and the rest of the strings corresponding to the 80 highest keys are in groups of two or three strings, depending on the instrument. The strings corresponding to the lowest 20 keys are wrapped with one or two layers of wire and the rest 68 sets of three strings are unwrapped. Wrapping makes the bass string more flexible compared to a solid string of the same diameter, which is advantageous since solid strings with the same diameter and mass would result in highly inharmonic sound (Fletcher and Rossing, 1991).

2.3.1 Inharmonicity

The inharmonicity in the piano strings is caused by stiffness. It makes the higher partials travel faster on the piano string, which means that their frequencies are a little higher compared to those of an ideal string. In the spectrum, the partial components are slightly shifted making the series of overtones “stretched” upwards.

A mathematic expression for a stiff string can be obtained when the wave equation is formulated as a combination of the equations for an ideal string and an ideal bar, see e.g. (Smith III, 2004). Solving this equation for small stiffness values, a relation between the modal frequencies and inharmonicity coefficient B can be solved. A more comprehensive study on the mathematical derivation of this relation is given in section 3.2.1.

The inharmonicity coefficient of a tone can be measured, although it is far from being an easy task. Galembo and Askenfelt (1999) presented a method which estimates the fundamental frequency and the inharmonicity coefficient at the same time. With this method they reached convincing results, although the fundamental frequency estimation problem is known to be challenging as well.

Usually, strong inharmonicity is considered to be an undesired feature of the piano sound. On the other hand, it is not desired to get totally rid of it, since a slight inharmonicity adds warmth to the sound (Fletcher et al., 1962). Interestingly, it has been shown that the human ear is more sensitive to inharmonicity when the fundamental frequency of the tone is low (Järveläinen et al., 2001). In this light, it is reasonable to use wrapped strings at the bass range of the piano. Although the treble strings are highly dispersive, they contain only a few partial components on the audio range. Therefore it is not necessary to eliminate the effects caused by inharmonicity by using wound strings at the treble end of the piano.

The inharmonicity is taken into account in the tuning process (Fletcher and Rossing, 1991). The tuning is usually done in a way that the beating between the notes is minimized. This will result in a stretched scale, where the fact that the upper partials of the lower tones are slightly sharp with respect to the upper tones, is compensated. In other words, the lowest strings are undertuned and the highest strings are overtuned.

2.3.2 Sound Decay

The sound generated by the piano strings is very complicated. The sound builds up rapidly and decays slowly. Moreover, partials decay at different rates. Some of the partials may sound even dozens of seconds whereas other partials decay in a few seconds. The spectrum varies over time and differs from key to key; at the bass end over 50 partials can be extracted while at the treble end the corresponding number is only about 3 or 4. As the strings preserve most of the energy transmitted by the hammer and the action, the decay rate of the tone depends on how rapidly the energy is leaking out from the strings.

The decay rate of the piano tone is twofold: The tone begins to decay rapidly, but after some milliseconds the decay rate changes and becomes slower. As stated by Weinreich (1977), the amount of aftersound, that is, the portion of the "slow" decay, depends on the degree of mistuning. He also pointed out that the hammer irregularities may have an effect on the two-stage decay as one string of the string group can be hit harder than another. There is yet another explanation to the two stage decay: Change in the predominant vibration of the strings. The vertical (perpendicular to the soundboard) vibration decays rapidly whereas the horizontal (parallel to the soundboard) vibration decays slowly (Weinreich, 1977).

Other important phenomenon in the piano sound is the beating, which results from unison groups of strings. When the hammer excites a tricord, that is, a set of three unison strings, the strings begin to vibrate in the same phase. Due to small differences in frequency between the strings in the tricord, the tone starts to beat soon. The strings can be by no means considered to be independent, since they are coupled to the bridge. This coupling allows energy leakage between the strings resulting in a highly complicated system. A comprehensive study about coupling of two strings and the bridge admittance is given in (Weinreich, 1977).

2.4 The Soundboard and the Frame

The soundboard is the main radiating structure of the piano. Usually it is made of spruce strips which are glued together edge-to-edge. Widely used spruce species are Sitka spruce or red spruce. Young's modulus, as well as the strength, varies from tree to tree. These values are much smaller in a direction at 90 degrees to the grain than in the grain direction. This difference in cross-gain stiffness is compensated with ribs, which are glued onto the bottom of the grand piano soundboard. In upright pianos, the ribs are glued onto the back of the soundboard. Conklin (1996b) stated that an unribbed solid spruce soundboard would cause splitting of the soundboard due to stresses caused by ambient factors, like changes in humidity and temperature. He also stated that the output of the instrument would be less than normal if there were no ribs on the soundboard.

The behavior of the soundboard can be examined with different kinds of methods. One of the most popular is the Chladni method (Conklin, 1996b), which gives visual results, but other powerful methods also exist. In addition, it is possible to investigate the vibrational behavior of the soundboard by numerical analysis, like the finite element method. This kind of approach was taken among others by Berthaut et al. (2003). The advantage of computer simulation is that a prototype model of a soundboard need not be constructed, on the understanding that the material parameters and the exact dimensions of the soundboard are known.

The bridge couples the strings to the soundboard. Modern pianos have generally two bridges, one for the bass strings and another for the middle and treble strings. Usually the bridges are made of hardwood instead of softwood (like the spruce) which is unsuitable for bridge material due to its low density. Too low a density allows extra losses at the string-terminating points. Also, the strength of the softwood is inadequate to resist the mechanical stresses around the metal bridge pins, where the strings are terminated. If the strings were terminated straight to the soundboard, the resulting sound would be louder but unpleasant and the decay times of the tones would be too short (Conklin, 1996b). Thus, the existence and the quality of the bridge affects the piano sound remarkably.

Since the mechanical impedance of the soundboard is greater than the mechanical impedance of the strings, the energy is leaking very slowly to the soundboard. An important parameter which affects the leaking rate is the driving-point impedance, which varies with frequency. The exact values can be determined by measurements. Logical places to measure impedances are in several places on the bridges.

In the late 1800s, cast iron frames replaced the wooden string plates. This renewal enabled the increase of the acoustical output and extension of the keyboard. At the same time, the quality of the piano sound improved. The cast iron plate provides proper strength to support the high tension of the strings (which can exceed 1000 N!) as well as sufficient stability (Fletcher and Rossing, 1991). Also the tunability improved as the tension of one single string became almost independent of the tension of the other strings.

The cast iron frame is not a good tone radiator. Due to this fact, the string energy dissipation within should be prevented. The plate should not vibrate during the playing and the mechanical impedance of the plate should be substantially higher than that of the strings.

2.4.1 Pedals

The use of the pedals was not marked in musical notes before the 1790s. In the beginning, it was considered gimmicky, but soon in the 19th century composers like Chopin and Liszt took advantage of the modern technique and employed the usage of the pedals actively in their work (Trustees of Dartmouth College, 2004).

With the help of the pedals, the player can add variation to the tones. Modern pianos have either two or three pedals. The right pedal is always the sustain pedal, which lifts all the dampers and sets all strings into vibration due to sympathetic resonance and the energy transmission via the bridge. The sound rings as long as the sustain pedal is held down, regardless of whether the key is released or not.

The left pedal is called the una corda pedal and its purpose is to make the output sound softer. Una corda is Italian and means “one string”. In grand pianos, the whole keyboard is shifted right causing the hammer to strike only one string instead of two, or two strings instead of three. In addition to the decrease in loudness, the una corda pedal affects the timbre as the coupling between the strings does not remain the same (Weinreich, 1977). On the upright piano, the una corda pedal moves the action closer to the strings.

The purpose of the third pedal, which is positioned between the una corda and the sustain pedal, can vary depending on the instrument. In some instruments, the center pedal can work as a bass sustain. In this case, depressing the middle pedal sustains only the notes in the bass section. In most grand pianos, the middle pedal is a "sostenuto" pedal which sustains only the notes played before the pedal is depressed. This effect gives impression of a third hand, because the player can employ both of his or her hands to the playing while the sostenuto pedal keeps the chosen notes ringing. In some upright pianos, the center pedal is a practice pedal. The tone is softened with a piece of felt which is lowered between the hammers and the strings. (Piano World, 2005; Fletcher and Rossing, 1991)

2.5 Grand Piano vs. Upright Piano

A question that has long exercised the minds of researchers is whether there is audible difference between the grand piano and the upright piano. It is clear that the structures and mechanisms between these two instruments are different, but how these differences correspond to the sounds of the instruments is still far from being completely obvious. In the following, the structural differences are investigated and some aspects on the distinction are given.

2.5.1 Structural Properties of the Upright Piano

Like grand pianos, there exist upright pianos of several sizes. The largest upright pianos stand 150 cm in height whereas some studio uprights are only 100 cm of height. In large upright pianos, the action is positioned some distance above the keyboard and in small uprights the action is partly or even completely below the keyboard.

The major difference in the upright piano action compared to the grand piano is that the hammers and the dampers move horizontally. The repetition lever, which enables fast repetitions in grand pianos, is absent from the upright piano action. This means that the key must be returned to its rest position before it can be depressed again. This is a clear feature that most pianists recognize; it may affect also the prevailing opinion among musicians that the grand piano is better than the upright piano. (Fletcher and Rossing, 1991)

The soundboard of an upright piano is rectangular. However, the vibrational area of the soundboard is trapezoidal, because of the trimming rims that are positioned in the two opposite corners of the soundboard. The two bridges lay diagonally on the soundboard, inside the vibrational area.

2.5.2 Is There Noticeable Difference Between the Sounds of the Two Instruments?

A few studies exist about the distinction of the two instruments. Wogram and Mori (2003) have listed the main properties of pianos with respect to their musical acoustical qualities. They have measured these parameters with subjective listening tests and resulted in several interesting observations. They used single tones, sequences of single tones, chords and musical pieces in their listening tests, but found out that the sounds of upright and grand pianos can only be distinguished reliably when musical pieces are played.

Probably the most interesting result is that according to their tests the two instruments cannot be distinguished by the inharmonicity. This is surprising since usually the upright piano is considered to have a more inharmonic sound than the grand piano as the strings are, in general, shorter than in the grand piano. In the test, the subjects found the differences in inharmonicity between the three strings of a choir in a grand piano greater than between one upright piano string and the grand piano strings!

A more intuitive result concerns the sound radiation. Because of the different positions of the soundboards of the two instruments, the listener perceives the sound fields in a different way. It turns out that in the case of a grand piano there is a better balance between the bass and treble sounds whereas in the case of an upright piano the bass seems to dominate the sound. Another factor that affects the perceived sound balance is the eigenmode density of the soundboard. In the case of the grand piano with higher eigenmode density, the sound is better balanced.

As already mentioned, the piano action plays a key role in the distinction of the two instruments for the player. Wogram and Mori (2003) found out that besides the ability for faster repetition, the grand piano also has more efficient dampers. In the upright piano action, the dampers are positioned beside the spot of attack, which results in insufficient damping of some sound components. In grand pianos, the dampers are positioned in a better way, and all modes will be damped properly.

2.6 Conclusion

In this chapter, an overview of the piano was given. This overview is by no means all-inclusive, but rather a compact glance at the structure and acoustics of the piano. As it can be noticed, the instrument is remarkably complex, but simple to control. This makes the modeling task attractive and challenging at the same time. It is also clear that perceptual issues can facilitate the modeling problem since there is no sense to model features that exist but are not perceptually significant. This fact calls for reliable investigations on the human perception and auditory system.

It is also worth noticing that the grand piano and the upright piano are two distinct instruments. Although they sound similar to many people, their structures and sound production mechanisms are different. Especially, skilled piano players and musicians distinguish the instruments by both sound and the response of the instrument.

Chapter 3

Digital Waveguide Modeling of the Piano

This chapter is divided into two major sections. In the beginning, a short overview of synthesis methods is given. Since the digital waveguide (DWG) modeling technique is the most important from the viewpoint of this thesis, it is presented in a more extensive way. The second part concerns the digital waveguide modeling of the piano.

3.1 A Short Overview of Synthesis Methods

Smith (1991) proposed a taxonomy which divides the sound synthesis methods into four different main classes: Abstract algorithms, processed recordings (and sampling), spectral models and physical models. According to Tolonen et al. (1998), the last two methods mentioned are the most promising when it comes to the future of sound synthesis. However, the two aforementioned methods give important basis to new, elaborate methods. For a more comprehensive overview of synthesis methods, see references (Roads, 1995) and (Tolonen et al., 1998). In the latter reference, also comparison between synthesis methods is performed.

3.1.1 Abstract Algorithms

Synthesis methods that produce sounds based on some *ad hoc* mathematical formula are considered to belong to the class of abstract algorithms. The most important synthesis methods belonging to this class are frequency modulation (FM) synthesis (Chowning, 1973), waveshaping synthesis (Arfib, 1979; Le Brun, 1979) and synthesis based on the Karplus-Strong (KS) algorithm (Karplus and Strong, 1983).

In general, with these methods it is almost impossible to produce natural sounds that resemble real instruments. On the other hand, the KS algorithm gives an important background to some physical sound synthesis methods (Smith, 1983; Jaffe and Smith, 1983).

3.1.2 Processed Recordings

Processing of recorded sounds (also called the sampling technique or wavetable synthesis), is a sound synthesis technique in which sounds are played back from recordings (Roads, 1995). Its advantages are its rather easy implementation and the fact that this technique offers a great variation of possibilities in terms of sound modifications. On the other hand, the drawback is that the amount of memory required for implementation is huge. At the moment, the best piano synthesizers are based on the sampling technique.

3.1.3 Spectral Models

Spectral models try to model and imitate the properties of sound waves. Well known methods of this class are, for example, additive synthesis, which is one of the oldest synthesis methods, the phase vocoder (Flanagan and Golden, 1966) and the source-filter synthesis, which is widely used especially in the area of speech synthesis (Moorer, 1985).

3.1.4 Physics-based Modeling Techniques

In the field of the sound synthesis of musical instruments, physical modeling is probably the most widely used method at the moment. The strength of physical modeling is that it provides intuitive approach to the control of the instrument models. The idea is to simulate the sound production mechanism of the instrument instead of the sound itself. Physical modeling techniques can be divided into six categories: Numerical solving of partial differential equations, source-filter modeling, vibrating mass-spring networks, modal synthesis, wave digital filters and digital waveguide synthesis (Välimäki et al., 2006).

Numerical solving of the wave equation gives base to the finite difference method. This approach was first taken by Hiller and Ruiz (1971a,b). This method is extensively used in the field of sound synthesis of musical instruments. Its drawback compared to the digital waveguide technique is that its computational complexity grows heavily with the size and complexity of the model.

The source-filter synthesis mentioned in Section 3.1.3, can be included to the category of physical models as well, because it models the sound source rather than the sound itself. The vibrating mass-spring systems are based on the idea of modeling the mechanical elements of a vibrating system, such as masses, dampers and springs. The idea was first presented by Cadoz et al. (1983).

The idea of modal synthesis is that any system that produces sound can be divided into vibrating substructures defined by modal data (Adrien, 1991). The energy flows between the substructures via couplings and the system can also respond to external excitations. The advantage of the method is that it can be applied to very complex structures.

Wave digital filters are initially developed for discrete-time simulation of analog electric circuits. They are especially well suited for wave-based mass-spring modeling. This technique complements the DWG technique, as the nonlinearities in the instrument models can be modeled with wave digital filters. The DWG technique is used for the wave propagation part of the model. This technique was first presented by Fettweis (1986).

The DWG synthesis technique has turned out to be the most efficient and feasible when it comes to the modeling of musical instrument sounds. The technique is especially well suited e.g. to the simulation of a vibrating string or an acoustic tube. The theory is mainly developed by Smith (1983, 1992, 1997). Also, in (Karjalainen et al., 1998) a very nice introduction to the topic is given, in addition to some extensions. The idea of this method originates from discretization of the traveling-wave solution of the wave equation. From this point of view, the DWG modeling has the same starting point as the finite difference method. However, it has certain advantages over the finite difference technique. For example, the losses and the dispersion of the model can be lumped in one point, which means, in turn that the whole system can be described with delay lines and linear filters.

In the following, the theory of DWG method is presented as a part of the piano string model presentation as it describes the mathematical background for the method sufficiently.

3.2 Digital Waveguide Modeling of the Piano

Several DWG piano models can be found in literature. The first one was presented by Garnett (1987). Later, in 1995, Smith and Van Duyne presented a piano model (Smith and Van Duyne, 1995; Van Duyne and Smith, 1995) based on commuted synthesis. A more recent study on DWG piano modeling was given by Bank (2000b). A good overview of prior research and extensions to the above-mentioned model was presented by Bank et al. (2003). A very detailed piano model partially based on DWG modeling was presented by Bensa (2003) and Bensa et al. (2003).

3.2.1 Modeling the String

The approach taken here follows the DWG modeling scheme starting from the time-domain solution of the wave equation. The one-dimensional wave equation for an ideal string can be written (Morse and Ingard, 1968)

$$T\frac{\partial^2 y}{\partial x^2} = \mu\frac{\partial^2 y}{\partial t^2} \tag{3.1}$$

where x is the position along the string, y is the transversal displacement, t is time, T stands for string tension and μ is linear mass density of the string.

The wave equation can be interpreted as the Newton's second law, $F = ma$, where F is force, m stands for mass and a for acceleration. In the case of an ideal string, the force is given as a product of the string tension and the curvature of the string, and the right side of the equation is the mass density times transverse acceleration. The traveling-wave solution of the wave equation can be written as a superposition of two waves, one traveling to the right and the other traveling to the left:

$$y(x,t) = f^+(ct - x) + f^-(ct + x), \tag{3.2}$$

where f^+ is the right-going wave and f^- is the left-going wave. It is assumed that both f^+ and f^- are twice differentiable functions. Their shape can be arbitrary.

The traveling-wave solution in Eq. 3.2 can be discretized by sampling the amplitudes of the traveling waves at every T seconds. This corresponds to the spatial sampling interval X, which is the distance the wave propagates in T seconds, that is $X \triangleq cT$, where c is the sound velocity. Since one spatial unit corresponds to one temporal unit and the whole wave propagates through the whole string, the wave propagating either right or left can be simulated with a single digital delay line. Respectively, both waves can be simulated with two parallel delay lines, see Fig. 3.1. This digital waveguide model of an ideal string can be written as

$$y(t_n, x_m) = y^+(n - m) + y^-(n + m). \tag{3.3}$$

In Eq. 3.3, the following changes of variables are made compared to Eq. 3.2: $x \to x_m = mX$ and $t \to t_n = nT$. Also, T is suppressed since it multiplies all arguments.

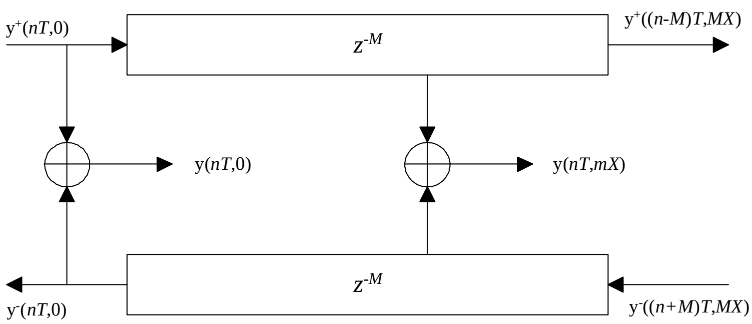

Figure 3.1: Two parallel delay lines for simulating two propagating waves.

The simulation made with a digital waveguide model is exact only at the sampling instants. However, it is possible to model a continuous waveform by interpolation between two adjacent points. Also, with the choice of the sampling period it is possible to contribute to the precision of the simulation. In general, the theoretical lower limit to the sampling frequency is twice the highest frequency occurring in the system. Equivalently, the waves propagating along the string must be bandlimited. Otherwise aliasing would occur and the reconstruction of the signal becomes impossible. This restriction is called the Nyquist condition, see e.g. Oppenheim and Schafer (1975), p. 29. In CD-quality audio, the sampling rate is 44.1 kHz.

It should be pointed out that so far only a one dimensional case is studied. In piano synthesis, at least three dimensions should be taken into account. That is, the two transversal vibration polarization, the horizontal and the vertical, and the third corresponds to longitudinal waves. Actually, when the hammer strikes three strings instead of one, nine waveguides per key should be included. However, this kind of approach is usually unnecessary, since the result is sufficient with one waveguide per coupled string (Bensa, 2003).

Modeling the Losses

In an ideal string, the string terminations are assumed to be rigid. This means that the waves reflect with the same amplitude; the force waves change their sign and the velocity waves reflect with the same sign. This is not the case with real piano strings, since losses occur due to energy transmission to the bridge and soundboard. Also sympathetic resonance between the strings causes energy losses, as well as the drag by the surrounding air. These losses have to be modeled in high-quality piano synthesis. This is traditionally done by inserting a loss filter to the string model (Van Duyne and Smith, 1995; Bank and Välimäki, 2003). Basically, the loss filter can reduce to one loop attenuation factor when all frequencies are decaying at the same rate. However, this is not the case with real strings, since high partials tend to decay more rapidly than lower ones. Moreover, the variations in the gain specification between two adjacent partials can be significant. The desired gain values can be determined from a real piano sound, since the relation between the decay time constant of the kth partial and the corresponding filter gain value g_k can be expressed as a closed-form formula:

$$g_k = e^{\frac{-1}{f_0 \tau_k}}, \tag{3.4}$$

where f_0 is the fundamental frequency, and τ_k is the decay time constant.

Van Duyne and Smith (1995) proposed a lowpass filter to be used as a loss filter. This lowpass filter is a part of a coupling filter, which also takes into account the coupling between the three strings of a tricord. This approach is simplified but efficient way of modeling the overall losses and coupling between the strings. The problem in the model is that it does not take the variation in the gain specification into account.

Bank (2000b) presented a transformation method for loss filter design. This technique comes from the idea of using a certain transformed specification, which minimizes the error of the decay times in a mean-square sense. The actual filter design process can be done by any least squares filter design algorithm. Later, Bank and Välimäki (2003) introduced a weighting function based on the first-order Taylor series approximation of the decay time errors which can be used in the design of high order loss filters. Generally, both of these aforementioned methods result in filters of relatively high orders.

Välimäki et al. (2004) proposed a combination of a one-pole filter and an FIR comb filter called the ripple filter in order to achieve a reduced-complexity loss filter for harpsichord synthesis. This technique allows exact matching of the decay rate of one partial and thus some variation in overall response. However, this technique seems to be insufficient for piano synthesis where the decay process is more complicated.

Rauhala et al. (2005) proposed an extension to the design method presented in (Välimäki et al., 2004) in a sense that more than one feedforward path is added in cascade with the one-pole filter. The design method is custom-made in addition to automatic smoothing of the gain specification. It is capable of high accuracy with a low computational cost.

Modeling the Dispersion

Another physical phenomena occurring in piano strings is dispersion. Due to the high stiffness of the strings, the resulting sound is inharmonic as the partials are slightly higher than those of the harmonic case. This is a feature that can not be ignored in high-quality synthesis. Unfortunately, for accurate simulation, a high-order filter is needed.

The inharmonicity of the strings is usually modeled with an allpass filter, which does not affect the magnitude response of the system (Smith, 1983; Garnett, 1987). As the target is to make the higher frequencies propagate faster along the string, a filter with a proper phase response is needed. The frequency of the kth partial of a tone can be computed as (Fletcher et al., 1962)

$$f_k = kf_0\sqrt{1 + Bk^2}, \tag{3.5}$$

where f_0 is the nominal fundamental frequency and B is the inharmonicity coefficient. The phase delay specification for the allpass filter can be written as

$$D_d(f_k) = \frac{f_s k}{f_k} - L - D_l(f_k), \tag{3.6}$$

where L is the delayline length of the string and $D_l(f_k)$ is the delay specification of the loss filter. From Eq. 3.6, the phase specification for the allpass filter can be derived.

Van Duyne and Smith (1994, 1995) proposed a bank of first-order allpass filters for the dispersion simulation. However, this approach does not allow completely accurate simulation of the dispersion phenomenon. A more accurate solution was introduced by Rocchesso and Scalcon (1996). The method computes the filter coefficients from an overdetermined system of equations by using the least-squared equation error criteria. In addition, the method takes the fine-tuning of the string into account in the design process. The disadvantage is the increased computational load.

Bank (2000b) introduced a multirate approach to decrease the computational load of the dispersion simulation significantly. He took advantage of the fact that the lowest tones of the piano can be simulated at half sampling rate. The filter simulating dispersion may thus be more complex for accurate simulation of the phenomenon. On the other hand, only interpolating filters in the outputs of the lowest tones are needed.

3.2.2 Modeling the Hammer

The interaction between the hammer and the string is highly nonlinear, which makes the modeling procedure more complicated. The nonlinearity comes from the felt covering the hammer, which exhibits a hysteretic behavior: The contact force during the compression is different from that during the decompression.

The contact force between the hammer and string can be described by the equation

$$F(t) = -m_h \frac{d^2 y_h(t)}{dt^2}, \tag{3.7}$$

where m_h is the hammer mass and y_h is the hammer displacement. The relation can be interpreted as the hammer being a lumped mass connected to a nonlinear spring. As can be found from literature, see e.g. (Chaigne and Askenfelt, 1994a), the hammer-string interaction can be described by the equation

$$F(t) = \begin{cases} \kappa \Delta y(t)^p & \Delta y(t) > 0 \\ 0 & \Delta y(t) \le 0 \end{cases} \tag{3.8}$$

where $\Delta y(t) = y_h(t) - y_s(t)$ presents the compression of the hammer felt and y_s is the string position. In Eq. 3.8, κ stands for hammer stiffness coefficient and p is the stiffness exponent. Both of these parameters are usually determined from experimental data. The condition $\Delta y(t) > 0$ applies when the hammer is in contact with the string and the latter condition $\Delta y(t) \le 0$ is valid when the hammer is not in touch with the string.

Eq. 3.8 does not describe the phenomenon fully satisfactorily, since in real piano the hammers indicate hysteretic behavior. This problem was studied by Stulov (1995). He solved the problem by replacing the first part of Eq. 3.8 by

$$F(t) = \kappa\Big(1 - h_r(t)\Big) \circledast \Big(\Delta y(t)^p\Big), \tag{3.9}$$

where $h_r(t) = (\epsilon/\tau)e^{-t/\tau}$ is a relaxation function, ϵ stands for hysteresis constant and τ is “nondimensional time”. As the relaxation function depends on time, it can be interpreted that it represents the “memory” of the material.

When it comes to modeling, the hammer model can be discretized and coupled to the string. There still exists a problem, though. There is a mutual dependence between Eq. 3.7 and Eq. 3.9, namely the hammer position should be known before computing the force and the force should be known before computing the hammer position.

The implicit relation between Eq. 3.7 and Eq. 3.9 can be made explicit by inserting a fictitious delay element in the model. This kind of approach is widely used in literature, see e.g. (Chaigne and Askenfelt, 1994a,b). However, this kind of approximation can make the model unstable.

To avoid the problem described above in addition to the nonlinearity problem, Smith and Van Duyne (1995) came up with an idea that the hammer-string interaction consist of a few discrete events during the hammer strike. These hammer strikes can be approximated with one or more impulses that are filtered with a lowpass filters. Taking advantage of the superposition of single lowpass filtered impulses the resulted signal approximates the force impulse in a very efficient way. The parameters of these filters depend on the collision velocity. That is, changing the collision velocity means that the filter parameters must be changed. In addition, with this method the hammer restrike cannot be simulated correctly. On the other hand, the advantage is that the linearized hammer can be included to the commuted piano model.

A more general method was presented by Borin et al. (2000). This method, called the “K method” maps the interaction force F as a function of the linear combination of the past values of the string and hammer positions as well as the interaction force. The advance is that the instantaneous dependencies of the variables are dropped. For a more extensive description of the method, see (Borin et al., 2000).

A more recent study is presented in Bank (2000b,a). The multi-rate hammer model overcomes the stability problem by doubling the sample rate in order to achieve smaller changes in the variables of interest. However, doubling the sample rate in the whole string model would double the computational load as well. To overcome this problem, Bank suggests that only the hammer would operate at the increased sampling rate.

3.2.3 Modeling the Soundboard

The soundboard is the main radiating part of the piano. Its task is to color and amplify the sound as well as to create the sensation of presence. Despite its importance, it is the least studied part of the piano.

The modeling is often done with feedback delay networks (FDN), which are known to be efficient for room reverberation simulations due to their abilities to achieve high modal density, see e.g. (Jot and Chaigne, 1991). The problem with FDN systems is that the choice of parameter values seems to have no correspondence in the real world. The delayline lengths are somewhat arbitrary, only the lengths are preferred to be prime numbers in order to avoid the unwanted coloration (Gardner, 1998). In addition, the modeling can be addressed as a filter design problem. Due to high modal density, high filter orders must be used.

The first DWG piano presented by Garnett (1987) included a soundboard model with six extra waveguides connected to the bridge at a single location. This simple, but efficient model can be interpreted as a predecessor for the feedback delay network solution for the soundboard modeling. Bank (2000b) proposed a soundboard model structure based on FDN with shaping filters. These filters match the system to imitate the overall magnitude response of a real piano as well as possible.

Välimäki et al. (2004) presented an efficient system for harpsichord soundboard modeling based on the algorithm presented in (Väänänen et al., 1997). The model consists of eight delay lines, loss filters and comb allpass filters. This structure results in an authentic impulse response of a harpsichord soundboard.

Chapter 4

A Novel Loss Filter Design Technique

The purpose of the loss filter is to model the decay rate of the partials accurately. Usually, the loss filter is a low order FIR or IIR filter (Smith, 1983; Bank and Välimäki, 2003). For good accuracy, a high-order filter is needed because the variations of the magnitude response from one partial to the next one cannot be matched with a low order filter. In loss filter design, a major problem is that the desired magnitude response is usually specified on a very narrow frequency band. For example, for the lowest keys of the piano, the partials below 2 kHz affect the sound most, and it is difficult to accurately estimate the decay rate of high-frequency partials. However, the sampling rate used in high quality piano synthesis is usually 44.1 kHz. The traditional filter design methods cannot face the challenge that the important frequency band is only 10 percent of the audio range.

Another major problem faced in the design process is that the gain values can vary significantly between two adjacent data points. Especially, the loss filters designed for the lowest key values would have to follow very dense and detailed gain specification. This is not possible for filters of low order.

The loss filter design technique presented here[1] is an extension to the multi-ripple loss filter design technique presented by Rauhala et al. (2005). The design technique proposed in (Rauhala et al., 2005) consists of two cascaded subfilters. A first-order all-pole filter takes care of modeling the general trend of the piano tone's decay rate and the Nth-order feedforward comb filter models the decay rate variations from one partial to the next one. The principal difference in the loss filter structure presented here is that in addition to the parallel feedforward structure it is possible to insert feedforward blocks in cascade. Also the design of the loss filter differs from the technique presented in (Rauhala et al., 2005). The method is related to the frequency sampling (Parks and Burrus, 1987) and the IFIR filter techniques (Neuvo et al., 1984). The structure of the proposed loss filter is based on a cascade of sparse

[1] also published in Lehtonen et al. (2005)

FIR filters, which are designed one after the other on subbands that are integer fractions of the audio range. Finally, the loss filter is up-sampled for implementation.

The proposed filter structure has three subfilters: The equalizer, the anti-imaging filter and the multi-ripple filter. The equalizer is responsible for the general trend of the piano tone's decay, the anti-imaging filter attenuates the image frequency responses caused by upsampling, and the multi-ripple filter is designed to fit the data as well as possible.

The purpose of the equalizer and the anti-imaging filter is basically the same as the anti-imaging filter in the IFIR technique. In the IFIR technique, it is important to eliminate the image frequency responses that result from the compression of the frequency response. In this case, it is not desired to completely get rid of the image frequency responses, since there are partials on higher frequencies as well. Instead, we need to be sure that these partials do not dominate the sound but are appropriately attenuated.

The block diagram of the Nth-order loss filter is shown in Fig. 4.1. The blocks H_1, H_2 and H_3 are the equalizer, the anti-imaging filter, and the multi-ripple filter, respectively. The system of Fig. 4.1 can be implemented in a time-reversed form, since then the delay blocks of the anti-imaging filter and the multi-ripple filter can share delay elements with the long delay line of length L in the waveguide model.

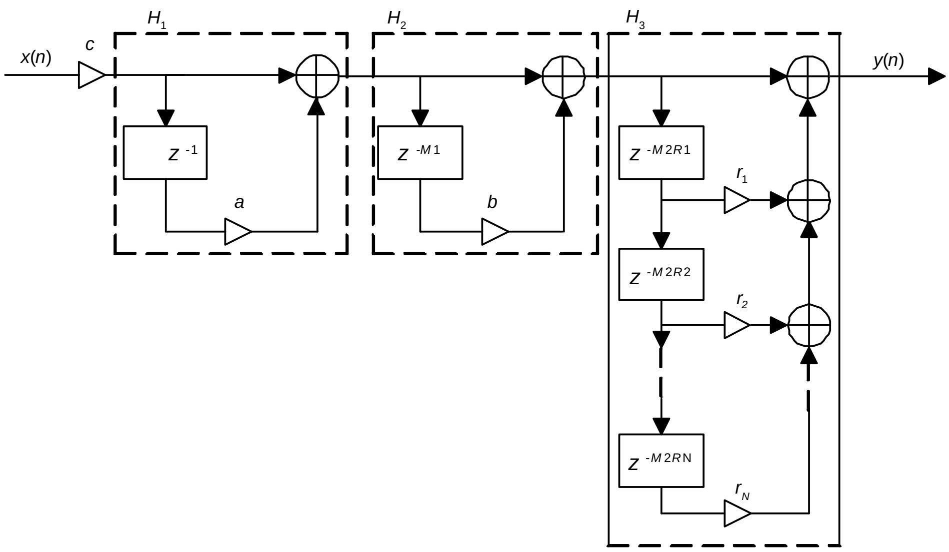

Figure 4.1: Block diagram of the Nth-order loss filter consisting of three subfilters: H_1 is the equalizer, H_2 is the anti-imaging filter, and H_3 is the multi-ripple filter.

4.1 Equalizer Design

The equalizer is designed as a first-order FIR filter in a way that the magnitude response matches two given data points, such as those at the fundamental frequency and at the Nyquist frequency. The result should imitate the general trend of the piano tone's decay rate in the band $0-22.05$ kHz. The phase is not linear, but since the changes in phase delay response are not significant, about 0.05 percent of the delay line length L with practical parameter values, this does not cause any problems. The dashed line in Fig. 4.2 presents the magnitude response of the equalizer.

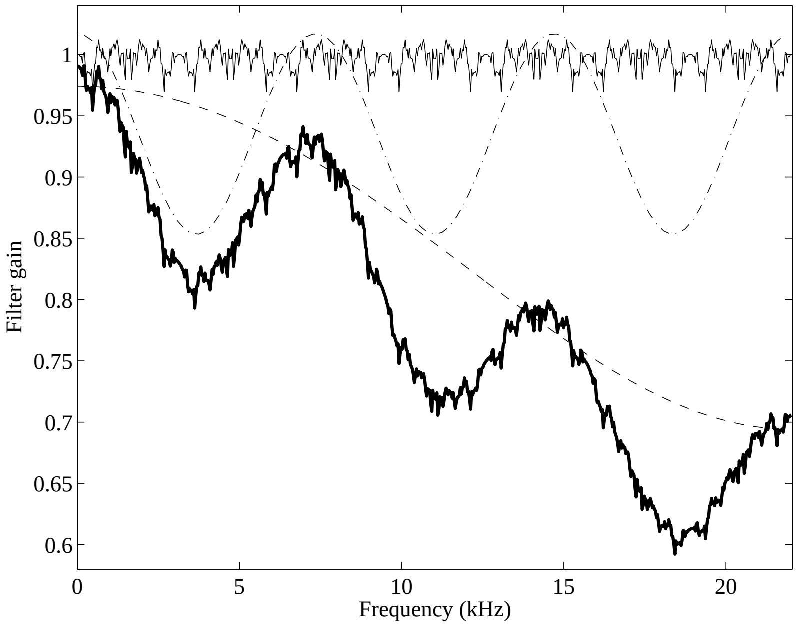

Figure 4.2: The magnitude responses of the subfilters presented in Fig. 4.1: The equalizer (dashed line), the anti-imaging filter with $M_1 = 6$ (dash-dotted line), the multi-ripple filter with $M_2 = 14$ (solid line), and the resulting loss filter (thick line).

4.2 Anti-Imaging Filter Design

The anti-imaging filter has two important tasks. Firstly, it takes care of attenuating the resulting image frequency responses and secondly, it can be used for easing the modeling of the important frequency band.

The filter is designed on a reduced frequency band in the same manner as the equalizer. It is important that the up-sampling factor M_1 is selected in a way that the anti-imaging filter attenuates the frequency band below 5 kHz appropriately. It has been found out that factor

6 is sufficient for the lowest piano tones (key index $1 - 10$) whereas for the key indices about $10-20$ factor 2 works well. In the middle range, where the harmonics cover a larger portion of the audio range, the anti-imaging filter is not essential. The resulting magnitude response is presented in Fig. 4.2 with a dash-dotted line.

4.3 Multi-Ripple Filter Design

The multi-ripple filter design process presented here results in the same multi-ripple filter structure as presented in (Rauhala et al., 2005). On the other hand, the design processes are somewhat different since the design method presented here uses standard filter design techniques.

The desired gain values can be determined from a real piano sound since the relation between the decay time constant of the kth partial and the corresponding filter gain value can be expressed with the closed form formula of Eq. 3.4.

The general trend of the obtained gain is decreasing at higher frequencies. This makes the task harder for the filter design algorithms. However, this problem can be made easier by multiplying the gain values with the inverse magnitude response values of the equalizer and the anti-imaging filter. Hence, the gain values are around one. This effect is compensated at the end with the equalizer and the anti-imaging filter.

The problem that the data covers only a small part of the audio range can be overcome by critical down-sampling. When the frequency of the largest partial is f_{max}, the up-sampling factor for critical down-sampling is $M_2 = \text{floor}\Big(44100/(2f_{max})\Big)$ and the new sampling rate is $f_s = 44100/M_2$.

The actual design method is based on frequency sampling (Parks and Burrus, 1987). First the impulse response (and thus the corresponding FIR filter coefficients) is obtained from the magnitude response by inverse discrete Fourier transform. After this, the N largest values are chosen and the other coefficients are set to zero. When N is chosen to be the exact number of the partials, it is possible to design a filter which models the given frequency response perfectly. When the impulse response is up-sampled, the result is a sparse FIR filter. In the frequency domain, the up-sampling means that $M_2 - 1$ image frequency responses follow the original frequency response. The magnitude responses of all three subfilters in the case of the exact fit at the lowest frequencies are presented in Fig. 4.2.

In the case of low filter orders, the fit in the data cannot be perfect, though, the N largest values of the impulse response fit the obtained magnitude response to the data best in the least-square sense. In the loss filter design, it is usually desired that especially the highest peaks are modeled accurately, since the gain values near the value one have the longest decay times. This fact can be taken into account in the design by emphasizing the largest gain values with a weighting function presented by Bank and Välimäki (2003).

4.4 Design Example: Perfect Match of 50 Partials

This example describes how to design a filter which matches the 50 lowest-order partials for the key index 3 ($f_0 = 30.8677$ Hz).

In the loss filter design, the case of an exact fit to the data has been considered particularly problematic. Laroche and Meillier (1994) have presented a method for designing a perfect match: When all harmonics are measured correctly, the loss filter can be implemented so that every partial is modeled with its own resonator. The computational cost of the implementation becomes large at low fundamental frequencies, when there are many partials to be modeled. The approach taken here is somewhat different and it leads to a computationally more efficient implementation.

The first phase of the design process is to determine the new, reduced sampling rate. When the target is to model 50 lowest-order partials accurately for the key index 3, the highest frequency in the data is 1543 Hz. This means that the up-sampling factor M_2 is chosen to be 14 and the new sampling rate used during the design is $44100 \text{ Hz}/14 = 3150$ Hz.

The second phase is to design the equalizer and the anti-imaging filter. The up-sampling factor M_1 for the anti-imaging filter is chosen to be 6.

In the next phase, inverse DFT is applied and the obtained impulse response is made minimum phase with Matlab's 'rceps' function (MAT, 2004). After this, 50 largest impulse response values are selected to be the FIR filter coefficients. At the end, the impulse response is up-sampled by factor 14 and a sparse FIR filter is obtained.

In Fig. 4.3 (a), it is seen that the resulted magnitude response (solid line) follows exactly the gain specification (dots). Fig. 4.3 (b) presents the corresponding T_{60}-times, i.e. the time it takes for each harmonic to decay 60 dB. The dashed line presents a response of the system without the anti-imaging filter. It can be seen that in the T_{60}-domain there are large values around 3 kHz, which do not follow the general trend of the gain specification (dots). This observation states that the usage of the anti-imaging filter is necessary especially in the case of low key indices. In the case the anti-imaging filter is used, the T_{60}-time is only about 3 seconds, and thus does not have a significant effect on the resulting sound.

4.5 Design Example: Low Order Sparse FIR Filter

When the selected filter order is low, some compromises in the fitting must be done. One solution is to smooth the original data so that a few important points are preserved. Smooth-

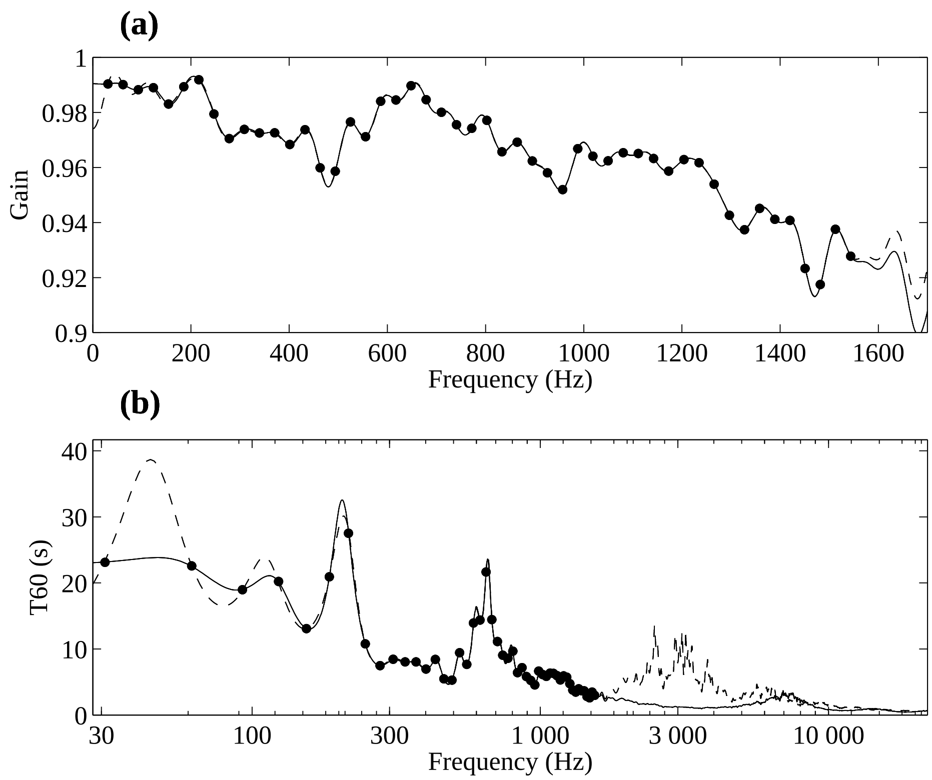

Figure 4.3: (a) The gain specifications (dots), the magnitude response (solid line) and the magnitude response without the anti-imaging filter (dashed line) of the designed system and (b) the corresponding reverberation time as a function of log frequency.

ing the data simplifies the design task significantly as most of the minor details are ignored. In the following, the data is smoothed in the same way as in (Rauhala et al., 2005): The amplitude maxima, loop gain maxima, and local maxima and minima are chosen as the points to be preserved, and these points are connected with a smooth polynomial function, which can be obtained by Matlabs' 'polyfit', 'polyder', and 'interp1' functions (MAT, 2004). The smoothing can be done also with Linear Predictive Coding (LPC), which is known to model the spectral peaks efficiently. However, neither one of the smoothing methods presented here is perfect. As the performance of the loss filter design relies heavily on the data smoothing, further investigations should be done in order to improve the expression power of this loss filter design technique.

After applying the inverse DFT, the four largest impulse response values are selected. These values can be optimized with the sparse weighted least squares technique presented in (Tarczynski and Välimäki, 1996). The optimized filter coefficients are calculated with (Parks and Burrus, 1987)

$$\boldsymbol{h} = (\boldsymbol{U}^{\boldsymbol{T}}\boldsymbol{W}\boldsymbol{U})^{-1}\boldsymbol{U}^{\boldsymbol{T}}\boldsymbol{W}\boldsymbol{H}, \tag{4.1}$$

where $\boldsymbol{U}$ is the DFT matrix (see e.g. (Mitra, 2002)), $\boldsymbol{W}$ is a diagonal weighting matrix whose kth diagonal element is w_k, and $\boldsymbol{H}$ is the target frequency response vector. The matrix $\boldsymbol{U}$ is modified in a way that only the elements corresponding to the selected impulse response values are nonzero (Tarczynski and Välimäki, 1996). The weighting function is the same as presented in (Bank and Välimäki, 2003), and it can be written as

$$w_k = (g_k - 1)^{-4}. \tag{4.2}$$

An example design is presented in Fig. 4.4 with a thick line. The up-sampling factors, the equalizer, and the anti-imaging filters of the previous design example are used, because the gain specifications are again for the key index 3.

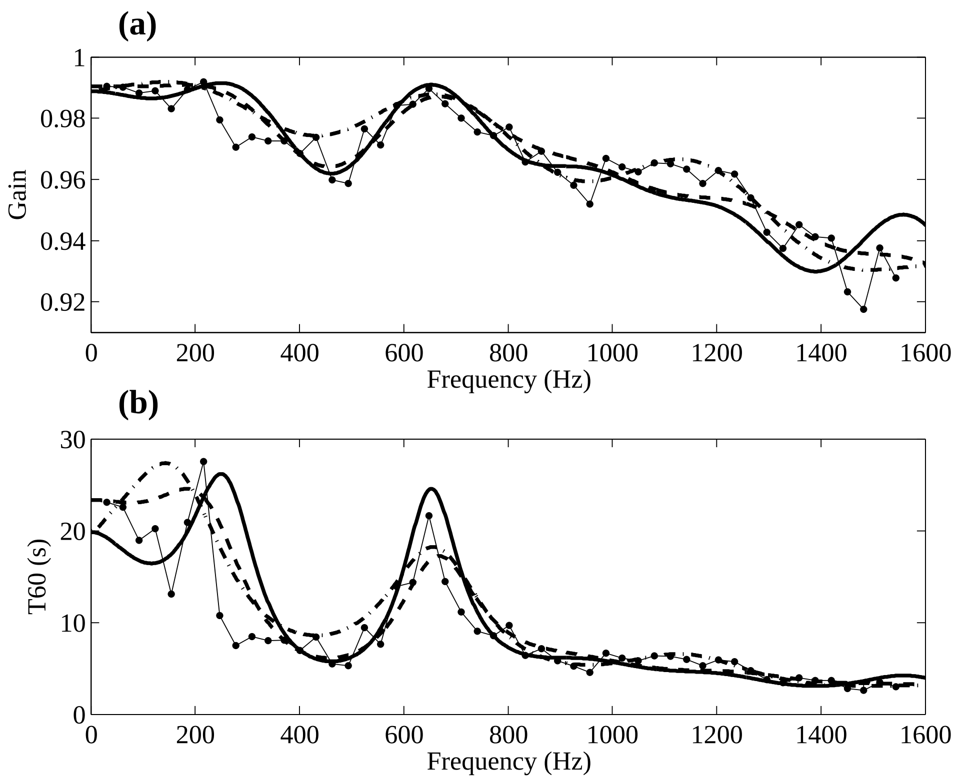

Figure 4.4: (a) The gain specification (dots) and the results of three design examples: Low-order loss filter presented here (thick), low-order loss filter presented in (Rauhala et al., 2005) (dashed), and high-order conventional filter (dash-dotted). (b) The corresponding reverberation times.

4.6 Results and Comparisons

In this section, the novel loss filter design technique is compared against the existing methods. The filters are the loss filter presented in (Rauhala et al., 2005) and a filter which is designed using Matlab's 'invfreqz' function and the weighting function in Eq. 4.2. The latter filter design process results in an IIR filter with a first order denominator and a numerator of order 201. The data is the same as used in Sections 4.4 and 4.5. The results of the comparison are presented in Fig. 4.4.

The characteristics of the proposed loss filter are shown in Fig. 4.4 with a thick line. It can be seen that the filter magnitude response follows the gain specification quite accurately at frequencies below 800 Hz. Also the highest peaks are modeled well. In the T_{60}-domain, which is more interesting from the perceptual point of view, the results are also quite sufficient.

The loss filter presented in (Rauhala et al., 2005) did also well in the comparison. Its magnitude response is presented in Fig. 4.4 with a dashed line. The high-order IIR filter is also practically as good as the two other filters presented, but the order of the filter is 40 times larger.

Obviously, the two loss filters presented here and in (Rauhala et al., 2005) are better for instrument synthesis purpose, since their computational cost is significantly smaller than the cost of the high-order filter with the same performance.

The differences of the proposed filter and the filter presented in (Rauhala et al., 2005) are minor when it comes to accuracy of the modeling. The computational costs are equally large. The major difference and improvement lies in the design. Since the filter proposed in this thesis uses standard filter design techniques, it is somewhat easier to design.

4.7 Conclusion

A new method of the loss filter design for piano synthesis purposes was presented. The proposed filter structure is based on a cascade of sparse FIR filters, which are designed one after the other on subbands and are then up-sampled for implementation. The strengths of this method are its simplicity and good performance even with low filter orders. It is also possible to design a filter, which models the given gain specification perfectly. In comparison against a traditional, non-sparse filter, the proposed loss filter performs well with a significantly smaller computational cost.

Chapter 5

Modeling of the Sustain Pedal

The importance of the sustain pedal can hardly be overestimated when it comes to professional piano performance. The decisions that the pianist makes about using the sustain pedal affects the whole style and nature of the music. At the same time, the effect is extremely complex, both from the musical and the physical point of view.

The art of using the sustain pedal is far from being an easy subject, since the proper use of the device is highly dependent on the era during which the music was composed. In the 18th century the use of the sustain pedal was considered to be mainly a special effect, whereas the piano music composed in the 19th and 20th centuries tends to be heavily pedaled. This fact, among others, needs to be taken into account when it comes to high quality piano music.

Despite the importance of the sustain pedal effect, the subject is much less studied than most other areas of the piano performance and physics. Only few studies can be found in literature (De Poli et al., 1998; Ambrosini et al., 1995), but they show interesting and beyond dispute important results. In Section 5.1.2, a brief overview of the research done in the field is given.

The use of the pedal device has primarily two purposes. First, it serves as "extra fingers" in situations where legato playing is not possible with any fingering. Second, since all strings are free to vibrate, a great enrichment of the tone is obtained.

5.1 Overview of Prior Work

In the first article that concerns the digital waveguide synthesis of the piano (Garnett, 1987), the author divides the model into five main sections: The hammer, the strings, the bridge, the soundboard and the pedals. There has been a considerable amount of work concerning the four aforementioned subjects, but the resonance pedal effect is much less studied. However, many authors have mentioned the pedal device in their articles and considered a way to model it (see e.g. Garnett (1987); Bank et al. (2003)) but very few implementations have been presented. Accordingly, there is a need for research in this particular field.

5.1.1 The Sustain Pedal Effect

The sustain pedal effect is twofold. Firstly, the strings corresponding to the depressed key continue to vibrate after the key is released. However, this does not affect the reverberation time significantly, as is shown in Section 5.2.1. Secondly, the other strings are set into vibration by sympathetic resonance. Also, the impulse that the hammer gives to the strings excites the string register (that is, the whole set of strings). This effect is clearly audible in the piano sound, especially when single tones are played.

Due to the sympathetic resonance of the strings, beating in the tone is increased. To analyze this phenomenon exactly is extremely difficult since the number of strings in a concert grand piano is nearly 250. The total amount of harmonic components in the whole system is huge, and every one of them can (at least, in theory) start to beat with one of the harmonics present in the played tone. The beating is thus by no means regular as in those cases the sustain pedal is not employed. This feature is shown in Figs. 5.5, 5.6 and 5.7.

In order to make the algorithm perform efficiently in the cases where many tones are played at the same time, these phenomena must be somehow simplified. In addition, the target is to find a simple solution which does not require any additional information other than the key index of the tone.

5.1.2 Studies Concerning the Modeling of the Sustain Pedal

Two studies about the sustain pedal modeling can be found in literature. One of them is a patent, which describes a pedal resonance effect simulation device for digital pianos (De Poli et al., 1998). The other study is given by Ambrosini et al. (1995). Both of these published solutions employ a certain number of simplified string models as a main part of the sustain pedal algorithm.

De Poli et al. (1998) presented a model with 28 ideal string models, of which 18 are fixed length and 10 are variable length. The output from these two junctions are lowpass filtered with a cutoff frequency of 1 kHz. The output from the whole system is multiplied with a coefficient, which determines the depth of the resonance pedal effect. Finally, the pedaled sound is added to the direct sound. With this device, an authentic resonance effect can be achieved.

Ambrosini et al. (1995) presented a little bit simpler model for sustain pedal simulations applied to recorded piano sounds. The effect is obtained with a bridge-string model which employs two parameters: One for controlling the decay time, that is, the loop gain, and the other for the characteristic admittance, which controls the energy leakage from the strings to the bridge and the soundboard. In this model, the 48 strings used are divided into two groups: The first set of 24 strings consists of fixed length strings simulating the “ensemble” of the strings whereas the other set consists of strings whose length is varied depending on the depressed key.

5.2 Reverberator Algorithm for Sustain Pedal Modeling

In this section, a simple but efficient algorithm for sustain pedal simulation is presented. It is based on 12 long delay lines which simulate the first 12 strings of the piano. In order to develop the algorithm further, more information about the phenomenon must be achieved. This analysis can be done by examining recorded tones and string register responses.

5.2.1 Signal Analysis

By analyzing the string register responses (that is, the response of all strings when the system is excited with the hammer) and tones with and without the sustain pedal depressed, two major differences can be found. The first one is caused by the freely vibrating strings, which are set into vibration by the hammer energy as they are not damped. The second phenomenon is noticed as increased beating in the tone, as the energy of the vibrating string set is leaking into other strings via the bridge. In addition, airborne sound has prominent effect on the excitation of the string register. In this section, the phenomenon is illustrated with the key indices 4 (note C1), 40 (note C4) and 76 (note C7), which represent examples of bass, middle and treble tones, respectively.

String Register Response

The first phenomenon can be illustrated by two examples. The first one is the impulse response of the string register system when the sustain pedal is depressed and the string group corresponding to the depressed key is damped. In the other case, the situation is similar except the sustain pedal is not depressed.

The string register responses are illustrated with 3D waterfall plots, see Figs. 5.1, 5.2 and 5.3. The plots are obtained by performing time-dependent frequency analysis with Matlabs’ ’specgram’ function (MAT, 2004) to the first second of the recorded string response (the effect of the hammer is excluded). The signals are scaled in order to avoid the possible differences in signal levels, which are due to recordings. The scaling is performed by dividing each signal by its power and setting the maximum value to 0.99. The frequency analysis is done with 256 point FFT applying a Hann window of length 64 samples with 32 samples overlap.

In the case of key index 4, the differences are shown in Fig. 5.1. In this figure, the magnitude of the first second of the string register response is presented as a function of frequency. As can be noted from Fig. 5.1 (a), after one second the overall response has decayed about 25 dB, whereas in the other case (b) the response has decayed about 40 dB within the same time. The situation is similar with the key indices 40 and 76, see Figs. 5.2 and 5.3. With these observations, it can be concluded that the reverberation time of the string register system is longer when the sustain pedal is used.

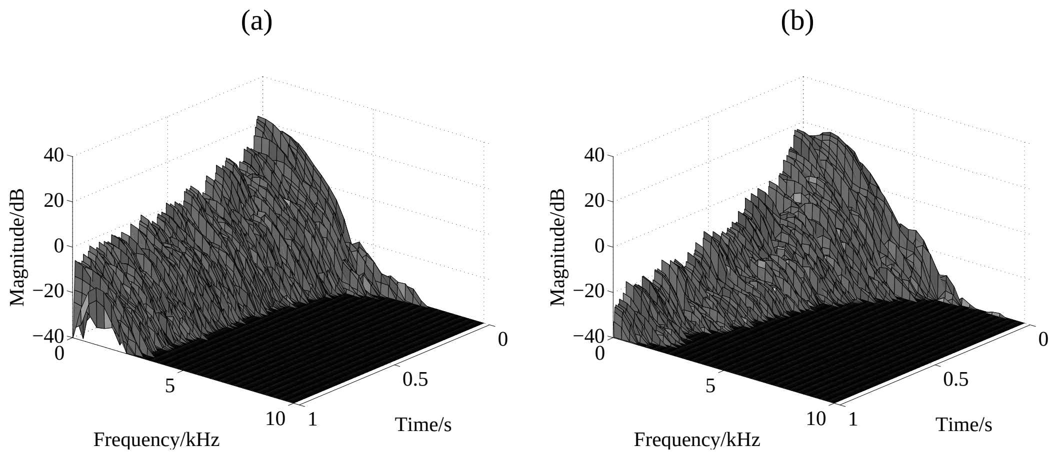

Figure 5.1: 3D plot of the string register response with (a) and without (b) the sustain pedal. The key index is 4.

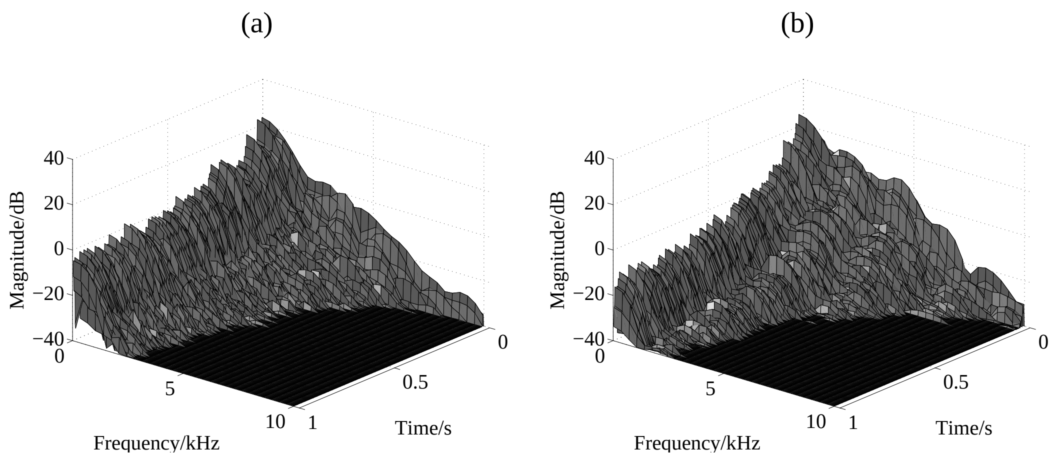

Figure 5.2: 3D plot of the string register response with (a) and without (b) the sustain pedal. The key index is 40.

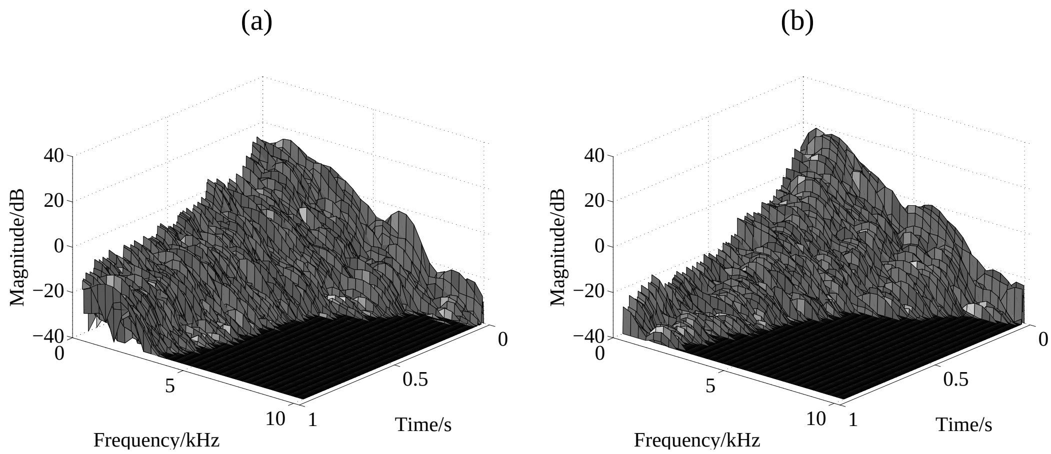

Figure 5.3: 3D plot of the string register response with (a) and without (b) the sustain pedal. The key index is 76.

Another way to illustrate the first phenomenon is to plot the energy of the string register response with and without the sustain pedal in octave bands. The energy is obtained from the spectrum of the signal by summing up the output corresponding to a certain octave band. The energy is measured from the first second of the signal, and the effect of the hammer hit is minimized by excluding it in order to obtain the energy of the true string register response. The signals are scaled in the same way in order to avoid possible differences between signal levels in the recordings. The center frequencies of the octave bands are shown in Table 5.1. The results are shown in Fig. 5.4 in the case of key indices 4 (a), 40 (b) and 76 (c). From the figure it can be seen that the energy is greater in those cases the sustain pedal is pressed down. In the cases (a) and (c), there is a clear difference between the signal energies but in (b) the phenomenon is not that clear. Generally, this comes from the input signal level adjustment during the recordings. In the case of key index 40, the levels vary significantly. The scaling process facilitates the problem, but as the signal is amplified, the background noise is also amplified. However, in general we can draw a conclusion that the energy of the string register response is greater when the sustain pedal is used.

Beating

The increased beating can be studied by separating the harmonics from the recorded signals by bandpass filtering. After a sufficient number of harmonics have been separated, the envelopes of the harmonics are calculated. In Figs. 5.5, 5.6 and 5.7 the results are shown in the case of three example tones, C1, C4 and C7, respectively. In the first case, harmonics with indices 2-7 are plotted, since the fundamental frequency was lost in the background noise. In the second case, that is, the note C4, first six harmonics are plotted. In the case

Table 5.1: Octave bands used in the string register response energy calculations.

Octave band index	Lower cutoff frequency/Hz	Center frequency/Hz	Upper cutoff frequency/Hz
1	22	31.5	44
2	44	63	88
3	88	125	177
4	177	250	355
5	355	500	710
6	710	1000	1420
7	1420	2000	2840
8	2840	4000	5680
9	5680	8000	11360

of the note C7, only two first harmonics are plotted as the others were lost in background noise.

As can be seen, the effect of the beating depends on the pitch of the depressed key. In the bass range, the differences are minor whereas in the middle range the differences are significant. Moreover, the increased beating seems to be dominating especially in the lowest harmonics. In the treble range, the beating is not prominent in either case. From these figures it can also be seen that the decay time of the tone does not depend on the usage of the sustain pedal in most cases. With some harmonics this assumption does not hold, as seen in Fig. 5.5 in the case of the fifth harmonic. This is probably because of energy leakage from some other strings that have a harmonic precisely at the same frequency. These harmonics amplify each other causing longer decay time.

To conclude, there are two significant effects resulting from the usage of the sustain pedal. The first one is a "stir" resulting from the string register system as it is excited by the hammer corresponding to the key that is played. Another important effect is the increased beating, as the energy is leaking to the string register via the bridge and the airborne sound.

5.2.2 Algorithm for Modeling the Sustain Pedal Effect

The basic idea of the reverberator algorithm for sustain pedal modeling to be presented lies on the imitation of the string register. Ambrosini et al. (1995) suggested the usage of a prominent number of string models. By using reduced amount of strings, the computational load can be facilitated. The authors suggest that the highest strings can be neglected as they have only minor influence on the sound, whereas the lowest keys contain more harmonics and are advantageous in that sense. According to the listening tests they conducted, they

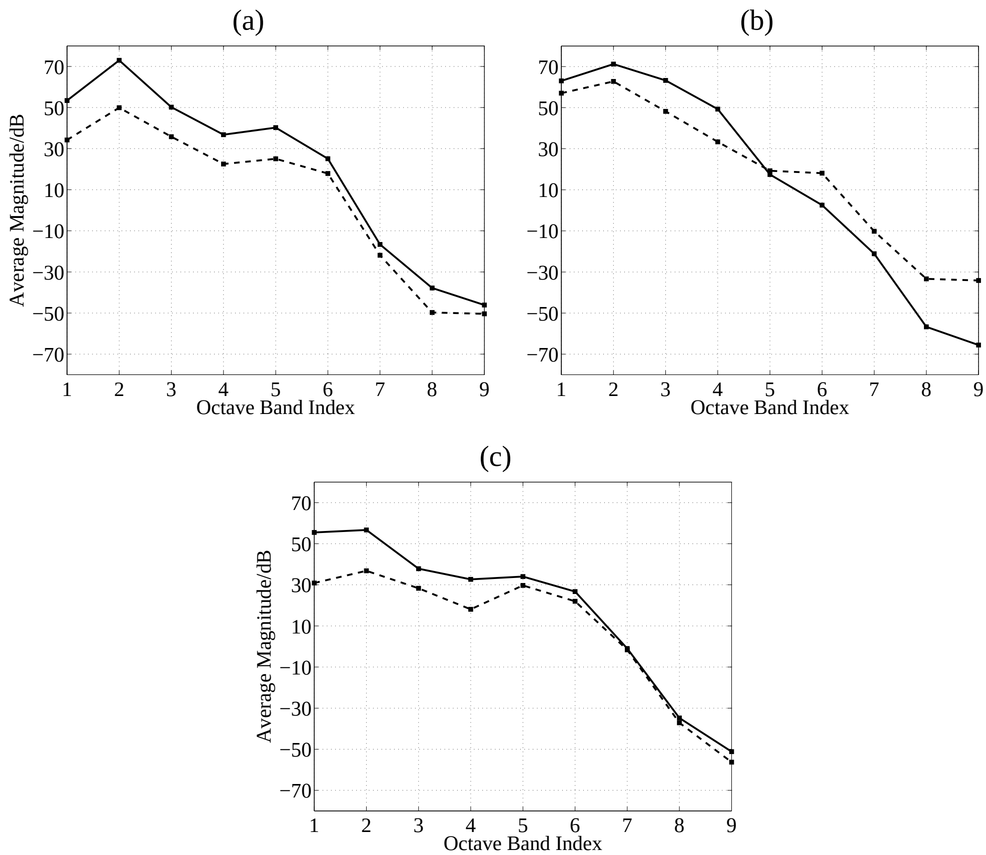

Figure 5.4: The energy of the string register response when the sustain pedal is depressed (solid line) and when the sustain pedal is not depressed (dashed line). The key indices are 4 (a), 40 (b) and 76 (c).

ended up with a model of 48 strings. However, the number of strings must be still reduced in order to get a computationally efficient algorithm.

In the model presented here, the amount of strings is determined to be 12, and their lengths correspond to 12 longest strings of the piano. By doing this, a reverberating component for every tone in the piano can be guaranteed, at least at some accuracy. In reverberation algorithm design for imitating room acoustics, usually the delay line lengths of incommensurate numbers are preferred in order to avoid unwanted coloration in the sound (Gardner, 1998). In the case of sustain pedal modeling, the objective is to design a “bad” algorithm in that sense, since some coloration and excited modes are wanted.

By inserting a comb allpass filter into the feedback loops, a diffusing effect can be brought about (Väänänen et al., 1997) in the system. In addition, Väänänen et al. (1997) pointed out that when the sum of the outputs from the strings is fed back to their inputs,

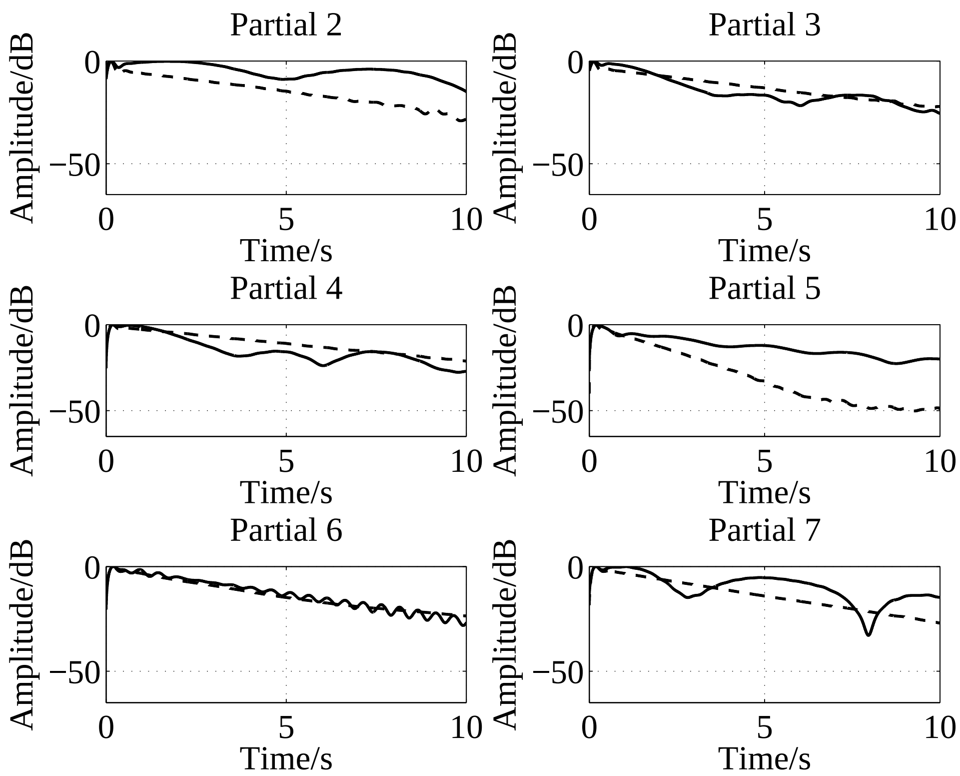

Figure 5.5: The envelopes of the harmonics 2-7 of the piano tone C1 (key index 4) when the sustain pedal is depressed (solid line) and without the sustain pedal (dashed line).

the modal density can be increased significantly. This is important, because in real situations the number of excited signal components is large, and thus a sufficient modal density must be achieved with the simulation model as well. The block diagram of the system is presented in Fig. 5.8. The choice of the parameters is discussed later.

In order to increase beating, a resonator can be added to the signal path. As noted in the previous section, the increased beating is dominating in the lowest harmonics. Bearing this in mind, it is reasonable to tune the resonator close to the fundamental frequency. If it is desired to have even more beating, several resonators can be included to the signal path. However, this increases the computational load as well. As already mentioned, the effects and the behavior of the beating phenomenon is extremely complex and calls for further studies.

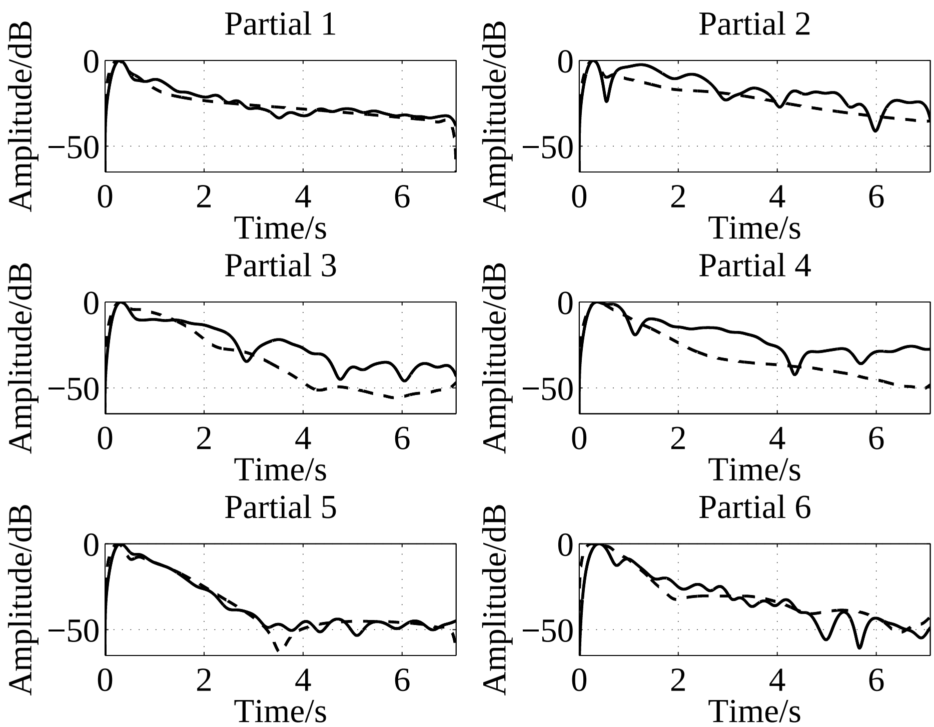

Figure 5.6: The envelopes of the first six harmonics of the piano tone C4 (key index 40) when the sustain pedal is depressed (solid line) and without the sustain pedal (dashed line).

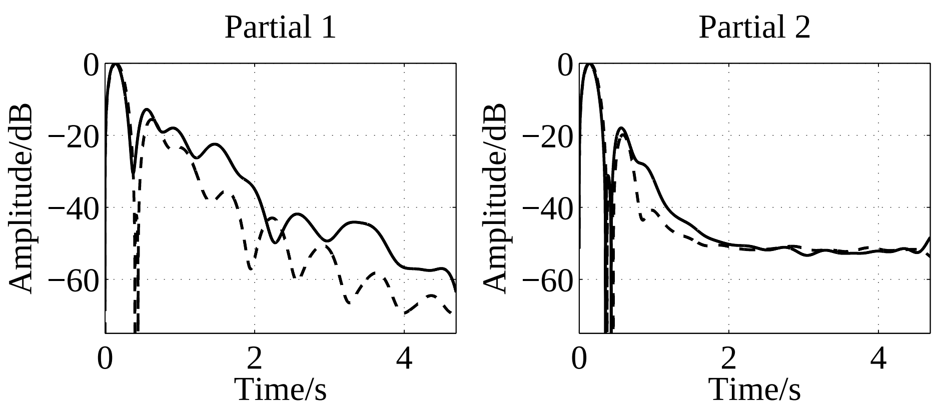

Figure 5.7: The envelopes of the first two harmonics of the piano tone C7 (key index 76) when the sustain pedal is depressed (solid line) and without the sustain pedal (dashed line).

Delay Line Lengths

The lengths of the delay lines are set to match the frequencies of the 12 lowest tones of the piano according to the formula 5.1.

$$L = \mathrm{round}\left(\frac{f_s}{f_0}\right) \tag{5.1}$$

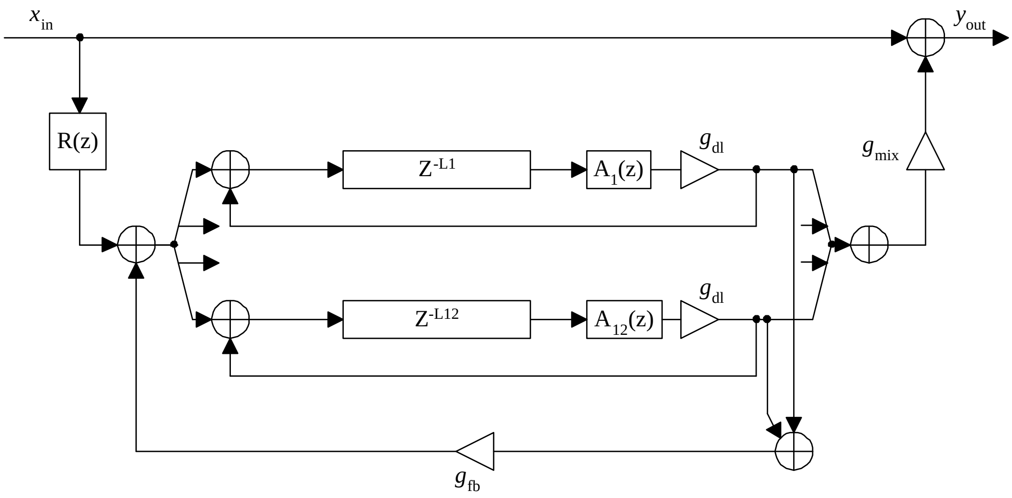

Figure 5.8: The block diagram of the sustain pedal algorithm.

where f_s is the sampling frequency and f_0 is the fundamental frequency of the tone. The rounding operation is performed, because without a fractional delay filter (see e.g. (Laakso et al., 1996) for more information about fractional delay filters) only integer delay lines can be implemented. It would be possible to insert a fractional delay filter into the model, but this does not improve the resulting sound significantly. Actually, a rough approximation is even better here, since in the real situation the resonating frequencies are likely to be only near the harmonics of the played tone.

Allpass Filter Parameters

The transfer function of the kth allpass filter can be written as

$$A_k(z) = \frac{a_{ap,k} + z^{-M_k}}{1 + a_{ap,k} z^{-M_k}}, \tag{5.2}$$

where all the filter coefficients $a_{ap,k}$ are set to 0.05 and the allpass delays M_k are chosen to be 8% of the delay lines L_k, as suggested in (Välimäki et al., 2004).

Other Parameters

The gain values g_{dl}, g_{fb} and g_{mix} in the algorithm must be chosen carefully, since they affect the stability of the algorithm and adjust the relations of the dry and reverberated sound. The coefficient g_{dl} determines the decay of the reverberated sound. In general, to guarantee the stability, the condition $|g_{dl}| < 1$ must be satisfied. In this case, the value g_{dl} was chosen to be 0.999.

Väänänen (1997) suggests that the sum of the outputs of the parallel comb filters can be multiplied with a coefficient which is a negative number whose absolute value is smaller than 1. This solution is a simplified structure of the multiple feedback delay network, originally presented by Jot and Chaigne (1991). The feedback matrix in the structure presented by Jot and Chaigne (1991) is now replaced by a single coefficient. In this case, the feedback coefficient g_{fb} was chosen to be -0.15.

The mixing coefficient g_{mix} is a very critical gain value, because it determines the relation between the dry and reverberated sound. Too large a value can make the sound unnatural while too a small value does not highlight the effect sufficiently. Practical studies concerning this coefficient show that the value of the mixing coefficient depends on the key index. In the bass range, the value is remarkably smaller than in the mid range, for example. At this point of the research process, the mixing coefficient is tuned by hand, but in the future some relation between the mixing coefficient and the key index could be derived. This can be done for example by analyzing the recordings, especially the case where the string group corresponding to the key to be pressed is damped while holding the sustain pedal down. For further description of the recordings, see Appendix A. In practice, the mixing coefficients for the example tones C1, C4 and C7 are set to be 0.4, 0.18 and 0.05, respectively.

Resonator Parameters

The purpose of the resonator is to increase the beating in the sound. It is tuned near the fundamental frequency f_0 of the tone. The coefficients for the second-order resonator $R(z)$ can be determined by formulas given in (Regalia and Mitra, 1987):

$$R(z) = \frac{0.5((1+a)+K(1-a)) + b(1+a)z^{-1} + 0.5((1+a)+K(1-a))z^{-2}}{1+b(1+a)z^{-1}+az^{-2}} \tag{5.3}$$

where

$$a = \frac{1-\tan(\frac{\Omega}{2})}{1+\tan(\frac{\Omega}{2})} \tag{5.4}$$

$$b = -\cos\theta \tag{5.5}$$

$$\theta = \frac{f_{beat}\pi}{f_s} \tag{5.6}$$

$$f_{beat} = f_0 + \gamma. \tag{5.7}$$

Ω is the notch bandwidth and θ is the center frequency of the notch in radians. K determines the gain at the center frequency. The choice of γ depends on the desired beating frequency. As in this case the aim is to increase beating, instead of modeling the beating exactly, the accurate value of the beating frequency is not critical. Examinations on this subject have shown that e.g. 0.5 Hz $< \gamma < 2$ Hz works sufficiently well.

5.2.3 Results and Analysis of the Algorithm

The proposed algorithm has been analyzed with recorded sounds by doing informal listening tests. The performance of the algorithm depends on the key index. At the moment, it works quite well for the bass and treble range. The most difficult part of the piano from the sustain pedal modeling point of view seems to be the middle range. In this range, both beating and the "stir" of the string register are clearly audible. At the same time, the quality of these effects are crucial elements for the authenticity of the synthetic sustain pedal. Especially, natural beating seems to be hard to implement with only one resonator. Increasing the amount of resonators will not solve the problem, as the computational load increases unnecessarily when many keys are depressed simultaneously, which is very a common situation when playing piano. A solution to this problem must be found in the future.

In the case of bass tones, the algorithm works sufficiently well. As a matter of fact, the differences between the recorded sounds with and without the sustain pedal are not prominent. However, there is a slight difference in nuance of these two cases, and this nuance is fairly well obtained with the proposed algorithm.

The quality of the synthetic pedal in the case of treble tones is sufficient. The stir from the string register sounds fairly authentic. The beating is not prominent in the real pedaled treble tones, and the result obtained with this sustain pedal algorithm is authentic also from this point of view.

At the moment, the algorithm does not work satisfactorily for the synthetic tones. This is probably due to the properties of the synthetic tones, as the results with recorded sounds are fairly good. However, to make these results reliable, a more formal listening test must be arranged in the future.

5.3 Conclusions

In this section an overview of an important feature of the piano sound, the sustain pedal, and a reverberator algorithm for modeling the pedal was presented. This gives a starting point to the future work, which contains some extensions to the model presented above as well as the model for the partial sustain pedal. In addition, the effect of the sustain pedal in the case of the lowest and highest tones must be analyzed more carefully in order to obtain high quality synthetic sustain pedal. At this point, the algorithm is tested mainly with recorded tones. In the future, more testing with synthetic tones must be performed. Also, a relation between the mixing coefficient and the key index must be derived.

Chapter 6

Conclusions and Future Directions

In this thesis, an overview of the piano was given. First, the acoustics of the grand piano was studied in order to give a good baseline for the physics-based modeling work. As the digital waveguide modeling is the most feasible modeling technique for plucked and struck string instruments, it was studied in more depth. The most important features of the piano sound were considered, and an overview of some modeling solutions that have been presented in literature was given. As there were two salient points in the piano sound synthesis that needed to be developed further, the solutions to these problems were selected to be the main target of this work. The first one was the modeling of the losses in the piano sound, and the other was the development of the algorithm for the sustain pedal modeling.

In Chapter 4 a novel solution for loss filter design for modeling the complicated, frequency-dependent decay of the piano tone with some practical design examples was presented. With this design technique it is possible to obtain computationally efficient solutions compared to the traditional design techniques, which usually result in high-order filters. In addition, this loss filter technique is easy to use, as it employs standard filter design techniques as a part of the solution.

In Chapter 5, the sustain pedal effect was studied. The analysis was done to the recorded string register responses and pedaled piano sounds. It was noticed that the usage of the sustain pedal affects the sound and this effect needs to be taken into account in the modeling procedure. A reverberator algorithm for the sustain pedal modeling was presented and the choice of the parameters was considered.

As a result, some important features of the piano sound were studied in more depth and modeling procedures for these features were presented. However, the presented models could be improved further, and more work related to these solutions needs to be done. This is a part of the future work.

In the future, the work related to the physical modeling of the piano will continue. There are many issues to be studied and learned about the instrument, and the synthesis model needs to be updated as new features are discovered.

Two main directions can be specified. Firstly, as the features of the usage of the sustain pedal are unknown, these must be studied more carefully, and maybe some new analysis techniques for the recorded sounds must be found. Another good option is to collaborate with professional pianists in order to study the control of the sustain pedal in more depth. This may be even a fundamental starting point, as the usage is only rarely an on-off process. How the output sound depends on the control technique is an essential question for the synthesis model. There is no point to make a sustain pedal algorithm that cannot be used in an authentic way, especially if the other parts of the model succeed in imitating the sound production mechanism of the instrument.

The second direction is to analyze how human auditory system perceives the features in the piano sound. Or, more precisely, what is important or even crucial to model and which features can be excluded in order to reduce the computational complexity of the model. Good examples are the criteria how the loss filter design should, or actually, can be modeled. For example, if the harmonic decay time analysis is performed to recorded sounds from two different instruments, it is likely that there are variations between the results obtained from the different instruments. Still, both instruments sound like piano. If the data are different, and still both cases sound good, there must be some clearance in the design limits. In this case, it would be possible to perform statistical analysis in order to determine limits for the decay times, and design filters based on these limits instead of measuring the harmonic decay times for all tones and imitate the obtained data precisely. It may simplify the model in addition that the design procedure may become more straightforward. These assumptions can be applied to other areas of the sound synthesis of the piano as well. As the perceptual issues become clearer it is likely that the physical model of the piano can be more efficient and result in more authentic sound.

Bibliography

J. M. Adrien. The missing link: Modal synthesis. In G. De Poli, A. Piccialli, and C. Roads, editors, *Representations of Musical Signals*, pages 269–297. The MIT Press, Cambridge, Massachusetts, USA, 1991.

M. Ambrosini, F. Campetella, F. Scalcon, and G. Borin. Simulazione dell'effetto del pedale di risonanze nei pianoforti digitali. In *Proceedings of the 2nd International Conference on Acoustic and Musical Research*, pages 101–106, Ferrara, Italy, May 1995. In Italian.

D. Arfib. Digital synthesis of complex spectra by means of multiplication of nonlinear distorted sine waves. *Journal of the Audio Engineering Society*, 27(10):757–768, 1979.

A. Askenfelt and E. Jansson. From touch to string vibrations. In A. Askenfelt, editor, *Five Lectures on the Acoustics of the Piano*. Kungliga Musikaliska Akademien, Stockholm, Sweden, 1990. http://www.speech.kth.se/music/5_lectures/contents.html.

B. Bank. Nonlinear interaction in the digital waveguide with the application to piano sound synthesis. In *Proceedings of the International Computer Music Conference*, pages 54–57, Berlin, Germany, September 2000a.

B. Bank. Physics-based sound synthesis of the piano. Master's thesis, Budapest University of Technology and Economics, Budapest, Hungary, May 2000b. Published as Report 54 of TKK Laboratory of Acoustics and Audio Signal Processing, available at http://www.acoustics.hut.fi/∼bbank/thesis.html.

B. Bank, F. Avanzini, G. Borin, G. De Poli, F. Fontana, and D. Rocchesso. Physically informed signal processing methods for piano sound synthesis: A research overview. *EURASIP Journal on Applied Signal Processing*, 10:941–952, 2003.

B. Bank and V. Välimäki. Robust loss filter design for digital waveguide synthesis. *IEEE Signal Processing Letters*, 10(1):18–20, 2003.

J. Bensa. *Analysis and Synthesis of Piano Sounds Using Physical and Signal Models*. PhD thesis, Université de la Méditérranée, Marseille, France, 2003. Available at http://www.lma.cnrs-mrs.fr/∼bensa/.

J. Bensa, S. Bilbao, R. Kronland-Martinet, and J. O. Smith. The simulation of piano string vibration: From physical models to finite difference schemes and digital waveguides. *Journal of the Acoustical Society of America*, 114(2):1095–1107, 2003.

J. Berthaut, M. N. Ichchou, and L. Jézéquel. Piano soundboard: Structural behaviour, numerical and experimental study in the modal range. *Applied Acoustics*, 64:1113–1136, 2003.

G. Borin, G. De Poli, and D. Rocchesso. Elimination of delay-free loops in discrete-time models of nonlinear acoustic systems. *IEEE Transactions on Speech and Audio Processing*, 8(5):597–605, 2000.

C. Cadoz, A. Luciani, and J. Florens. Responsive input devices and sound synthesis by simulation of instrumental mechanisms: The CORDIS system. *Computer Music Journal*, 8(3):60–73, 1983.

A. Chaigne and A. Askenfelt. Numerical simulations of piano strings. I. A physical model for a struck string using finite difference methods. *Journal of the Acoustical Society of America*, 95(2):1112–1118, 1994a.

A. Chaigne and A. Askenfelt. Numerical simulations of piano strings. II. Comparisons with measurements and systematic exploration of some hammer-string parameters. *Journal of the Acoustical Society of America*, 95(3):1631–1640, 1994b.

J. M. Chowning. The synthesis of complex audio spectra by means of frequency modulation. *Journal of the Audio Engineering Society*, 21(7):526–534, 1973.

H. A. Conklin. Piano design factors — their influence on tone and acoustical performance. In A. Askenfelt, editor, *Five Lectures on the Acoustics of the Piano*. Kungliga Musikaliska Akademien, Stockholm, Sweden, 1990. http://www.speech.kth.se/music/5_lectures/contents.html.

H. A. Conklin. Design and tone in the mechanoacoustic piano. Part I. Piano hammers and tonal effects. *Journal of the Acoustical Society of America*, 99(6):3286–3296, 1996a.

H. A. Conklin. Design and tone in the mechanoacoustic piano. Part II. Piano structure. *Journal of the Acoustical Society of America*, 100(2):695–708, 1996b.

G. De Poli, F. Campetella, and G. Borin. Pedal resonance effect simulation device for digital pianos. United States Patent 5,744,743, April 1998. Appl. No. 618379, filed: March 1996.

A. Fettweis. Wave digital filters. *Proceedings of the IEEE*, 74(2):270–327, 1986.

J. L. Flanagan and R. M. Golden. Phase vocoder. *The Bell System Technical Journal*, 45: 1493–1509, 1966.

H. Fletcher, E. D. Blackham, and R. Stratton. Quality of piano tones. *Journal of the Acoustical Society of America*, 34(6):749–761, 1962.

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

N. Fogwall. Modelled digital pianos, October 2004. http://www.af.lu.se/∼fogwall/piano.html.

A. Galembo and A. Askenfelt. Signal representation and estimation of spectral parameters by inharmonic comb filters with application to the piano. *IEEE Transactions on Speech and Audio Processing*, 7(2):197–203, March 1999.

W. G. Gardner. Reverberation algorithms. In M. Kahrs and K. Brandenburg, editors, *Applications of Digital Signal Processing to Audio and Acoustics*, pages 85–131. Kluwer Academic, Boston, Mass, USA, 1998.

G. E. Garnett. Modeling piano sound using waveguide digital filtering techniques. In *Proceedings of the International Computer Music Conference*, pages 89–95, Urbana, Illinois, USA, 1987.

L. Hiller and P. Ruiz. Synthesizing musical sounds by solving the wave equation for vibrating objects: Part 1. *Journal of the Audio Engineering Society*, 19(6):462–470, 1971a.

L. Hiller and P. Ruiz. Synthesizing musical sounds by solving the wave equation for vibrating objects: Part 2. *Journal of the Audio Engineering Society*, 19(7):542–551, 1971b.

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

J.-M. Jot and A. Chaigne. Digital delay networks for designing artificial reverberators. In *Proceedings of 90th AES Convention*, Paris, France, February 1991.

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

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

K. Karplus and A. Strong. Digital synthesis of plucked-string and drum timbres. *Computer Music Journal*, 7(2):43–55, 1983.

T. I. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine. Splitting the unit delay — Tools for fractional delay filter design. *IEEE Signal Processing Magazine*, 13(1):30–60, 1996.

J. Laroche and J.-L. Meillier. Multichannel excitation/filter modeling of percussive sounds with application to the piano. *IEEE Transactions on Speech and Audio Processing*, 2(2): 329–344, 1994.

M. Le Brun. Digital waveshaping synthesis. *Journal of the Audio Engineering Society*, 27 (4):250–266, 1979.

H.-M. Lehtonen, J. Rauhala, and V. Välimäki. Sparse multi-stage loss filter desing for waveguide piano synthesis. In *Proceedings of the 2005 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 331–334, New Paltz, NY, USA, October 2005.

MAT. *MATLAB version 7.0*. The Mathworks Inc., 2004.

S. K. Mitra. *Digital Signal Processing. A Computer Based Approach*. McGraw-Hill, 2 edition, 2002.

J. A. Moorer. Signal processing aspects of computer music: A survey. In J. Strawn, editor, *Digital Audio Signal Processing: An Anthology*, chapter 5, pages 149–220. William Kaufmann, Inc., 1985.

P. M. Morse and K. U. Ingard. *Theoretical Acoustics*. McGraw-Hill, 1968.

Y. Neuvo, C.-Y. Dong, and S. K. Mitra. Interpolated finite impulse response filters. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 32(3):563–570, 1984.

A. V. Oppenheim and R. W. Schafer. *Digital Signal Processing*. Prentice-Hall, New Jersey, USA, 1975.

T. W. Parks and C. S. Burrus. *Digital Filter Design*. John Wiley and Sons, Inc., 1987.

Piano World. What are the different pedals used for, 2005. http://www.pianoworld.com.

J. Rauhala, H.-M. Lehtonen, and V. Välimäki. Multi-ripple loss filter for waveguide piano synthesis. In *Proceedings of the International Computer Music Conference*, pages 729–732, Barcelona, Spain, September 2005.

P. A. Regalia and S. K. Mitra. Tunable digital frequency response equalization filters. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 35(11):118–120, 1987.

C. Roads. *The Computer Music Tutorial*. The MIT Press, Cambridge, Massachusetts, USA, 1995.

D. Rocchesso and F. Scalcon. Accurate dispersion simulation for piano strings. In *Proceedings of the Nordic Acoustical Meeting*, pages 407–414, Helsinki, Finland, June 1996.

J. O. Smith. *Techniques for Digital Filter Design and System Identification with Application to the Violin*. PhD thesis, Stanford University, California, USA, 1983.

J. O. Smith. Viewpoints on the history of digital synthesis. In *Proceedings of the International Computer Music Conference*, pages 1–10, Montreal, Canada, 1991.

J. O. Smith. Physical modeling using digital waveguides. *Computer Music Journal*, 16(4): 74–91, 1992.

J. O. Smith. Acoustic modeling using digital waveguides. In C. Roads, S. T. Pope, A. Piccialli, and G. De Poli, editors, *Musical Signal Processing*, pages 221–264. Swets and Zeitlinger, Lisse, The Netherlands, 1997.

J. O. Smith and S. A. Van Duyne. Commuted piano synthesis. In *Proceedings of the International Computer Music Conference*, pages 335–342, Banff, Canada, 1995. Available at http://www-ccrma.stanford.edu/∼jos/cs.html.

J. O. Smith III. *Digital Waveguide Modeling of Musical Instruments*. http://ccrma.stanford.edu/∼jos/pasp/, 2004.

A. Stulov. A hysteretic hammer-string interaction model for physical model synthesis. *Journal of the Acoustical Society of America*, 97(4):2577–2585, 1995.

A. Tarczynski and V. Välimäki. Modifying FIR and IIR filters for processing signals with lost samples. In *Proceedings of the IEEE Nordic Signal Processing Symposium*, pages 359–362, Espoo, Finland, September 1996.

T. Tolonen, V. Välimäki, and M. Karjalainen. *Evaluation of Modern Sound Synthesis Methods. Report no. 48.* Helsinki University of Technology, Department of Electrical and Communications Engineering, Laboratory of Acoustics and Audio Signal Processing, TKK, Espoo, 1998. Available at http://www.acoustics.hut.fi/publications/reports/sound_synth_report.pdf.

Trustees of Dartmouth College. The history of piano pedals, March 2004. http://www.dartmouth.edu/∼music33/Mus33projects/nodes/PianoHistory/pedals.html.

S. A. Van Duyne and J. O. Smith. A simplified approach to modeling dispersion caused by stiffness in string and plates. In *Proceedings of the International Computer Music Conference*, pages 407–410, Århus, Denmark, September 1994.

S. A. Van Duyne and J. O. Smith. Developments for the commuted piano. In *Proceedings of the International Computer Music Conference*, pages 319–326, Banff, Canada, 1995. Available at http://www-ccrma.stanford.edu/∼jos/cs.html.

V. Välimäki, J. Pakarinen, C. Erkut, and M. Karjalainen. Discrete-time modelling of musical instruments. *Reports on Progress in Physics*, 69(1):1–78, 2006.

V. Välimäki, H. Penttinen, J. Knif, M. Laurson, and C. Erkut. Sound synthesis of the harpsichord using a computationally efficient physical model. *EURASIP Journal on Applied Signal Processing*, 7:934–948, 2004.

R. Väänänen. Efficient modeling and simulation of room reverberation. Master's thesis, Helsinki University of Technology, 1997. Available at http://www.acoustics.hut.fi/publications/files/theses/vaananen_mst.pdf.

R. Väänänen, V. Välimäki, J. Huopaniemi, and M. Karjalainen. Efficient and parametric reverberator for room acoustics modeling. In *Proceedings of the International Computer Music Conference*, pages 200–203, Thessaloniki, Greece, September 1997.

G. Weinreich. Coupled piano strings. *Journal of the Acoustical Society of America*, 62(6): 1474–1484, December 1977.

K. Wogram and T. Mori. Is a grand piano always better than an upright piano. In *146s ASA Meeting*, Austin, Texas, November 2003. http://www.acoustics.org/press/146th/Wogram.htm.

Appendix A

Recordings

Basics

– The grand piano: Steinway & Sons Mod. C Sn. 559907

– The upright pianos: Yamaha Mod. U3AS Sn. 4558497 and Yamaha Hybrid piano, Mod. MPX1Z Sn. 5895823

– Place: Espoo Music School, Espoo, Finland

– Dates: February 23, 2005 (Session 1), and August 19, 2005 (Session 2)

Recording equipment

- Microphones: AKG C 480 B
- Amplifiers: Behringer Eurorack MX602A
- PC
- Sound card: Digigram VXpocket v2
- Impulse hammer

Description of the recordings

The recording sessions were both performed in Espoo Music School in the class Beethoven, which is their biggest rehearsal class. In the first recording session, the sounds of Steinway & Sons grand piano and Yamaha upright piano were recorded whereas in the second recording session, sound of Yamaha hybrid upright piano was also recorded. In all cases, the sounds were recorded with two channels. An example of the microphone positioning is given in Fig. A.1. The figure is from the second recording session, where the microphones are positioned 32.5 cm away from the strings of the grand piano, and the microphone corresponding to the left channel is in 60 ° angle and the microphone corresponding to the right channel is in 90 ° angle. In the case of the upright piano, the microphone positionings are shown in Fig. A.2. Both microphones are positioned 35 cm away from the soundboard. In the first recording session, the setups were similar, except in the case of the upright piano, where the microphone corresponding to the left channel was positioned above the instrument. However, this setup proved to give insufficient results, as the sound was colored in a different way when listening the channels separately, and the microphone positions were changed in the second session.

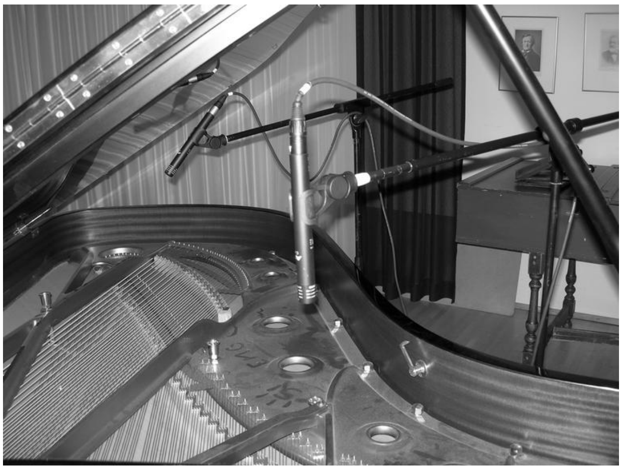

Figure A.1: Microphone positioning in the case of the grand piano in the second recording session.

Performed Measurements: Recording session 1

1. Soundboard and string responses (Grand piano and upright piano)

   - Air radiation response: The bridge was hit with the impulse hammer in the bass, middle and treble range, 3-5 times/position. The recordings were carried out with and without the sustain pedal. The air radiation response recording is illustrated in Fig. A.3. The impulse hammer is hit to the bass end of the bridge.

2. Dynamics (Grand piano and uprigth piano)

   - Selected notes (C1, C4, C7) with four different dynamics: piano (*p*), mezzoforte (*mf*), forte (*f*), forte fortissimo (*fff*). Two cases: Long and short playing styles.

3. Pedal (Grand piano and upright piano)

Figure A.2: Microphone positioning in the case of the upright piano in the second recording session.

Figure A.3: Recording the soundboard response. The bass end of the bridge is hit with the impulse hammer.

- Selected notes (C1, C4, C7) with sustain pedal depressed with four different dynamics (*p*, *mf*, *f*, *fff*).
- Selected notes (C1, C4, C7) with half sustain pedal depressed with four different dynamics (*p*, *mf*, *f*, *fff*).

4. Staccato

- Grand piano: Selected notes (C1, C4, C7) with staccato with four different dynamics (*p*, *mf*, *f*, *fff*). Three cases: Without, with and with half sustain pedal.

- Upright piano: Selected notes (C1, C4, C7) with staccato with four different dynamics (*p, mf, f, fff*). Two cases: Without and with half sustain pedal. In the case 'full sustain pedal', only one dynamic level (*fff*) was played.

5. Chords, scales and musical pieces (Grand piano and upright piano)

   - Chords: C major, E minor, G major and B major. Left hand plays an octave of the fundamental note and right hand plays the full chord: Octave of the fundamental note, third, and quint. Two cases: With and without the sustain pedal.
   - Chromatic scale through the whole keyboard.
   - Musical piece: Jean Sibelius: Neilikka, Op. 85 Nr. 2.

6. Hammer (Grand piano and upright piano)

   - Selected notes (C1, C4, C7) with the corresponding string group damped with four different dymanics (*p, mf, f, fff*).

7. Dispersion (Grand piano and upright piano)

   - All notes played with one dynamic level (*mf*).

8. Special cases (Grand piano) (Fogwall, 2004)

   - Damper effect: The sustain pedal was depressed, after that a loud chord was played, the chord was released, but the pedal was held down.
   - While the sustain pedal was held down, the chord was silently pressed down. After a while the pedal was released, but the chord was still held down.
   - Loud chords were played and immediately after releasing the keys, the sustain pedal was depressed.
   - Sympathetic resonance: An up-down scale was played while holding the sustain pedal down.
   - String resonance: A chord was layed down without the hammer touching the strings and some keys were pressed and released.

Performed Measurements: Recording session 2

1. Pedal, Part I

   - (Grand piano and upright piano) Selected notes (C1, C4, C7) played with four different dynamics (*p, mf, f, fff*) with and without the sustain pedal.

- (Grand piano) Selected notes (A0, D1, G1, C2, F2, B2, ES2, AS2, CIS4, F4, H4, E5, A5, D6, GIS6, C7, F7, H7) played with and without the sustain pedal.

2. Pedal, Part II: Special Cases

   - (Grand piano and upright piano) Selected notes (C and G from every octave) played while the sustain pedal was depressed and the string group corresponding to the key was damped.
   - (Grand piano) Selected notes (G2, E3, C4, G5, E6) played in three different cases: Simulated sustain pedal with 12 lowest keys silently depressed, without the sustain pedal and with the sustain pedal.
   - (Grand piano) Selected notes (G2, E3, C4, E5, E6) played with long and short playing style when three nearest neighbor from the both sides were silently pressed down.
   - (Grand piano) Selected notes (C3, C4, C5) played when all other C-notes were silently depressed.
   - (Grand piano and upright piano) Selected notes (C and G from every octave) played while the sustain pedal was depressed and the string group corresponding to the key was damped right after the key was depressed.
   - (Grand piano) The string response was recorded while pressing the sustain pedal down quietly and heavily.

3. Pedal, Part III: Chords and Scales

   - (Grand piano) Selected chords (C-major, E-minor, G-major and B-minor) were played with and without the sustain pedal in bass, middle and treble range of the piano.
   - (Grand piano) Chromatic scale through the whole keyboard.

4. Staccato

   - (Grand piano) Selected notes (C1, C4, C7) with staccato with four different dynamics (*p*, *mf*, *f*, *fff*). The sustain pedal was not employed.

5. Hammer

   - (Grand piano) Selected notes (C1, C4, C7) with the corresponding string group damped with four different dymanics (*p*, *mf*, *f*, *fff*).

- (Grand piano) Air radiation response: The bridge was hit with the impulse hammer in the bass, middle and treble range, 3-5 times/position. The recordings were carried out with and without the sustain pedal.