Digital audio antiquing – Signal processing methods for imitating the sound quality of historical recordings

Article *in* Journal of the Audio Engineering Society · March 2008

Vesa Välimäki
Aalto University

Ossi Kimmelma
nLIGHT

Jukka Parviainen
Aalto University

Digital Audio Antiquing—Signal Processing Methods for Imitating the Sound Quality of Historical Recordings

VESA VÄLIMÄKI,[1] *AES Member*, **SIRA GONZÁLEZ,**[1] **OSSI KIMMELMA,**[2] **AND JUKKA PARVIAINEN**[3]

[1]*Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Espoo, Finland*

[2]*Helsinki University of Technology, Micro and Nanosciences Laboratory, Espoo, Finland*

[3]*Helsinki University of Technology, Laboratory of Computer and Information Science, Espoo, Finland*

vesa.valimaki@tkk.fi, sigonzal@cc.hut.fi, ossi.kimmelma@tkk.fi, parvi@cis.hut.fi

Digital signal processing methods to modify sound files to appear aged by imitating historical disturbances are proposed. The opposite activity, audio restoration, has been cultivated during the past few decades to improve the sound quality of old recordings. Case studies of audio antiquing to render a music file to sound like a phonograph, gramophone, and LP recording are presented.

0 INTRODUCTION

Since the sound waves were first stored on a piece of paper using the Phonautograph in 1857, the evolution of recording technology has reached unthinkable levels. The sound quality of music recordings has improved much along the way. Meanwhile, the recorded music has been undergoing a constant transformation, which has attempted to get the most out of the possibilities offered by technology. During the last few decades, much work has been devoted to restoring old recordings [1]. The goal has been to improve the sound quality of historical and nostalgic material to meet the ever-increasing expectations of the audience.

In this paper, we discuss the opposite of audio restoration, which we call audio antiquing. There has been some interest in the simulation of the degradations in audio signals previously for various purposes. In fact, several such methods have been developed and used for testing audio restoration algorithms. By adding disturbances to an original error-free sound file, it has been possible to make objective comparisons between the original and the processed signal, see, e.g., [2], [3], and [4]. Skritek has reported a transmission channel simulator for audio signals for testing noise reduction techniques [5]. Speaker auralization, a method introduced by Klippel, refers to real-time simulation of a loudspeaker output signal to demonstrate large-signal distortion [6]. Zacharov has reported on a virtual audio prototyping project at Nokia to simulate the sound output of a mobile phone including degradations caused by coding and other processing [7].

Other related themes of research include software simulation of traditional electroacoustic systems and effects. Recent examples of such studies include the modeling of guitar amplifiers [8], [9] and effects pedals [10], the rotating Leslie speaker [11], an analog phaser [12], the reverberation plate [13], and analog music synthesizers [14], [15].

The closest prior work that we have found is a free plug-in called iZotope Vinyl, which simulates the degradations of an LP recording [16]. The company does not disclose the details of how the plug-in operates. However, according to their representative, they use some of the same ideas that are discussed in this paper, but many of the signal processing techniques presented here are very different from their approach [17]. It appears that different noise sources are simulated separately in iZotope Vinyl, since for example the mechanical and electrical noise can be adjusted in the user interface.

The initiative for the work presented in this paper was made by museums. In 2004, the Heureka Science Center in Vantaa, Finland, started planning a “Music Exhibition”, where they wanted to present new technical demonstrations related to musical sound and music technology. One idea was to highlight how music listening has changed over the last hundred years or so. The following concept was made up: exhibit the history of recording by playing the same musical excerpt in the phonograph, the gramophone, the LP, and the CD quality. This way, the audience could listen and thus easily grasp the progressive evolution made from one technique to another. Two different quality versions of all songs were also produced: a new, high-quality one and a time-worn, low-quality one. The original song was taken from a flawless CD and seven different versions were then produced using signal processing methods. This demonstration was open for public for one year at the Heureka Science Center in 2005–2006. In 2006, the Sibelius Museum in Turku, Finland, ordered a similar set of different quality versions of a piece of music. For these two projects, several signal processing techniques were designed to reproduce the degradations that appear in old recordings. This paper summarizes the results of these two projects. Sound examples of audio antiquing are available online [18].

This paper is organized as follows. In Section 1, an overview of the history of recordings and of the different degradations present in old recordings is presented. Section 2 explains the simulation of degradations, and in Section 3 the parameters used in example audio files are presented. The conclusions and directions for future work are given in Section 4.

1 DEGRADATIONS IN HISTORICAL MUSIC RECORDINGS

In 1877 the first useful device for storing and reproducing sound was invented, the phonograph [19]. Created by Thomas Edison, this device was able to store sound in a grooved tin foil rotating cylinder using a steel point stylus which cut vertically into it. An inverted horn, whose final part was closed by a metallic diaphragm, transmitted the movements caused by the sound to the stylus. For the reproduction, a system similar to the storing one in reverse order was used. The needle was placed at the beginning of the recorded groove and when turning the cylinder, the oscillations of the needle arouse vibrations, which were turned into sound waves amplified and channeled by a horn.

The gramophone invented by Emile Berliner in 1888 solved some of the problems the phonograph had [19]. The gramophone has a flat disc record instead of a cylinder, using an acid etching process to cut grooves into the surface of a polished zinc plate. Berliner discovered a method for creating a metal negative master disc from which hundreds of shellac discs were pressed, which lead to easy mass production of recordings. Moreover, the flat disc allowed more precise mechanic arrangement as the discs were controlled by the fitting of the center hole. The sound was registered in the spiral grooves of the flat record, but instead of using vertical movements of the needle tip, lateral movements were used, making the friction between the needle and the groove minimized and the membrane vibrations more precise.

The disc recordings were much improved during the next several decades, and in 1948 the first vinyl LP, a 33 rpm (revolutions per minute) long playing record, was introduced. At the same time, the seven-inch 45 rpm discs were introduced for popular music. These together replaced the 78 rpm disc, although at the end the 33 rpm long playing records were the predominant ones. The CD (Compact Disc) was developed by Sony and Philips since the late 1970s and it was available on the market in late 1982. The stereo sound is recorded on a plastic disc in which the digital information is stored as a series of tiny indentations read by a small laser.

Many different kinds of degradations can be frequently heard in LP, gramophone, and phonograph recordings. They can be divided into two general groups: global and localized degradations [1]. Global degradations affect all samples of the waveform and contain effects such as background noise, wow and flutter, and certain types of nonlinear distortion. Localized degradations are discontinuities which affect only certain samples of the waveform [1].

1.1 Global Degradations

Global degradations can be classified into frequency response deviation, limitation of the signal bandwidth, dynamic range, distortion, pitch variation defects, and hiss. The frequency response deviation refers to the imperfect frequency response of the loudspeaker or the horn used in music reproduction. In audio, a high-quality system should have an approximately flat frequency response in the audio range from about 20 Hz to about 20 kHz. Although the quality of high-quality loudspeakers can be quite good, the same cannot be said about the horns used for the gramophone or phonograph music reproduction [20]. The emphasis and attenuation that some frequencies gain with these horns gives the output sound a characteristic tone.

Another important feature of the recording technology is the signal bandwidth it can store and reproduce. Nowadays the CD allows the storage of all the frequencies below 20,000 Hz, which corresponds to the full range of human hearing. Nevertheless, in old recordings the bandwidth of the stored audio signal was narrower, with the lowest and the highest frequencies being completely lost. It is difficult to notice the differences between a CD and an LP, in which the highest stored frequency is about 15 kHz, because of our limited sensitivity to high frequencies. However, listening to an older recording device, like a phonograph or a gramophone, where the bandwidth is much narrower, the loss of high and low frequencies is clearly noticeable. This restriction of the bandwidth in the analog recordings (LP, gramophone, and phonograph) is set by the size and speed of the record and the size and shape of the stylus used for recording and reproduction process. Moreover, the frequency response of the media is degraded by frequent playback, if the cartridge is set to the track too heavily.

The dynamic range is defined as the ratio of the maximum signal value to the softest one or to the noise level. The dynamic range of human hearing is about 120 dB, but less can be sufficient for a good quality (the CD system has the theoretical dynamic range of about 96 dB). In the case of old recordings, the noise level is higher than in the CD and it gets worse with time. For example, the best dynamic range for a gramophone disc can be around 70 dB [22].

Distortion is a different kind of global degradation. It includes nonlinear defects, such as amplitude related overload (e.g., clipping) or groove wall deformation and tracing distortion [1]. There are multiple sources from which the distortion can stem from. The stylus used in the cutting and reproduction process is an important part for the audible distortion. In this way, the distortion caused by the different stylus can be classified into cutting distortion, tracing distortion, or tracking error [23]. Overload, a kind of cutting distortion, has important effects in the final sound quality and it is made when the stylus can no longer trace the modulation in the groove [21].

In historical recordings, pitch variation defects, which were not part of the original music performance, can be found. Wow refers to slow shifts in pitch (i.e., playback speed) while flutter refers to faster fluctuations. These fluctuations are usually periodic in both cases. They can occur during the recording or playback process, leading to undesirable changes of all frequency components [24]. A possible mechanism by which wow occurs is the variation of the rotational speed of the recording medium [25]. Another cause can be the eccentricity during the disc and cylinder sound recordings or when copying or reproducing, a wrongly punched gramophone disc being a typical example of this kind of degradation [1]. The audible effect of wow and flutter is very noticeable on signals which contain pure tones, where the threshold of audibility for speed variation is about 0.5%. For musical signals, the threshold of audibility is between about 0.1% and 0.5% depending on the type of material [26].

Lastly, when listening to an analog system, a characteristic kind of noise, usually perceived as ‘hiss’, is always present, in a more or less noticeable way. This random additive background noise stems from different sources: electrical circuit noise, irregularities in the storage medium, and ambient noise from the recording environment [1]. Despite its different origins, this degradation can be considered as one single noise process, even though the ambient noise from the recording environment could be considered part of the original audio recording [1]. The random nature of this noise makes it be present in all the frequencies.

1.2 Localized Degradations

Localized degradations are those which only affect certain samples of the waveform: tracking errors, clicks, and low-frequency pulses (thumps). Tracking errors occur when a large discontinuity in the played groove makes the needle jump to an adjacent groove. If the needle goes back to a previous groove, there is the possibility that a short part of a song is repeated many times until the needle is able to continue by itself or external help is supplied to stop the repetition. The length of the repeated part is naturally related to the rotational speed of the device. In case the needle continues playing the next groove, a part of the audio signal will be missed.

One of the most common problems in historical music recordings are clicks. They are impulsive disturbances random in time and amplitude and their duration is usually less then 1 ms. Although this degradation can be quite irritating, it usually affects less than approximately 10% of the audio samples, which facilitates a successful restoration [27]. The sources by which clicks occur can be diverse. The most common reasons in analog disc recordings are the specks of dirt and dust adhered to the grooves (Fig. 1) and granularity in the material used for pressing the disc. Small scratches in the surface are common, too. An example of a click-degraded signal is shown in Fig. 2. The sharp positive and negative peaks correspond to audible clicks.

When severe damage is done on the groove walls of a disc or cylinder, the effect is the degradation known as 'low-frequency pulses' or 'thumps', one of the most annoying disturbances in historical musical recordings. Thumps are related to deep scratches on the disc surface or the joints resulting when parts of a broken disc are fixed with glue. During the reproduction of the musical recording, the stylus-arm is excited by these discontinuities and the slowly fluctuating impulse response of the arm is additively superimposed upon to the undistorted audio signal (see Fig. 3) [1]. In this way, this type of distortion can be generally described by a short and strong discontinuity similar to a click of duration shorter than 2 ms followed by a long and decaying transient of low frequency content, which normally lasts longer than 50 ms [4].

The various types of degradations that appear in historical recordings are summarized in Table 1.

2 SIMULATION OF DEGRADATIONS

The distinct types of degradations can be implemented using different techniques. In the following, the most suitable simple method that we have found is described for each of them.

2.1 From Stereo to Mono

In order to reproduce the quality of the old recordings from the era before the stereophonic sound, the first step in audio antiquing is to convert the music signal into the monophonic form. Even though some old disc recordings were stereo, the earliest ones (even for the LPs) were monophonic.

Various issues must be taken into account when a music file is reduced from stereo to mono. The two available channels allow the use of audio effects, which are fairly popular in recordings. These effects include the inversion of the phase in one of the channels or the introduction of a delay between the channels. In these cases, the conversion from stereo to mono should be done carefully, knowing where the effects are and deducing in each case the best method not to lose information (not more than strictly necessary). For example, a way for the detection of a delayed channel could be done with the evaluation of the peak in the correlation of the two signals. Nowadays, most of the music is produced "mono-compatibly" to avoid the problems that some devices with only one speaker (e.g., a simple radio) would have with its reproduction. However, in general the stereo-to-mono conversion cannot be considered a solved process, since in some cases part of the audio signal can be suppressed or colored, when the two channels are combined [28].

The most broadly extended method is working out the average of the two channels, which is an effortless way to implement the stereo-to-mono conversion. It yields good results in a high percentage of cases.

2.2 Frequency Response Deviation

The frequency response of the horn used in gramophone and phonograph recordings reproduction has an important role in the final sound. There are programs available to compute the frequency response of an acoustic horn, such as ‘Hornresp’ and ‘AJ-Horn’. ‘Hornresp’ was used for the simulation [29]. The main reasons for this election were that it was free and easy to use. Having the measurements of the desired horn and the place in which it would be positioned (radiating into a free space, a half space, a quarter space or an eighth space) the acoustical impedance of the horn can be calculated by this program [29].

Once the impedance is estimated, the next step is to work out the transmission coefficient to know how much sound will be transmitted through the horn at different frequencies. From [30] the formula to calculate the transmission coefficient is

$$t = 1 + r = \frac{2Z(f)}{Z(f) + Z_0} \tag{1}$$

where r is the reflection coefficient, $Z(f)$ is the acoustical impedance of the horn at frequency f, and Z_0 is the characteristic impedance of the pipe to which the horn is attached, expressed as

$$Z_0 = \frac{\rho c}{S} \tag{2}$$

where ρ is the density of the medium inside the pipe (in this case air), c the velocity of sound, and S the cross-sectional area of the pipe. The formula for a lossless cylindrical pipe has been chosen, since it simplifies the calculation yet providing sufficiently reliable results.

When the transmission coefficient of the horn is known, i.e., when the frequency response has been calculated, it is implemented as a digital filter. The ‘Hornresp’ program gives the output data on a logarithmic frequency scale. A technique for calculating FIR filter coefficients from the frequency response is to use the IFFT (Inverse Fast Fourier Transformation). In this case, the frequency response data obtained are non-uniformly spaced, so interpolation is required to obtain uniformly spaced values in frequency.

In Fig. 4 the steps for obtaining the frequency response are shown. Starting from a schematic diagram of the horn, its impedance is calculated by 'Hornresp', and finally the transmission coefficient is worked out with the explained formulas. The FIR filter whose magnitude response approximates the transmission coefficients can be used for antiquing music signals.

2.3 Signal Bandwidth

The signal bandwidth of an audio signal recording represents the range of frequencies, which the recording device can store and reproduce. To achieve the desired limited bandwidth, lowpass or bandpass filters can be used to suppress the appropriate frequency components.

To implement this degradation two different Butterworth filters are used. Although the order of this kind of filter is higher than the one obtained for Chebychev or Cauer filters, the smoothness it provides is worthy for the final purpose. Moreover, since the audio antiquing is an offline process, there is no need for saving time in the signal processing, unless the processing takes an unreasonable time.

The first Butterworth filter is used at the beginning of the process, just after the stereo-to-mono conversion. It is either a lowpass or a bandpass filter, depending on the historical recording to be simulated. The second filter, a lowpass filter in all cases, is used in the final steps, to reduce the high-frequency components introduced by the different simulated degradations, which would be impossible in old recordings.

2.4 Distortion

Distortion can be caused for various reasons in a recording. Some of the mechanisms are the tracing distortion on stereo vinyl disks, deformation of the groove walls of a disk, nonlinear behavior of the pickup of the player, and amplifier nonlinearities [31]. We find that detailed modeling of these several distortion mechanisms would be unnecessarily complicated, since the end results are clear: usually harmonic distortion and the clipping of the waveform at large signal levels. One way to simulate the nonlinear distortion is to use two different functions: one to create the nonlinearity for the loud passages and another for the soft ones.

The hyperbolic tangent is used for the first purpose because of its linearity for low values of the signal and its saturation at high signal values. By introducing a base parameter B_{loud} and by normalizing the function, the amount of distortion in the output signal $y(n)$ can be controlled in the following way:

$$y = \frac{\tanh(x(n)B_{\mathrm{loud}})}{\tanh(B_{\mathrm{loud}})} \tag{3}$$

where $x(n)$ is the input signal. For the soft passages Eq. (4) is used. The higher the B_{soft}, the more distortion is introduced (see Fig. 5):

$$y = \operatorname{sign}(x(n))\operatorname{abs}(x(n))^{B_{\text{soft}}} \tag{4}$$

This technique for creating the distortion has acceptable results and the hard studies needed for having reliable data about the distortion sources and functions are avoided. The combination of the two non-linearities and its effect over a sinusoidal signal are presented in Fig. 6. Distorted musical examples affecting soft (B_{soft} = 2) and loud (B_{loud} = 5) parts of a recording and both the soft and the loud parts (B_{soft} = 2, B_{loud} = 5) are available online [18].

When a large amount of harmonic distortion is generated with the above hyperbolic tangent waveshaper by allowing the signal waveform to be clipped, the appearing new spectral components can lead to aliasing. This is a known problem in nonlinear audio signal processing, for example in tube amplifier modeling, where the usual solution is to oversample the signal by factor 2 to 8 [8], [9]. In audio antiquing, heavy distortion can be used in gramophone or phonograph simulation, but in these cases the music signal is lowpass (or bandpass) filtered before distorting it. Since this suppresses aliasing in a similar way as oversampling, we have found it unnecessary to increase the sampling rate in our simulations.

2.5 Wow and Flutter

The process of introducing wow or flutter into a music file can be divided into two fundamental parts: creation of the time warping function and resampling of the audio waveform. In resampling, the same signal processing techniques can be used as in the restoration of pitch variation defects, such as a time-varying FIR filter with coefficients taken from a long truncated sinc function, as proposed in [2], by using high-order Lagrange interpolation [32], for example. In the following we focus on the time warping function.

The first item to account for when creating the time warping function is the nature of the type of degradation in disc and cylinder recordings. In these cases the distortion is usually periodic and has smooth variation with time. Without taking into account the noise, the frequency components of the audio signal can be expressed as [1]:

$$F = p(n)F_0 \tag{5}$$

where F_0 is the centre frequency, $p(n)$ the time-varying pitch variation factor with sample index n, and F is the final frequency.

It is possible to derive a physically based model for the wow caused by an eccentric disc or a cylinder. Consider a disc with radius b that is fixed at its center point O and that rotates with constant angular speed ϕ, as illustrated in Fig. 7. The samples s_i obtained from the output waveform of the disc correspond to the points that are uniformly spaced along the disc, as shown by short lines between points B and C. Consider then the case in which the center point is displaced by a. The displacement causes the wow effect because sequence s_i is not read in the same constant speed compared to a normal disc. This is caused by the fact that the radius c is larger than radius b in Fig. 7 while the speed of rotation remains the same. In order to create the pitch variation function for the wow effect, the new sampling times for samples s_i must be computed, or equivalently, new indices that are generally not integers must be found. We make the following simplifying assumptions: the angular frequency of the player is constant, the cartridge and stylus follow the track fluently, and we consider the track as a circle (although it is spiral in reality).

We next consider the triangle AOB in Fig. 7. We know the angle β from the constant angular frequency, the displacement a, and radius b of the track. We need to solve the time-varying angle ω, which can be solved from $\pi - \gamma$ (see Fig. 7). Using the law of sines

$$\frac{\sin(\beta)}{b} = \frac{\sin(\alpha)}{a} \tag{6}$$

we obtain

$$\alpha = \arcsin\left[a \sin(\beta) / b\right] \tag{7}$$

Using the sum of triangles $\gamma = \pi - \beta - \alpha$ we can finally solve:

$$\omega = \pi - \gamma = \beta + \alpha = \beta + \arcsin\left[a \sin(\beta) / b\right] \tag{8}$$

When β is incremented with small fixed steps that correspond to the rotation speed of the disc, this equation yields the angle ω at which the samples are taken. This model can yield a regular wow that is typical of an eccentric LP or a gramophone disc. The radius b is normally much larger than the distance a. Therefore, the pitch variation curve derived from Eq. (8) will be approximately sinusoidal in practice.

Next we propose a more generic model for the pitch variation curve, which we find more useful than the above model that only simulates a single source or wow. Knowing the smoothness and periodic nature of this degradation, a sinusoidal function with variable frequency $f(n)$ and envelope $A(n)$ can simulate a pitch variation curve:

$$p(n) = 1 + A(n)\sin\left[2\pi f(n)n / f_s\right] \tag{9}$$

This simulation model is split up into two different items: frequency variation and envelope variation of the pitch curve. The two curves are formed in the same way: having their average value and their deviations (assuming normal probability density functions) a random function is created for each one. Since the changes should be smooth, some time is needed for the transition from one frequency and envelope value to another. That transition was decided to last one fifth of the mean period of the pitch variation curve (a time of revolution of the recording in the case of wow). To obtain all the needed points (the sampling rate is 44100 Hz), a 'spline' interpolation is used to join all the data in a smooth way. Now that the corresponding frequency and envelope values are known for every sample, the sinusoidal signal, see Eq. (9), can be implemented (see Fig. 8).

The resampling for the creation of the defect is done using spline interpolation. As can be seen in [33], this technique is one of the best ones. It is also easy to implement with functions provided by Matlab. Sound examples demonstrating the synthetic wow and flutter effects are available on our web site [18].

2.6 Hiss

In the historical recordings there are usually moments when music is not played (silence), so the heard sound is due to the noise the recording system, the storage medium, and the playback system. This noise is mostly formed by hiss, clicks, and thumps. While hiss is a kind of global degradation, clicks and thumps are localized degradations, being thus not present all the time the without musical signal. For this reason, if a silent part is present in the audio file, usually at the beginning or at the end, it will be a period when the hiss is the most predominant degradation.

An important aspect to explain is that only stationary hiss is considered in the simulation. This assumption is made due to the historical audio sources provided for the realization of the project, where few important changes can be observed looking at the spectral properties of the hiss at different moments (see Fig. 9). Moreover, it simplifies the recovering data procedure and its performance being at the same time quite realistic.

The first step for hiss reproduction is to have the characteristics which represent it in the different historical recordings to imitate. For this, some silent extracts have been obtained from the given audio files and their frequency characteristics have been studied, noticing at which frequencies the hiss has peaks or dips. Starting from white noise, a filter with a similar frequency response as the corresponding spectral properties of the hiss is used (see Figs. 9 and 10). After this process, a noise similar to hiss is achieved. The filters are designed using linear prediction (LPC).

2.7 Tracking Errors

A tracking error is produced when a considerably large discontinuity is found by the needle when following the groove. The result is usually a jump to an adjacent groove, repeating or skipping a part of the audio signal. If the original recording is available the simulation of this effect is a simple editing task. Knowing the revolution speed of the playback system, the effect starts with a strong thump and then it goes back or forward a period of the revolution system. For example, the revolution speed of an LP is 33 rpm, so its period is 60/33 s or about 1.8 s. In the case of repetition of a period, the effect can be repeated a few times until the needle finally follows the correct groove.

2.8 Clicks

The parameters to define this impulsive disturbance are the duration of the burst, the time span between them, and their amplitude. If the probability distributions of these parameters are known, clicks can be reproduced in a trustworthy way with the help of random numbers. Thus, the simulation can be divided into three different parts: harvest of data, modeling of the clicks, and reproduction. For the harvest of data, the corrupted and the restored (restoration of clicks only) music files are needed. The clicks can be extracted by subtraction of the two signals, because it is known that clicks can be modeled by an additive model [1].

The most difficult part is to obtain the restored audio file, but luckily many techniques have been developed to remove the clicks, see, e.g., [1], [27], [34] and [35]. The restoration of the audio file can be split into the detection of clicks and the interpolation to replace the corrupted samples. Reconstruction of the audio signal samples can in this case be implemented for example using sinusoidal modeling [36] or autoregressive extrapolation [37], [38]. An easy solution is to use a piece of click/pop removing software, such as the one available in the Adobe Audition. An audio file can be restored and the final result can be listened to verify that the restoration was successful. It should be taken into account that undistorted samples can suffer from some small changes with this process, so a threshold must be used after the subtraction of the corrupted and restored audio file to define the minimum level of a click. Furthermore, it is important to notice that usually not only the clicks but also longer disturbances called pops are removed. Defining that the duration of the pops is much larger than the duration of the clicks (around 50 ms in comparison with 1 ms), it is easy to filter out the pops. Critical listening and visual inspection of the waveform helps to decide the correct parameters for the harvest of data.

In the modeling of the clicks, three different parameters are important: the duration, the gap between the clicks, and the amplitude. In the following, we characterize statistically these parameters to be able to reproduce them in a reliable way. With the available data, histograms can be produced to obtain the parameter distributions. Then, a known distribution similar to the shape of the histogram is chosen, and the parameters which define it are deduced. Matlab was used for the statistical analysis. The probability density functions with which the data is compared are the exponential, the gamma, the Weibull, the normal (Gaussian), the lognormal, the extreme value, and the Poisson distribution. The most adequate one is chosen by selecting the highest Probability Plot Correlation Coefficient (PPCC).

Finally, once all the distributions and their parameters are known the clicks can be simulated and inserted in the desired audio file. When implementing this part, an additional problem was observed that compromise the credibility of the synthetic clicks. They appear to be very static, while in a real historical recording the timbre of the clicks is constantly changing. To solve this problem, a Butterworth lowpass filter with a variable cut-off frequency is used to make the clicks sound more dynamic (see Fig. 11). A good principle is to change the cut-off frequency parameter for each click by picking it from a random function. However, this can increase the computing time considerable, so a simpler solution was selected in this work: Before adding the clicks to the audio waveform, the signal with the clicks is framed (1000 samples per frame at 44.1 kHz sampling rate) and then each frame is filtered with a filter in which the cut-off frequency is decided by a random function.

2.9 Low-Frequency Pulses

For simulating low-frequency pulses, the structure of artificial and real extracted thumps [4] was studied. It is important to stress the introduced errors in the thump due to the technique used for its extraction. The ideal case of a low-frequency pulse is a short discontinuity followed by a long and decaying transient of low-frequency content. When an idealized thump like this is introduced in an audio file and then extracted, the errors in the detection algorithm make it seem less smooth than it really is. Fig. 12 shows the synthetic thump introduced and the one detected.

Knowing the errors introduced by the detection algorithm helps in interpreting the results obtained when low-frequency pulses are extracted from historical recordings. In ref. [4], a few extracted pulses from a historical recording are available, so a general study can be based on them. After the analysis (see Fig. 13), the conclusion is that a low-frequency pulse, as in theory, can be divided into two parts: the initial discontinuity and a long tail. The discontinuity is modeled as a strong click, which is followed by a time where the excited stylus starts an oscillation that decays exponentially towards zero.

Having as parameters the length, the amplitude, and its deviation of the two parts and the frequency and its deviations for the tail (assuming in all the cases a normal probability density function), a thump can be simulated. In case of the tail, a pitch variation curve method with a slight modification is implemented. It should be remembered that the thump is usually caused by deep scratches, so it is repeated for all revolutions in which the scratch is crossing the groove with a period equal to the time of revolution of the recording medium. The length of the scratch is another parameter needed for the correct simulation of this degradation.

3 CASE STUDIES

In this section, the steps to make a CD-quality recording sound as if it were reproduced from an LP, a gramophone, or a phonograph are explained. Since we have discussed previously in Section 2 how the simulated degradations are implemented, only specific data for each simulation are given here.

3.1 LP Disc Simulation

The sound quality of an LP depends on many distinct factors, ranging from when and how it was manufactured to its maintenance. In this section, the quality of an early monophonic LP from the 1950s is simulated. The processing steps of the simulation are shown in Fig. 14. The first step, common to all our simulations, is the stereo-to-mono conversion. Due to the example material which was provided to us, it was decided to render the LP file monophonic too, despite the fact that some simple changes in the Matlab code would allow the required degradations for a stereo file. A single-channel version of the sound file is obtained by averaging the left and the right channel.

The spectrum of a historical vinyl recording is shown in Fig. 15. By analyzing old recordings, the missing frequency range, i.e., the frequencies to be removed during simulation, can be deduced. In this case, it is only necessary to suppress the high frequencies, so a lowpass Butterworth filter is used. Its parameters are shown in Table 2, where W_p is the pass-band corner frequency, W_s is the stop-band corner frequency, R_p is the pass-band ripple (the maximum permissible passband loss), and R_s is the stop-band attenuation.

The next step in Fig. 14 is the addition of clicks. The histograms of the three different model parameters (the duration of the clicks, the duration of the gap between the clicks and their amplitude) and the parametric distributions, which best fit with these data are shown in Fig. 16. The selection of the suitable statistic distribution is made in terms of the Probability Plot Correlation Coefficient (PPCC or R). For the duration of the clicks (Fig. 16(a)), a Weibull distribution is chosen for the simulation. Its pdf (probability density function) is defined by:

$$f(x/a,b) = ba^{-b}x^{b-1}e^{-(\frac{x}{a})^b}I_{(0,\infty)}(x) \tag{10}$$

In the case of the time between the clicks the statistical distribution which best represents it is the Gamma distribution, whose pdf is:

$$f(x/a,b) = \frac{1}{b^a\Gamma(a)}x^{a-1}e^{\frac{x}{b}} \tag{11}$$

As can be observed, while the distributions of the click duration and of the gaps between the clicks are discrete probability distributions, the ones with which they are modeled are continuous. For solving this problem, the obtained sample index at which the click is to be inserted is rounded to the nearest integer. If the nearest integer is 0, 1 is chosen instead.

The amplitude of the clicks is simulated with a lognormal distribution. Its pdf is expressed as:

$$f(x/\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}}e^{\frac{-(\ln x-\mu)^2}{2\sigma^2}} \tag{12}$$

The modeling of the amplitude is based on the absolute value, so when reproducing it a random function with an equal probability for –1 and 1 is used to randomly pick the sign. Moreover, the loudness of the clicks must be adjusted, because the obtained results are dependent on the volume of the sound file from which they were extracted. Knowing that they are modeled by a lognormal distribution, the expected value is expressed by:

$$E(X) = e^{\mu+\sigma^2/2} \tag{13}$$

If the desired mean is *m*, the obtained samples can be multiplied by *m/E(X)* and the volume of the clicks increases as desired. For the vinyl recordings, the average value is set in this case to $m = 0.2$, when the original signal values vary between –1 and +1.

The filters used to provide the dynamic variations for clicks have a minimum normalized cut-off frequency of 0.1 and a maximum of 0.5. The cut-off frequency is uniformly distributed, so all the values between the minimum and the maximum have the same probability. The frame duration is 1000 samples and the order of the Butterworth filter is 3. In Fig. 11 different possible frequency responses of the dynamic lowpass filter are shown. An audio example where synthetic vinyl disk clicks have been added to a music file is available online on the companion web page [18]. This example does not include any other degradations.

The next degradations to be introduced are the low-frequency pulses. There were supposed to be eight different types of scratches of different lengths (crossing over 5 to 9 grooves). Distinct thumps were simulated for each scratch. As the rotation speed of a vinyl record is 33 rpm, the time of revolution is T_d = 60/33 s, i.e., the time elapsed between thumps created by a single scratch. The companion web page contains an example sound file which has several bursts of synthetic thumps [18].

The filter with which the hiss is produced is estimated with a second-order linear predictor, and its frequency response is shown in Fig. 10. The spectrogram of the silent part in which the design of filter is based can be seen in Fig. 9, where the presence of clicks and hiss is easily observed. The signal-to-noise ratio (SNR = $10\log_{10}(P_{signal}/P_{noise})$, where P refers to a short-term power estimate) is set to 37 dB.

If the disc is punched wrongly, wow will be generated during reproduction. In this case, the pitch variation curve used for the addition of this degradation (see Fig. 17) is a sinusoidal function with the same period as the time of revolution (T_d = 60/33 s).

At this point, after the degradations simulation some non-desired high frequencies are present in the music file, so a lowpass filter with a smooth transition is used for slightly attenuating them. The parameters are shown in Table 3. To finish this simulation and to make it sound as a vinyl disc with severe damage, some tracking errors are included. The final result that has gone through all the processing stages is available on the companion web page [18]. Some of these errors are introduced at the beginning of the music file, as if the needle jumped to the previous groove repeating the same short part of the song. In this example, after three repetitions the needle is allowed to continue in the correct groove.

3.2 Gramophone Simulation

To simulate the sound quality of a gramophone more degradations than used for the vinyl simulation are needed, such as distortion and the use of a horn for the music reproduction (see Fig. 18). After the stereo-to-mono conversion, the signal bandwidth is reduced according to the characteristics of a gramophone recording. Fig. 19 shows frequency components of a gramophone recording, where the poor response at high frequencies can be observed. Additionally, the gramophone system is unable to reproduce very low frequencies. Thus, unlike with the LP simulation, a bandpass filter is used with the parameters shown in Table 4.

Two parameters adjust for the distortion simulation. The parameters used in this case are B_{loud} = 2.5 for the loud passages and B_{soft} = 1.8 for the soft passages. The final distortion function is represented in Fig. 20.

For the click simulation, the procedure is the same as the one for the LP, but in this case the specific statistic functions shown in Fig. 21 are used. The mean value used for the click amplitude is 0.1. In the lowpass filter used for providing different spectral components to the different clicks the minimum normalized cut-off frequency is 0.2 while the maximum is 0.4. The order of the Butterworth filters is 3 and the cut-off frequency is changed every 1000 samples.

In this case, for the addition of thumps, more scratches on the surface of the disc than in the LP are expected, maximum ten. Their length varies between 4 and 9 grooves and their amplitude is higher than in the vinyl case (see Fig. 22). The rotation speed of a gramophone disc is 78 rpm, which corresponds to a time of revolution of $T_d = 0.77$ s. This is the time between the thumps produced by a single scratch.

The next task is the hiss simulation. The filter for hiss production (see Fig. 23) is estimated from a signal extracted from a silent passage on a gramophone disc (see Fig. 24) using a linear predictor of order 4. The SNR is set to 30 dB.

In gramophone recordings, the pitch variation defect can be quite annoying. In this case, the pitch variation curve, although smooth, is changing along the time (see Fig. 25). As in the LP case, a lowpass filter is needed for removing the high-frequency components introduced along this process, which could not appear on the gramophone recording (see Table 3).

The horn model used for the frequency-response deviation is a Victor petalled horn appropriate for Victor II, III, IV or similar machines. It is approximately 56 cm long with a 48 cm bell. Although it is made with 8 petals, the simulation is for a one-piece exponential horn due to the limitations of the program ('Hornresp'). The frequency response of the horn is presented in Fig. 26.

A sound file that imitates the sound quality of a gramophone recording by including all the processing techniques mentioned above, is available on the web page [18]. In comparison to the LP disk simulation described in Sec. 3.1, the sound quality is clearly worse. Particularly the wow and hiss are now more easily observed, and the low frequencies are much suppressed in comparison to the LP disk simulation.

3.3 Phonograph Simulation

The phonograph is the oldest successful system for reproducing and storing sound, and the degradations are in this case more severe than in the previous ones. The steps for achieving the desired quality are the same as those for the gramophone simulation (see Fig. 18), but with more harmful parameter settings.

An example spectrogram of a phonograph recording is given in Fig. 27. As can be observed, the bandwidth is quite narrow, missing both low and high frequencies. The filter used in this case is a bandpass filter with parameters given in Table 5. An example of a bandpass filtered musical signal with these parameters, but without other degradations, is also available on our web site [18]. In comparison with the original audio signal, the sound quality is dull, lacking both bass and treble.

The distortion in a phonograph cylinder is an important feature. When listening to a phonograph recording, nonlinear effects, like clipping, can be identified. The parameters for the distortion are higher than for the gramophone, with $B_{\text{loud}} = 3$ for loud passages and $B_{\text{soft}} = 2$ for soft passages. In Fig. 20, the final distortion curve is plotted.

Next it is time for the simulation of the clicks. The distinct statistical functions for each parameter are presented in Fig. 28. The mean value for the amplitude is set to 0.07, the same as the expected value of the distribution. This time there is no need for changing the obtained amplitude of the clicks. Examples of the filters used for adding dynamic variations to the clicks have the normalized cut-off frequency ranging from 0.1 to 0.4.

For the phonograph cylinder simulation two different kinds of scratches are expected: soft and severe. Thirteen scratches of the first type were simulated with a length varying from 4 to 9 grooves and four strong scratches with 10 to 13 grooves length. Fig. 29(a) shows a soft thump while Fig. 29(b) represents a strong one. The phonograph cylinder is revolving with a velocity of about 120 rpm, so the time between consecutive thumps is $T_{\text{d}} = 0.5$ s.

The SNR for the hiss is around 23 dB, and the filter (see Figs. 30 and 31) is obtained with the eighth-order linear predictor. A resonance at about 1 kHz is observed in the calculated filter response. A sound example demonstrating phonograph-like hiss (SNR = 5.6 dB) is available on our web site [18].

The pitch variation curve in this recording medium is caused by wow and flutter. While the mean period of the defect is the same as the time of revolution ($T_{\text{d}} = 60/120$ s $= 0.5$ s), the flutter variations are faster, and the chosen value for its mean frequency is 10 Hz. The final pitch variation curve is formed as the combination of these two degradations (wow and flutter). Some examples of these curves are shown in Fig. 32.

At this point, a lowpass filter is used to limit the final frequency component of the file. In this case, the stop-band and pass-band corner frequencies are much lower than in the vinyl or gramophone simulations due to the narrow bandwidth provided by the phonograph cylinders (see Table 3).

An antique horn of the Columbia Standard Phonograph is used as the model for the frequency response deviation simulation. This horn is built with eight scalloped, metal panels. It is approximately 41 cm long and 44.5 cm in diameter. The opening at the tip measures 2.7 cm (diameter). As previously, the simulation of the transmission coefficient (Fig. 33) is conducted for an exponential horn constructed in one piece.

An example of a phonograph cylinder simulation realized according to the above description can be heard on our web page [18]. In this audio file, the disturbances, such as hiss and fast thumps, are louder than music itself. Wow and flutter are strong making the pitch of musical tones shaky. Low and high frequencies are severely suppressed, and the music generally sounds distorted to the extent that it is not easy to hear all notes present in the original recording.

4 CONCLUSIONS AND FUTURE WORK

Audio antiquing was introduced as an approach to simulate the sound quality of historical recording media, such as phonograph, gramophone, and LP recordings. In the work presented in this paper, the source signal is a modern digital recording, such as one taken from a new CD. In this paper, the different kinds of degradations in these historical recordings were first studied on a theoretical level, analyzing their causes and effects. After that, the best yet simple way to simulate each of them was studied. The most emphasis was put on in the implementation of the disturbances that are clearly audible, such as clicks and wow. Finally, the suitable parameters of the synthetic degradations were estimated for simulating the sound quality of each old device. When all the data was recollected, the required degradations were added to the source audio file. In general, the results were convincing, although some manual adjustment may be necessary in some cases.

Future research in audio antiquing can include the simulation of recording devices not discussed in this article, such as an open reel tape recorder or a C cassette player. Moreover, it would be of interest to go further to the modification of musical instrument sounds one at a time, which would require source separation. An example is to separate and process the sound of an electric guitar to modify its timbre, because modern electric guitars sound very different in comparison to early ones. Other future projects include cancellation of effects processing, such as de-compression of pop music files, or imitation of the live recording quality, which may be a rock concert on a stadium including additional noise of the audience.

5 ACKNOWLEDGMENTS

The authors are grateful to Prof. Perry R. Cook who first suggested the use of the term ‘antiquing’ in this context, and to Prof. Matti Karjalainen and to Mr. Jyri Pakarinen for helpful discussions.

6 REFERENCES

[1] S. J. Godsill and P. J. W. Rayner, *Digital Audio Restoration – A Statistical Model Based Approach*. London, UK: Springer-Verlag, 1998.

[2] S. J. Godsill, "Recursive Restoration of Pitch Variation Defects in Musical Recordings," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing* (Adelaide, Australia, Apr. 1994), vol. 2, pp. 233–236.

[3] P. A. A. Esquef, V. Välimäki, and M. Karjalainen, "Restoration and Enhancement of Solo Guitar Recordings Based on Sound Source Modeling," *J. Audio Eng. Soc.*, vol. 50, no. 4, pp. 227–236 (2002 April).

[4] P. A. A. Esquef, L. W. P. Biscainho, and V. Välimäki, "An Efficient Algorithm for the Restoration of Audio Signals Corrupted with Low-Frequency Pulses," *J. Audio Eng. Soc*., vol. 51, no. 6, pp. 502–517 (June 2003).

[5] P. Skritek, "Unified Compandor Measurements by Using a Channel Simulator," *J. Audio Eng. Soc*., vol. 35, no. 5, pp. 317–336 (1987 May).

[6] W. Klippel, "Speaker Auralization – Subjective Evaluation of Nonlinear Distortion," presented at the 100th Convention of the Audio Engineering Society (Amsterdam, The Netherlands), preprint 5310, April 2001.

[7] N. Zacharov, "VirtualPhone – A Rapid Virtual Audio Prototyping Environment," presented at the 122nd Convention of the Audio Engineering Society (Vienna, Austria), May 2005.

[8] M. Karjalainen and J. Pakarinen, "Wave Digital Simulation of a Vacuum-Tube Amplifier," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing* (Toulouse, France, May 2006), vol. 5, pp. 153-156.

[9] M. Karjalainen, T. Mäki-Patola, A. Kanerva, and A. Huovilainen, "Virtual Air Guitar," *J. Audio Eng. Soc.*, vol. 54, no. 10, pp. 964-980 (2006 Oct.).

[10] D. T. Yeh, J. S. Abel, and J. O. Smith, "Simplified, Physically-Informed Models of Distortion and Overdrive Guitar Effects Pedals," in *Proc. Int. Conf. on Digital Audio Effects* (DAFx-07) (Bordeaux, France, Sept. 2007), pp. 189-196.

[11] J. O. Smith, S. Serafin, J. Abel, and D. Berners, "Doppler Simulation and the Leslie," in *Proc. Int. Conf. on Digital Audio Effects* (Hamburg, Germany, Sept. 2002), pp. 13-20.

[12] A. Huovilainen, "Enhanced Digital Models for Analog Modulation Effects," in *Proc. Int. Conf. Digital Audio Effects* (Madrid, Spain, Sept. 2005), pp. 155-160.

[13] S Bilbao, "A Digital Plate Reverberation Algorithm," *J. Audio Eng. Soc*., vol. 55, no. 3, pp. 135-144 (2007 Mar.).

[14] A. Huovilainen, "Non-linear Digital Implementation of the Moog Ladder Filter," in *Proc. Int. Conf. Digital Audio Effects* (Naples, Italy, Oct. 2004), pp. 61-64.

[15] V. Välimäki and A. Huovilainen, "Oscillator and Filter Algorithms for Virtual Analog Synthesis," *Computer Music J.*, vol. 30, no. 2, pp. 19-31 (2006).

[16] iZotope, Inc. *iZotope Vinyl*. WWW page: http://www.izotope.com/products/audio/vinyl/, June 2007.

[17] iZopote, Inc. Personal e-mail communication, Nov. 2007.

[18] S. González, *Digital Audio Antiquing (Audio Examples)*. Online at: http://www.acoustics.hut.fi/publications/papers/antiquing/, June 2007.

[19] R. T. Beyer, *Sounds of Our Times – Two Hundred Years of Acoustics*. Springer-Verlag, New York, 1999.

[20] J. C. Fesler, "Electrical Reproduction of Acoustically Recorded Cylinders and Disks, Part 2," *J. Audio Eng. Soc.*, vol. 31, no. 9, pp. 674-694 (1983 Sept.).

[21] J. H. Kogen, "Tracking Ability Specifications for Phonograph Cartridges," *J. Audio Eng. Soc.*, vol. 12, no. 2, pp. 186–190 (1968 April).

[22] Encyclopedia Britannica, *Compact Disc: Dynamic Range*. Online: http://www.britannica.com/eb/article-92850/compact-disc, June 2007.

[23] T. Muraoka, H. Onoye, and J. M. Eargle, "On the Measurement of Phonograph Cartridges by the Pulse-Train Method, Part 2," presented at the Convention of the Audio Engineering Society, preprint 1088, 1975.

[24] A. Czyzewski, A. Ciarkowski, A. Kaczmarek, J. Kotus, M. Kulesza, and P. Maziewski, "DSP Techniques for Determining "Wow" Distortion," *J. Audio Eng. Soc.*, vol. 55, no. 4, pp. 266–284 (2007 Apr.).

[25] U. R. Furst, "Periodic Variations of Pitch in Sound Reproduction by Phonographs," *Proc. IRE*, vol. 34, pp. 887–895 (1946 Nov.).

[26] H. Sakai, "Perceptibility of Wow and Flutter," *J. Audio Eng. Soc.*, vol. 18, no. 3, pp. 290–298 (1970 June).

[27] P. A. A. Esquef, L. W. P. Biscainho, P. S. R. Diniz and F. P. Freeland, "A Double-Threshold-Based Approach to Impulsive Noise Detection in Audio Signals," in *Proc. X European Signal Processing Conf.* (EUSIPCO 2000, Tampere, Finland, Sept. 2000), pp. 2041–2044.

[28] M. M. Goodwin and J.-M. Jot, "Multichannel Surround Format Conversion and Generalized Upmix," in *Proc. AES 30th International Conference: Intelligent Audio Environments* (Saariselkä, Finland, March 2007), paper no. 30.

[29] D. J. McBean, *Horn Loudspeaker Response Analysis Program*. Online: http://www.users.bigpond.com/dmcbean/, 2007.

[30] N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*. Springer-Verlag, 1991.

[31] E. C. Fox and J. G. Woodward, "Tracing Distortion—Its Causes and Correction in Stereodisk Recording Systems" *J. Audio. Eng. Soc.*, vol. 11, no. 4, pp. 294–301 (1963 Oct.).

[32] 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*, vol. 13, no. 1, pp. 30–60 (1996 Jan.).

[33] P. Maziewski, "Wow Defect Reduction Based on Interpolation Techniques," in *Proc. IV Krajowa Konferencja Elektroninki*, 2005.

[34] I. Kauppinen, “Methods for Detecting Impulsive Noise in Speech and Audio Signals,” in *Proc. 14th Int. Conf. Digital Signal Processing* (July 2002), vol. 2, pp. 967–970.

[35] C. A. Luszick, *Digital Audio Restoration of Vinyl Records and Implementation on a Digital Signal Processor Based System*. Digitale Signalverarbeitung Universitäat Kaiserslautern, Diploma Thesis, Nov. 2003.

[36] R. C. Maher, “A Method for Extrapolation of Missing Digital Audio Data,” *J. Audio Eng. Soc*., vol. 42, no. 5, pp. 350–357 (1994 May).

[37] I. Kauppinen and J. Kauppinen, “Reconstruction Method for Missing or Damaged Long Portions in Audio Signal,” *J. Audio. Eng. Soc.,* vol. 50, no. 7, pp. 594–602 (2002 Jun.).

[38] P. A. A. Esquef and L. W. P. Biscainho, “An Efficient Model-Based Multirate Method for Reconstruction of Audio Signals Across Long Gaps,” *IEEE Trans. Audio, Speech and Language Processing,* vol. 14, no. 4, pp. 1391–1400, (2006 Jul.).

7 FIGURES

Fig. 1. Dirty surface of a vinyl disc.

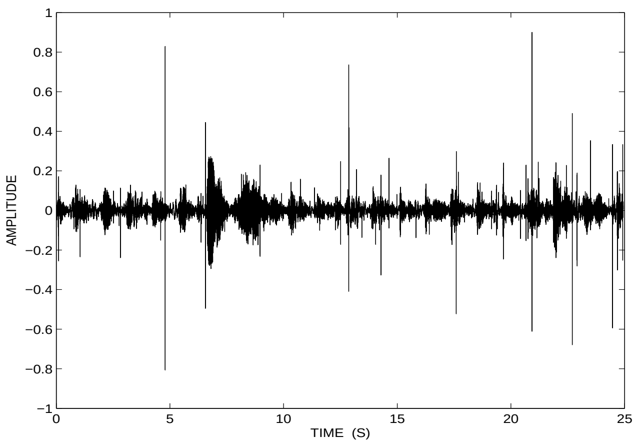

Fig. 2. Click-degraded audio waveform taken from a vinyl disc.

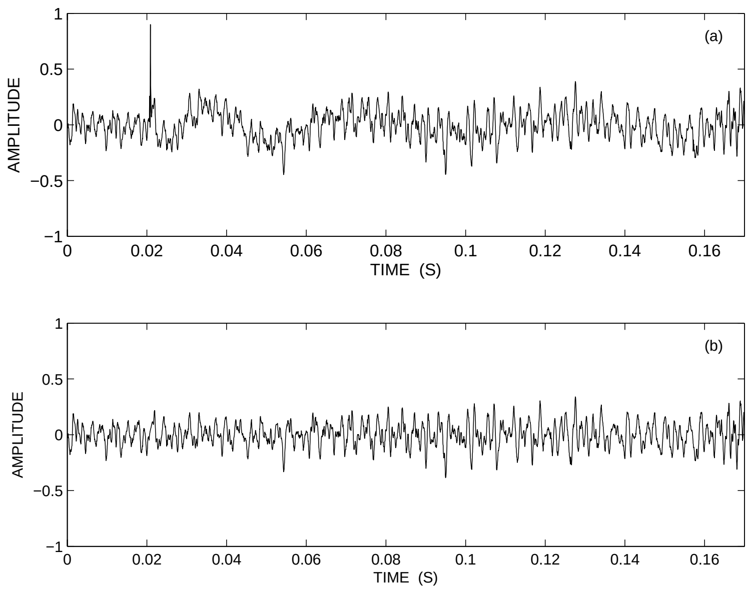

Fig. 3. (a) Audio waveform corrupted by a low-frequency pulse. (b) Restored audio waveform of the corrupted one presented in (a).

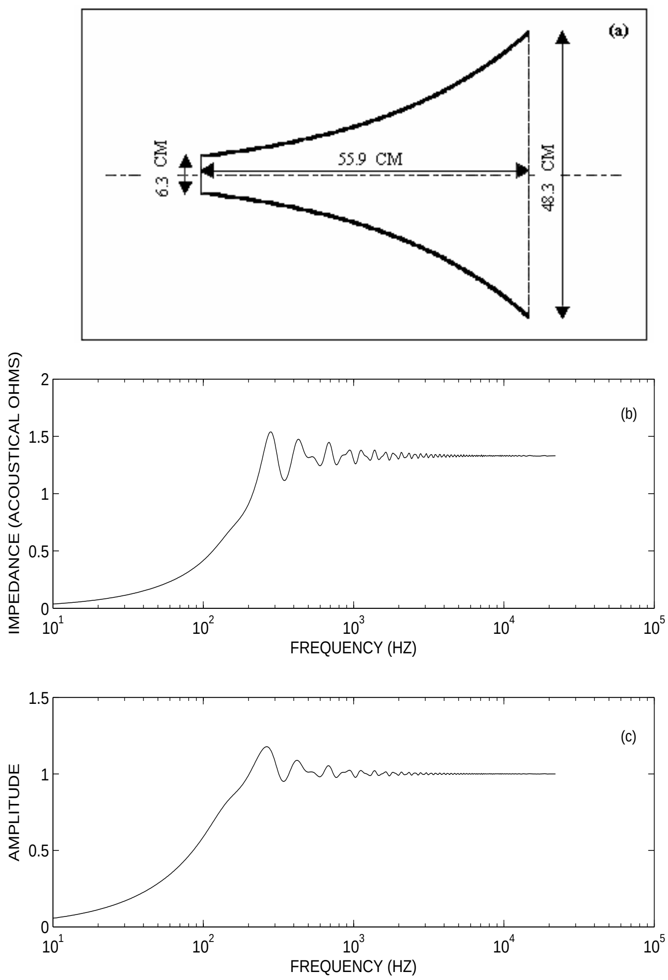

Fig. 4. (a) Schematic diagram of a gramophone horn. (b) Acoustic impedance of the horn represented in (a).

(c) Transmission coefficient of the horn represented in (a).

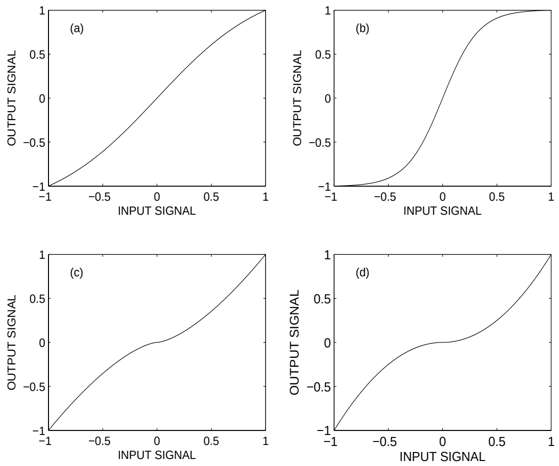

Fig. 5. Examples of functions used for introducing the distortion. (a) Distortion function with Eq. (3) when $B_{\text{loud}} = 1$. (b) Distortion function with Eq. (3) when $B_{\text{loud}} = 3$. (c) Distortion function with Eq. (4) when $B_{\text{soft}} = 1.5$. (d) Distortion function with Eq. (4) when $B_{\text{soft}} = 2$.

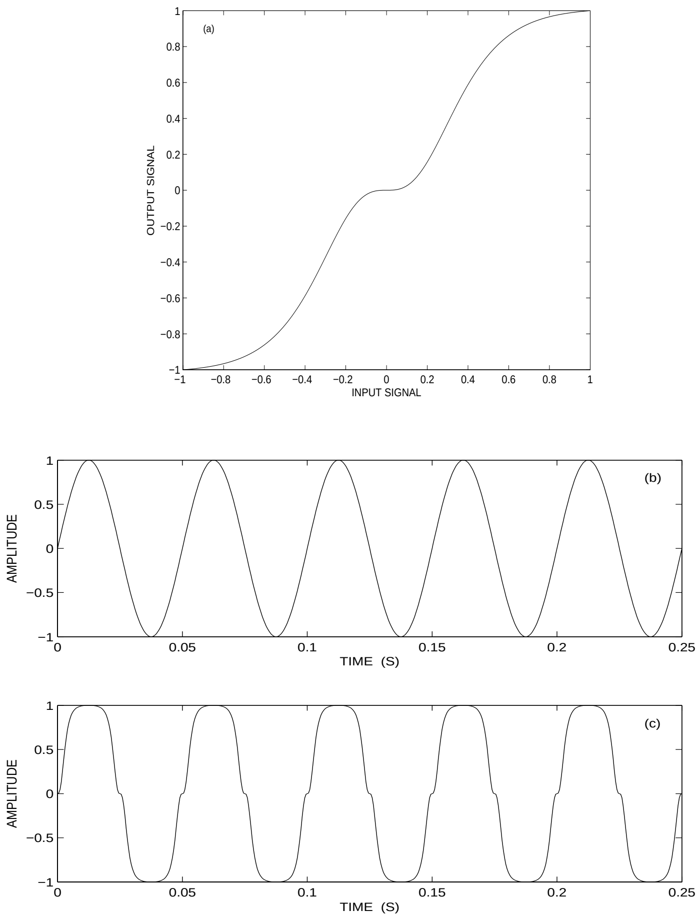

Fig. 6. (a) Total distortion function ($B_{\text{loud}} = 2.7$ for Eq. (3) and $B_{\text{soft}} = 2$ for Eq. (4)). (b) Undistorted sinusoidal signal. (c) Distorted output signal.

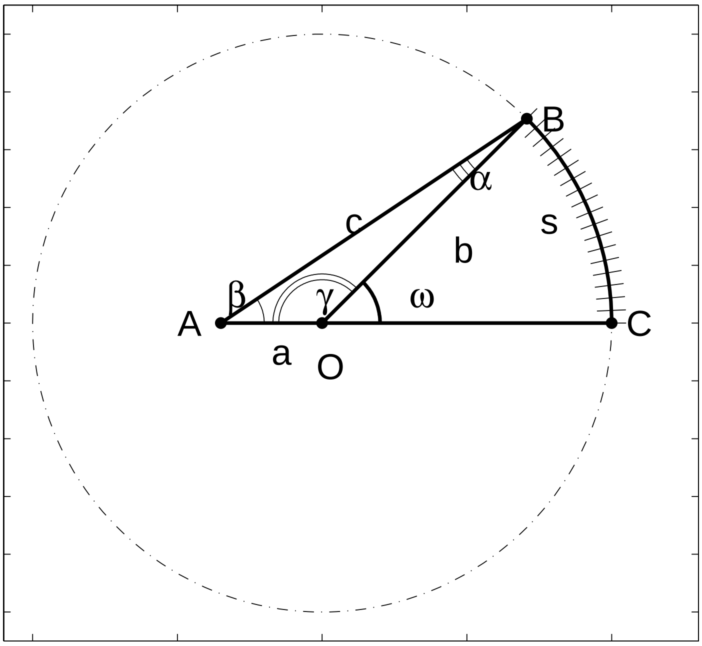

Fig. 7. Geometry of a disc with a displaced center point (A) in comparison to the correct center point (O).

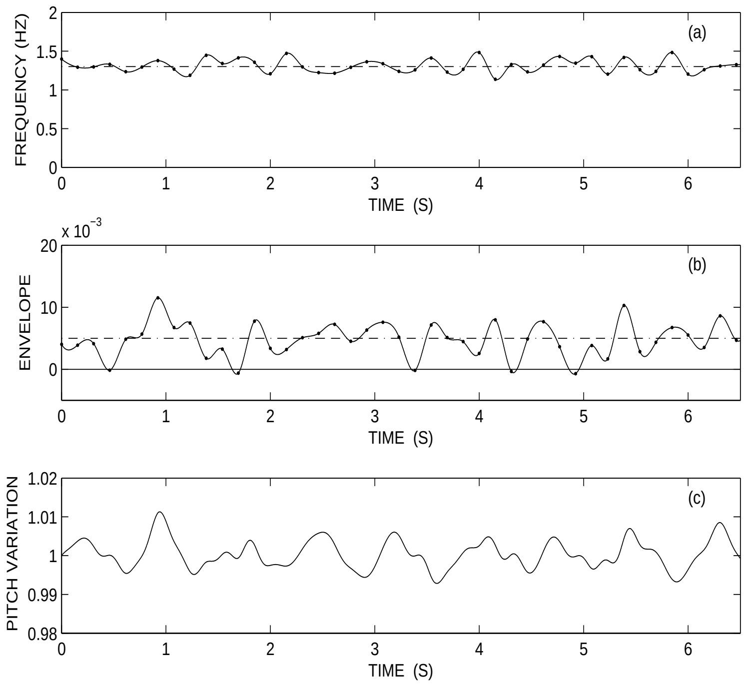

Fig. 8. (a) Frequency variation curve. (b) Envelope of the pitch variation. (c) Synthetic pitch variation curve.

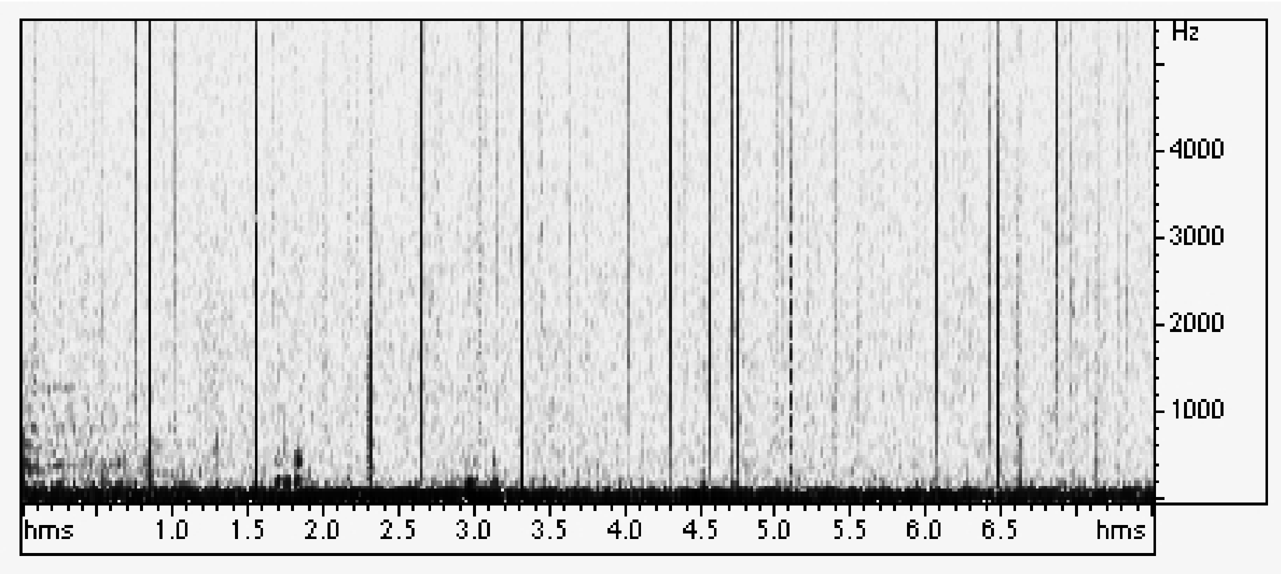

Fig. 9. Spectrogram of a silent part (no music is played) of a vinyl recording. Clicks appear as dark vertical lines.

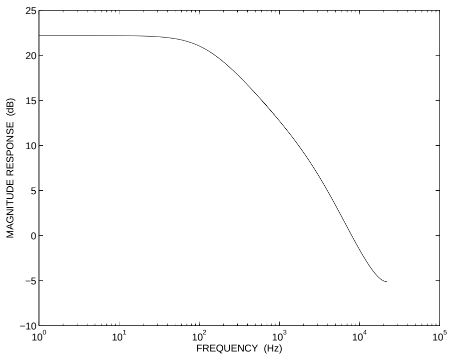

Fig. 10. Magnitude response of the second-order all-pole hiss filter used in vinyl disc simulation.

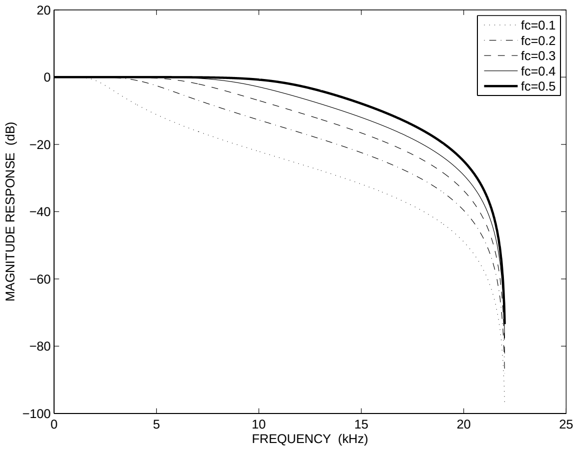

Fig. 11. Magnitude Response of third-order Butterworth lowpass filters with different cut-off frequency. In the legend the normalized cut-off frequency for each filter is shown. The sampling rate is 44100 Hz.

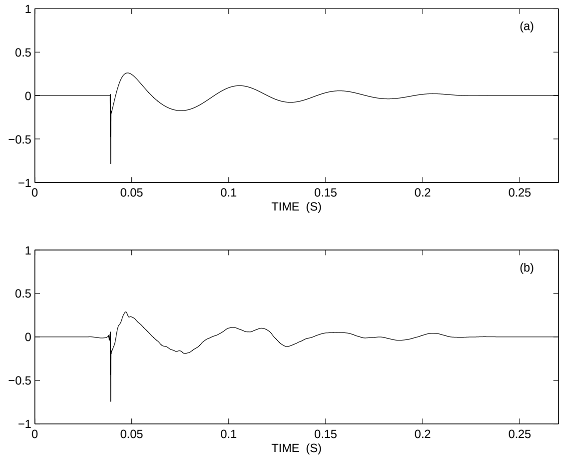

Fig. 12. (a) Artificial thump that was added in a test signals and (b) the same thump after is has been extracted from the audio signal using the method proposed by Esquef *et al.* [4].

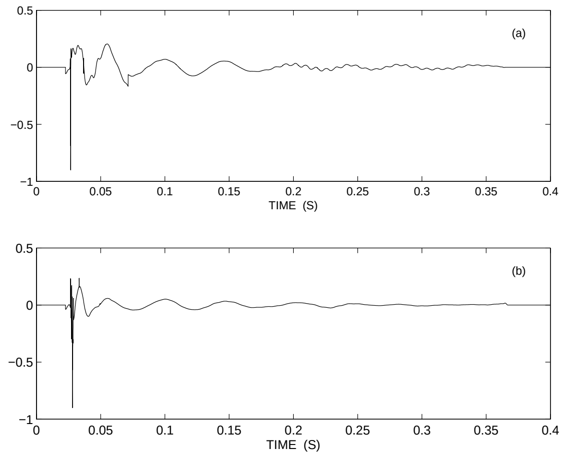

Fig. 13. (a), (b) Two examples of real thumps extracted from old recordings.

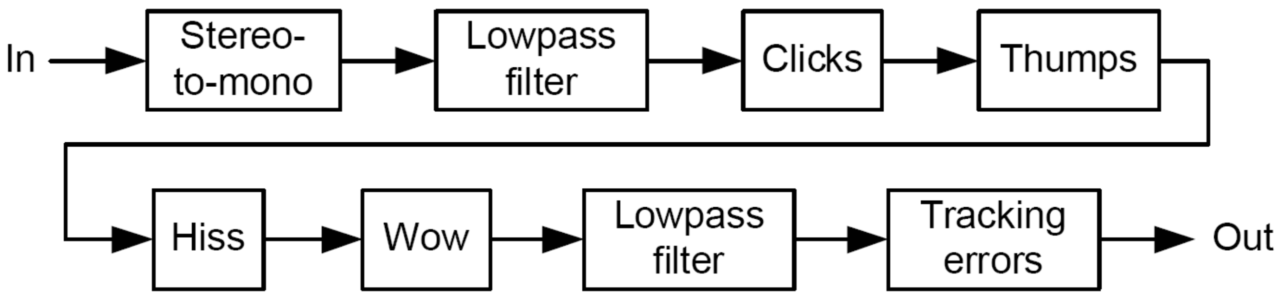

Fig. 14. Processing steps for LP quality simulation. It is possible to skip one or several of these processing steps to avoid specific degradations and to obtain better quality.

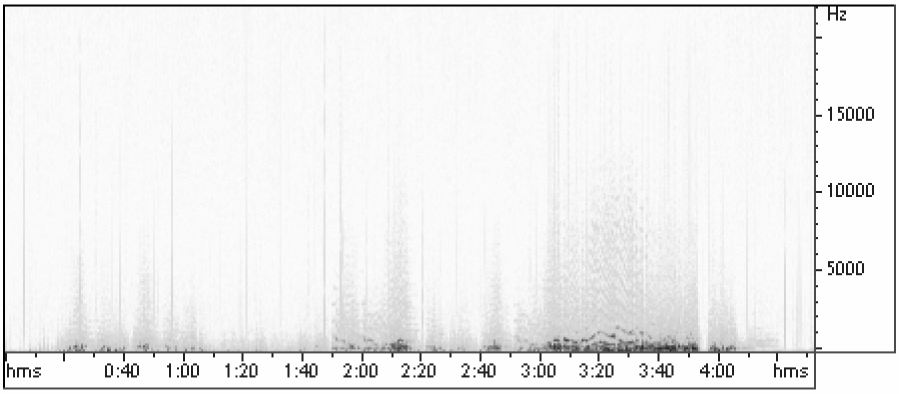

Fig. 15. Spectrogram of an orchestral music vinyl recording released in 1955.

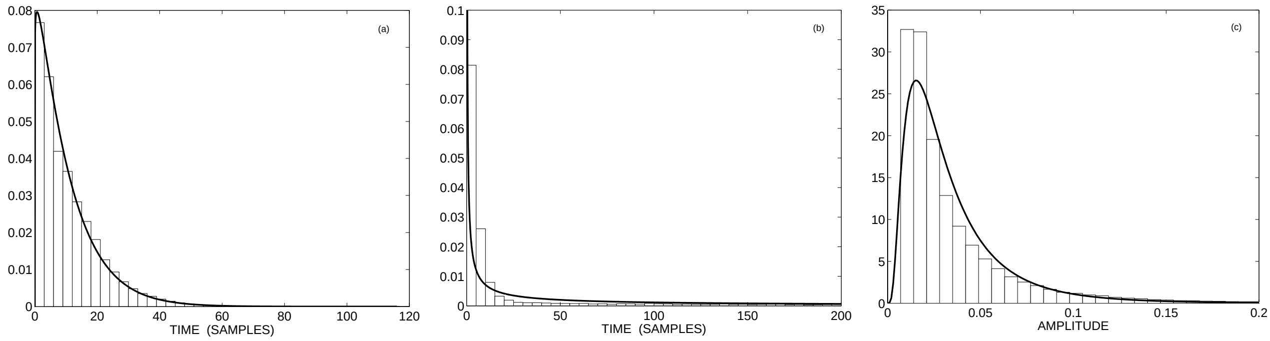

Fig. 16. (a) Duration of the clicks histogram vs. Weibull distribution (R = 0.99853). Parameters: a = 10.6907 and b = 1.0606 (see Eq. (10)). (b) Time between the clicks histogram vs. Gamma distribution (R = 0.97029). Parameters: a = 0.2 and b = 2433.8 (see Eq. (11)). (c) Amplitude of the clicks histogram vs. Lognormal distribution (R = 0.97397). Parameters: μ = -3.6267 and σ = 0.7421 (see Eq. (12)).

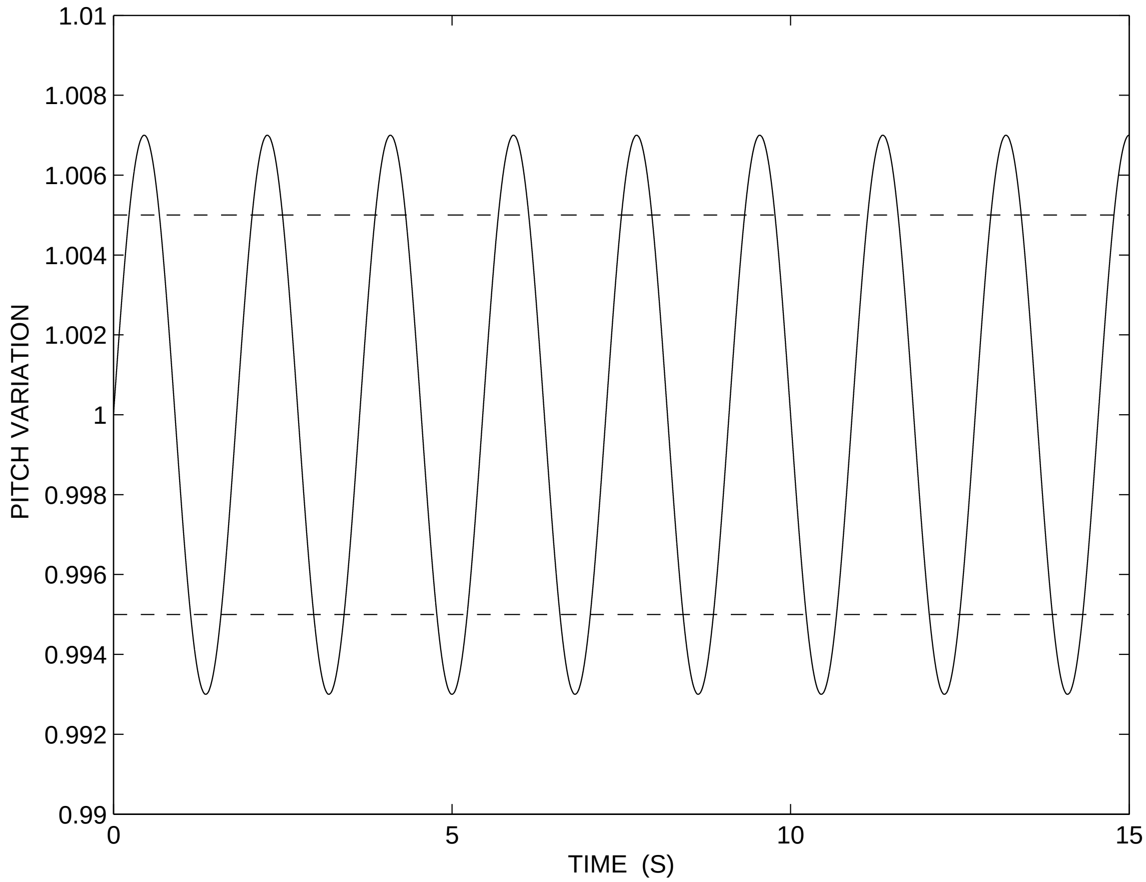

Fig. 17. Pitch variation curve of a wrongly punched LP disc (time of revolution is T_d = 60/33 s = 1.8 s).

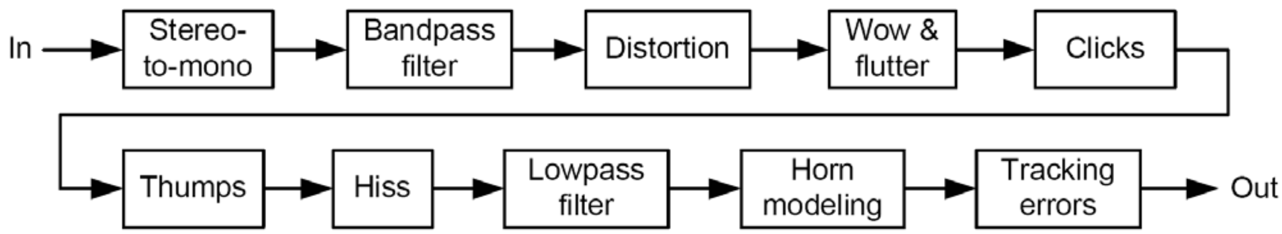

Fig. 18. Processing steps for gramophone and phonograph quality simulation.

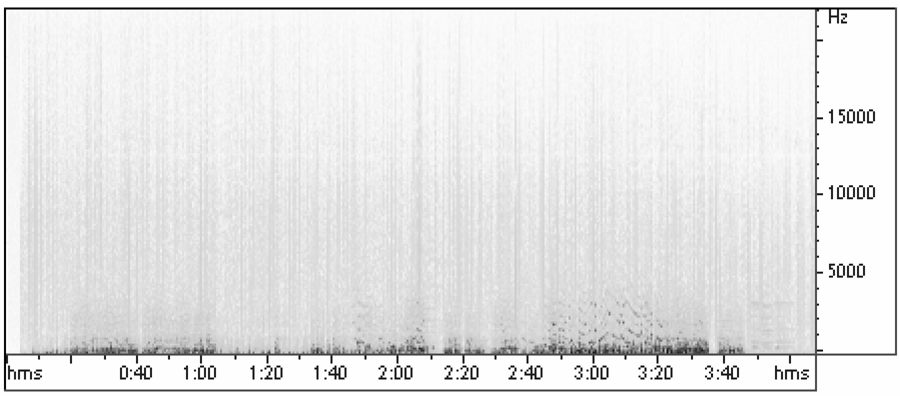

Fig. 19. A gramophone recording spectrogram.

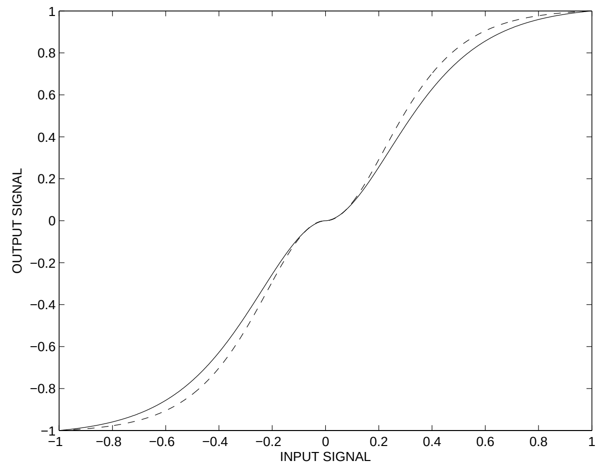

Fig. 20. Total distortion function for gramophone (solid line) and phonograph (dashed line) recording simulation.

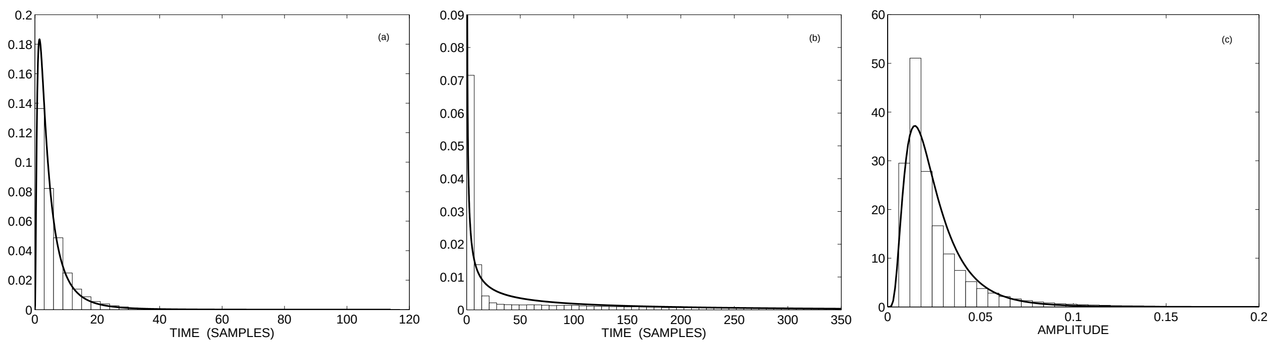

Fig. 21. (a) Clicks duration histogram vs. Lognormal distribution (R = 0.98428). Parameters: μ = 1.2811 and σ = 0.9387 (see Eq. (12)). (b) Time between clicks histogram vs. Gamma distribution (R = 0.99656). Parameters: a = 0.3378 and b = 276.6830 (see Eq. (11)). (c) Clicks amplitude histogram vs. Lognormal distribution (R = 0.94819). Parameters: μ = -3.8530 and σ = 0.6086 (see Eq. (12)).

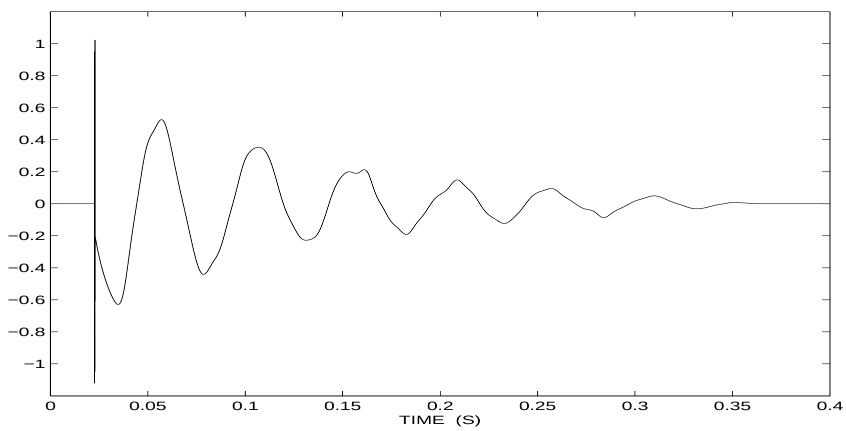

Fig. 22. Possible gramophone thump.

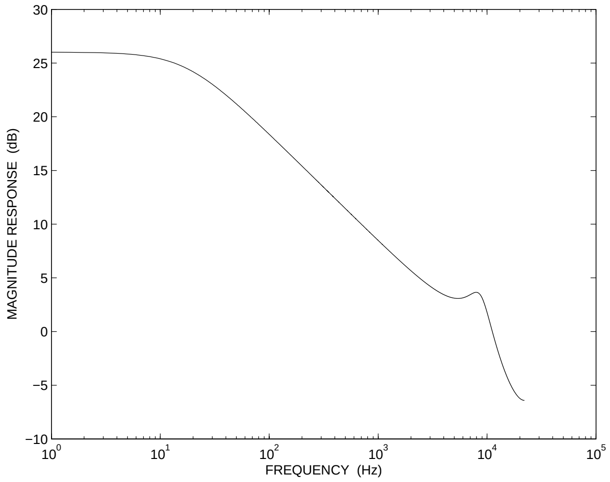

Fig. 23. Magnitude response of the filter used for gramophone hiss simulation.

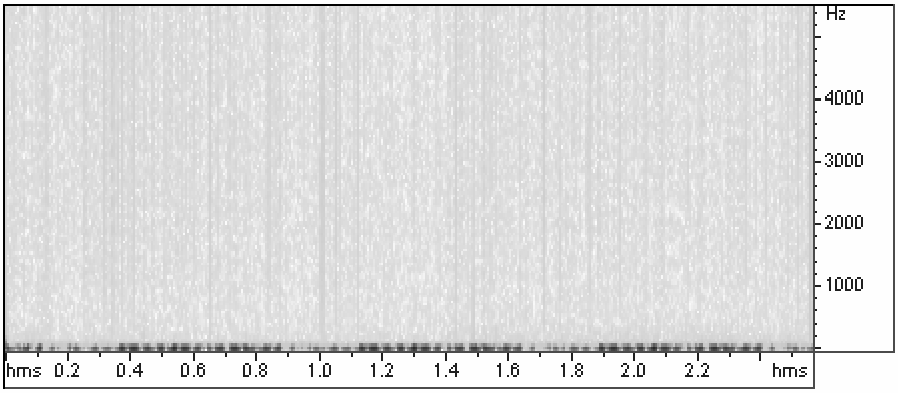

Fig. 24. Spectrogram of a silent part (no music is played) of a gramophone recording.

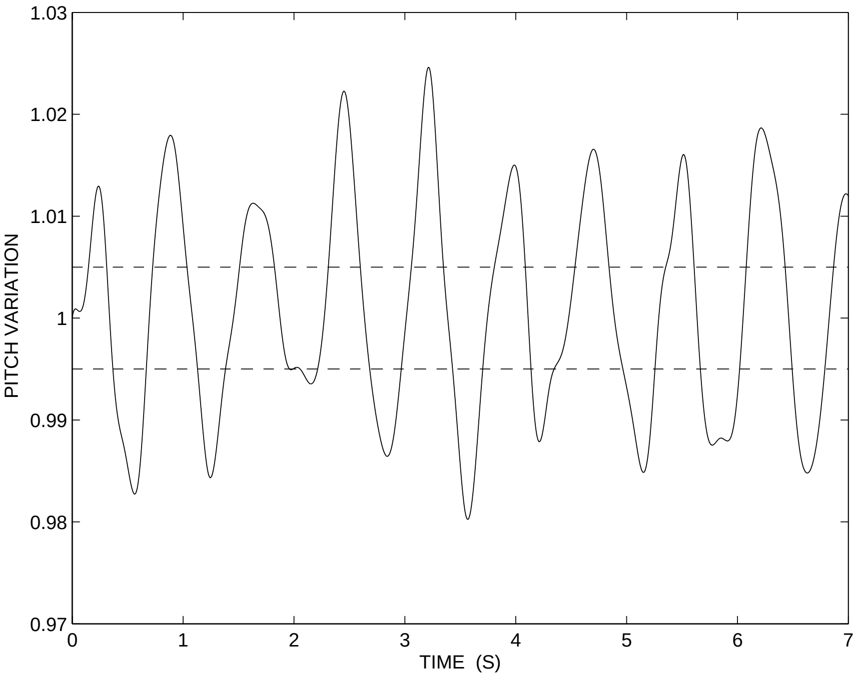

Fig. 25. Synthetic pitch variation curve for a gramophone disc simulation.

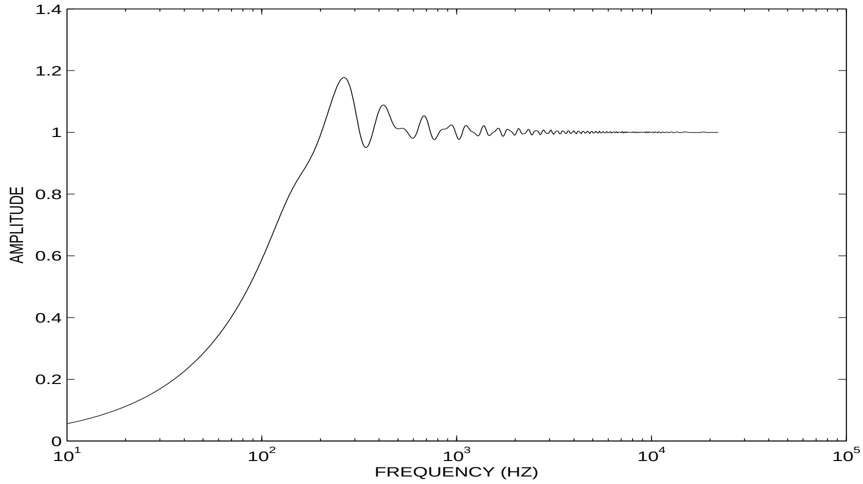

Fig. 26. Simulated transmission coefficient of a Victor petalled horn.

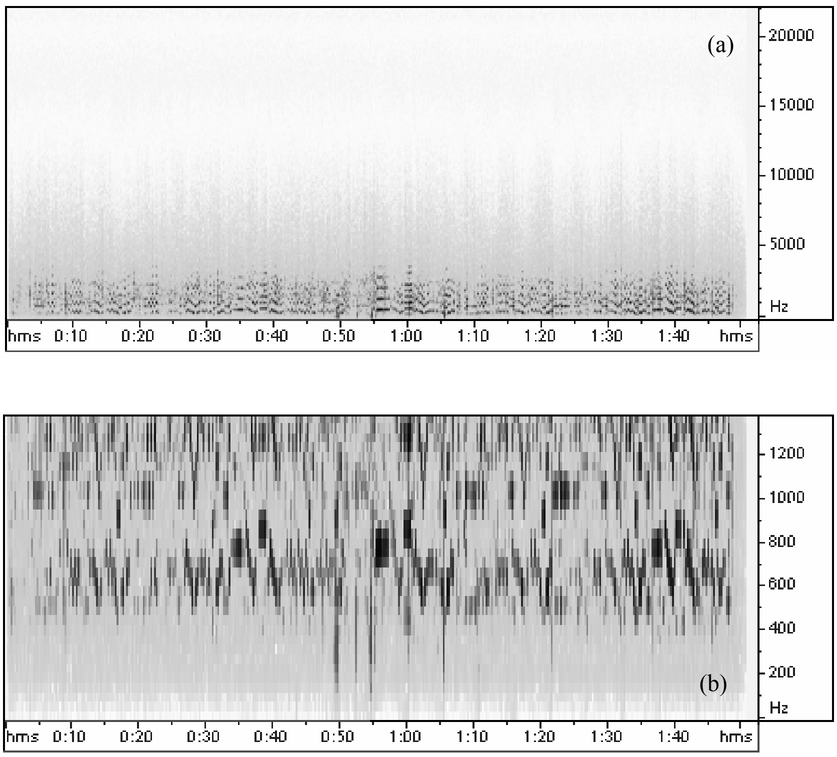

Fig. 27. Phonograph recording spectrogram. (a) All the frequency components. (b) Low frequencies zoomed.

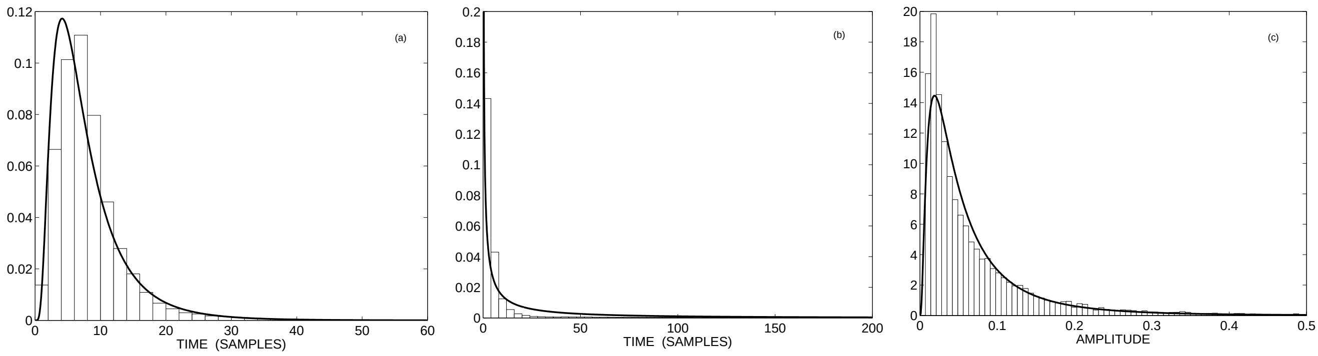

Fig. 28. (a) Clicks duration histogram vs. Lognormal distribution (R = 0.9939). Parameters: μ = 1.8561 and σ = 0.6617 (see Eq. (12)). (b) Time between clicks histogram vs. Weibull distribution (R = 0.98748). Parameters: a = 17.1571 and b = 0.3975 (see Eq. (10)). (c) Clicks amplitude histogram vs. Lognormal distribution (R = 0.9903). Parameters: μ = -3.0870 and σ = 0.9410 (see Eq. (12)).

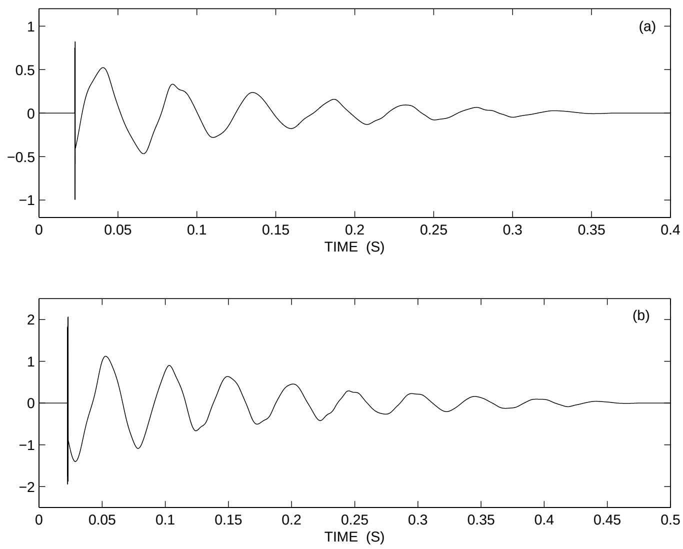

Fig. 29. (a) Normal synthesized phonograph thump. (b) Strong synthesized phonograph thump.

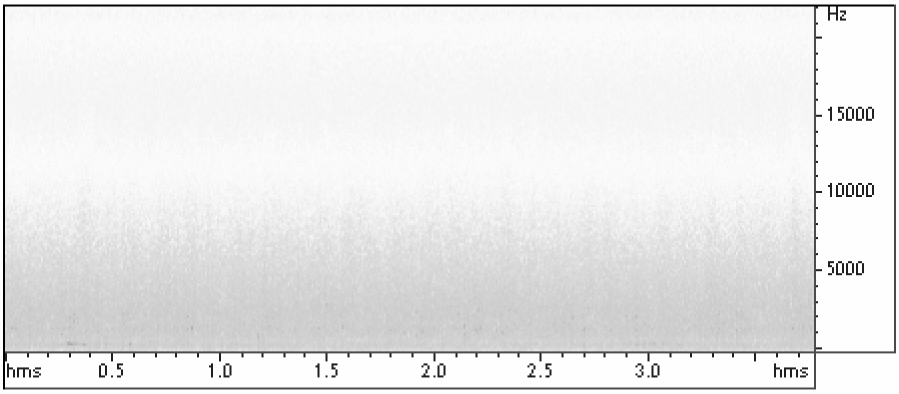

Fig. 30. Spectrogram of a silent part (no music is played) on a phonograph cylinder.

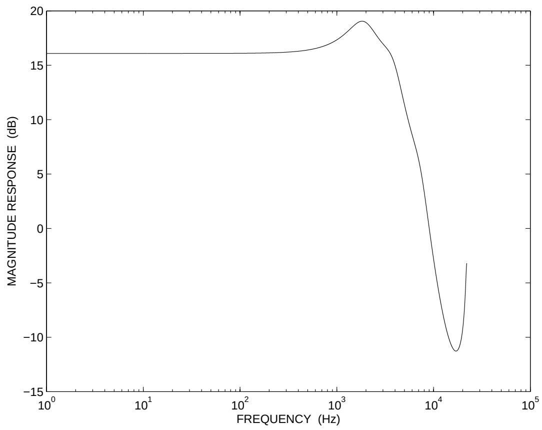

Fig. 31. Magnitude response of the filter used for phonograph hiss simulation.

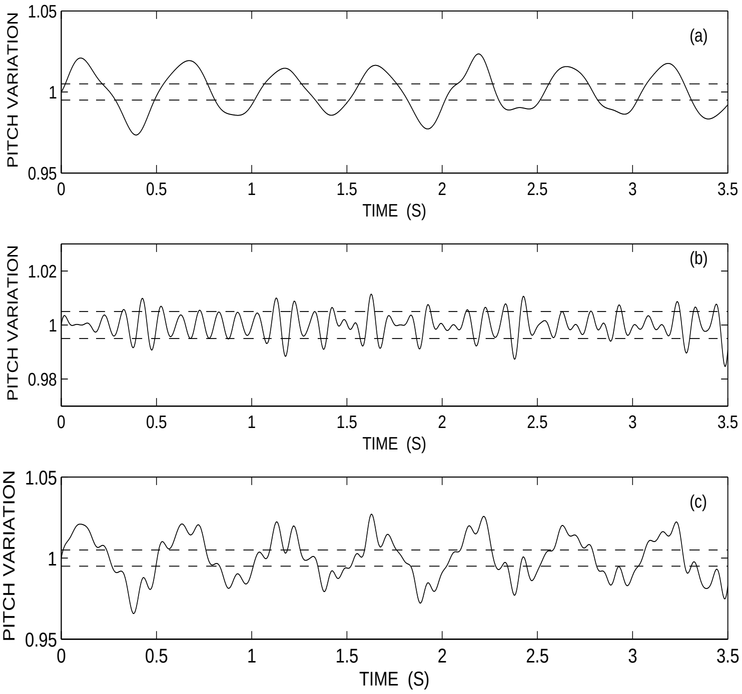

Fig. 32. (a) Simulated wow, (b) simulated flutter, and (c) simulated total pitch variation curve for phonograph recordings.

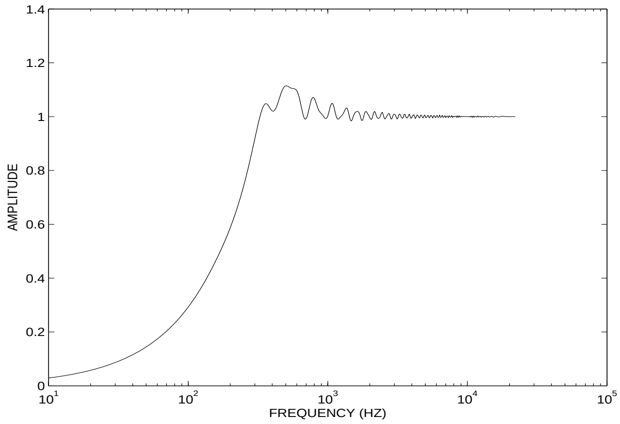

Fig. 33. Simulated transmission coefficient for a phonograph horn.

Table 1. Summary of degradations types in old recordings.

Global degradations	Localized degradations
• Frequency response deviation • Limited signal bandwidth • Dynamic range limitations • Distortion • Pitch variation defects • Hiss	• Tracking errors • Clicks • Low-frequency pulses

Table 2. Coefficients for the lowpass filter in LP bandwidth simulation.

Wp (Hz)	Rp (dB)	Ws (Hz)	Rs (dB)
9000	0.45	12000	13

Table 3. Coefficients of the filter for attenuating or removing undesired high frequency components in LP, gramophone, and phonograph quality simulation.

	Wp (Hz)	Rp (dB)	Ws (Hz)	Rs (dB)
LP	4000	0.46	18000	10
Gramophone	3000	0.46	19000	20
Phonograph	2000	0.46	7500	20

Table 4. Coefficients for the bandpass filter in gramophone bandwidth simulation.

Ws1 (Hz)	Rs1 (dB)	Wp1 (Hz)	Rp1 (dB)	Wp2 (Hz)	Rp2 (dB)	Ws2 (Hz)	Rs2 (dB)
100	20	200	0.46	3000	0.46	5000	20

Table 5. Coefficients for the bandpass filter in phonograph bandwidth simulation.

Ws1 (Hz)	Rs1 (dB)	Wp1 (Hz)	Rp1 (dB)	Wp2 (Hz)	Rp2 (dB)	Ws2 (Hz)	Rs2 (dB)
400	23	1000	0.46	2000	0.46	4000	20