Gramophone Noise Reconstruction
A Comparative Study of Interpolation Algorithms for Noise Reduction

Christoph F. Stallmann and Andries P. Engelbrecht
Department of Computer Science, University of Pretoria, Pretoria, South Africa

Keywords: Gramophone Records, Interpolation, Audio Reconstruction, Polynomials, Signal Modelling.

Abstract: Gramophone records have been the main recording medium for seven decades and regained widespread popularity over the past few years. Records are susceptible to noise caused by scratches and other mishandlings, often making the listening experience unpleasant. This paper analyses and compares twenty different interpolation algorithms for the reconstruction of noisy samples, categorized into duplication and trigonometric approaches, polynomials and time series models. A dataset of 800 songs divided amongst eight different genres were used to benchmark the algorithms. It was found that the ARMA model performs best over all genres. Cosine interpolation has the lowest computational time, with the AR model achieving the most effective interpolation for a limited time span. It was also found that less volatile genres such as classical, country, rock and jazz music is easier to reconstruct than more unstable electronic, metal, pop and reggae audio signals.

1 INTRODUCTION

Gramophone records were the first commercial audio storage medium with the introduction of Berliner's turntable in 1889 (Wile, 1990). Records continued to be widely used for more than seven decades, until they were replaced by the compact disc (CD) in the late 1980s. Although downloadable digital music has become the forerunner in the 20$^{\text{th}}$ century, gramophone sales have surged in the past few years, achieving record sales since they were discontinued as main music medium in 1993. Approximately six million records were sold in the United States alone in 2013, an increase of 33% from the previous year (Richter, 2014). However, the Nielsen Company indicated that only 15% of the sales were logged, since most gramophone records do not have a bar code which is used to track sales (CMU, 2011). Besides an increase in sales of new records, most recordings prior to the 1960s are only available on gramophone. Efforts are made by both commercial music labels and private audiophiles to digitize these historic records. The digitization process is a tedious and time consuming process, since most records are damaged and need to be refurbished.

This paper examines twenty different interpolation algorithms that are utilized for the digital restoration of audio signals that are distorted by scratches and other physical damage to the record. This research is part of a larger project aimed at automating the detection and reconstruction of noise on damaged gramophone records, removing the burden of manual labour. Although the research focuses on gramophones, the algorithms can be directly applied to other areas of audio signal processing, such as lost packets in voice over IP (VoIP) or the poor reception in digital car radios. The methodology, test dataset and the measurement of the reconstruction performance and execution time used during the empirical analysis is discussed. Finally, the algorithms are benchmarked on music from eight different genres and the results are presented in the last section.

2 ALGORITHMS

This section briefly discusses the theoretical background and mathematics of various interpolation algorithms used to reconstruct gramophone audio signals. The algorithms are categorized into duplication approaches, trigonometric methods, polynomials and time series models.

2.1 Duplication Approaches

A simple method for reconstructing missing values is to copy a series of samples from another source or from somewhere else in the signal. Some duplication algorithms make use of equivalent sources to re-

Stallmann C. and Engelbrecht A..
Gramophone Noise Reconstruction - A Comparative Study of Interpolation Algorithms for Noise Reduction.
DOI: 10.5220/0005461500310038
In *Proceedings of the 12th International Conference on Signal Processing and Multimedia Applications* (SIGMAP-2015), pages 31-38
ISBN: 978-989-758-118-2

construct the audio, where at least one of the sources is not subjected to noise (Sprechmann et al., 2013). However, multiple copies are mostly not available and copying samples from different parts of the same source is a more practical solution. A smart copying algorithm was proposed, able to replicate a similar fragment from the preceding or succeeding samples using an AR model with mixed excitation and a Kalman filter (Niedźwiecki and Cisowski, 2001). This section discusses four duplication algorithms, namely adjacent and mirroring window duplication, followed by a nearest neighbour approach and similarity interpolation.

2.1.1 Adjacent Window Interpolation

Adjacent window interpolation (AWI) reconstructs a gap of size n at time delay t by simply copying the preceding n samples from the signal y, that is,

$$y_{t+i} = y_{t-n+i} \quad i \in \{0, 1, \ldots, n-1\} \tag{1}$$

This approach relies on the idea that if a certain combination of samples exists, there is a likelihood that they might be repeated at a later stage. The interpolation accuracy is improved by using bidirectional processing and taking the average between the forward and backward interpolation process.

2.1.2 Mirroring Window Interpolation

Volatile signals that are interpolated with AWI can cause a sudden jump between sample y_t and y_{t-1} and sample y_{t+n-1} and y_{t+n}, that is, where the gap of missing samples starts and ends respectively. By mirroring the samples during mirroring window interpolation (MWI), the signal is smoothed between the first and last sample of the gap as follows:

$$y_{t+i} = y_{t-1-i} \quad i \in \{0, 1, \ldots, n-1\} \tag{2}$$

Similarly, the average between the forward and backward mirrored windows increases the interpolation accuracy.

2.1.3 Nearest Neighbour Interpolation

The nearest neighbour interpolation (NNI) reconstructs a point by choosing the value of the closest neighbouring point in the Euclidean space. NNI for a sequential dataset at time delay t is defined as

$$y_t = \sum_{i=t-k}^{t+k} h(t - i\triangle_t) y_i \tag{3}$$

where k is the number of samples to consider at both sides of y_t, $\triangle_t$ the change in time, and h the rectangular function.

2.1.4 Similarity Interpolation

If the interpolation gap shares little characteristics with the preceding and successive samples, a duplication-based approach is inaccurate and is improved using similarity interpolation (SI) which searches for a sample sequence that is similar to the samples on each side of the gap. This can be done by constructing a set of vectors $\mathbf{d}_i$ by calculating the deviation between the amplitudes of neighbouring samples in a moving window as follows:

$$\mathbf{d}_i = [(y_i\text{-}y_{i+1}), (y_{i+1}\text{-}y_{i+2}), \ldots, (y_{i+n-1}\text{-}y_{i+n})] \tag{4}$$

where y is the series of observed samples with a moving window size of $n+1$. The goal of similarity interpolation is to find the vector in $\mathbf{d}_i$ that shares most of its characteristics with the samples elsewhere in the signal. Note that, by using the amplitude deviation, the algorithm will not just find a sequence similar in amplitude, but also sequences similar in direction and gradient.

2.2 Trigonometric Approaches

Smoothing between two groups of samples can also be achieved through trigonometric functions such as the sine and cosine functions. This section examines Lanczos and a cosine reconstruction approach.

2.2.1 Lanczos Interpolation

Lanczos interpolation (LI) is a smoothing interpolation technique based on the sinc function (Duchon, 1979). The sinc function is the normalized sine function, that is, $\frac{\sin(x)}{x}$ (Gearhart and Shultz, 1990). The LI is defined as

$$l(x) = \sum_{i=\lfloor x \rfloor - n+1}^{\lfloor x \rfloor + n} y_i L(x-i) \tag{5}$$

where $\lfloor x \rfloor$ is the floor function of x, n the number of samples to consider on both sides of x and $L(x)$ the Lanczos kernel. The Lanczos kernel is a dilated sinc function used to window another sinc function as follows:

$$L(x) = \begin{cases} \text{sinc}(x)\text{sinc}(\frac{x}{n}) & \text{for} - n < x < n \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

2.2.2 Cosine Interpolation

A continuous trigonometric function like cosine can be used to smoothly interpolate between two points. Given a gap of n missing samples starting at time delay t, the cosine interpolation (CI) is defined as

$$c(x) = y_{t-1}(1 - h(x)) + y_{t+n} h(x) \tag{7}$$

where $h(x)$ is calculated with cosine as follows:

$$h(x) = \frac{1 - \cos\left(\frac{\pi(x+1)}{n+1}\right)}{2} \quad (8)$$

The cosine operation can be replaced with any other smoothing function $f(x)$, as long as $f(0) = 1$ and $f'(x) < 0$ for $x \in (0,1)$. Alternatively, if the function has the properties $f(0) = 0$ and $f'(x) > 0$ for $x \in (0,1)$, the points y_{t-1} and y_{t+n} in equation (7) have to be swapped around.

2.3 Polynomials

A polynomial is a mathematical expression with a set of variables and a set of corresponding coefficients. This section discusses the standard, Fourier, Hermite and Newton polynomials. Additionally, the standard and Fourier polynomials are applied in an osculating fashion and are utilized in spline interpolation.

2.3.1 Standard Polynomial

A standard polynomial (STP) is the sum of terms where the variables only have non-negative integer exponents and is expressed as

$$m_{stp}(x) = \alpha_d x^d + \cdots + \alpha_1 x + \alpha_0 = \sum_{i=0}^{d} \alpha_i x^i \quad (9)$$

where x represent the variables, α_i the coefficients, and d the order of the polynomial. x represents the time delay of the samples y in audio signals. According to the unisolvence theorem, a unique polynomial of degree n or lower is guaranteed for $n+1$ data points (Kastner et al., 2010). STP are typically approximated using a linear least squares (LLS) fit.

2.3.2 Fourier Polynomial

Fourier proposed to model a complex partial differentiable equation as a superposition of simpler oscillating sine and cosine functions. A discrete Fourier polynomial (FOP) is approximated with a finite sum d of sine and cosine functions with a period of one as follows:

$$m_{fop}(x) = \frac{\alpha_0}{2} + \sum_{i=1}^{d} \left[\alpha_i \cos(i\pi x) + \beta_i \sin(i\pi x)\right] \quad (10)$$

where α_i and β_i are the cosine and sine coefficients respectively, approximated using LLS regression.

2.3.3 Newton Polynomial

Newton formulated a polynomial of least degree that coincides at all points of a finite dataset (Newton and Whiteside, 2008). Given $n+1$ data points (x_i, y_i), the Newton polynomial (NEP) is defined as

$$m_{nep}(x) = \sum_{i=0}^{n} \alpha_i h_i(x) \qquad h_i(x) = \prod_{j=0}^{i-1} (x - x_i) \quad (11)$$

where α_i are the coefficients and $h_i(x)$ is the i^{th} Newton basis polynomial. An efficient method for calculating the coefficients is using a Newton divided differences table.

2.3.4 Hermite Polynomial

Hermite introduced a polynomial closely related to the Newton and Lagrange polynomials, but instead of only calculating a polynomial for $n+1$ points, the derivatives at these points are also considered. The Hermite polynomial (HEP) using the first derivative is defined as

$$m_{hep}(x) = \sum_{i=0}^{n} h_i(x) f(x_i) + \sum_{i=0}^{n} \overline{h}_i(x) f'(x_i) \quad (12)$$

where $h_i(x)$ and $\overline{h}_i(x)$ are the first and second fundamental Hermite polynomials, calculated using

$$h_i(x) = \left[1 - 2l_i'(x_i)(x - x_i)\right] [l_i(x)]^2 \quad (13)$$

$$\overline{h}_i(x) = (x - x_i) [l_i(x)]^2 \quad (14)$$

$l_i(x)$ is the i^{th} Lagrange basis polynomial and $l_i'(x_i)$ the derivative of the Lagrange basis polynomial at point x_i. In the original publication, Hermite used Lagrange fundamental polynomials. However, the concept of osculation can be applied to any polynomial as long as the derivatives are known. This paper also examines the osculating standard polynomial (OSP) and the osculating Fourier polynomial (OFP).

2.3.5 Splines

Splines are a set of piecewise polynomials where the derivatives at the endpoints of neighbouring polynomials are equal. Given a set of $n+1$ data points, n number of splines are constructed, one between every neighbouring sample pair. The splines are created as follows:

$$m_s(x) = \begin{cases} s_1(x) & \text{for } x_0 \le x < x_1 \\ \vdots & \\ s_{n-1}(x) & \text{for } x_{n-2} \le x < x_{n-1} \\ s_n(x) & \text{for } x_{n-1} \le x < x_n \end{cases} \quad (15)$$

To ensure a smooth connection between neighbouring splines, the first derivatives at the interior data points x_i have to be continuous, that is, $s_i'(x_i) = s_{i+1}'(x_i)$ for $i \in \{1, 2, \ldots, n\}$ (Wals et al., 1962). As the order of the individual splines increases, higher order derivatives also have to be continuous. The individual splines can therefore be established using any other kind of polynomial function whose derivatives are known, such as standard polynomial splines (SPS) and Fourier polynomial splines (FPS). In practice, cubic splines or lower are mostly used, since higher degree splines tend to overfit the model and reduce the approximation accuracy for intermediate points (Cho, 2007).

2.4 Time Series Models

This section provides an overview of some widely used time series models. The autoregressive (AR), moving average (MA), autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA), autoregressive conditional heteroskedasticity (ARCH) and generalized autoregressive conditional heteroskedasticity (GARCH) models are discussed.

2.4.1 Autoregressive Model

The AR model is an infinite impulse response filter that models a random process where the generated output is linearly depended on the previous values in the process. Since the model retains memory by keeping track of the feedback, it can generate internal dynamics. In recent years a number of AR-based algorithms for gramophone noise removal were proposed which in general provide a good reconstruction accuracy (Niedźwiecki et al., 2014a; 2014b; 2015) . Given y_i as a sequential series of $n+1$ data points, the AR model of degree p predicts the value of a point at time delay t with the previous values of the series, defined as

$$y_t = c + \varepsilon_t + \sum_{i=1}^{p} \alpha_i y_{t-i} \tag{16}$$

where c is a constant, typically considered to be zero, ε_t a white noise error term, almost always considered to be Gaussian white noise, and α_i the coefficients of the model. A common approach is to subtract the temporal mean from time series y before feeding it into the AR model. It was found that this approach is not advisable with sample windows of short durations, since the temporal mean is often not a true representation of the series' mean and can vary greatly among subsets of the series (Ding et al., 2000). The series y is assumed to have a zero mean, whereas non-zero mean series require an additional parameter α_0 at the front of the summation in equation (16). The model coefficients are typically solved using Yule-Walker equations with LLS regression.

2.4.2 Moving Average Model

The moving average is a finite impulse response filter which continuously updates the average as the window of interest moves across the dataset. A study on applying the moving average on random events lead to the formulation of what later became known as the MA model where univariate time series are modelled with white noise terms (Slutzky, 1927). The MA model predicts the value of a data point at time delay t using

$$y_t = \mu + \varepsilon_t + \sum_{i=1}^{q} \beta_i \varepsilon_{t-i} \tag{17}$$

where μ is the mean of the series, typically assumed to be zero, β_i the model coefficients of order q and $\varepsilon_t, \ldots, \varepsilon_{t-q}$ the white noise error terms. Since the lagged error terms ε are not observable, the MA model can not be solved using linear regression. Maximum likelihood estimation (MLE) is typically used to solve the MA model, which in turn is maximized through iterative non-linear optimization methods such as the Broyden-Fletcher-Goldfarb-Shanno (BFGS) (Broyden, 1970) or the Berndt-Hall-Hall-Hausman (BHHH) (Berndt et al., 1974) algorithms.

2.4.3 Autoregressive Moving Average Model

The ARMA model is a combination of the AR and MA models. The ARMA model is based on Fourier and Laurent series with statistical interference (Whittle, 1951) and was later popularized by a proposal describing a method for determining the model orders and an iterative method for estimating the model coefficients (Box and Jenkins, 1970). The ARMA model is given as

$$y_t = c + \varepsilon_t + \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{i=1}^{q} \beta_i \varepsilon_{t-i} \tag{18}$$

where p and q are the AR and MA model orders respectively. The ARMA coefficients are typically approximated through MLE using BFGS or BHHH.

2.4.4 Autoregressive Integrated Moving Average Model

The ARIMA model is a generalization of the ARMA model. ARIMA is preferred over the ARMA model if the observed data shows characteristics of non-stationarity, such as seasonality, trends and cycles (Box and Jenkins, 1970). A differencing operation is added as an initial step to the ARMA model to remove possible non-stationarity. The ARMA model in equation (18) can also be expressed in terms of the lag operator as

$$\alpha(L)y_t = \beta(L)\varepsilon_t \tag{19}$$

where $\alpha(L)$ and $\beta(L)$ are said to be the lag polynomials of the AR and MA models respectively. The ARIMA model is expressed by expanding equation (19) and incorporating the difference operator, $y_t - y_{t-1} = (1-L)y_t$, as follows:

$$\left(1 - \sum_{i=1}^{p} \alpha_i L^i\right)(1-L)^d y_t = \left(1 + \sum_{i=1}^{q} \beta_i L^i\right)\varepsilon_t \tag{20}$$

where p is the AR order, q the MA order and d the order of integration. ARIMA coefficients are approximated with the same method used by the ARMA model.

2.4.5 Autoregressive Conditional Heteroskedasticity Model

ARMA models are the conditional expectation of a process with a conditional variance that stays constant for past observations. Therefore, ARMA models use the same conditional variance, even if the latest observations indicate a change in variation. The ARCH model was developed for financial markets with periods of low volatility followed by periods of high volatility (Engle, 1982). ARCH achieves non-constant conditional variance by calculating the variance of the current error term ε_t as a function of the error terms ε_{t-i} in the previous i time periods. Therefore, the forecasting is done on the error variance and not directly on the previously observed values. The ARCH process for a zero mean series is defined as

$$y_t = \sigma_t \varepsilon_t \qquad \sigma_t = \sqrt{\alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2} \tag{21}$$

where ε_t is Gaussian white noise and σ_t is the conditional variance, modelled by an AR process. Since ARCH makes use of an AR process, the coefficients can be estimated through LLS fitting using Yule-Walker equations. However, since the distribution of ε_{t-i}^2 is naturally not normal, the Yule-Walker approach does not provide an accurate estimation for the model coefficients, but can be used to set the initial values for the coefficients. An iterative approach, such as MLE, is then used to refine the coefficients in order to find a more accurate approximation.

2.4.6 Generalized Autoregressive Conditional Heteroskedasticity Model

The GARCH model is a generalization of the ARCH model which also uses the weighted average of past squared residuals without the declining weights ever reaching zero (Bollerslev, 1986). Unlike the ARCH model which employs an AR process, GARCH uses an ARMA model for the error variance as follows:

$$\sigma_t = \sqrt{\alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2} \tag{22}$$

where α_i and β_i are the model coefficients and p and q the GARCH and ARCH orders respectively. Since GARCH makes use of the ARMA model for the error variance, the model can not be estimated using LLS regression, but has to follow the same estimation approach used by ARMA.

3 METHODOLOGY

This section discuses the optimal parameters of the algorithms, the methodology applied during the analysis, the performance measurement used to compare the algorithms and the evaluation of the execution time.

3.1 Parameter Optimization

All algorithm parameters were optimized using fractional factorial design (Fisher, 1935). Ten songs in each genre were used to find the optimal parameters. The parameter configurations that on average performed best over all 80 songs were used to calculate the reconstruction accuracy of the entire set of 800 songs. The optimal parameters of the benchmarking are given below in the format [w,o,d], where w is the windows size, o the order and d the derivatives.

- NNI: [2, -, -]
- SI: [284, -, -]
- STP: [2, 1, -]
- OSP: [6, 2, 1]
- SPS: [4, 1, -]
- FOP: [250, 1, -]
- OFP: [270, 10, 9]
- FPS: [2, 1, -]
- NEP: [2, -, -]
- HEP: [2, -, -]
- AR: [1456, 9, -]
- MA: [4, 1, -]
- ARMA: [1456, 9-2, -]
- ARIMA: [1440, 9-1-4, -]
- ARCH: [8, 1, -]
- GARCH: [8, 1-1, -]

3.2 Empirical Dataset

The dataset was divided into eight genres, namely classical, country, electronic, jazz, metal, pop, reggae and rock, consisting of 100 tracks each. The songs were encoded in stereo with the Free Lossless Audio Codec (FLAC) at a sample rate of 44.1 kHz and a sample size of 16 bits. Figure 1 shows typical distortions in the sound wave caused by scratches on the gramophone record. The duration of these disruptions, which will be referred to as *gap sizes*, is typically 30 samples or shorter. To accommodate longer distortions, gap sizes of up to 50 samples were analysed. Although the algorithms are able to reconstruct a gap of any duration, disruptions longer than 50 samples rarely occur and were therefore omitted from the results.

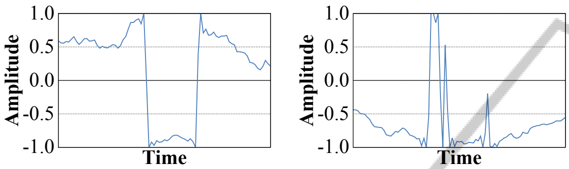

Figure 1: Disruptions caused by scratches on gramophones.

3.3 Reconstruction Performance

Each song was recorded in its original state without any disruptions. The records were physically damaged and rerecorded to generate the noisy signals. A detailed discussion on the noise acquisition, detection and noise masking processes is given in (Stallmann and Engelbrecht, 2015a; 2015b) . The reconstructed signal was compared to the original recording to determine the quality of interpolation. The reconstruction performance was measured using the normalized root mean squared error (NRMSE), defined as

$$\text{NRMSE} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\tilde{y}_i - y_i\right)^2}}{\hat{y} - \check{y}} \tag{23}$$

for a set of n samples, where y_i are the samples of the original signal and $\tilde{y}_i$ is the samples of the reconstructed signal. $\hat{y}$ and $\check{y}$ are the maximum and minimum amplitudes of the original signal respectively. A perceptual evaluation of the audio quality was also conducted,

3.4 Execution Time

In addition to the reconstruction accuracy, the algorithms were also compared according to their execution time. The time was measured as the number of seconds it takes to process a second of audio data, denoted as s\s. Hence, a score of 1 s\s or lower indicates that the algorithm can be executed in real time. Based on the concept of the scoring metric in (Sidiroglou-Douskos et al., 2011), in order to evaluate the tradeoff between the reconstruction accuracy κ and the execution time τ, the speed-accuracy-tradeoff (SAT) is calculated using

$$\text{SAT} = \left(\frac{\kappa}{\hat{\kappa} - \check{\kappa}} + \frac{\tau}{\hat{\tau} - \check{\tau}}\right)^{-1} \tag{24}$$

$\hat{\kappa}$ and $\check{\kappa}$ are the NRMSEs of the best and worst performing algorithms respectively. $\hat{\tau}$ and $\check{\tau}$ are the computational times of the fastest and slowest algorithms respectively. Benchmarking was conducted on a single thread using an Intel Core i7 2600 at 3.4 GHz machine with 16 GB memory.

4 EMPIRICAL RESULTS

Figure 2 shows the reconstruction accuracy of the duplication and trigonometric interpolation approaches for an increasing gap size. AWI and MWI achieved a good interpolation for small gap sizes, but quickly declined as the gap increased. LI struggled to reconstruct small gaps, but still outperformed AWI and MWI for gaps of five samples and larger. CI had the best accomplishment for gaps of two samples and greater. The reconstruction performance of the algorithms in figure 2 for different genres is given in figure 3. CI and AWI achieved the best and worst results respectively for all genres. NNI and SI outperformed LI in all genres, except classical music.

Figure 4 illustrates the interpolation NRMSE for the examined polynomials over an increasing gap size. The STP, SPS, NEP and HEP achieved almost identical results. Since the best performing spline interpolation utilizes linear piecewise polynomials, the interpolation accuracy of the STP and SPS are equivalent. The STP does not benefit when employed in an osculating fashion. However, the FOP improved with the inclusion of derivatives and clearly benefited when applied as splines. An opposite trend is observed for the FOP and OFP, where smaller gaps were more difficult to interpolate than larger gaps. This trend is caused by a high frequency FOP fitted over smaller gaps. As the gap size increases, the frequency of the sine and cosine waves decreases, providing a smoother interpolation. Figure 5 shows the polynomials' interpolation performance for the different genres. FOP and OFP had a clear inflation compared to the other algorithms. OSP also had a slight surge, with the rest of the algorithms achieving a similar interpolation over all genres.

The time series models' interpolation performance for a growing duration is illustrated in figure 6. The AR and ARMA models had a similar performance, whereas the ARIMA model started deviating from the trend with gaps wider than eight samples. The ARCH and GARCH models had an identical trend, indicating that there is no difference between using and AR or ARMA process to predicting

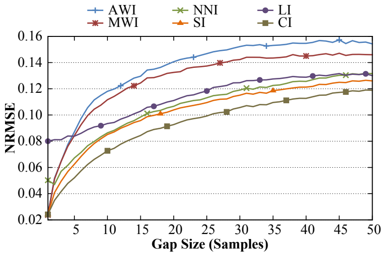

Figure 2: The reconstruction of duplication methods for different gap sizes.

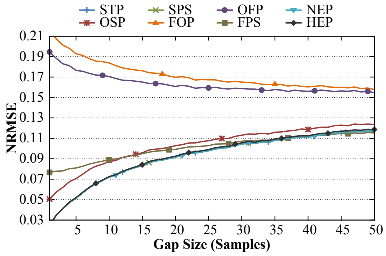

Figure 4: The reconstruction of polynomials for different gap sizes.

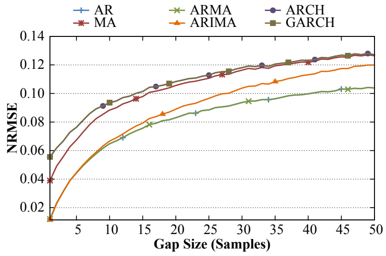

Figure 6: The reconstruction of time series models for different gap sizes.

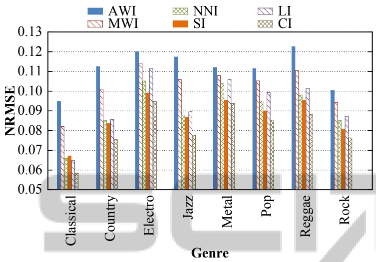

Figure 3: The reconstruction of duplication methods for different genres.

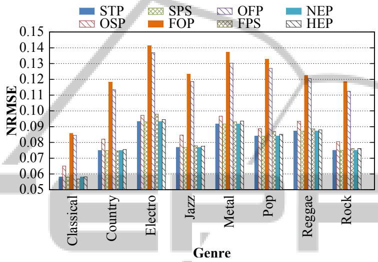

Figure 5: The reconstruction of the polynomials for different genres.

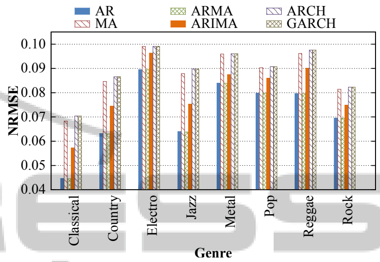

Figure 7: The reconstruction of the time series models for different genres.

a music signal's variance. On average a music signal's variance is comparatively low compared to high volatile financial markets for which the ARCH and GARCH models were originally intended and therefore do not perform as well as the AR and ARMA models. The corresponding genre comparison for the models in figure 6 is given in figure 7. The AR and ARMA models performed best for all genres, followed by the ARIMA, MA and then the ARCH and GARCH models. Table 1 shows the overall reconstruction accuracy, execution time and the tradeoff between the accuracy and computational time calculated using equation (24). The ARMA model performed the best on average, with just a minor improvement over the AR model. All algorithms, except the OFP and ARIMA model, can be executed in real time using a single thread. CI was on average the fastest. The tradeoff in the last column shows that the AR model achieved the most effective results for the given execution time. Although not a major improvement over the ARMA model, the AR had a considerable lower computation time, since it was estimated with a LLS fit and not an iterative gradient-based algorithm.

The reconstruction process was also perceptually evaluated. The refurbished songs had a pleasant listening experience with little noise in the background, mostly restricted to the song segments with a narrow dynamic range, such as rests which are prominent in classical music. Due to the subjective nature, different hearing ranges and the participants' difficulties to distinguishing between the interpolation of some algorithms, the NRMSE was used as the principal measurement of the reconstruction accuracy.

Table 1: The average reconstruction accuracy, execution time and tradeoff for the interpolation algorithms.

Algorithm	NRMSE	Time (s\s)	SAT
AWI	0.111371	0.049794	1.098892
MWI	0.104806	0.051691	1.167703
NNI	0.090748	0.027313	1.348714
SI	0.087269	0.054413	1.402265
LI	0.093215	0.027908	1.313032
CI	0.081218	**0.027128**	1.506959
STP	0.080057	0.031490	1.528750
OSP	0.086014	0.034523	1.422878
SPS	0.080057	0.038588	1.528683
FOP	0.122412	0.068523	0.999724
OFP	0.117923	5.325502	1.015416
FPS	0.081493	0.033035	1.501807
NEP	0.080058	0.027329	1.528749
HEP	0.081066	0.027557	1.509767
AR	0.071764	0.092778	**1.704671**
MA	0.087952	0.029541	1.391575
ARMA	**0.071709**	2.435243	1.678909
ARIMA	0.080201	6.808781	1.464877
ARCH	0.089057	0.062245	1.374061
GARCH	0.089057	0.062378	1.374060

5 CONCLUSION

Twenty interpolation algorithms were analysed and benchmarked against each other in order to determine their reconstruction ability on disrupted gramophone recordings. Different approaches to interpolation were considered, including duplication and trigonometric methods, polynomials and time series models. It was found that the ARMA model performed the best with an average NRMSE of 0.0717. The CI had the fastest execution time at 0.0271 s\s. The AR model was the most effective approach by achieving the best interpolation for a given time limit.

Future work includes the analyses of more complex models, such as neural networks, that may increase the interpolation accuracy. Further research has to be done in other areas of audio processing, such as VoIP, in order to determine how well the examined algorithms perform with other types of noise and different audio sources, such as speech instead of music.

REFERENCES

Berndt, E. K., Hall, B. H., Hall, R. E., and Hausman, J. A. (1974). Estimation and Inference in Nonlinear Structural Models. In *Annals of Economic and Social Measurement*, volume 3, pages 653–665.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.

Box, G. E. P. and Jenkins, G. M. (1970). *Time Series Analysis: Forecasting and Control*. Holden-Day Series in Time Series Analysis. San Francisco, CA.

Broyden, C. G. (1970). The Convergence of a Class of Double-rank Minimization Algorithms. *IMA Journal of Applied Mathematics*, 6(1):76–90.

Cho, J. (2007). *Optimal Design in Regression and Spline Smoothing*. PhD thesis, Queen's University, Canada.

CMU (2011). SoundScan may be under reporting US vinyl sales. www.thecmuwebsite.com/article/soundscan-may-be-under-reporting-us-vinyl-sales. 2015-04-16.

Ding, M., Bressler, S., Yang, W., and Liang, H. (2000). Short-window Spectral Analysis of Cortical Event-related Potentials by Adaptive Multivariate Autoregressive Modeling. *Biological Cybernetics*, 83:35–45.

Duchon, C. E. (1979). Lanczos Filtering in One and Two Dimensions. *Journal of Applied Meteorology*, 18(8):1016–1022.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4):987–1007.

Fisher, R. A. (1935). *The Design of Experiments*. Oliver and Boyd, Edinburgh, United Kingdom.

Gearhart, W. B. and Shultz, H. S. (1990). The Function sin x/x. *The College Mathematics Journal*, 21(2):90–99.

Kastner, R., Hosangadi, A., and Fallah, F. (2010). *Arithmetic Optimization Techniques for Hardware and Software Design*. Cambridge, United Kingdom.

Newton, I. and Whiteside, D. T. (2008). *The Mathematical Papers of Isaac Newton*, volume 1 of *The Mathematical Papers of Sir Isaac Newton*.

Niedźwiecki, M. and Ciołek, M. (2014a). Elimination of Impulsive Disturbances from Stereo Audio Recordings. In *European Signal Processing Conference*, pages 66–70.

Niedźwiecki, M. and Ciołek, M. (2014b). Localization of Impulsive Disturbances in Archive Audio Signals using Predictive Matched Filtering. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2888–2892.

Niedźwiecki, M., Ciołek, M., and Cisowski, K. (2015). Elimination of Impulsive Disturbances From Stereo Audio Recordings Using Vector Autoregressive Modeling and Variable-order Kalman Filtering. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23(6):970–981.

Niedźwiecki, M. and Cisowski, K. (2001). Smart Copying - A New Approach to Reconstruction of Audio Signals. *IEEE Transactions on Signal Processing*, 49(10):2272–2282.

Richter, F. (2014). The LP is Back! www.statista.com /chart/1465/vinyl-lp-sales-in-the-us. 2015-04-16.

Sidiroglou-Douskos, S., Misailovic, S., Hoffmann, H., and Rinard, M. (2011). Managing Performance vs Accuracy Trade-offs with Loop Perforation. In *ACM SIGSOFT Symposium*, pages 124–134. ACM.

Slutzky, E. (1927). The Summation of Random Causes as the Source of Cyclic Processes. *Econometrica*, 5(2):105–146.

Sprechmann, P., Bronstein, A. M., Morel, J.-M., and Sapiro, G. (2013). Audio Restoration from Multiple Copies. In *International Conference on Acoustics, Speech and Signal Processing*, pages 878–882.

Stallmann, C. F. and Engelbrecht, A. P. (2015a). Digital Noise Detection in Gramophone Recordings. In *ACM International Conference on Multimedia*, Brisbane, Australia. In review.

Stallmann, C. F. and Engelbrecht, A. P. (2015b). Gramophone Noise Detection and Reconstruction Using Time Delay Artificial Neural Networks. *IEEE Transactions on Systems, Man, and Cybernetics*. In review.

Wals, J. L., Ahlberg, J., and Nilsson, E. N. (1962). Best Approximation Properties of the Spline Fit. *Journal of Applied Mathematics and Mechanics*, 11:225–234.

Whittle, P. (1951). *Hypothesis Testing in Time Series Analysis*. PhD thesis, Uppsala University, Uppsala, Sweden.

Wile, R. R. (1990). Etching the Human Voice: The Berliner Invention of the Gramophone. *Association for Recorded Sound Collections Journal*, 21(1):2–22.