```
/*****************************************************************/
/* cadence(L, Ton, Mod) - L forms a cadence of Ton - Mod          */
/*****************************************************************/

cadence([H|_], X, Y) :- theor(H, [X, [c], Y]).
cadence([_|T], X, Y) :- cadence(T, X, Y).

theor([H|T], [X, F, Y]) :- theor(T, [X, L, Y]),
    prove(H, [X, L, Y], [X, F, Y]).
theor([[X, F, Y]], [X, F, Y]).

prove([X, L1, Y], [X, [F|T2], Y], [X, L2, Y]) :-
    invert(L1, [F, d|T1]), invert(T1, T1i), append(T1i, T2, L2).
prove([X, L1, Y], [X, [F, e|T2], Y], [X, L2, Y]) :-
    invert(L1, [F|T1]), invert(T1, T1i), append(T1i, T2, L2).

invert([H|T], L) :- invert(T, Ti), append(Ti, [H], L).
invert([], []).

/*****************************************************************/
```

Figure 4: A Theorem Prover for Functions of Chords

An Optimized Method for Storage, Transmission and Display of Digital Audio Data

CARLOS AUGUSTO PAIVA DA SILVA MARTINS
RUGGERO ANDREA RUSCHIONI
SÉRGIO TAKEO KOFUJI
[capsm, roger, kofuji]@lsi.usp.br
Escola Politécnica da Universidade de São Paulo
LSI - Laboratório de Sistemas Integráveis
Divisão de Sistemas Digitais
Av. Professor Luciano Gualberto, 158, Travessa 3
05508-900 - São Paulo - SP - Brasil

Abstract

This paper analyses the application of a new reconstruction method for digital audio signals. The method is called Normalized Sampled Finite Sync Reconstructor (NSFSR) and its behavior is closer to the ideal reconstructor used in the original sampling theorem. When compared to the reconstruction method currently used, the Zero Order Reconstructor (ZOR), it substantially reduces storage size, transmission bandwidth and synthesis time. Both qualitative and quantitative analysis of the methods are presented.

Introduction

This paper analyses the application of a new method of digital audio signal reconstruction. The goal is to reduce the storage size and consequently the transmission bandwidth the synthesis time. The reconstruction method was designed with the intent of surpassing the quality of the popular Zero Order Reconstructor (ZOR), which produces jag effects in one dimensional signals [Martins, 1994].

Digital audio signals are currently used in many computer applications, including entertainment and telecommunications. In addition our research is important for future applications such as computer music, virtual reality, distributed multimedia, scientific visualization and digital television, among others.

Most existing signals of importance to the humans, such as audio or static and dynamic images, may be considered analog at the macroscopic level. Nonetheless, for these signals to be manipulated and processed by digital computers, they need to be transformed into digital signals through sampling, quantizing and codification. Since our final objective is the display of the analog signal, after all the digital signal processing the data must be converted back to the analog domain.

According to the sampling theorem [Shannon, 1949], the minimum sampling frequency must be twice the frequency of the highest component in the analog signal. However it is common practice to use sampling frequencies well above this minimum. Reconstructors often fail to meet theoretical performance levels, and higher sample rates simplify the design of the analog lowpass filter used after digital to analog conversion.

Currently almost all systems that work with digital signals use digital to analog converters (reconstructors) connected to the output devices, that can be modeled as a zero order reconstuctor (ZOR). Since this is not the ideal reconstructor used in the sampling theorem, errors occur in the reconstruction.

Our main objective is to show it is possible to use sampling frequencies closer to the minimum and obtain better results than by employing the ZOR, using the NSFSR.

We must remember that audio signals are a subclass of the one dimensional signals which present unique features related to human perception. What is an essential characteristic of audio signals is not essential in other one dimensional signals. In this paper we will focus more on the general characteristics of one dimensional signals which are also valid for audio signals, such as storage size requirements, bandwidth for transmission and the quality of the generated signal.

We compare the use of the ZOR and the NSFSR on audio signals through the analysis of the reconstructed signals in the time and the frequency domains.

Results are analyzed qualitatively, comparing the reconstructed signals' wave form and sound, and quantitatively comparing size and the complexities of the reconstruction methods.

Reconstruction Method

The reconstruction method presented here is named "Normalized Sampled Finite Sync Reconstructor" (NSFSR). It is a new reconstructor which matches more precisely the ideal reconstructor proposed in the sampling theorem (Shannon, 1949) than the Zero Order Reconstructor (ZOR) used in practice.

It is very important to remember that the sampling process corresponds to a time domain multiplication between the original signal and the sampling impulse train. In an equivalent form it corresponds to a convolution between the Fourier transforms of the original signal and the sampling impulse train. In this way the sampled signal in the frequency domain is formed by copies of the original analog signal in the frequency domain. This copies are centered at points which are multiples of the sampling frequency employed.

The reconstruction process from the sampled signal corresponds to a time domain convolution between the sampled signal and the reconstruction filter. Equivalently it corresponds to a multiplication between the Fourier transforms of the sampled signal and of the reconstruction filter. According to the sampling theorem, using a sampling frequency above Nyquist limits and using the ideal reconstructor the reconstructed sampled signal should be identical to the original. The demonstration of this affirmation is the proof of the sampling theorem. Detailed mathematical analyses may be found in [Brigham, 1988], [Oppenheim, 1983], [Oppenheim, 1989] and [Shannon, 1949].

The ideal reconstructor, named time domain sync, has infinite support, and cannot be implemented in practice. The NSFSR is generated by sampling the ideal sync reconstructor in the time domain with a finite number of points; it is later on normalized by dividing each sample by the sum of the values of all samples. Once the number of samples is defined in the reconstructor, on each side of the interpolated point will fall half the points which form the ideal sync reconstructor, sampled and normalized. A detailed analysis of this reconstructor is presented in [Martins, 1994], chapter 7.

Results

We applied the ZOR and NSFSR in one dimensional signals and analyzed them. For the reconstructors implementation, the software package Matlab was used; for the analyses of the results of the reconstructors in the signals Matlab and the Khoros environments were used. Here we present the analyses of a low frequency sine wave. Nevertheless it is important to mention that the conclusions are also valid for other complex signals as presented in more extensive tests involving different signals in [Martins, 1994], chapter 6. In the analyses we considered both subjective quality factors and quantity factors in time and frequency domain.

Qualitative analyses

Proceeding with the quality analyses, disregarding sound perception factors, we may observe some important features in the reconstructed signals employing the ZOR and the NSFSR. Figures 1 to 4 show us the zoom of the original analog signal and the time domain reconstruction. We see that the signals generated by the reconstructors NSFSR 2 points and NSFSR 6 points do not show any perceivable jag effect as in the one generated by ZOR. Analyzing visually the images 3 and 4 we see that there is no jag effect for the NSFSR 2 points in figure 3 and for the NSFSR 6 points in figure 4, seeming that NSFSR 2 points shows better results.

Analyzing figures 5 to 8, which show the original signal and the reconstructed signal in the frequency domain, we notice that in the signal generated by the ZOR the frequency components in the repetition points generated by sampling, have magnitude approximately 100 times greater than the ones generated by the NSFSR 2 points and approximately 20 times greater than the ones generated by the NSFSR 6 points. These high frequency components, of quite significant magnitude in the replication points generated by sampling, are what characterizes the existence of jag effects in the time domain.

The reconstruction error introduces high frequency components not present in the original signal, changing the spectrum of the original signal and eventually its timbre. The additional frequency components are a consequence of a non ideal lowpass filtering, because we used a zero order reconstructor, while sampling errors (aliasing) introduces low frequencies.[Brigham, 1988; Oppenheim, 1975].

Differences between the original signal and the reconstructed with the family of NSFSR in the frequency domain is smaller than between the original and the ZOR signal. The NSFSR 2 points present the smaller magnitude in the replication points.

Observing figures 9 to 11, which show the absolute difference signal between the original and reconstructed signals, we notice that the best results are presented by the NSFSR 2 points which yields a maximum value of about 1.2×10^{-3} We should warn that by itself this measurement is not sufficient to select a better reconstructor.

We concluded by analyzing visually the wave forms that best results are obtained through the NSFSR 2 points, and the analyses by the difference signal also point to this as the best reconstructor.

Also should be stressed that either NSFSRs are far better than the ZOR and apparently do not present oscillation problems (Gibbs effect).

A qualitative analysis based on the sound of the reconstructed signals presented by the ZOR and the two NSFSRs shows that the introduction of high frequency partials on the spectrum of the original signal by the ZOR changes the timbre considerably, while the NSFSR reconstruction preserves the original sound characteristics. Both the original sound and the reconstructed through the NSFSR 2 points are audibly indistinguishable from each other, while the signal reconstructed with the ZOR presents complex spectra due to the artifacts (high partials) introduced by the reconstruction error of the ZOR. So we conclude by perceptive analyses that the NSFSR 2 points also yield the best results.

Quantitative analysis

We carried on some tests to analyze the transmission of audio signals quantitatively on an actual network. The aim of such tests is to verify, in broad terms, the times expended in the transmission of signals with different samples. We simulated the transmission of an audio signal with a number of samples far greater than the minimum established by the sampling theorem and the signal critically sampled. The smallest signal is later interpolated [Oppenheim, 1989] using the ZOR and the NSFSR.

The files used in the process had sizes of 262.144 bytes and 16.384 bytes, the latter is 16 times smaller than the former.

Tests were realized using a local area network configuration (same building) and long distance (two cities 240 Km apart). We transmitted both files under two net traffic conditions, heavy load and light load. We should warn that due to the illustrative purposes of the measurements we didn't monitor the net load during the transmissions, the classification was based on the transmission's time.

Based on the obtained results, we calculated the sample average and using the mean value estimate method calculated the confidence interval for the distributed average, supposed normal. For the calculation of the confidence interval it was used an uncertainty coefficient of $\alpha = 0.05$.

In spite of the limited amount of transmissions, the sampled average obtained is very similar to the probable distribution average. The values for the standard deviation for local transmission were greater than the ones obtained for long distance. This happens because the transmission time in local networks are considerably faster and so more influenced by load variations. For the same reason, on a net under the same load, the transmission times for smaller files present greater deviation than average when compared to files of greater size. The average of the transmission times had a very small variation under conditions of heavy and light load, this suggests that the net did not present definite periods of heavy and light load or that these periods do not follow a known pattern.

The gain values obtained during transmission time, considering the average, were: 5.01 for local net with light load, 5.03 for local net with heavy load, 18.8 for long distance net with light load and 19.5 for long distance net with heavy load.

We noticed the gains should equal 16 if the behaviors of the system were deterministic, which is not true for the net configurations used in this tests. Nevertheless the storage gains are deterministic and always equal 16.

Comparing quantitatively the complexity of the reconstruction algorithms tested, we notice that the NSFSR presents greater complexity in the order of approximately 8 operations per sample against 1 for the ZOR. In software implementations this is a considerable drawback.

At present we are developing an implementation of the NSFSR in hardware, in this case the complexity drawback is unimportant because we use high speed dedicated hardware and optimized parallel structures to operate in real time even for very high sample rates.

Conclusion

The reconstruction error introduces high frequencies in the reconstructed signal, while sampling error (aliasing) introduces low frequencies (Oppenheim, 1975; Oppenheim, 1983; Proakis, 1992).

The quality analyses of the reconstructed signals show that applying the new method of NSFSR in digital audio signals yields superior results when compared to the currently used ZOR. At least for this particular signal the timbre differences were noticeable, we still don't know how a ZOR with a properly designed lowpass filter might affect the reconstruction of complex signals such as those found in recordings of natural instruments. This study provides an indication, though, that there might be timbre distortion introduced by the reconstructor. We plan to do further tests on recordings of natural instruments and compare the quality of the two reconstructors.

The transmission of digital audio data is a crucial point in the above mentioned applications, the gain obtained in the transmission of files with lesser samples are considerable. For most current multimedia applications and net transmissions which use sampling rates as low as 8KHz to 16KHz, the gain in sound quality is expressive. We conclude that the transmission of signals with the minimum number of samples drawn up by the sampling theorem - which is the method used - reduces the importance of this crucial parameter, allowing the practicability of applications which demand greater flux of digital audio signals.

Considering the advantage of transmitting a digital audio signal with a smaller number of samples, which does not violate the minimum sampling value, we still need to choose the method to be used for reconstruction. Even though the NSFSR has a greater algorithmic complexity in relation to the ZOR, the qualitative and quantitative analyses show that the NSFSR is still the best solution. For transmission purposes also we should note that the approach here presented is independent of any compression scheme. Further compression techniques could well be employed over the signal minimizing even further the file size.

So we reached our main objective, showing that by using the NSFSR it is possible to use a lower sampling frequency and obtain better results than the ones we currently obtain using the ZOR.

Furthermore we should also point that using the NSFSR we might not need to use the lowpass filters, which have a complex design. In the ideal case we would not need to use such a lowpass filter after the digital to analog conversion.

Based on all the results, which are expressive and promising, we believe that using the NSFSR on audio signals is desirable and presents various advantages over the current approach. However many tests and development are necessary, this is the beginning of a new research of medium-long term.

Future Works

As said above, audio signals present unique aspects related to perception. In this paper we focused on general aspects valid for most one dimensional signals. It is our intent to carry on deeper research on the psychoacoustic effects of the reconstructor, particularly the effects on timbre. We believe to have enough evidence that the current approach using the ZOR might be introducing unwanted high frequency components in some audio applications. We will proceed on the comparative study of both reconstructors and their effect on recordings of complex sounds such as natural instruments.

The algorithm employed for the NSFSR has greater complexity than the ZOR. In the tests presented in this paper we used a software implementation. We have already started working on a hardware implementation of the NSFSR, in this case the computing drawback is irrelevant since we are using high speed dedicated hardware and optimized parallel structures for real time applications even for high sampling rates.

Lastly, regarding the reconstruction method, we are continuing our research developing other reconstructors better optimized than the NSFSR used in this work.

References

Brigham, E. O. (1988). *The fast Fourier transform and its applications*. Englewood Cliffs, NJ: Prentice-Hall Inc.

Oppenheim, A.V. & Schafer, R. W. (1975). *Digital signal processing*. Englewood Cliffs, NJ: Prentice-Hall Inc.

Oppenheim, A.V. & Willsky, A. S. (1983). *Signals and Systems*. Englewood Cliffs, NJ: Prentice-Hall Inc.

Oppenheim, A.V. (1989). *Discrete-time signal processing*. Englewood Cliffs, NJ: Prentice-Hall Inc.

Martins, C. A. P. S. (1994). *O efeito de serrilhamento em dispositivos matriciais de exibição de imagens como erro de reconstrução*. Dissertação (Mestrado), PPGEE-UFMG, Belo Horizonte.

Proakis, J. G. & Manolakis, D. G. (1992). *Digital signal processing: principles, algorithms, and applications*. 2nd Edition, New York, Macmillan Publishing Company.

Shannon, C. E. (1949). Communication in the Presence of Noise. *Proceedings of the I.R.E.*, jan 1949, 10-21.

Acknowledgments

The reconstruction method used in this paper was developed as part of the research necessary to obtain the title of Master of Science of one of the authors at PPGEE-UFMG. This work was sponsored by CAPES. The authors would like to thank the researchers Luis Gustavo G. Kiatake and Marcelo H. Cintra for their collaboration on the digital audio transmission tests. This research is being developed at LSI-USP, the authors wish to express their gratitude for the support of this laboratory and its staff.

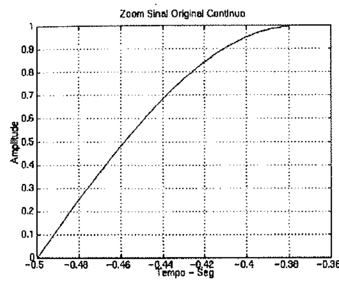

Fig. 1 Original signal zoom

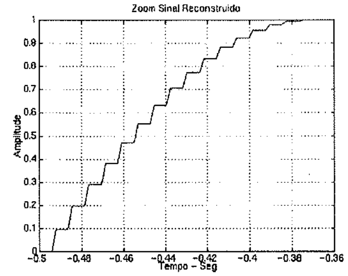

Fig. 2 ZOR signal zoom

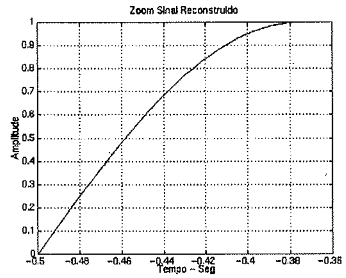

Fig. 3 NSFSR 2p signal zoom

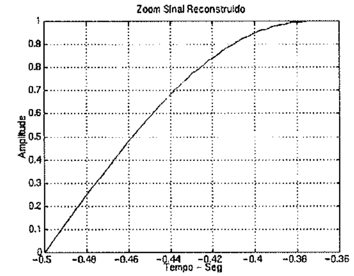

Fig. 4 NSFSR 6p signal zoom

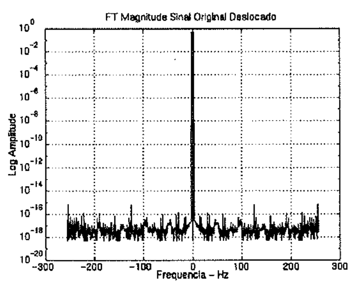

Fig. 5 Original signal FT mag log

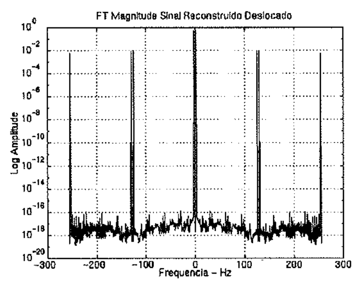

Fig. 6 ZOR signal FT mag log

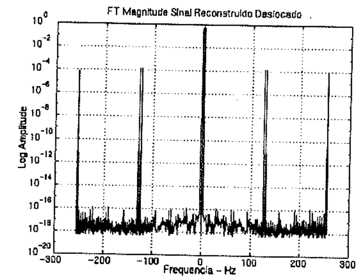

Fig. 7 NSFSR 2p signal FT mag log

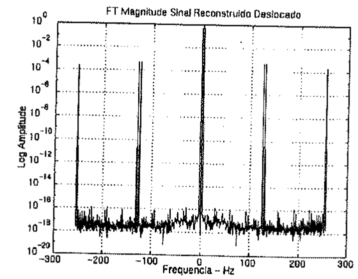

Fig. 8 NSFSR 6p signal FT mag log

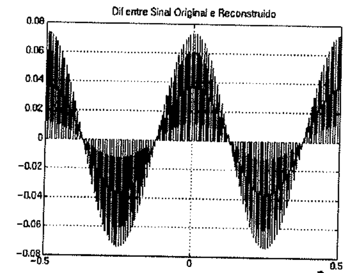

Fig. 9 Original - ZOR dif signal

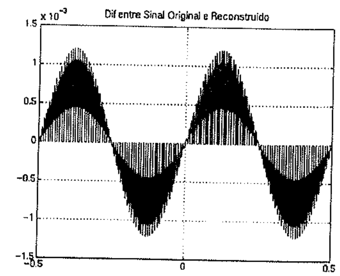

Fig. 10 Original - NSFSR 2p dif signal

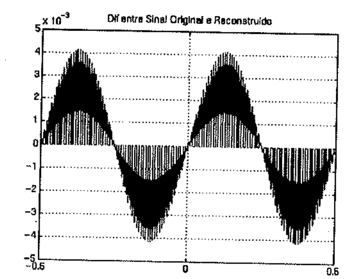

Fig. 11 Original - NSFSR 6p dif signal