High-frequency Noise Removal of Vintage Songs in Audio Cassettes using Signal Processing Techniques

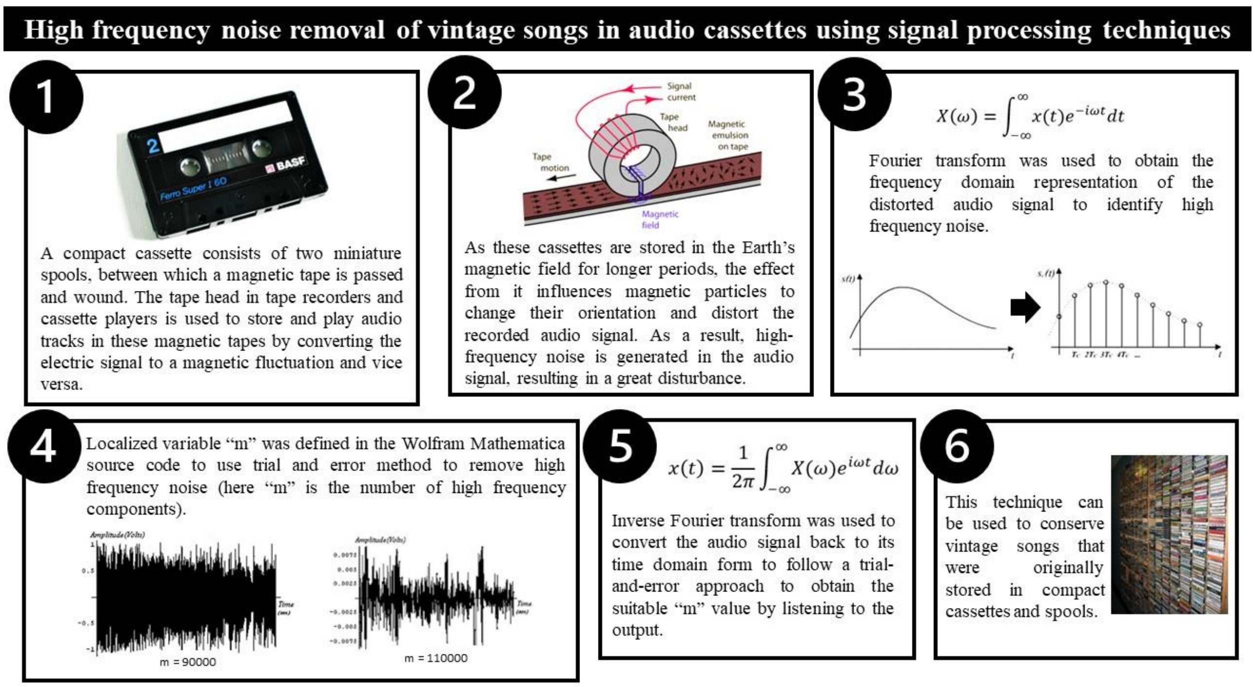
M. M. P. M. Fernando*, N. K. Kularathne

Department of Physics, University of Sri Jayewardenepura, Sri Lanka

**Corresponding Author: pmadhuranga@sjp.ac.lk* +940714478729

Date Received: 26-09-2022 Date Accepted: 17-04-2023

Graphical abstract



Abstract

From 1960 to the early 2000s, the audio cassette was one of the three most used storage devices for prerecorded music, alongside gramophone records (LPs) and compact discs (CDs). A compact cassette consists of two miniature spools, between which a magnetic tape is passed and wound. The tape head in tape recorders and cassette players is used to store and play audio tracks in these magnetic tapes by converting the electric signal to a magnetic fluctuation and vice versa. The sound quality of an audio track recorded on a compact cassette is decent. However, as these cassettes are stored in the Earth's magnetic field for longer periods, the effect from it influences magnetic particles to change their orientation and distort the recorded audio signal. As a result, high-frequency noise is generated in the audio signal, resulting in a great disturbance when listening to songs and other audio recordings on vintage cassettes.

Therefore, the removal of such noises while minimizing the impact on the original audio signal will make sure the listener gets to enjoy the original quality of the vintage audio recordings. Fourier transform of a given distorted audio file (in “wav” format) was obtained to identify the high-frequency noise using Wolfram Mathematica 12.3. Then a variable “m” was introduced to the program to identify magnitudes of high-frequency components and remove their effect on the original audio recording. A trial-and-error approach was taken to identify the adequate “m" value for each channel of a given audio signal. Vintage songs were used to evaluate the accuracy of the developed method in this study. It was identified that the suitable “m” value changes for each channel in each song. Therefore, it is impossible to state a constant value for the variable “m”. However, the accuracy of the output also depends on the hearing frequency of the person, and the audio equipment being used. As an application, this technique can be used to conserve vintage songs that were originally stored in compact cassettes and spools, because it allows the user to control the number of high-frequency components removed from the distorted audio signal.

Keywords: *Signal Processing Techniques, Fourier Transforms, Mathematica, Noise Removing, Audio Cassettes*

1. Introduction

1.1 Compact Cassette (Audio Cassette)

The compact cassette was first introduced by the “Philips” cassette company for audio storage in Europe (1963) and the United States (1964). From 1960 to the early 2000s, the cassette was one of the three most used storage devices for prerecorded music, alongside gramophone records (LPs) and compact discs (CDs). A compact cassette consists of two miniature spools, between which a magnetic tape is passed and wound. These spools and their attendant parts are held inside a protective plastic shell. The magnetic tape inside the compact cassette consists of two stereo pairs of tracks (four total) or two monaural audio tracks. A stereo pair or a single monophonic track is played or recorded when the tape is moving in one direction, while another one does the same when the tape moves in the other direction (Andriessen, 1999; Bohlman, 2017)

Compact cassettes can be categorized based on the magnetic material used in the tape. The gamma ferric oxide (γ-Fe_2O_3) was used as the original magnetic material in compact cassettes and its performances were enhanced by the 3M company using a cobalt volume doping process combined with a double-coating technique (Daniel et al., 1999). Later, chromium dioxide (CrO_2) was also introduced as a magnetic material for tapes, and it was coated with magnetite ($Felon_4$) to enhance the sound quality by TDK. In

1979 pure metal particles were also introduced as a magnetic material by 3M under the trade name Metafine (Daniel et al., 1999).

The tape head in tape recorders and cassette players is used to store and play the audio tracks in these magnetic tapes by converting the electric signal to a magnetic fluctuation and vice versa. Generally, audio recorders and players have a fixed tape head, which consists of a core of magnetic material arranged into a toroid with a narrow gap of insulating material. This narrow gap is used to spill the magnetic flux out of the material to magnetize the magnetic tape used in the cassette. As shown in figure 1, there is a coil of wire wrapped around the core opposite the narrow gap in the tape head to vary the magnetic flux with the receiving electric signal when recording an audio track to a compact cassette. In the case of playback, an electric current is induced in the coil by the varying magnetic field in the gap (Barnard, 1972; Van Bogart, 1995).

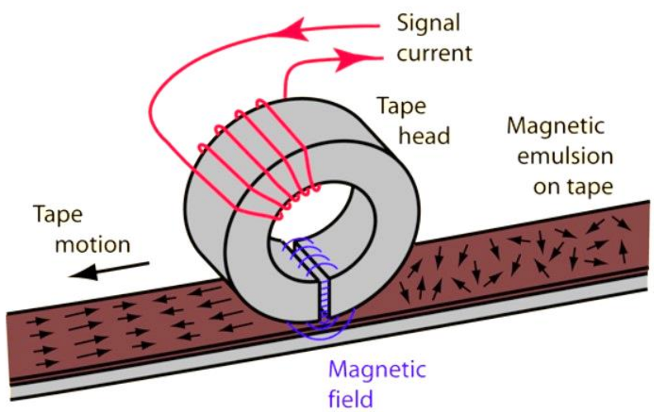


Figure 1: Tape head action

The sound quality of an audio track recorded on a compact cassette is decent. However, as these cassettes are stored in the Earth's magnetic field for longer periods, the effect from it influences magnetic particles to change their orientation and distort the recorded audio signal (Nasuta, 1981). Therefore, after a while, it is common to hear a noise (somewhat like "S… s… s…") in audio tracks recorded on compact cassettes and spools. This added noise is a great disturbance when listening to vintage songs and other audio recordings. Therefore, the removal of such noises while minimizing the impact on the original audio signal will make sure the listener gets to enjoy the original quality of vintage audio recordings (Chen et al., 2008). Digital signal processing techniques can be applied to fulfill this requirement.

1.2 Digital Signal Processing (DSP)

The application of signal processing techniques in digital devices is referred to as digital signal processing (DSP). A wide variety of digital signals including audio, sonar and radar, sensor readings, biomedical signals, etc. can be processed using DSP techniques (Oppenheim, 1978; Vaseghi, 2008). Digital signals can be studied in the time domain (1D signals), spatial domain (multidimensional signals), frequency domain, autocorrelation domain, and wavelet domain. The frequency domain representation should be obtained when analyzing the properties of a signal originally in the time or spatial domain (Hassanpour, 2008). Fourier transform can be used to complete this conversion and obtain the magnitude and phase component information of each frequency in such a signal. However, to apply Fourier transform to an analog signal (for example a song), first, it should convert into its digital form. For that sampling is used to obtain a discrete signal from a continuous signal (Rabiner & Gold, 1975) as shown in figure 2.

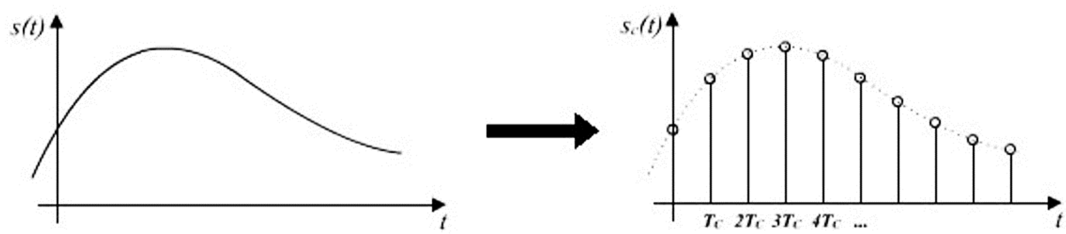


Figure 2: Analog signal and resulting sampled signal

The selection of an adequate sampling rate is vital when converting an analog signal to its digital form. Therefore, the Nyquist-Shanon sampling theorem is used. It states the perfect reconstruction of a signal is possible when the sampling frequency at least exceeds the highest frequency of the original signal by a factor of two (Landau, 1967). After converting the analog signal to its digital form in the time or spatial domain, the Fourier transform can be applied to obtain the frequency domain (K – space) representation of it (Harčarik et al., 2012).

1.3 Fourier Transform

A signal represented in the time domain can be converted into the frequency domain using the Fourier transform. This mechanism was introduced by French mathematician Joseph Fourier. As no information is added or lost during the Fourier transform process, the original time domain signal can be reobtained after performing the necessary analysis on the frequency domain signal using the inverse Fourier transform (Smith, 2008; Ashwin & Manoharan, 2018; Zhang, 2019).

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-i\omega t}dt \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)e^{i\omega t}d\omega \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots)$$

2. Methodology

The distorted song was obtained in the “wav” file format. Since the processing time for an audio track with a duration of over three minutes is very high, the “Jet-Audio 8.1.9” software was used to trim and extract a 20 s sample from the audio file to apply digital signal processing techniques. First, the sample “wav” file was read into Wolfram Mathematica 12.3 and converted into a data file using the inbuilt function “AudioData”. The accuracy of digitized data in Wolfram Mathematica was verified using the open-access SoX (Sound Exchange) software. Since there are two separate channels in audio waves used in this experiment, channel 1 and channel 2 were separated in the form of 2D arrays to remove the noise components. Then fourier transform of each channel was obtained to identify and remove high-frequency noise from the audio file.

The 1-D array generated by Fourier transform for each channel is symmetric around its midpoint when considering the magnitude of frequencies. These magnitudes decrease from the midpoint in both directions. Therefore, a variable “m” was defined in the program to remove high-frequency components corresponding to added noise. Depending on the number assigned to “m”, the program replaces the amplitudes of high-frequency components with zeros from the midpoint of the array in both directions to remove the effect of those. High-frequency noise in each channel was removed from the given audio signal in that manner. Then the data file was converted back to “wav” file format using the inverse Fourier transform to check the effectiveness of the selected “m” value. A trial-and-error approach was taken to identify the ideal “m” value for each channel of the audio track. After obtaining the suitable “m” value for a 20 s sample, it was applied to the entire audio track. The accuracy of this method was examined for a diverse collection of songs, and a few of those samples and their corresponding "m" values are mentioned in Table 1.

3. Results

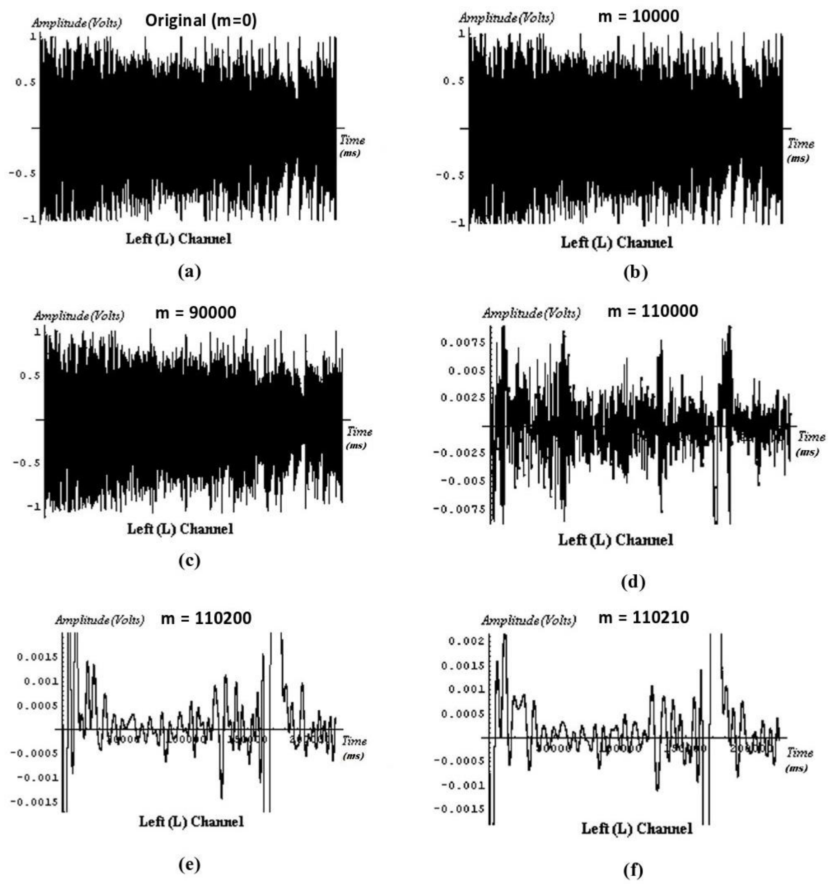


Figure 3: (a) Left channel of the distorted sound file of "Sandagalathenna" tele drama them song by Amarasiri Peris and Nelu Adakai (b) m = 10000, (c) m = 90000, (d) m = 110000, (e) m = 110200, (f) m = 110210.

Table 1: adequate “m” values obtained for each channel of the selected 10 different songs in the experiment

Song, Artist	m value (Channel 1)	m value (Channel 2)
andagalathenna” tele drama theme song Amarasiri Peris and Nelu Adikari	50000	60000
ſage Mandri Nam” Amarasiri Peiris	55000	55000

radhana” W D Amaradeva	55000	56000
eri Bindiya Re” Lata Mangeshkar and Mohammed Rafi	65000	55000
dunil” Indrachapa Bulathsinghala	65000	65000
asak Kota Ethi” Gunadasa Kapuge	70000	65000
bama Maada Viya Maada Obama Viya” Malani Bulathsinghala	65000	60000
edho oru Pattu” Unidathil Ennai Koduthen and Karthik	50000	50000
nna Mada Nale” Shanthi Geethadeva	65000	65000
ello” Lional Richie	40000	45000

4. Discussion

Since SoX is a globally recognized open-access software used by audio engineers to digitize audio signals, it was used in the experiment to check the accuracy of digitization of audio signals in Wolfram Mathematica 12.3. It was observed that digitized datasets obtained from both methods were equal for four decimal points.

The number of high-frequency components (indicated by the value assigned to “m”) dropped from the original audio signal by replacing their amplitudes with zeros can be varied as shown in figure 4. The suitable value for “m” change for each channel in each audio file because the level of high-frequency noise varies from song to song. Therefore, it is difficult to state a universal threshold value for “m”. Because of that, a trial-and-error approach was taken in the experiment to obtain the ideal value for “m”, to reduce the noise in each channel of a given audio file, while minimizing the impact on high-frequency components of the original song. The use of a trial-and-error method is vital to minimize the negative impact that could cause by high-frequency components of the original audio file being cut off (figure 3(d), 3(e), and 3(f)). Apart from properly analyzing the frequency components of the output signal, the recognition of a suitable “m” value for a given signal also depends on the hearing sensitivity of the listener and the equipment being used (earphones, speakers, etc.).

Results in Table 1 show that the suitable "m" value (least upper bound) varies for each song within the 50000 to 70000 range, except for "Hello" by Lionel Richie, which has the lowest range of 40000 - 45000. Among the others, both “Mage Mandri Nam” by Amarasiri Peiris and “Aradhana” by W. D. Amaradeva display an "m" value of around 55000 for both channels. Similarly, “Teri Bindiya Re” by Lata Mangeshkar and Mohammed Rafi, “Indunil” by Indrachapa Bulathsinghala, “Obama Maada Viya Maada Obama Viya” by Malani Bulathsinghala and “Enna Mada Nale” by Shanthi Geethadeva show "m" value of around 65000 for high-frequency noise removal. Among the samples shown in Table 1, “Pasak Kota Ethi” by Gunadasa Kapuge has the highest "m" value of 70000. Therefore, in general, high-frequency noise can be removed from a given vintage song by varying the "m" value defined in the developed program between 50000 to 70000.

A 20 s sample was extracted from the song to minimize the processing time when applying the trial-and-error method to find the suitable value for “m”. As stated earlier, the program replaces magnitudes of high-frequency components in the digitized dataset with zeros based on the value assigned to “m” and removes the effect of those on the original audio signal. It was observed that the suitable “m” value identified through the trial-and-error method for the left and right channels of a particular song were also different from each other. Digital filters can be used to further improve the results obtained using the method developed in this study.

This developed technique can be applied to remove the added noise and conserve vintage songs stored in compact cassettes and spools. That includes the original recordings of songs sung by Sunil Shantha, Rukmani Devi, … etc. stored at Sri Lanka Broadcasting Corporation (SLBC). Even though this task can also be completed using third-party noise removal software, such software does not allow the user to control the number of high-frequency components removed from the distorted audio file. Since the “m” value can be changed, the user has control over that in the technique developed in this project.

5. Conclusion

The digital signal processing procedure tested in this study can be applied to remove high-frequency noise in audio files recorded in cassettes and spools. However, the level of accuracy of the output generated using this method varies on the hearing frequency of the listener, and the quality of the audio equipment being used. Based on the results of this study, the approximate range for “m” value that can be applied to remove high frequency noise from a given vintage song is between 50000 and 70000. Technique developed in this research can be used to conserve vintage songs that were originally stored in compact cassettes and spools.

References

(1) Andriessen, W., 1999. The Winner compact cassette. A commercial and technical look back at the greatest success story in the history of AUDIO up to now. Journal of magnetism and magnetic materials, *193*(1-3), pp.11-16.

(2) Ashwin, J.S. and Manoharan, N., 2018. Audio denoising based on short time fourier transform. Indonesian Journal of Electrical Engineering and Computer Science, 9(1), pp.89-92.

(3) Barnard, W. R., 1972. Audio tapes and cassette tape recorders. Journal of Chemical Education, 49(2), 136–138.

(4) Bohlman, A.F., 2017. Making tapes in Poland: The compact cassette at home. Twentieth-Century Music, 14(1), pp.119-134.

(5) Chen, J., Benesty, J., Huang, Y., Diethorn, E.J., 2008. Fundamentals of Noise Reduction, In: Benesty, J., Sondhi, M.M., Huang, Y.A. (eds.), Springer Handbook of Speech Processing. Springer Handbooks. Springer, Berlin, Heidelberg.

(6) Daniel, E. D., Mee, C. D., & Clark, M. H., 1999. Instrumentation Recording on Magnetic Tape, In: Magnetic Recording: The First 100 Years, pp. 316–332. IEEE.

(7) Harčarik, T., Bocko, J. and Masláková, K., 2012. Frequency analysis of acoustic signal using the Fast Fourier Transformation in MATLAB. Procedia Engineering, *48*, pp.199-204.

(8) Hassanpour, H., 2008. A Time-Frequency Approach for Noise Reduction. Digital Signal Processing. 18, no. 5: 728–38.

(9) Landau, H. J., 1967. Sampling, Data Transmission, and the Nyquist Rate. Proceedings of the IEEE, *55*(10), 1701–1706.

(10) Nasuta, A. T., "Floppy Disk and Magnetic Tape Cassette Magnetic Field Susceptibility," 1981 IEEE International Symposium on Electromagnetic Compatibility, 1981, pp. 1-3

(11) Oppenheim, A. V., 1978. Applications of digital signal processing. Physics.

(12) Rabiner, L. R., & Gold, B., 1975. Theory and application of digital signal processing. Tads.

(13) Smith, J.O., 2008. Mathematics of the discrete Fourier transform: with audio applications. Julius Smith.

(14) Van Bogart, J.W., 1995. Magnetic Tape Storage and Handling: A Guide for Libraries and Archives. Commission on Preservation and Access, 1400 16th St., NW, Suite 740, Washington, DC 20036-2217.

(15) Vaseghi, S.V., 2008. Advanced digital signal processing and noise reduction. John Wiley & Sons.
(16) Zhang, D., 2019. Wavelet transform. In Fundamentals of image data mining, pp. 35-44. Springer, Cham.