Identification of A Music Instrument In A Superimposed Tone

1st Dony Armstrong DSouza, 2nd Veena Devi Shastrimath V
[1]Senior Assistant Professor/Research Scholar, [2]Professor,
[1]Electronics & Communication Engineering Department,
[2]Electronics & Communication Engineering Department,
[1]MITE, Moodabidri, INDIA
[2]NMAMIT, Nitte, INDIA

Abstract - Music has been a very fascinating aspect since the beginning human civilization with improvements in it. Since the inventions of different musical instruments belonging to percussions, string etc. the people willing to learn to play music have revolved around and also hard to find teachers. Hence most of the musicians have also a problem of getting resources to learn such as the background or the single instrument role in an audio clip. Hence finding the music track for practicing is a very difficult task and most of the musicians struggle to find it. In order to overcome this problem many have tried to extract the musical instrument signal from an audio track. However, the identification of one instrument mainly violin is considered. Different instrument data is considered in this paper. There are many possibilities in this broader area and many methods have been evolved and tried upon by many researchers. Some have computed with statistical models and others have tried with different filter models. By considering the spectrogram the features to be extracted are considered for identification of the musical instrument which is violin taken for the work. However there are a lot of opportunities in this regard by different methods to be employed in future.

Index Terms - MFCC, GFCC, tonal, Multitonal, Chords, Spectrogram

I. INTRODUCTION (HEADING 1)

The most common way to play the violin is by moving a bow across its four or five strings, that are usually having tuning tuned notes G3, D4, A4, and E5[1] with the intervals being in perfect fifths. Occasionally, it is and may be played by strumming and pulling back the strings with the fingers (pizzicato), and in few and rare circumstances, the wooden side of the bow may be used to strike the strings based on the effects required.

In most of the Different types of music throughout the world Violin is the main instrument or can be said as the essential instrument. They are most dominant in the western classical music legacy as lead instrument and combination of many as ensembles ranging from chamber music to symphonies. In addition, violins play a significant role in a number of folk music genres, such as jazz, country, and bluegrass. The challenge of learning violin is a task which very often tend to let go learning violin. Since Violin is used in its many forms to play such as for symphonies and melodies it's widely used. Also, there were many attempts and they separating out violin tune from tracks in order to concentrate only on one single instrument to while learning the instrument. Some electric violinists employ solid-body instruments with piezoelectric pickups.

This Violin holds a very special place among instrument because of its wide range of dynamic characteristics. violin sound generated from instrument can be varied a lot by its bowing techniques used [2]. The violin instrument sound when played in tandem with other instruments it blends perfectly with the music and its genre [7]. The musician who plays violin faces the challenge of identifying the position of fingers to be placed in the fingerboard of the instrument. Hence for a young and new learner it becomes a very difficult process to identify the violin sound from a track. Hence identifying and extracting the sound of violin from a recorded track becomes very difficult [6]. The violin instrument consists of the sound that is having rage from lower as G3 with frequency component 296 Hz and highest practical limit as E7 that is of 2637Hz [3]. The violin is a tremendously varied and practical instrument when you consider the large range of notes it can play, the decibel range, the variety of textures, timbres, and articulations it can produce. Hence the identifying the violin sound with related to its features results in a vast amount of typical and diverse sound for classifications.

II. RELATED WORKS

The violin features are mainly consisting with respect to its architecture [4] and its set of manufactured way. The physical structure also plays an important role in making the violin to sound very distinct. The violin also consists of the different aspects of bowing techniques and other information basically by considering the wood and hull measurements [5]. The instrument itself is having a very dynamic range of sounds [6]. Most of the instruments have their distinct features with respect to the device characteristics of Timbre and instrument which are essential in defining the instrument characteristics for identification.

The element that can be most easily distinguished from timbre is duration, which also serves as an example of the reasoning behind timbre's definition (although there is not actually a standard definition of the perception of duration). A discernible shift in sound results from a vocalist holding a note for a longer time than for a shorter time; however this is not a change in timbre. Once the duration is well beyond the temporal frame utilized to construct the magnitude spectrum, duration has no impact on the magnitude spectrum of a sound. Usually, music's sustained notes are longer [7], [8].

A similar work is done by Wang, Qi Bao have worked on recognition of Individual instruments of the similar type which is very difficult and it is not seen upon much. The violins were explored to be detected and recognized separately by employing source filter model. After the spectral components are plotted they divided into tonal and non-tonal content[16]. This is very often seen in relation to timbre of the instrument. The content similar to MFCC a GFCC i.e. gammatone is considered for describing the spectrum. A gaussian mixture model was employed for parameterization [7]. Other music signal processing employed are based on ADSR curves to emulate musical instruments and effects with filters are employed in many of the instances [9], [10]. The Gaussian mixture models are considered by many in the [11] is seen however here the complexity of the system is more and the Support vector machine method also yields high amount of calculations.

III. Methodology

The system proposed mainly deals with acquiring the sample violin tone and processing it with the aspects of signal processing. The basic block of the system is shown in Figure 1.

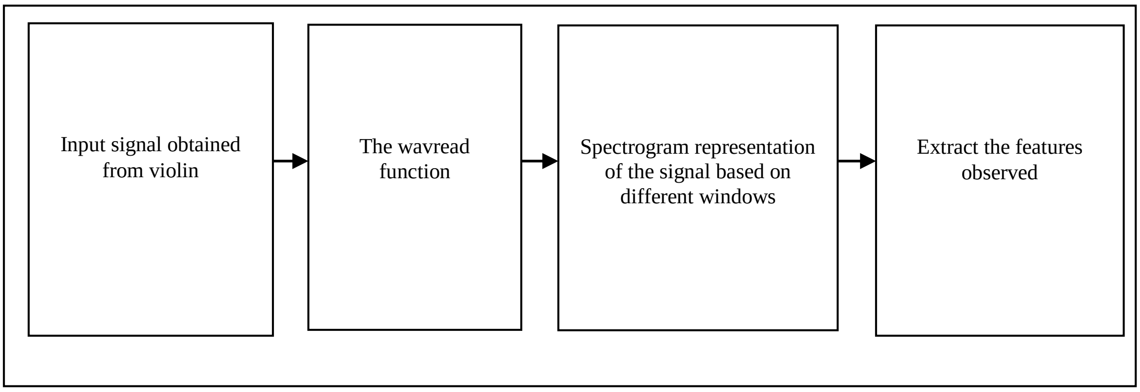

Figure 1. Basic block diagram of the System

a. Obtaining the Violin sound signal

The violin considered here is a standard 4/4 violin and the set up for data acquisition of the note A4 is shown in Figure 3.The note played which is A in notation for a beat is shown in the Figure 2.

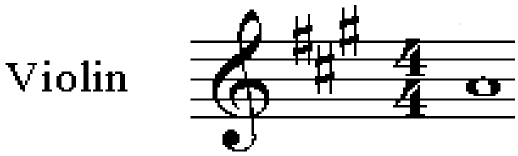

Figure 2. Musical Notation representation of Note A4.

The setup consists of the violin being connected via the piezo electric pick up instead of a standard coil wound pickups as used for guitar pickups [9]. The guitar pickups however will have a more specific range to pick only the sound produced from metal strings. The piezo electric pick up of violin has a wider dynamic frequency range and hence, it picks up the surrounding sounds with it. Extracting these sounds and only allowing the violin tune to be accepted becomes a bit difficult. The violin sound is recorded with a sampling rate of 44.1KHz and stored in as information by the wavread command in MATLAB with a sampling rate of 8KHz. The violin A4 note is the frequency accepted universally for tuning instruments with a 440Hz signal.

Figure 3. Setup for Data Acquisition

The violin tone is played for a common timing of 4/4 with the note A and recorded. The recorded note is then read into the MATLAB environment in the WAV format[10]. The violin sound is taken as the reference A4 of frequency 440Hz and is an open string which will certainly have some more dynamics included in it. Also, the open string will surely have the limitations of many other disturbances and harmonics being added in the process of recording. The bowing used is the upward pressure to obtain the violin sample wave. The Violin sound signal may be recorded using wavrecord function or recorded in any general-purpose recording tools. For the purpose the recording is done using wave recorder i.e. audio grabber. The bowing used is the upward pressure to obtain the violin sample wave. The musical score of the audio file is shown in Figure 4. The score consists of polyphonic notes which forms the chord as seen as E and E5th triad chords.

The other recording of a guitar being played is recorded in the same manner and is the just rhythm chord progression in the key of A. The guitar used is an electric guitar with magnetic pickups being used. The signals obtained are the result of sampling at a frequency of 44,100Hz for the sampling frequency Fs. The flow chart of the proposed system is shown below in Figure 5

Figure 4. Musical Notation representation of the audio file.

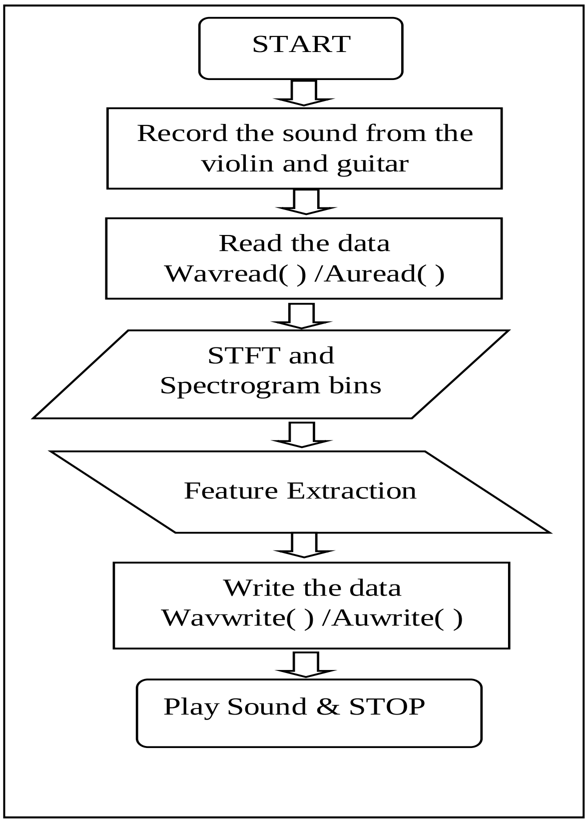

Figure 5. Flowchart depicting the methodology of the proposed system

a. Processing of the signal

The signal obtained after sampling is taken as a vector or matrix of mono audio signal. This signal is sampled by the frequency Fs and the processing is carried out by MATLAB software. The Frequency spectrum is considered by having absolute FFT of the observed input violin audio sample in order to have a deeper look at the features for extraction spectrograms. [12]. The spectrogram of the signal is considered for viewing the frequency components in the frequency domain. Violin single note is considered and another polyphonic note of guitar chord is taken in order to identify the components in spectral domain. The spectrogram consists of the frequency components with respect to time and magnitude. The signal is processed by the equations considering the STFT for a window size of 256 and 512 in some cases[13]. The spectrogram procedure is followed for many experiments and has yielded good results with log frequency plots and have been basically used for identification. The same approach is used in order to obtain the separation of different instruments used.

b. Features of the Audio to be taken from the instrument-Timbre.

According to the definition of timbre, it is the characteristic of auditory perception that allows a listener to distinguish between two sounds that are similarly presented and have the same loudness and pitch [6].The idea is strongly related to the ability to identify the source of a sound. For instance, although having similar loudness and pitch characteristics, flute and violin sounds can nevertheless be easily recognized from one another. Additionally, when listening to polyphonic music, humans may typically perceptually group the individual sounds according to their sources using timbre information[14].

A multifaceted notion with many underlying acoustic elements is timbre. There are five main acoustic elements that affect timbre: 1) the distinction between a sound's tonal and noise-like characteristics; 2) the spectral envelope; 3) the temporal envelope; 4) the variations in the fundamental frequency and spectral envelope; and 5) the point at which the sound first diverges significantly from the vibration that has been sustained [15].

In order to compute Fourier, Transform for a short and finite duration parts of a long sequence of signal for obtaining Short Time Fourier transform, small duration samples are taken. Normally, short duration sequence of long signal is achieved by considering a window function w[n] of bandwidth very small and moved across the signal or multiplied by the longer signal. Commonly used Windows to get this response is usually a short duration windows such as Rectangular Window. This method is used as the desired short sequence only is extracted by not altering the signal further. Additionally, it would be appropriate to have tapered ends for the effective and improvised representation in the domain of frequency, a window i.e., Hamming window is preferred and employed. If the continuous frequency variable, the STFT can be described as

$$X[n,w] = \sum_{m=-\infty}^{\infty} w[n-m]x[m]e^{-jwm} \quad \text{----(1)}$$

Here by employing a window of finite length we can obtain the equation as

$$X[n,k] = \sum_{m=n-(N_w-1)}^{n} w[n-m]x[m]e^{-j\omega_k m}$$

$$= \sum_{m=n-(N_w-1)}^{n} w[n-m]x[m]e^{-j2\pi mk/N}$$

By this, it can be observed that the variables of frequency as well as time and a function of two variables consisting of Time and Frequency is X [n, k] which is a discreet function. By considering the signal to be band limited to have an effective frequency distribution in order to separate the frequencies which are not required.

c. By considering the Band pass filters

STFT expression obtained is not a systematic representation. Hence in order to have a equivalent and proper mathematical equation a very small alteration to the expression is done. This is shown as a block diagram in Figure 6. Here the input sequence is x[n] and the window function with an addition of tie function is taken in exponential form to obtain the STFT based output as x[n,k].

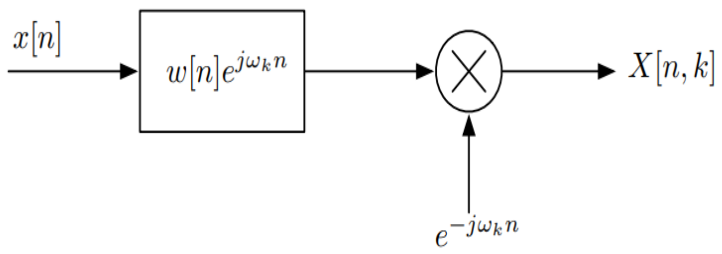

Figure 6. Block representation of the expression

The expression results in bandpass implementation of filter which is originally a low pass filter window.

$$X[n,k] = \sum_{m=n-(N_w-1)}^{n} w[n-m]x[m]e^{-j\omega_k m}$$

This results in the equation to be,

$$= \left(\sum_{m=n-(N_w-1)}^{n} \left(w[n-m]e^{j\omega_k(n-m)} \right) x[m] \right) e^{-j\omega_k n}$$

IV. RESULTS AND CONCLUSIONS

For the input signal the violin signal of a single note is considered. In this case the note d5 is taken for the purpose of viewing the spectrum of it. The violin signal considered and taken is shown in time domain in Figure 7. The violin string 1a5 is considerd and its corresponding dynamics is shown with respect to time and amplitude. The frequency is 440Hz. Which is the standard accepted universally for tuning of musical instruments and is the reference frequency.

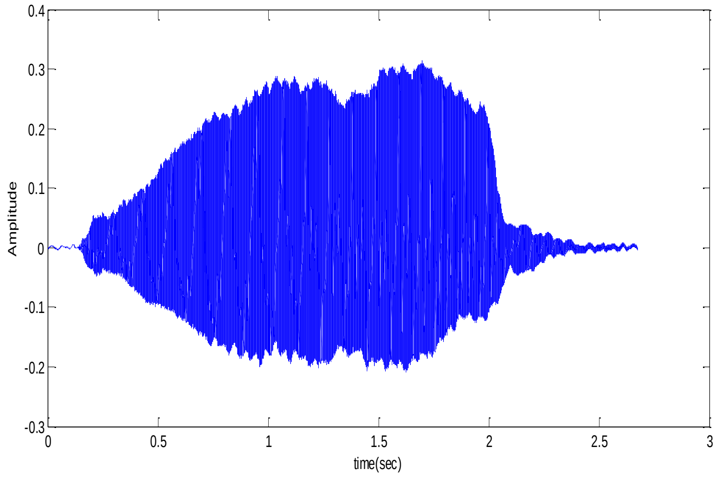

Figure 7. Sample Violin note played for d5

The audio clip of electric guitar of chords is shown in Figure 8. The initial first peak is the sound before strumming the strings. From the graph the information seen as the strumming pattern of the guitar for different notes held together which is the chord patterns.

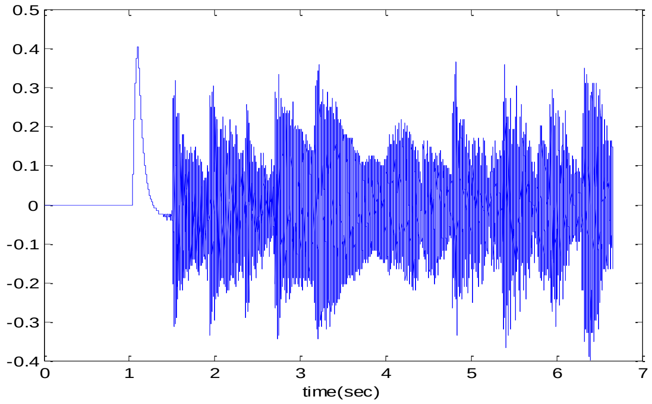

Figure 8. Audio clip played on electric guitar

The mixed signal is shown in the Figure 9. The two audios of violin and guitar are seen and they are appearing with their characteristics. We can easily observe the guitar patterns and violin tune. The violin and guitar audio is blended in fractions or ratio of 60%.

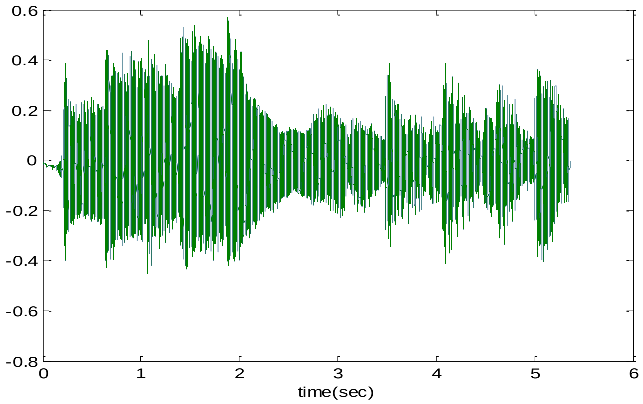

Figure 9. Mixed signals in time domain

The results obtained through spectrogram of Violin d5 sample is show in Figure 10. The violin signal

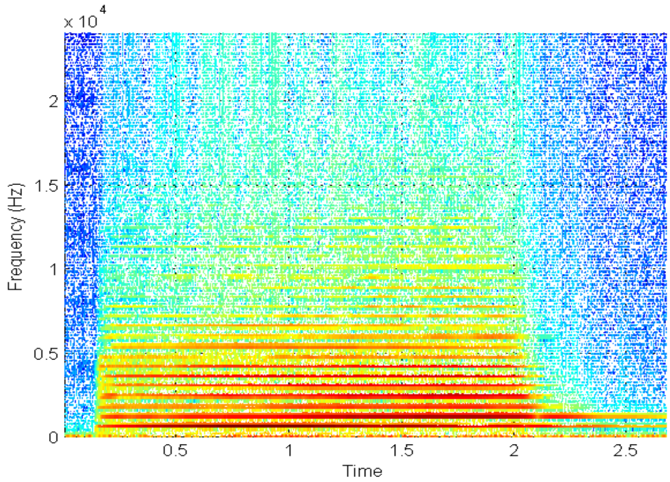

Figure 10. Spectrogram of violin signal played for note d5

The results obtained through spectrogram of Audio sample is show in Figure 11. The spectrogram in Figure 11 can be viewed as in the beginning there is no signal and hence it is just a blue spectrum up to one second. The spectrum overall consists of the harmonics generated in the guitar chords and it is played in the upper parts of the guitar fret board. As a result, the signal is seen in that with lot of traces of frequency in the vertical columns which is nothing but frequencies with different ratios to form the chords.

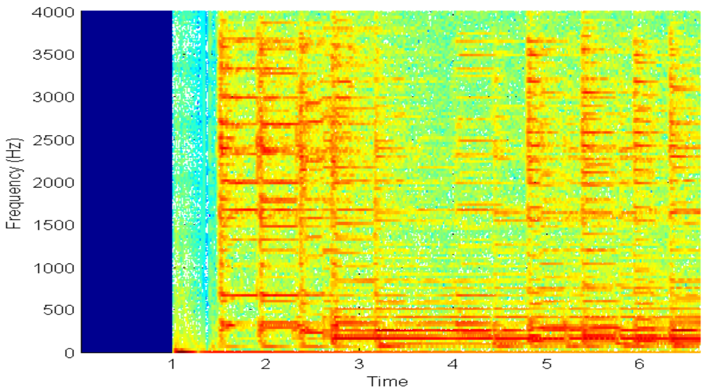

Figure 11. Spectrogram of the Audio sample.

The results obtained through mixing of the signals are shown in Figure 12.

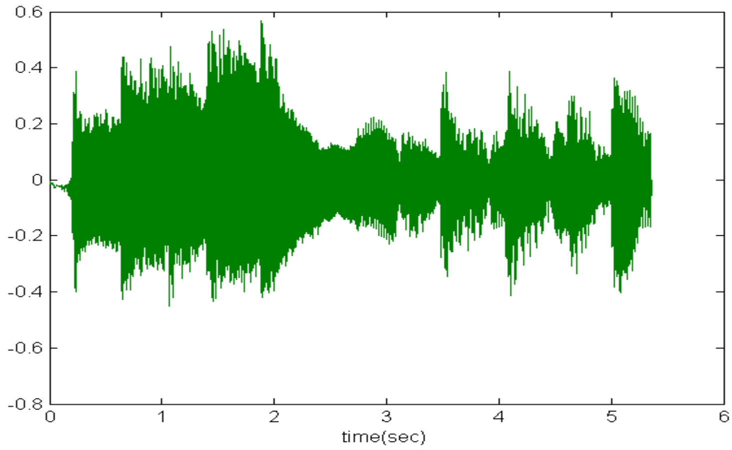

Figure 12. Results of mixing the two signals in time domain.

The results obtained through spectrogram of mixing of the signals is show in Figure 13.The results obtained through spectrogram of the mixed signal is shown in Figure 13. By considering the waveform and obtaining the spectrogram of it over a small-time interval we can clearly distinguish the values of different notes and the four strings mainly E, A D G strings of violin which overlays on it in a musical clip. By observing the obtained waveform for the different or tonal quality of the instrument with a very small variation we can get the estimation from the spectrogram.

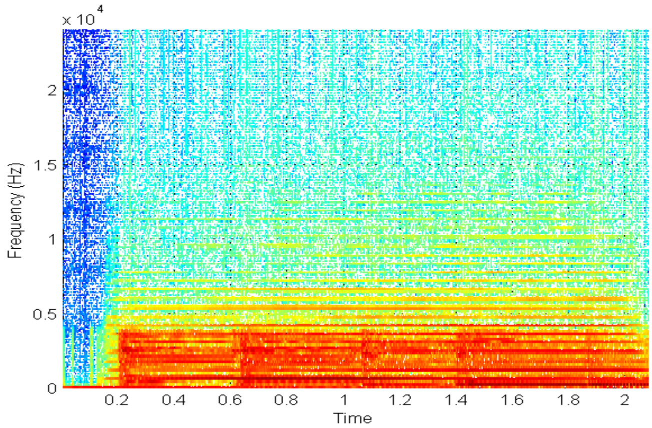

Figure 13. The spectrogram of the two mixed signals.

However, in an orchestration where there are multiple instruments with same tonal qualities and overlapping frequency and power bands with same magnitudes, it requires other modes of observation as to finding frequency-power formants and in some cases pattern in the signals and since it is non-periodic in nature and are not predictable it becomes difficult to identify. However, the STFT method forms the base for all research pertaining to most of the recognition and identification of sources. However, in this paper the spectrogram is considered and the values after spectrogram plotting are seen and the distinguishing characteristics of the instruments are seen and this can be seen are separated out to form a separate track. Future works include adding some more instruments and considering the mixture models with gaussian etc. which include even considering Gabor transforms and chirping in order by considering the detection and perception capabilities of our human ears.

TABLE 1. IDENTIFICATION OF INSTRUMENTS

Sample No	**Comparison with different instruments**		
	Instrument Considered	***Second instrument considered***	***Recovery possibility***
1	Violin	Guitar(polyphonic)	Identifiable but no accuracy
2	Violin	Flute	identifiable
3	Violin	Viola	Very difficult
4	Mouth organ	Violin	Identifiable but no accuracy, Difficult

The experimentation done by some musical instruments is shown in Table 1. Here the spectrogram of instrument signal is taken and they are studied as per the required frequency components and they were found to be identifiable and extraction was involved and the instrument identification and separation was observed as per the required instrument and within the given constraints of computational complexities. By considering the short window of time to find the amount frequencies and their variations with respect to the formants we can up to certain level achieve the separation.

V. REFERENCES

[1] M. Mannone, "*Introduction To Gestural Similarity In Music. An Application Of Category Theory To The Orchestra*," *Journal of Mathematics and Music*, vol. 12, no. 2, pp. 63–87, May 2018, doi: 10.1080/17459737.2018.1450902.
[2] S. Giraldo *et al.*, "*Automatic Assessment Of Tone Quality In Violin Music Performance*," *Front Psychol*, vol. 10, no. MAR, 2019, doi: 10.3389/fpsyg.2019.00334.
[3] V. Ronchetti, "*Violin Timbre Analysis with Mel-Frequency Cepstral Coefficients*." University of Rochester.
[4] A. Livshin, "*Automatic Musical Instrument Recognition And Related Topics*." .fr/tel-00810688.
[5] J. Charles, D. Fitzgerald, and E. Coyle, "*Violin Sound Quality Detection*." At arrow.tudublin.ie/argcon
[6] G. Agostini, M. Longari, and E. Pollastri, "*Musical Instrument Timbres Classification With Spectral Features*," 2003.
[7] Q. Wang and C. Bao, "*Individual Violin Recognition Method Combining Tonal And Nontonal Features,*" Electronics (Switzerland),vol. 9, no. 6, pp. 1–18, Jun. 2020, doi: 10.3390/electronics9060950..
[8] Y. Wu, Q. Wang, and R. Liu, "*Music Instrument Classification Using Nontonal* " MFCC," 2017.
[9] D. Armstrong D'Souza and V. Veena Devi Shastrimath, "*Modelling Of Audio Effects For Vocal And Music Synthesis In Real Time*," in 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), Mar. 2019, pp. 1–4. doi: 10.1109/ICCMC.2019.8819852.
[10] D. Dony Armstrong, S. v. Veena Devi, and V. N. Ganesh, "*Pedal Effects Modeling For Stringed Instruments By Employing Schemes Of DSP In Real Time For Vocals And Music*," 2020, pp. 97–113. doi: 10.1007/978-981-15-0626-0_9.
[11] Sumit Kumar Banchhor, Arif Khan "*Musical Instrument Recognition Using Spectrogram And Autocorrelation*" International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2 Issue-1, March 2012".
[12]Y.-N. Hung and Y.-H. Yang, "*Frame-Level Instrument Recognition By Timbre And Pitch*," Jun. 2018, [Online]. Available: http://arxiv.org/abs/1806.09587
[13] Z. Otčenášek and J. Štěpánek, "*Directional Timbre Spaces Of Violin Sounds*."
[14] J. A. Charles, D. Fitzgerald, and E. Coyle, "Violin Sound Quality Detection," in IET Conference Publications, 2008, no.539 CP,

pp. 282–287. doi: 10.1049/cp:20080676.

[15] Y. Qi, J. W. Paisley and L. Carin, "*Music Analysis Using Hidden Markov Mixture Models*," in IEEE Transactions on Signal Processing, vol. 55, no. 11, pp. 5209-5224, Nov. 2007, doi: 10.1109/TSP.2007.898782.

[16] *Douglas A. Reynolds, Thomas F. Quatieri, And Robert B. Dunn."* Speaker Verification Using Adapted Gaussian Mixture Models" Digital Signal Processing 10, 19–41 (2000) doi:10.1006/dspr.1999.0361,