6th Asia Pacific Conference on Advanced Research (APCAR-2018), Melbourne, Australia
ISBN :978-0-6481172-1-6
Asia Pacific Institute of Advanced Research (APIAR)
www.apiar.org.au

SYNTHESIS OF NONLINEAR ARTICULATIONS OF SITAR USING TIME VARYING AUTO REGRESSIVE MOVING AVERAGE MODEL

Neel Gunasekera[a], Janaka Wijayakulasooriya[b], Pantalion Perera[c]
[a]The Open University, Nugegoda, Sri Lanka
[bc]University of Peradeniya, Sri Lanka
Corresponding email: sngun@ou.ac.lk

Abstract

The effects of articulation techniques that engage in decorating musical compositions in North Indian Classical Music (NICM) are extremely exotic and acoustically highly nonlinear. Sitar is the only instrument in NICM on which all the articulation techniques (Meend, Andolan, Gamak, murki, Ghazeth, Krinthen, Zam-Zam, etc) can be performed. Owing to the fact that most of the other articulation techniques can be derived from the meend technique, this research focuses on modelling the articulation technique, mend employed on Sitar. However, the standard mathematical and physical models used in synthesis of musical instruments fails to capture the essence of sitar tone due to the non-linear vibrations caused by the curved bridge of the sitar. In this research, the sitar tone is modelled as an Auto Regressive Moving Average(ARMA) process and the system model of the sitar is estimated using Steigtz-McBride (stmcb) algorithm. Then, the model is factorized to identify the properties of individual frequency partials and the dynamic behaviour of the frequency partials during articulation was analyzed. The signal reconstruction is achieved by super-positioning the most dominant frequency partials with the micro tonal transition due to the meend effect modelled. The output waveform of the sitar tone is compared with natural sitar tones of different genres on timbre space and the effects of the micro tonal transition due to the meend effect is validated using spectrograms.

Keywords: Sitar, Articulation Techniques, Synthesis, Meend

1. Introduction

North Indian Classical Music (NICM) is primarily homophonic (Alves, 2012; Janaki, 2005), and in contrast, the magic of Western Classical Music (WCM) lies to a great extent in polyphonic composition. Though melodies exist in WCM, from a broad perspective, it is not the singular or defining focus, most of the time. While WCM is composed, NICM is improvised. All compositions which can be encountered in WCM are formally presented in staff notations, and performers have literally no latitude for improvisation. Contrarily, performers of NICM practice an improvisation called *meend* which is primarily an articulation effect, by modulating a frequency of a tone in a continuous fashion to another frequency via continuum of frequencies (Janaki, 2005; Bandyopadhyaya,1988). The articulation techniques performed on the Sitar have not been sufficiently explored or successfully modeled. In this endeavor, an attempt has been made in successfully modelling the articulation techniques carried out on Sitar by viewing it as a stochastic process.

1.1 Time variant timbre of Sitar

The simple first-order linear theory of musical instruments is remarkably successful in explaining their acoustic behavior. The vibrating elements of musical instruments are chosen to have normal modes in harmonic frequency relationship and the mode frequencies all being integral multiples of the lowest or fundamental mode frequency (Janaki, 2005). Simple impulsively excited instruments such as guitars and bells have nearly linear behavior, with all modes simply decaying exponentially with time (Juan, 2010).

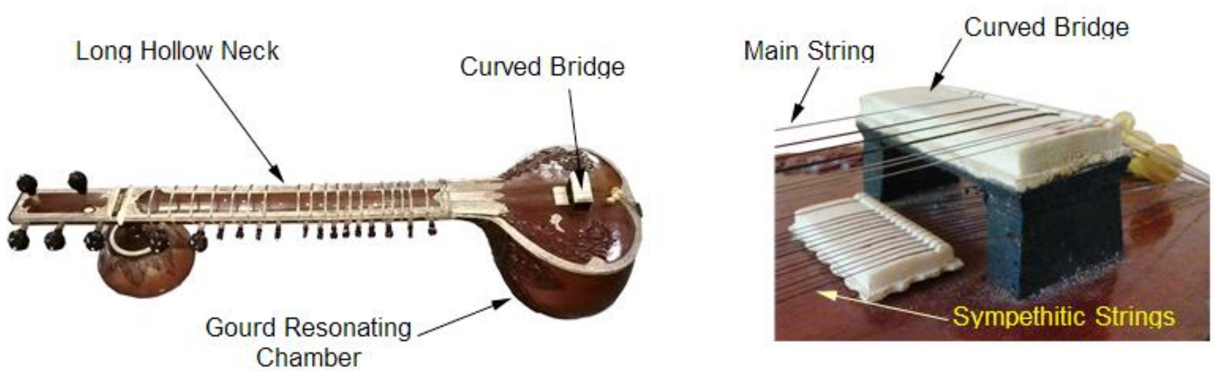


Figure 1:Sitar and the curved bridge

Unlike the guitar string, one end of a sitar string is rested on a sharp edge while the other end of the string is suspended on a curved bridge called *jawari*, as shown in Figure 2. This arrangement allows us to vary the length of the vibrating part of the string in the period of oscillation by wrapping and unwrapping the string around the curved bridge (Alsahlani &Mukherjee , 2010; Kartofelev et al, 2003)as depicted in Figure 2.

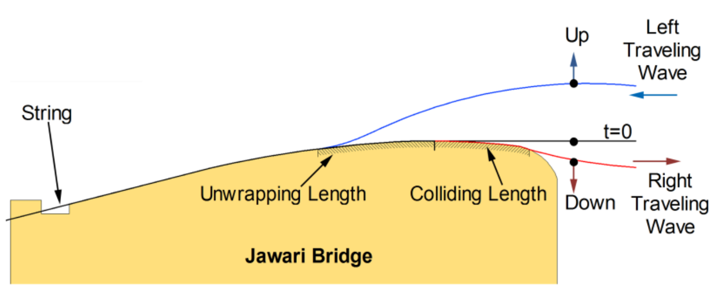


Figure 2: Wrapping and unwrapping the string around the curved bridge

This unique feature, resulting in a harmonically rich sound intertwined with buzzing effect which cannot be grasped with the clutches of first-order linear theory of musical instruments and making it extremely challenging in modelling and synthesizing the timbre of sitar (Dmitri K et al, 2013; Raman,1921). The existing digital synthesizing techniques such as additive, subtractive and frequency modulation are incapable in modelling the timbre of sitar successfully due to its time varying nature. Though the physical modelling approach exhibits favorable results in generating a simple tone, this strategy is inefficient in the case of modelling articulation effects due to its high processing time and the memory overload.The above argument suggests that it requires a better approach in synthesizing the timbre together with the articulation techniques at the same time.

The vibration of the sitar string can be viewed as the response of the vibrating (strings) a resonating (body) parts of the sitar to an impulse as shown in Figure 3.

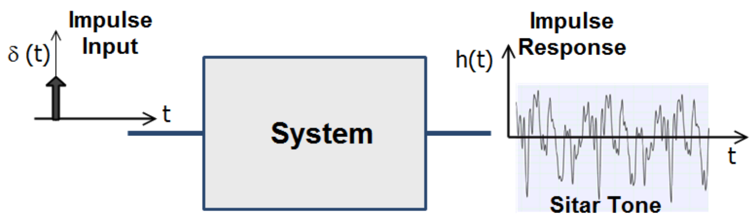


Figure 3: Impulse response system

This can be modeled as an ARMA process as shown in Figure 4. Thus, a ARMA can be used to model the timbre of a time variant instrument (Niedzwiecki, 1993;Hanck, 2008;Grenier, 2003)like sitar, together with articulation techniques widely used in NICM (Meillier, 1991; Esquef at al, 2003). This paper describes the modelling of sitar timbre using ARMA followed by the modelling of articulation techniques such as *Meend, Andolan, Gamak, Ghazeet, Murkhi*etc.

2. Modelling of Timbre of Sitar Using Auto Regressive and Moving Average (ARMA) Process

The output waveform generated by a musical instrument is known to be a continuous time solution*f(t)* to a Stochastic Differential Equation (SDE) where *f(t)*є *Ĺ*. The sampled solution to a SDE is necessarily an ARMA process [15][16]. In particular, the sampled solution to a p^{th} degree SDE is an *ARMA(p,p-1)* process. This arises the need to investigate the ARMA process.

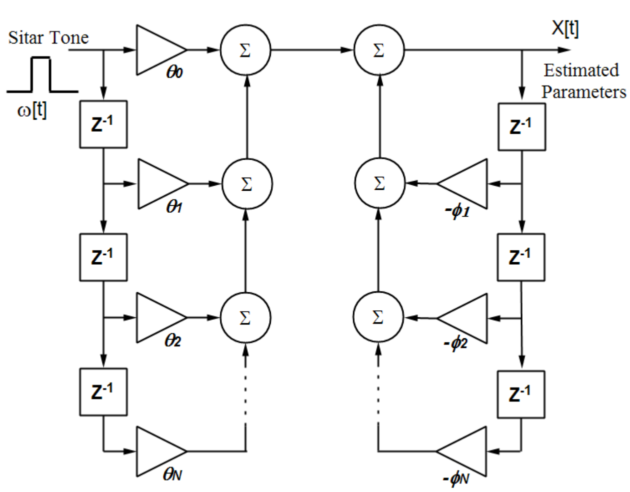


Figure 4: ARMA process

ARMA(p,q) process is defined by,

$$x_t = \theta_0\omega_t + \theta_1\omega_{t-1} + \theta_2\omega_{t-2} + .. + \theta_q\omega_{t-q} - \phi_1 x_{t-1} - \phi_2 x_{t-2} - \phi_3 x_{t-3} - .. - \phi_p x_{t-p} \quad \text{----(1)}$$

Using backshift operator B, the above can be expressed as,

$$x_t = \frac{\theta(B)}{\phi(B)}\omega_t \quad \text{-----(2)}$$

Where,

$$\theta_0 = 1$$

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \theta_3 B^3 + \ldots + \theta_q B^q$$

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \ldots - \phi_p B^p$$

Special cases:
if*p=0*, the process is a moving average of order *q* abbreviated as *MA(q)*

$$x_t = \omega_t + \theta_1\omega_{t-1} + \theta_2\omega_{t-2} + .. + \theta_q\omega_{t-q} \quad \text{------(3)}$$

This is a general linear process having a finite number of nonzero coefficients. Contrarily, if *q=0*, the process is an auto regression of order *p*, abbreviated by *AR(p)*

$$x_t = \omega_t + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \phi_3 x_{t-3} + .. + \phi_p x_{t-p} \quad \text{-------(4)}$$

Factoring the polynomial φ(z) as a product,

$$\phi(z) = \left(1 - \frac{z}{z_1}\right) + \left(1 - \frac{z}{z_2}\right) + \left(1 - \frac{z}{z_3}\right) + \ldots + \left(1 - \frac{z}{z_p}\right) \quad \text{-----(5)}$$

Where

$$|z_j| > 1$$

Now, let us refocus on the fact that the sampled output waveform generated by a musical instrument is an *ARMA(p,p-1)* process. The coefficients of the ARMA process pertaining to the output waveform of a musical instrument can be estimated by Steigtz-McBride iterative process as shown in Figure 5.

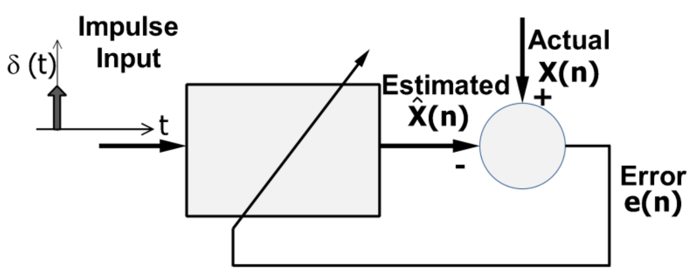


Figure 5:Iterative process

The output of an *ARMA(p,q)* process is a discrete signal. Therefore, naturally it is equivalent to the impulse response of an IIR filter. In Steigtz-McBride algorithm (STMCB) (Lindgren, et al 2013; Bosq, 2012), the filter coefficients of corresponding IIR filter is computed for a given

ARMA process while minimizing square sum error. The algorithm is relatively fast and the convergence is guaranteed for lower order systems (Bosq, 2012; Miranda, 2002; Pollock, 1990).

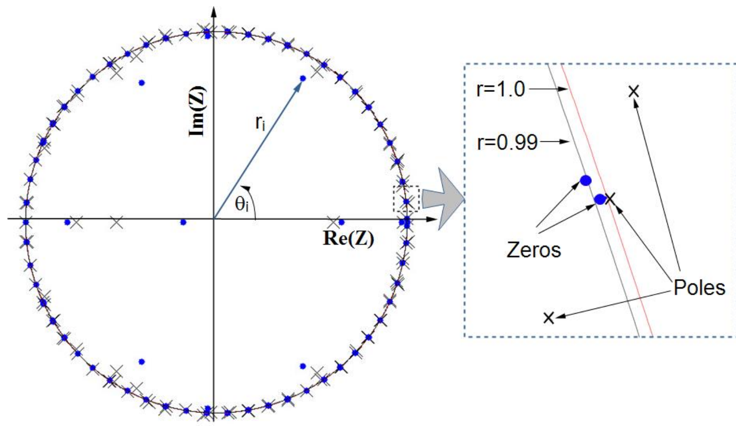


Figure 6: Representation of sitar sound spectrum in Z-plane

The poles corresponding to the fundamental frequency and inharmonics due to the effect of the bridge curvature can be clearly seen in Figure 6. Although, the frequency components $\theta_i$of the signal including inharmonics and the corresponding damping r_i can be seen on z-plane, the amplitude Υ_i and phase information cannot be directly found. As it is required to model the evolution of frequencies of each frequency component, the amplitude and phase information is also needed. In order to extract this information, the *equation (*Alsahlani & Mukherjee, 2010)is factorized using partial fractions given by *equation (*Kartofelev, et al, 2003).

$$x(t) = \frac{\theta(z)}{\phi(z)}\omega_t \qquad \text{--------(6)}$$

$$\frac{\theta(z)}{\phi(z)}\omega_t = \frac{r_1}{1-p_1 z^{-1}} + \frac{r_2}{1-p_2 z^{-2}} + .. + \frac{r_n}{1-p_n z^{-n}} \qquad \text{--------- (7)}$$

The STMCB algorithm is applied to the signal recorded by playing note Sa of a *Pt. Ravi Shankar* style sitar and the estimated zeros and poles of the system is shown in Figure 7. The angle of the poles are representing the frequency of each frequency component θ_i including inharmonics and given by equation (8). The damping of each frequency component is reflected by the amplitude of the pole (Υ_i) wherer_i< 1 for damped oscillations.

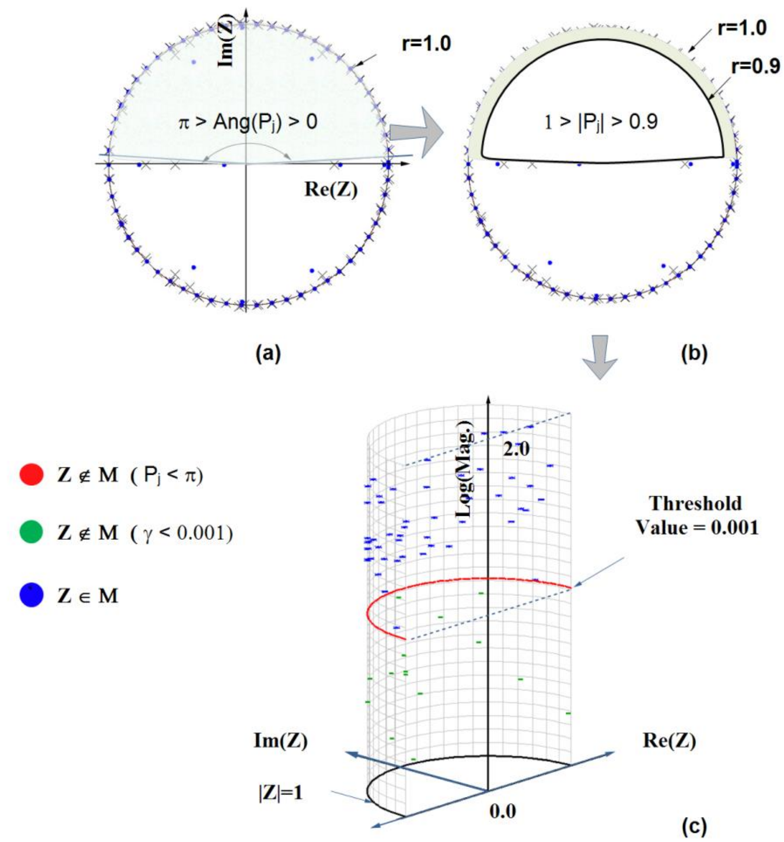


Figure 7: Filtering process

$$f_i = \theta_i \frac{F_s}{2\pi} \qquad \text{------(8)}$$

where, θ is the angle, F_s is the sample frequency Here, it is attempted to construct the waveform of a single tone using a digital IIR filter whose coefficients are obtained using an ARMA process. Recall that the sampled solution to ap^{th} degree SDE is an *ARMA(p,p-1)* process.

Thus, degθ(z)<degϕ(z)from equation (Juan, 2010), it follows that the filter, which is corresponding to an *ARMA(γ,γ-1)* process has the transfer function in z-domain,

$$H(z) = \frac{r_1}{1-p_1 z^{-1}} + \frac{r_2}{1-p_2 z^{-2}} + .. + \frac{r_n}{1-p_n z^{-n}} \qquad \text{-----(9)}$$

Here, it is worth noting thatγ is used to denote the order of process, for the sake of notational simplicity. Now, it is required to obtain the constants r_i and p_i for $i \in \{1,2,...\gamma\}$ .However, in constructing the amplitude envelopes for each harmonic, the frequency components of the spectrum appearing in the right hand-side are taken into account.

Thus, consider $M_1 \subset \{1,2,...\gamma\}$

such that, $\arg(P_j \in \{0,\pi\}$ for j_1, as shown in Figure 7-a

Also, in achieving a decaying pattern or attenuation for each harmonic of a given tone, only the terms corresponding to p_i slaying inside the unit circle are considered.

Define a set $M_2 \subset \{1,2,\ldots\gamma\}$ such that,0.90< | P_j| < 1 $j \in M_2$ as shown in Figure 7-(b)

In order to discard the terms in the series of partial fractions with negligible magnitudes, define a threshold value for minimum permissible r_i in the series, namely, γa sufficiently small positive number.

Define a set $M_3 \subset \{1,2,\ldots\gamma\}$ such that, $|r_j| > \gamma$ for $j \in M_3$, as shown in Figure 7-(c)

Now, define the intersection $M = M_1 \bigcap M_2 \bigcap M_3$

here, $M = \{1,2,..\hat{\gamma}\}(\hat{\gamma} \le \gamma)$

Now, define $ARMA(\hat{\gamma}, \hat{\gamma}-1)$ process.

3. Modelling the Waveform of the Fundamental Note (Sa)

As we observed earlier, the equation of the waveform of a single note can be generated by

$$f(t) = \sum_{j=1}^{N} M_j e^{(F_s \ln \alpha_j)t} \cos\left(2\pi \int_0^t p_j(u)du + \phi_j\right) \qquad \text{-------(10)}$$

Where Fs is the sample frequency.

By comparing the Figure 8-(a) and Figure 8-(b) it can be observed that the harmonically rich sound of sitar can be successfully synthesized by means of the ARMA process.

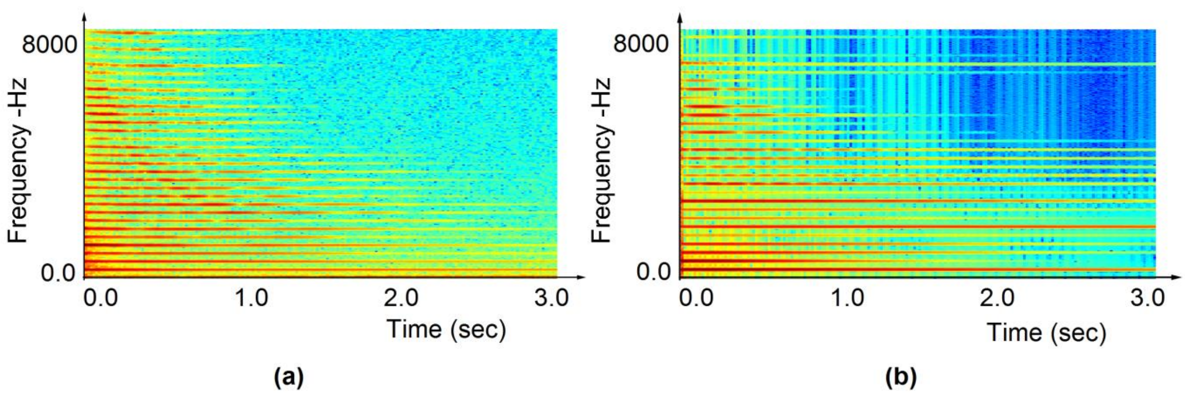


Figure 8: Spectrogram of waveforms of a single note (Sa)
(a):Real sitar (b):Synthesis sitar tone using ARMA process

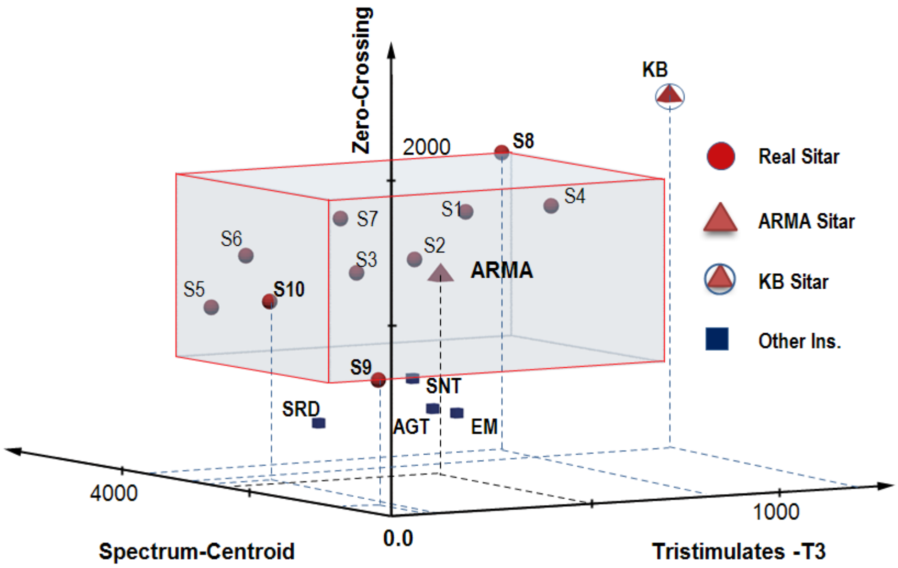


Figure 9:Timbre space for sitar

The Figure 9 depicts a comparison of different real sitar tones, sitar tone of YAMAHA PSRE06 keyboard, the sitar tone generated by ARMA processes with the other plucked string instruments such as Santoor (SNT), Sarod (SRD), English Mandolin (EMD) and Acoustic Guitar (AGU).As shown in the Figure 9, the timbre measures Zero Crossing Rate, Spectrum Centroid and Tristimulus-3 been selected (Lutkepohl, 2005; Dimitris & Manolakis, 1996)to quantify the timbral qualities buzziness, brightness and harmonic richness respectively, which have provide of the color of sitar tone.

4.Meend Articulations Technique

Meend is the technique in which frequency is modulated in continuous fashion through continuum of notes from one frequency to another (Lartillot & Toiviainen, 2007). When it is produced on a fretted plucked instrument like sitar, a note is struck and the string is pulled outwards or transversely across a fret to reach a higher note or several higher notes from that single stroke by varying the tension of the string. Proper rendering *meends* not only depends on the accuracy of the starting and ending *swaras,* but also on the exacting knowledge of the *kan-swaras* of the ragas, the transition speed of these *meends* and the accents on intermediate *swaras*. *Andolan, Gamak, murki, Ghazeth* can be considered as sibling of the *meend* family (Bandyopadhyaya, 1988). The frequency variation of the articulations isshown in the Figure 10.

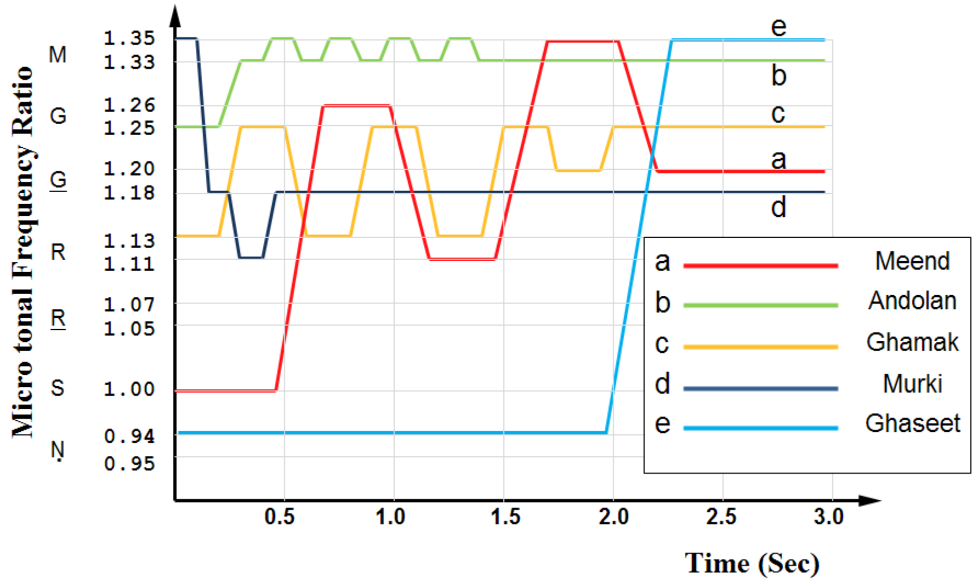


Figure 10:Temporal behaviour of articulations

4.1 Construction of Mathematical Model for MeendTechnique

The modelling process of articulation technique have been explained using a *meend* example *SGRMG*, where S , R, G and M are the NICM musical notes equivalent to the C, D, E and F notes in the western musical scale respectively, as shown in the Figure 11. The meend *SGRMG* is produced by placing the middle finger (usually) on fret Sa, exciting the string and followed by a pulling of the string outward to reach the higher notes by increasing tension of the string.

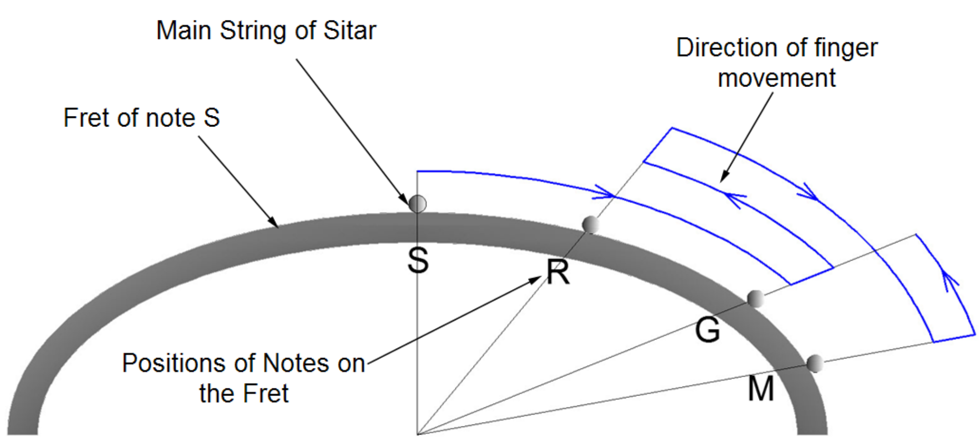


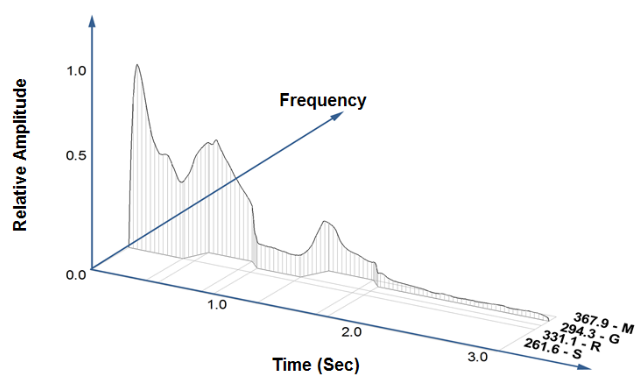
Figure 11: Variation of string position for meend*SGRMG*

The rest of the paper describes the analysis of the meend produced by a real sitar and then explaining the modelling process used to create the same meend sound.

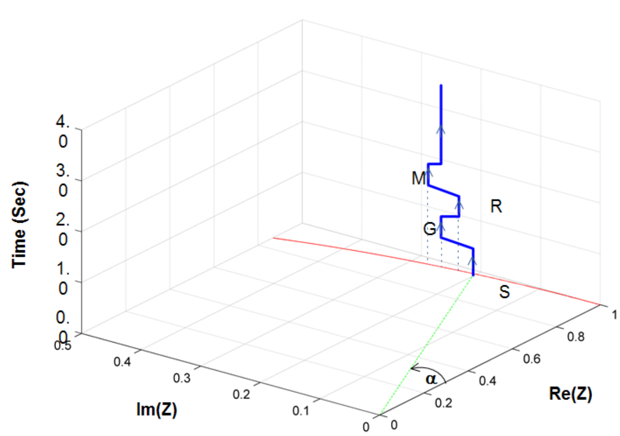
The Figure 12-(a) shows the nonlinear behaviour of the articulation technique *meend* in time and frequency space for the fundamental frequency component. The effect of meend action on position of the poles and zeros can be seen in the Figure 12-(b).

The variation of all partials due to meend action is shown in Figure 14-(a). The parameters of the waveform of a single tone are affected when an articulation technique is carried out. The effects are twofold. The first is its direct effect on the frequency. The second is its affect of variation of amplitude for each partials. However, this is more subtle due to its associated complexity. It is worth noting that, at the onset of a generated *meend* effect, a tension modulation takes place as shown in Figure 14-(b). Although, the amplitude of the waveform is affected due to tension modulation (B.S, 2014; Tolonen & Karjalainen, 1999), it is found to be negligible in practical synthesis scenario.

The Figure 13 shows the graphical user interface provided to develop the *meend* technique by adjusting the duration for the microtones or *sruthi* engage in a particular *meend* action (in this case SGRMG).



(a)



(b)

Figure 12:Details of the meend SGRMG
(a):Frequency and amplitude evolution
(b):Frequency evolution in Z-plane

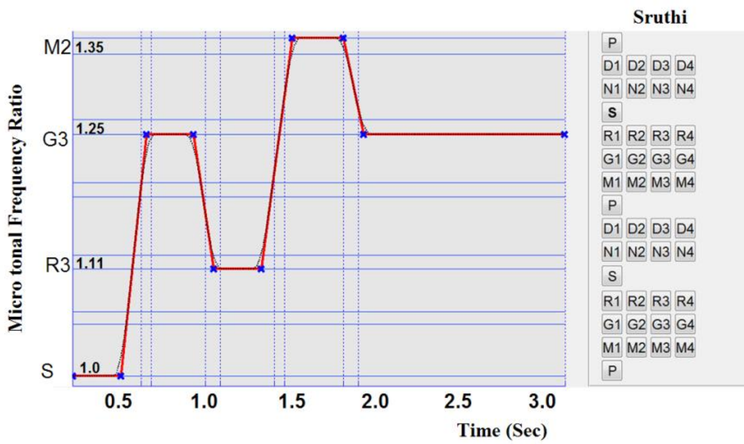


Figure 13:Interface for modelling articulations

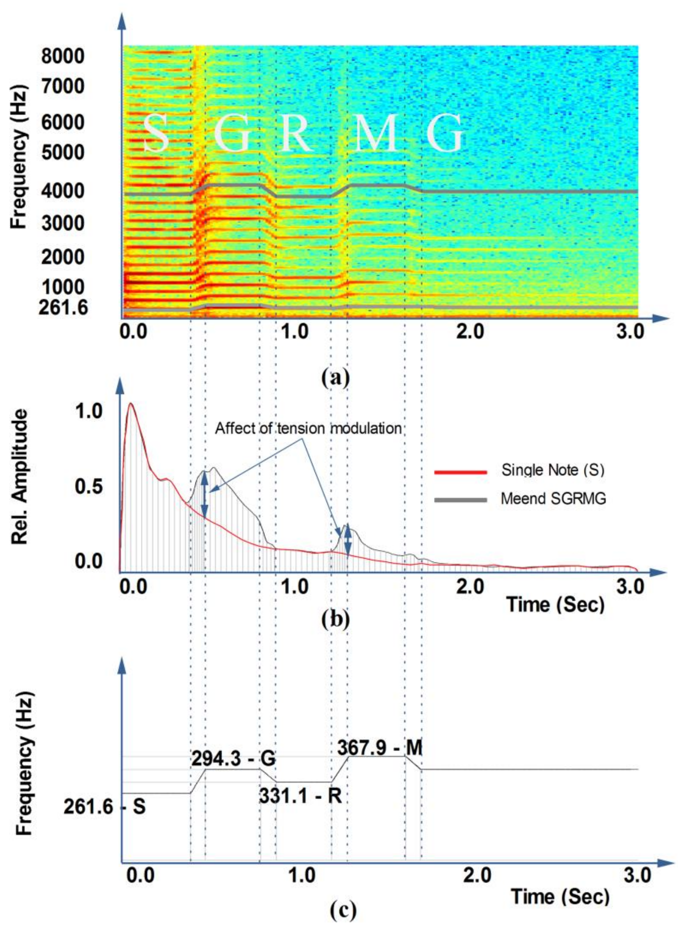


Figure 14: Details of a meend*SGRMG* in time domain
(a):Spectrogram of meend SGRMG in real sitar
(b):Comparison of amplitude envelops of meend SGRMG and single note}
(c):Frequency variation of meend SGRMG

Let the duration of articulation be T. Also, the number of frequencies of the distinct microtones attached to the articulation of our interest be q and the corresponding frequencies be $f_1,\ldots,f_q$. Define the set $Q=\{ v_1,\ldots,v_q\}$ where $v_i=f_i/f_o$ for $i=1,2,\ldots,q$ and f_o is the base frequency. It is worth noticing that the base frequency is not included in this set. Define two monotonically increasing finite sequences

$\{S_n\}_{n=0}^{k+1}$ and $\{t_n\}_{n=0}^{k}$ with the property that $s_k<t_k<s_{k+1}$ for $i=1,\ldots k$, $s_1=0$ and $s_{k+1}=T$

Define: $\hat{f}:[0,T]\rightarrow R$ such that

$$\hat{f}(t)=\begin{cases} \dfrac{(\hat{f}(t_1)-1)}{t_1-s_1}t+1 & \text{for } t \in [s_1,t_1] \\ \dfrac{(\hat{f}(t_i)-\hat{f}(t_{i-1}))}{t_i-s_i}(tt1-s_i)+\hat{f}(t_i-1) & \text{for } t \in [s_i,t_i] \text{ and } i\geq 2 \\ \hat{f}(t_i)\rightarrow & \text{for } t \in [t_i,s_{i+1}] \end{cases} \quad \text{-------------} (11)$$

By virtue of equation (10) and equation (11), the required equationcan be formulated of the waveform by taking the effect on the frequency into account. Let the resulting function $g:[0,T]\rightarrow R$ be defined by

$$f(t)=\sum_{j=1}^{N} M_j e^{(F_s \ln\alpha_j)t}\cos\left(2\pi\hat{f}(t)\int_0^t p_j(u)du+\phi_j\right) \quad \text{-----------}(12)$$

$$h(t)=\begin{cases} \sum_{j=1}^{N} M_j e^{L(\ln\alpha_j+\gamma_1 f(t_1))t}\cos\left(2\pi\hat{f}(t)\int_0^t p_j(u)du+\phi_j\right) & \text{for } t \in [s_1,t_1] \\ \sum_{j=1}^{N} \hat{f}(s_i)e^{Fs(\ln\alpha_j+\gamma_2(\hat{f}(t_i)-\hat{f}(s_i))t}\cos\left(2\pi\hat{f}(t)\int_0^t p_j(u)du+\phi_j\right) & \text{for } t \in [s_i,t_i] \text{ and} \\ \sum_{j=1}^{N} \hat{f}(t_i)e^{Fs(\ln\alpha_j+\gamma_3(f(t_i)-\hat{f}(t_i))t}\cos\left(2\pi\hat{f}(t)\int_0^t p_j(u)du+\phi_j\right) & \text{for } t \in [t_i,s_{i+1}] \end{cases} \quad \text{---}(13)$$

A *meend* techniques can be model by using equation (12) and equation(13). To investigate the applicability of the model it is required to compare the frequency and the amplitude evolution of a *meend*created by a real sitar with the synthesis of the same *meend*. The notes *SGRMG* has been synthesized based on the equation (12) and equation (13) and also naturally generated using a sitar. For the purpose of comparison single note (Sa-261.3Hz) is excited on a sitar and recorded. The spectrogram corresponding single note excitation, the *SGRMG meend* generated on a real sitar and the Synthesized *SGRMG meend* using ARMA are shown in Figure 15-(a)-(b)-

(c) respectively.The Figure15-(d) compares the amplitude envelop of the fundamental frequency component obtain by above three methods.

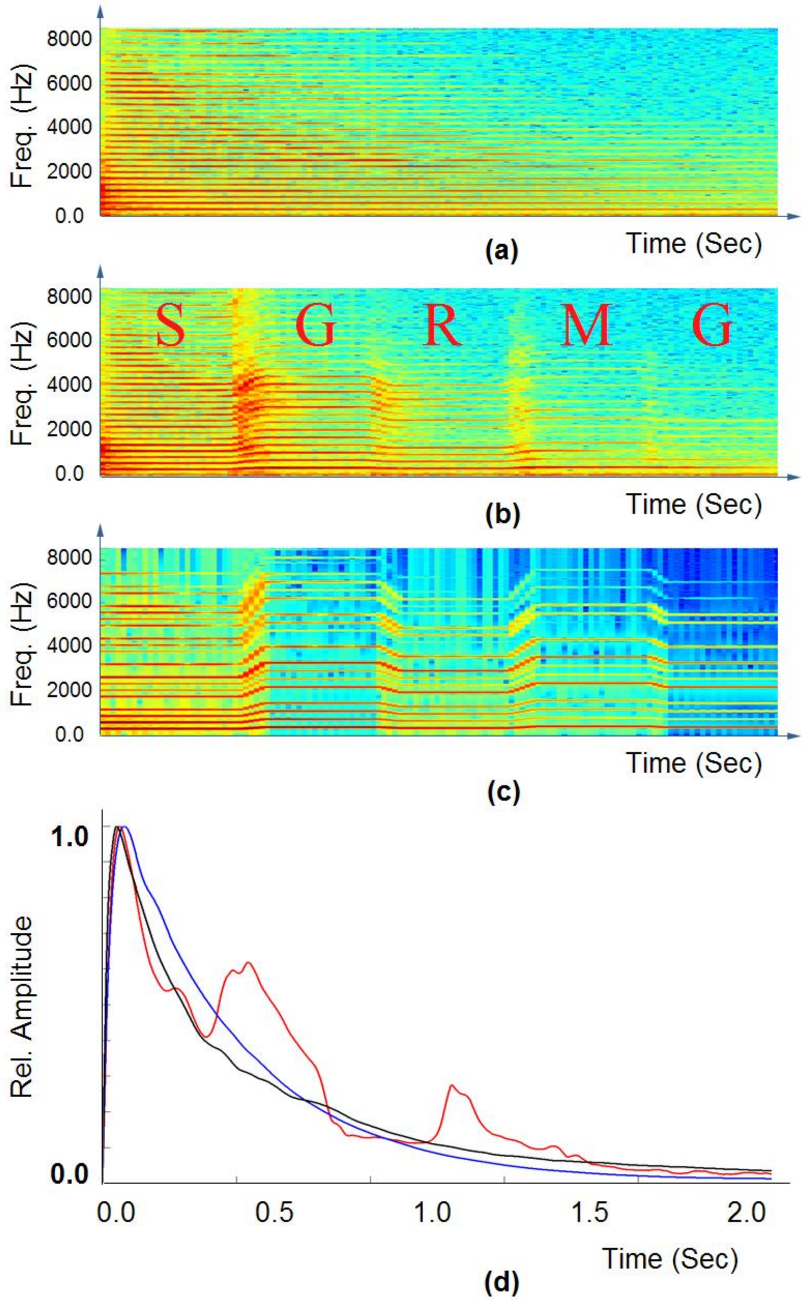


Figure 15: Comparison of Spectrogram
(a):Single (S-260.3Hz) note of real sitar
(b):Meend*SGRMG* performed in real sitar
(c):Meend*SGRMG* synthesized using ARMA
(d):Comparison of amplitude envelops

Conclusion

The curvature of sitar bridge makes its output nonlinear. In this effort, the nonlinear characteristics of sitar tone is approximated by a linear time variant stochastic process ARMA.
The validation of the results has turned out to be a challenging task. Thus, a metric space called timbre space of sitar is constructed by using Rate of Zero Crossing, Spectrum Centroid and Tristumulus-3 which quantify the salient properties of the sitar tone, buzziness, harmonic richness and brightness respectively. Figure 9 compare and contrast the positions of the real sitar tones and the ARMA model output within the timbre space and Figure 15 compares the spectrograms of articulation sound of the real sitar and the ARMA model.The degree of proximity indicates how close an actual sitar tone and the tone synthesized by the ARMA model are. It was evident from the relative positions of the model and the standard sitars that the ARMA model is sufficiently close to the actual.

References

i. Alsahlani A., & Mukherjee R., 2010. Vibration of a string wrapping and unwrapping around an obstacle. *Journal of Sound and Vibration*, Volume 329.

ii. Alves W., 2012. *Music of the People of the World*. 3rd ed. s.l.: Schirmer

iii. Bandyopadhyaya S., 1988. *Technique of Sitar*. Delhi: D.K. Fine Art Press.

iv. Bosq D., 2012. *Mathematical Statistics and Stochastic Processes*. s.l.: John Wiley.

v. B. S., 2014. The changing picture of nonlinearity in musical instruments modeling and simulation. *International Symposium on Musical Acoustics*.

vi. Dimitris J.G, &Manolakis G., 1996. *Digital Signal Processing, Principles, Algorithms and Applications*. s.l.:Prentice-Hall.

vii. Dmitri K., Anatoli S., Lehtonen H.M, &Vesa V., 2013.*Modeling a vibrating string terminated against a bridge*. Stockholm, Sweden: Proceedings of the Stockholm Music Acoustics onference 2013, SMAC.

viii. Esquef P.A.A, Karjalainen M., &Valimaki V., 2003. Frequency zooming arma modeling for analysis of noisy string instrument tones. *EURASIP Journal on Applied Signal Processing,*Volume 31.

ix. Grenier Y., 2003. Time dependent arma modeling of nonstationary signals. *IEEE Transaction on Acoustics, Speech and Signal Processing*, Volume 31, pp. 899-911.

x. Hanck C., 2008. *Stochastic Integration und Zeitreihen model lierung Learning Sequence in Music*. Springer Berlin Heidelberg: Springer.

xi. Janaki B., 2005. *Two Men and Music*. Chap 1-3. s.l.: Oxford University Press.

xii. Juan G. R., 2010. *The Physics and Psychophysics of Music-An Introduction*. 4th ed. s.l.: Springer.

xiii. Kartofelev D., Stulov A., & Valimaki V., 2003. Pitch Glide Effect Induced by a Nonlinear String-Barrier Interaction, Applications of Signal Pro- cessing to Audio and Acoustics.*IEEE Workshop*.

xiv. Lartillot & O. Toiviainen, P., 2007. *Mir in matlab (ii) a toolbox for musical feature extraction from audio*. Vienna: International Conference on Music Information Retrieval, Vienna.

xv. Lindgren G., Rootzen H., &Sandsten M., 2013. *Stationary Stochastic Processes for Scientists and Engineers*. s.l.: CRC.

xvi. Lutkepohl H., 2005. *New Introduction to Multiple Time Series Analysis*. s.l.: Springer.

xvii. Meillier J.L.,& Chaigne A., 1991. Ar modeling of musical transients. *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, Toronto, Canada*, pp. 3649-3652.

xviii. Miranda E., 2002. *Computer Sound Design Synthesis techniques and programming*. s.l.: Focal Press.

xix. Niedzwiecki M., 1993. Statistical reconstruction of multivariate time series. *IEEE Transaction on Signal Processing*, 44(3), 528-537.

xx. Pollock D.S.G., 1990. *A Handbook of Time-Series Analysis, Signal Processing and Dynamics*. s.l.: Academic Press.

xxi. Raman C. V., 1921. On some indian stringed instruments, in Proc. Indian string terminated against a bridge. *Proceedings of the Stockholm Assoc. Cultiv.* , Volume 7 pp. 29-33.

xxii. Sharma R., 2010.*Bhartiya Sangeet Ka Itihasa*. s.l.: Vani Prakashan.

xxiii. Stephen M. S., 1989. *Sitar Technique in Nibaddh Forms*. s.l.: University of Texas Press.

xxiv. Tolonen T., & Karjalainen, M., 1999. Plucked string synthesis algorithms with tension modulation nonlinearity, Acoustics, Speech, and Signal Processing. *IEEE International Conference*, Volume 2.

xxv. Vnner M., 2005. *The MPEG home page , MPEG doc No: N7708*. [Online] Available at:http://mpeg.chiariglione.org/standards/mpeg-7/audio