OBJECTIVE ANALYSIS OF SHRUTIS FROM THE VOCAL PERFORMANCES OF HINDUSTANI MUSIC USING CLUSTERING ALGORITHM

A. K. Datta[1], R. Sengupta[2] and N. Dey[3]

[1]BOM Public Charitable Trust,
3/3 Girish Ghosh Street, Kolkata 700035, India
dattashoke@yahoo.com

[2]Sir C V Raman Centre of Physics and Music
Jadavpur University, Kolkata 700 032, India
sgranjan@gmail.com

[3]Scientific Research Department, ITC Sangeet Research Academy,
1, N S C Bose Road, Tollygunge, Kolkata 700 040, India
srasrd@vsnl.net

Abstract

An objective analysis of shrutis from the 'alap' from 150 songs performed by 53 eminent musicians and scholars covering 21 different ragas are presented in this paper. The total time of singing analysed is nearly 8.5 hours. The audio samples are taken from recorded performances. Pitch extraction is done using a specially developed PDA/VDA (phase-space algorithm). The experiment handled 1.36 GB of data to extract the steady pitch states. The steady states are determined from the extracted pitch sequences. The ratio of the steady states with the tonic of each artist in each song was then found out. The complete data of 150 songs was then folded into the middle octave and distributed over 1200 bins of one-cent width each. The shruti positions were then ascertained using a clustering algorithm. The non-cognitive approach utilized here for the extraction of shrutis is unique in nature.

Introduction

Musicological research has long existence since ancient times. The present state of science and technology can provide ample scope for quantitative as well as qualitative approach to investigate shrutis, swaras, intervals, octaves (*saptak*), consonance (*vaditya*), musical quality, rhythm etc. [1]. The performance of the 'ragas' in Indian Classical music is based on both explicit knowledge (which forms the so called grammar and is transmitted through the tradition of oral teaching and also through documented analysis e.g. old Shastra's, Bhatkhande's works etc.) and non-explicit knowledge which is born out of the musician's sensibilities and normally defies identification, categorisation etc. in the manual realm.

Shruti appears to be the basic interval on which the Indian musical scale rests [2]. The concept of shruti is as old as the classical music itself, at least four millennia [3]. According to Sarangadeva[2] shrutis are pleasing and discernible. However scanty and vague this definition may appear in the objective context of modern science, this definition together with the 'shastriya' concept of pleasingness as 'vaditya' (consonance) [2] in Indian classical music led to the development of a psycho-perceptual objective model for Indian musical scale [4,5]. There have been various opinions about the number of shrutis during this long span of development. References are found for the number to be 66, 53[2] and 22[6,7]. Though the number 22 is now broadly agreed upon controversies on the exact number of these still persist.

It should be borne in mind that objective attention from an experimentalist point of view in this area is a comparatively recent phenomena. Furthermore psycho-perceptual categoricity in perception of pitch in music mode together with a vague notion of shruti boundaries make the understanding of the scale and its use both in performance and appreciation quite complex an affair. Some studies in this area have been reported earlier [4,5].

Technological advancement has made it feasible to provide objective assessment of the measurement and usage of the shrutis [8]. It is of prime importance to a music practitioner that he/she is equipped with an objective and standard shruti position for every note he/she employs in every raga. Research by Indian Musicologists over the last 70 years has concentrated on this area with insufficient success, perhaps owing to confusion between theoretical tonal framework & exigencies of practice and aesthetics [9]. This entire area is in dire need of rationalization.

In this paper we present an objective analysis of 'alap' from 150 songs performed by 43 eminent musicians and scholars covering 21 different ragas. The total time of singing analysed is nearly 8.5 hours. The audio samples are taken from recorded performances. The experiment handled 1.36 GB of data to extract the steady pitch states. The steady states are determined from the extracted pitch sequences. The ratio of the steady states with the tonic of each artist in each song was then found out. The complete ratio-data of 150 songs was then folded onto the middle octave and distributed over 1200 bins of one-cent width each. The shruti positions were then ascertained using a clustering algorithm. The results obtained were statistically significant.

Methodology

A. Pitch Period Extraction from Signal

A method based on Phase-Space Analysis (PSA) [10] was used for extracting pitch periods. The pitch pattern files, extracted from all the 116 signal files using the above method, is referred to as '.cep file'.

B. Smoothing

The .cep file contains pitch extracted only in the quasi-periodic region of the signal. The PD algorithm above uses a predefined specific range for possible pitch values e.g., in the present case the default range is fixed between 70 Hz to 700 Hz. The smoothing operation for the first type of error is simply detection of it and doubling or halving the value as required. For the two other types of error are detected through an examination of local pitch values and the erroneous values are replaced using linear interpolation from the neighbouring valid data. The resulting pitch profiles are put into 116 'pit files'.

C. Steady State Detection

Even in a perfectly perceptible steady note all the pitch values are never identical. There are always some involuntary small variations in pitch, which has no relevance in the perception of the corresponding note. To remove these we simply replace the $(i+1)^{th}$ pitch x_{i+1} by the i^{th} pitch x_i when $|(x_{i+1} - x_i)| > x_i * 0.1$ & $x_{i+1} > 0$. This creates a p-file. The definition of steady state is not a trivial problem. We define a steady state in the pitch file as that subsequence of pitch data where all the pitch values lies within a predetermined band around the mean value. From p-file steady state sequences are created with all consecutive pitch in a sequence, which is terminated when $| x_{i+1} - M | > M/30$ where $M = (1/i) \Sigma x_i$.

Furthermore as the signal considered in the present study is taken from '*alap*' portion of the total singing we restricted our steady states to minimum duration of 60 milliseconds. This same minimum duration has also been used for detecting the validity of a silence. If two steady states are separated by silence for duration less than this minimum time, the silence is deleted. Similarly if two silences are separated by a steady state for duration less than this minimum duration, the two silences are clubbed together to get one bigger silence, ignoring the steady state in between. The resulting steady states (derived from the pit files, stated above) are put into total 116 'std files'.

D. Tonic (Sa) Measurement

A skilled musician was requested to listen to the signal files one after another to detect the position of Sa in the file. By frequency domain analysis at the detected region, Sa for the respective signal file is calculated. All these Sa values (in Hz.) are put into a file, 'safile.txt'.

E. Finding Ratio-Intervals

The pitch values of the steady states in 'std files' are first divided by the pitch value of the corresponding 'Sa' obtained from the 'safile.txt'. This gives the frequency ratios. The complete data of 150 songs was then folded into the middle octave and distributed over 1200 bins of one-cent width each. K-Means clustering algorithm was then applied to find out the 22 shrutis used by the musicians in actual singing.

F. Clustering Algorithm (K-Means)

This nonhierarchical method initially takes the number of components of the population equal to the final required number of clusters. In this step itself the final required number of clusters is chosen such that the points are mutually farthest apart. Next, it examines each component in the population and assigns it to one of the clusters depending on the minimum distance. The centroid's position is recalculated every time a component is added to the cluster and this continues until all the components are grouped into the final required number of clusters. This is an algorithm [11] for partitioning (or clustering) N data points into K disjoint subsets S_j containing N_j data points so as to minimize the sum-of-squares criterion

$$J = \sum_{j=1}^{K} \sum_{n \in S_j} |x_n - \mu_j|^2 ,$$

where x_n is a vector representing the nth data point and μ_j is the geometric centroid of the data points in S_j.

Experimental Details

Twenty-one ragas sung by fifty-three (53) singers from five different gharanas of Hindustani music were taken for analysis. The database constituted of 8.5 hours of recording for four ragas sung by 53 singers of 5 gharanas (viz. Agra, Patiala, Kirana, Gwalior and Mixed). In most cases the notes extended on both sides of the middle octave. The digitisation of the signal was done at the rate of 22050 samples /sec (16 bits/sample). For our analysis only the aalap part of each singer was selected from each raga. Pieces of alap for each singer for a raga were taken out from the complete alap deleting the bandish part. These constituted the alap signal files for a singer for each raga. For each raga a singer had one alap signal of ~8 to 10 minutes. One hundred fifty (150) alap signal files, stored as 'wav files', were thus selected for analysis, which constituted our database. The Mixed gharana constitutes those contemporary singers who do not belong to a particular traditional gharana. They are known to receive training from teachers of different traditional gharanas.

Pitch periods were extracted using a method based on Phase-Space Analysis (PSA), as described in the earlier section. The resulting pitch pattern files, extracted from all the 150 signal files using the above method were stored as 150 'cep files'.

The smoothing operation, as detailed in the earlier section, was done on the above pitch pattern files. The resulting pitch profiles were put into 150 'pit files'. Then steady state detection using minimum duration restriction and other criteria, as outlined in the earlier section, was done. The resulting steady states (derived from the pit files, stated above) were put into total 150 'std files'. The experiment handled 1.36 GB of data to extract the steady pitch states. The steady states are determined from the extracted pitch sequences.

A skilled musician listened to the signal files one after another to perceptually detect the position of Sa in the file. By FFT analysis, using a standard software package, at the detected region, Sa for the respective signal file was calculated. By suitably resolving the ambiguity, in some cases, total 53 Sa values for 53 singers were finalized. All these Sa values (in Hz.) were put into a file, 'safile.txt'.

Results and Discussion

Figure 1 shows the 3D spectra, waveform and raw pitch profile of a song-sample (Mian-ki-Mallhar) extracted by PSA algorithm. The pitch data are then smoothed and the steady state was detected. Figure 2 shows the frequency distribution of one song of raga Bhairav spread over 1200 bins of 1-cent interval. The peak shows the ratios, which correspond to shruti positions R1, R2, G4, M1, D1 and N4, utilized by the artist in this song. For individual songs of each artist peak positions revealed various shruti positions those the singers have used in the corresponding songs. When the steady states of all the 150 songs are pooled together and put into the 1200 bins the shruti positions get lost and we get the distribution, which shows peaks at the 12 note ratios (Figure 3). In this sense one can say that the twelve notes are the receptacle of the shrutis. The open circles in figure 3 represents the values of the shruti positions extracted using the clustering algorithm. The clustering algorithm described in an earlier section is used to find the mean positions of the original shruti clusters from this pooled data. The shruti positions obtained from different songs are used as initial seed positions for the

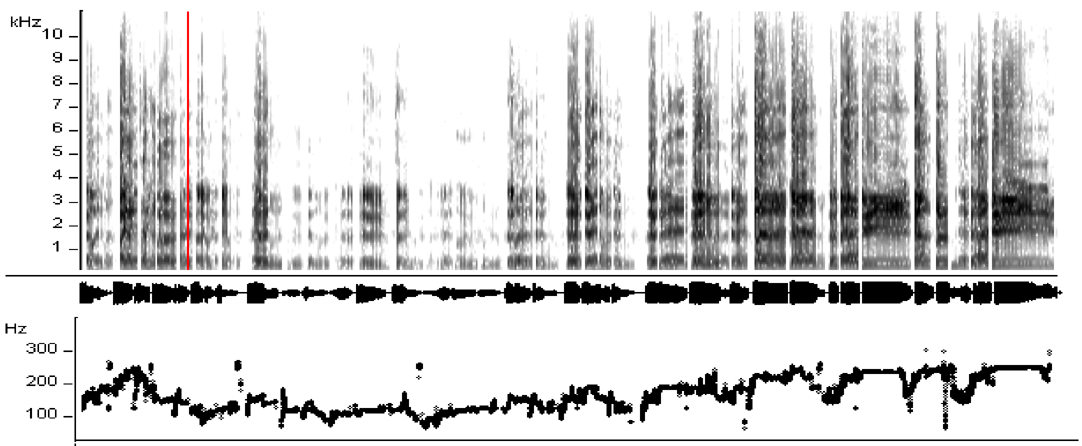

Figure 1: 3D spectra, waveform and raw pitch profile for raga Mian-ki Mallhar

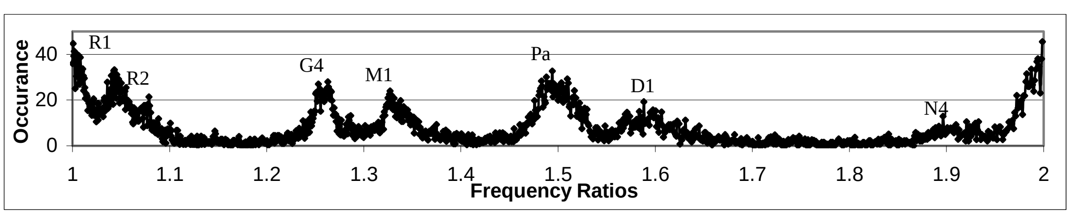

Figure 2: Distribution of the frequency ratios in 1200 bin of a song-sample (Bhairav)

clustering algorithm. This algorithm uses these seed points and finds the means of optimum cluster from the pooled data. These mean values are given in column 2 of Table

1. This table also presents the values in cents of different shrutis from Deval [12], Bhatkhande [7] as well as those compiled by the western musicologist Daniloue and

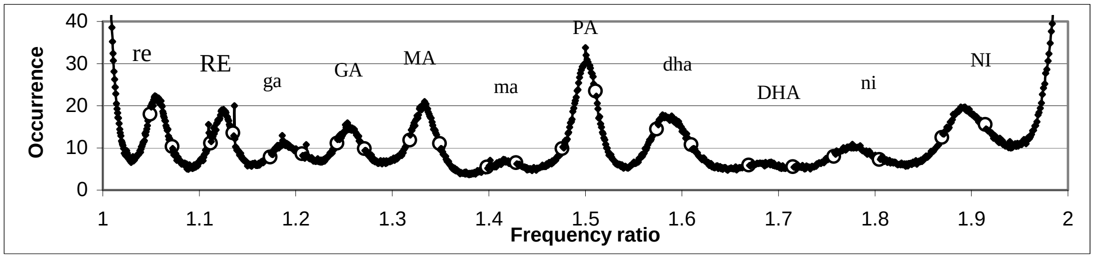

Figure 3: Distribution of the frequency ratios of all data (150 songs) pooled together

Lentz [13]. The ratios provided by Deval is closest to the algorithmic values closely followed by Daniloue and Lentz, the average discrepancy being close to10 cents. R4, G1, D4, N1, N4 of Debal and Western only show difference to be noticeably large. The ratios proposed by Bhatkhande (column 9) are a little far off.

Table I: Ratios obtained from Clustering algorithm and its variation from the standard scales in cents. [**(A) -**Algorithmic values, **(D)-** Debal, **(W)-** Western, **(B)-** Bhatkhande]

Shruti	**(A)**	**Nearest Ratios**	**Values in Cents**				**Diff. in cents from Alg_value**		
			(A)	**(D)**	**(W)**	**(B)**	**(D)**	**(W)**	**(B)**
R1	1.04671	22/21	79.03414	84.46719	90.225	58.69152	5.43305	11.19085	20.34262
R2	1.070762	15/14	118.3654	111.7313	111.7313	111.7313	6.634129	6.634129	6.634129
R3	1.111183	10/9	182.5157	182.4037	182.4037	165.0042	0.112007	0.112007	17.51149
R4	1.134694	17/15	218.7639	203.91	203.91	216.6867	14.85395	14.85395	2.077253
G1	1.17311	27/23	276.406	294.135	294.135	277.5907	17.72904	17.72904	1.184698
G2	1.205566	29/24	323.6528	315.6413	315.6413	327.6222	8.011475	8.011475	3.969431
G3	1.243383	26/21	377.1249	386.3137	386.3137	386.3137	9.188803	9.188803	9.188803
G4	1.271147	14/11	415.3571	407.82	407.82	435.0841	7.537051	7.537051	19.72704
M1	1.316667	25/19	476.2686	470.7809	498.045	498.045	5.487714	21.77638	21.77638
M2	1.348594	27/20	517.7473	498.045	519.5513	543.0146	19.7023	1.803988	25.26734
M3	1.398649	7/5	580.8407	590.2237	590.2237	596.9996	9.38297	9.38297	16.15885
M4	1.428004	10/7	616.8	609.7763	609.7763	656.9854	7.02374	7.02374	40.18533
P1	1.475638	28/19	673.6066						
P2	1.510041	3/2	713.5053	701.955	701.955	701.955	11.55026	11.55026	11.55026
D1	1.573047	11/7	784.2741	786.4222	792.18	764.9159	2.148062	7.905864	19.35823
D2	1.612613	29/18	827.2803	813.6863	813.6863	813.6863	13.59402	13.59402	13.59402
D3	1.667811	5/3	885.547	884.3587	884.3587	863.8705	1.188257	1.188257	21.67647
D4	1.709233	29/17	928.0189	905.865	905.865	924.6218	22.15389	22.15389	3.397107
N1	1.758579	7/4	977.2922	996.09	996.09	983.3133	18.79782	18.79782	6.02113
N2	1.803481	9/5	1020.941	1017.596	1017.596	1034.996	3.344781	3.344781	14.0547
N3	1.869026	28/15	1082.744	1088.269	1088.269	1088.269	5.524748	5.524748	5.524748
N4	1.913853	23/12	1123.776	1109.775	1109.775	1141.308	14.00102	14.00102	17.53245

Column 3 of table I give the nearest small fractions to the shruti positions given in column 2. The average difference is less than 3 cents. These ratios may be considered as the shruti positions used in actual performances. It may be mentioned here that the data from individual songs revealed the shruti positions clearly (figure 2) establishing beyond doubt the use of shrutis in performances of Indian Classical singing. The analysis of individual songs revealed two separate clusters for Pa, one at 28/19 in addition to the normal Pa at 3/2. The first cluster point i.e. P1 was found only in six out of 150 songs.

Conclusion

The non-cognitive approach utilized here for the extraction of shrutis reveals:

1. The existence of 22 shrutis

2. The extracted ratios are generally in good agreement with the ratios given by Debal and the Western sources.

Reference

[1] A. K. Datta, R Sengupta, N Dey and D Nag, "On Scientific Approaches to the Study of Vaditya in Indian Music, Proc. Fifth Int. Workshop on Recent Trends in Speech, Music and Allied Signal Processing, Thiruvanathapuram, India, 14-15 Dec., 2000.

A.K. Datta, R Sengupta, N Dey and D Nag, "Pitch Analysis of Recorded Vocal Performances in Hindustani Music : Evidence of a Personal Scale", J. Acoust. Soc. Ind., XXV, VI-1.1, 1997.

A.K. Datta, R. Sengupta, N. Dey and D. Nag ,"Multiple States in a Note in Hindustani Music and their Relevance to Consonance", J. Acoust. Soc. Ind., XXVI (3-4), 394,1998.

R. Sengupta, "Study on some Aspects of the Singer's Formant in North Indian Classical Singing", J. of Voice, 4 (2), 129, Raven Press, New York, 1990.

A.K. Datta, R Sengupta, N Dey and D Nag, "Studies on Identification of Raga using Short Pieces of Taan : A Signal Processing Approach", Proc. Sixth Int. Workshop IWSMSP, New Delhi, India, 19-21 Dec., 2001.

B.M. Banerjee, et. al., "Spectral Characteristics of Vowels Sung in Hindustani Sastriya Sangeet", J. Acoust. Soc. Ind., XI (4), 47, 1983.

2. Sarangadeva,"Sangeet Ratnakar with Comments of Kalinath", Anandashram ed. 1987.
3. A K Coomarswamy, "The Transformation in nature of Arts", Harvard University Press, Cambridge, Massachussetts, 1935.
4. A.K. Datta and R. Kundu. "A Psycho-Perceptual Hypothesis for Indian Musical Scales", Proc. IWSMSP, Indian Statistical Institute, New Delhi, pp. m31, 1991.
5. A.K. Datta, R Sengupta, N Dey and D Nag, "Relevance of Consonance in Indian Musical Scale : Theory and Practice", J. Acoust. Soc. Ind., XXIII, 110, 1995.

   A.K. Datta, R Sengupta, N Dey and D Nag, "Ratios and intervals used in north Indian classical music-a study based on recorded performance", J. Acoust. Soc. India, XXIV, pp. III-4.1 to III-4.6, 1996
6. Matanga,"Brihaddesi", Samvasiva Sastri(ed), Trivandram Sanskrit Series, 1928
7. Pt. Bhatkhande, "Hindusthani Sangeet Padhati", Bombay, 1937
8. B C Deva, "Music and Science", J.Indian Musicological Society, Vol.5 (1), 1974.

   Ibid,"Some Problems in Science and Music", J.Sang.Nat.Acad.(April-June), 1970

   R Sengupta, N Dey , D Nag and A. K. Datta, "Automatic Extraction of *Swaras* and *Srutis* from Khyal Renditions", J. Acoust. Soc. Ind., 30, 139, 2002.

   A. K. Datta, R Sengupta, N Dey, D Nag & A Mukerjee "Study of srutis in Indian musical scales and relevance of schools in recent times using signal processing techniques", Proc. FRSM-2003, IIT Kanpur, February 15-16, 2003.

   A.K. Datta, R. Sengupta, N. Dey, D. Nag and A. Mukerjee, "Sruti usage by old and contemporary singers in Khyal: An objective approach", Proc. FRSM-2004, Annamalai University, 2004.

   A K Datta et.al., "Evaluation of srutis in Hindustani Music from recorded performances", Proc. Frontiers of Research on Speech and Music (FRSM-2005), 6-7 January, 2005, Utkal University, Bhubaneswar, India.

   A K Datta et.al., "A Psycho-Perceptual hypothesis for srutis in Indian Music", Ninaad (J. ITC Sangeet Research Academy), vol.19, 2005.

   A K Datta et.al., "Objective Analysis of the Interval Boundaries and Swara-Shruti relations in Hindustani vocal music from actual performances", Ninaad (J. ITC Sangeet Research Academy),vol.20, 2006.

9. Matanga, “Brihaddesi”, Samvasiva Sastri (ed.), Trivandrum Sanskrit Series, 1928.
Pandit Bhatkhande, “Hindustani Sangeet Padhati”, Bombay, 1937.
S. M. Tagore, “Hindu Music”, Hindu Patriot, Sept. 7, 1874.
Pandit Abhola, ”Sangita Parijata”, Kalivara Vedanta-Vagisa (ed.), Calcutta ed., 1884.
B C Deva, “Music and Science”, J.Indian Musicological Society, Vol.5 (1), 1974.
B C Deva, “Some Problems in Science and Music”, J.Sangeet Natak Academy, (April-June),1970.
10. A. K. Datta, “Generation of Musical Notations from Song using State-Phase for Pitch Detection Algorithm”, J. Acoust. Soc. India, Vol. XXIV, 1996.
11. E. W. Weisstein. "K-Means Clustering Algorithm." From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/K-MeansClusteringAlgorithm.html
J Hartigan and M Wang, “A K-Means clustering Algorithm”, Applied Statistics,28, 100-108,1979
12. Rao Sahib M.Abraham Pandither, “Karunamirtha Sagaram – On srutis”, Asian Educational Services, New Delhi, 1984.
13. A. Daniloue, “Introduction to the Study of Musical Scales”, Oriental Books Reprints Corpn., New Delhi, 1979.
D A Lentz, “Tones and Intervals of Hindu Classical Music”, Nebraska Univ. Studies, New Series No. 24, Lincoln Univ. Press, 1961.