



# **Rhythmic Structure based segmentation for Hindustani music**

**T.P.Vinutha  
Prof. Preeti Rao**

**Department of Electrical Engineering  
Indian Institute of Technology  
Bombay**



# OUTLINE

- Introduction
  - Hindustani Music concert
  - Bases of segmentation, our approach
  - Problem addressed, motivation
- Rhythm
  - Percussion rhythmic structure as cue
- Implementation
  - Database
  - Audio segmentation system
- Conclusion & Future scope
- References





# Bases for segmentation



- Musical Structure
  - Repetitions, contrast and homogeneity of musical aspects such as
    - Melody
      - Steady in *Alap* and rendered around 'sa', few gamaks and spread throughout the octave in *sthayi*, traversing up to upper octave in *Antara*, more dynamic in *Tan* and *Sargam*
    - Timbre
      - Instrument change or change in singer
      - Vowels are used in *Alap*, *Aakar vistar* and *Tans*, but consonants in *Bol-alap*, *Sargam* sections
    - Rhythm
      - Rhythm refers to all aspects of musical time patterns
      - the inherent tempo of the melodic piece, the way syllables of the lyrics are sung or the way strokes in instruments are played
- *Rhythm is used as cue for segmentation*
  - *As it is more explicitly changing in most of the sections*



# Problem addressed , Motivation



- Problem addressed
  - Automatically locate the points of significant change in the rhythmic structure of the Hindustani music concert by analyzing the local self-similarity of a rhythm representation computed with a sliding window over the audio recording
- Motivation
  - Efficient navigation of audio recordings to different sections



## Different temporal scales of segmentation

- Khyal vocal concert
  - Major segments are, Alap, Bada khyal and Chota khyal
  - Bada khayal or chota khyal can be further segmented into
    - Sthayi, Antara, Vistar, Bol-alap, Sargam, Tan, Bol-baat sections
- Tabla solo
  - Major segments are, Peshkar, Kaidas, Gats, Relas and Rouns
  - Each of these can have slow and fast sections within it



# RHYTHM OR LAYA

- Rhythmic pattern represents the periodicities of events in the audio and their relative strengths
  - These events could be percussion strokes, syllable onsets in lyrics, note onsets of melodic instruments
- ***Percussion rhythmic structure is taken as cue for segmentation***
  - Hindustani concerts use tabla as an accompaniment
  - an explicit representation of the rhythm is available in terms of the pattern and timing of tabla strokes.
    - Eg., no tabla in alap, tempo is either slow or medium in bada khyal section and increases drastically in chota khyal section



# Tal in Hindustani music



- Tal
  - Is the cyclically recurring metrical pattern of fixed length imposed by the accompanying percussion instrument
- Theka is the defined structure of a tal
  - representing particular pattern of strokes (identified by bols or names)

The Theka for Tintal is,

X	2				0				3						
dha	dhin	dhin	dha	dha	dhin	dhin	dha	dha	tin	tin	ta	ta	dhin	dhin	dha
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.				.				.				.			
.															

**Matra**  
**Vibhag**  
**Avart**

## Three metrical levels of tal

- **Matra**
  - equivalent to beat
- **Vibhag**
  - Section or measure consisting of 2or5 matras
- **Avart**
  - Tal cycle,
  - theka is repeated in each avart

In a performance, the tabla player will deviate from the theka



# Prakars of Kaherva tal



- Prakar is an alteration of the basic form of tala

*Baisc theka  
of Kaherva  
tal*

<b>X</b>					<b>0</b>			
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	
dhā	ge	na	ti	na	ka	dhin	na	



**Prakar1**

<b>X</b>					<b>0</b>			
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	
dhā	ti	ti	Tā	tā	dhin	dhin	dhā	



**Prakar2**  
2 pause bols

<b>X</b>					<b>0</b>			
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	
gha		ti	Tā		ga	dhin	nā	



**Prakar3**  
Matra 2  
with 2bols

<b>X</b>					<b>0</b>			
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	
dhā	ti	ra	ki	ti	tā	ge	dhin	nā

**Prakar4**  
Matra 2 & 6  
with 2bols

<b>X</b>					<b>0</b>				
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>		
dhā	ti	ra	ki	ti	tā	ti	ra	ghi	dhin



# Relationship of Laya with tempo



- **Tempo**

- Represents the rate of succession of one of the pulse rates, selected according to the human perception. Expressed in beats per minute (bpm)

Laya	Tempo (bpm)	Vibhag duration (in sec)
Ati Vilambit	Very slow	8
Vilambit	Slow	4
MadhyaVilambit	Medium slow	3
Madhya	Medium	2
MadhyaDrut	Medium fast	1.5
Drut	fast	1
AtiDrut	Very fast	0.75

- **Tempo in Hindustani music**

- Indian music has traditionally three main tempos or laya
  - » *vilambit* (slow)
  - » *madhya* (medium)
  - » *drut* (fast)
- matra rate defines the tempo in madhya laya

- **Surface Rhythm**

- In a performance, pulse rate deviates from madhya laya
- the listener perceives the surface rhythm but within the context of tal framework

Gottlieb, Robert S. *“Solo tabla drumming of North India: its repertoire, styles, and performance practices”*, Motilal Banarsidass Publishers, 1993.

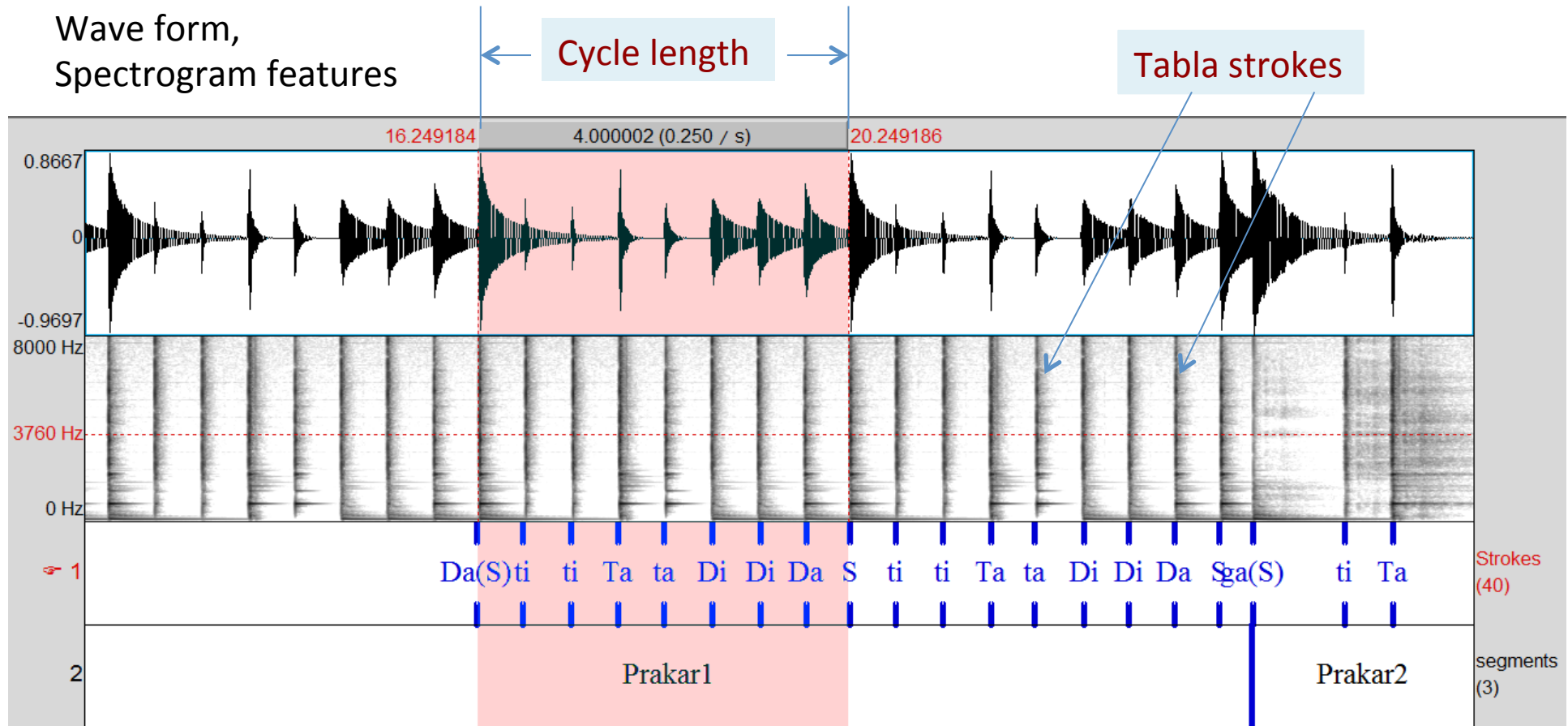


# Rhythm structure change

- *Tala change*
  - In some concerts, bada khyal section will be in a particular tal, while chota khyal section will be in different tal
    - Eg., In raag Marwa performance by RK, tal changes from Ektal to Tintal
- *Tempo change*
  - In a performance, bada khyal section will be presented in either madhaya laya or vilambit laya and chota khyal in drut laya.
- *Improvisation*
  - Increasing the rhythmic density
  - Replacing some bols in theka by other bols
  - Introducing pause bols



# Audio signal characteristics



- Onsets are wide band, transient events representing the location of tabla stroke.

- **Rhythmic structure from the audio can be derived by the pattern of derived onsets**

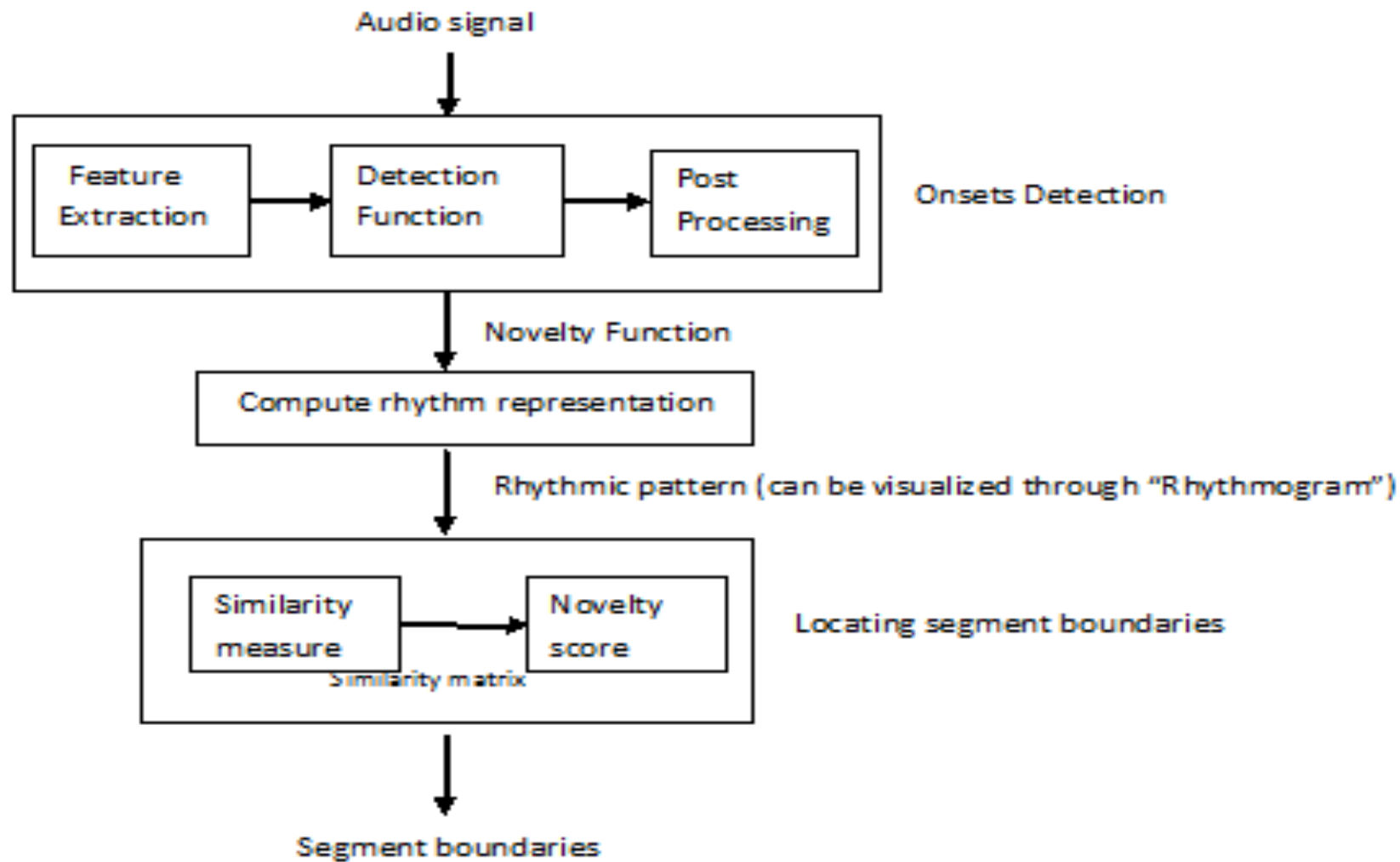


# IMPLEMENTATION

- Database
  - Prakarās of Kaherva tal (used for bhajans)
    - To do signal dependent parameter selection and to have controlled variability in rhythm
  - Tabla solo performances
    - Having lead instrument as tabla
    - having highest rhythmic improvisation
  - Khyal Vocal Concerts in Tintal and Ektal
    - Complexity due to Polyphonic audio
    - ensuring that the performances are rendered in different laya
- Ground truth
  - segment boundaries in these performances have been marked in PRAAT with careful hearing

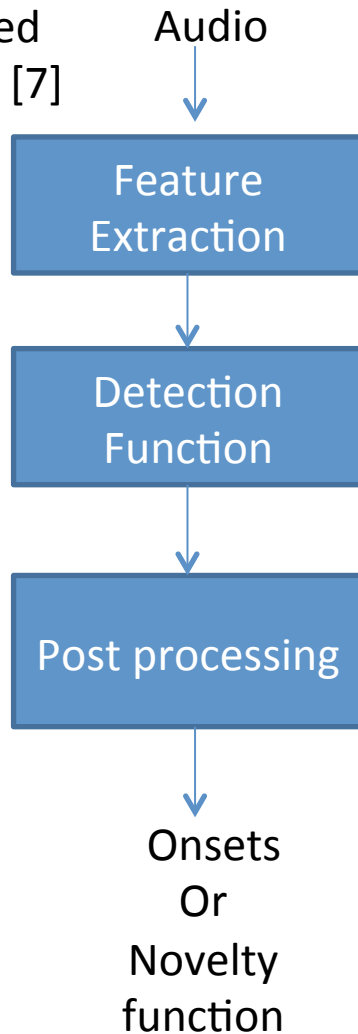


# Audio Segmentation system





scheme  
suggested  
by Bello [7]



# Onsets detection



- **Acoustic features** (suggested by Bello[7])
  - **Temporal Features**
    - Full band energy
    - Sub band energy
  - **Spectral features** (preferred by Dixon[8])
    - **Sub band spectral amplitude**
    - **Spectral amplitude**
- **Detection Function**
  - **Derivative**
  - **Smoothed Differentiator** (suggested by Hermes[7])
- **Post Processing**
  - **Normalization**
  - **Thresholding**



# Detection function



- As a derivative of adjacent frames
  - SF(n) represents the rectified spectral flux that is summed for all the bins.

$$SF(n) = \sum_{k=1}^K H[ (|X(n, k)| - |X(n-1, k)|) W(k) ]$$

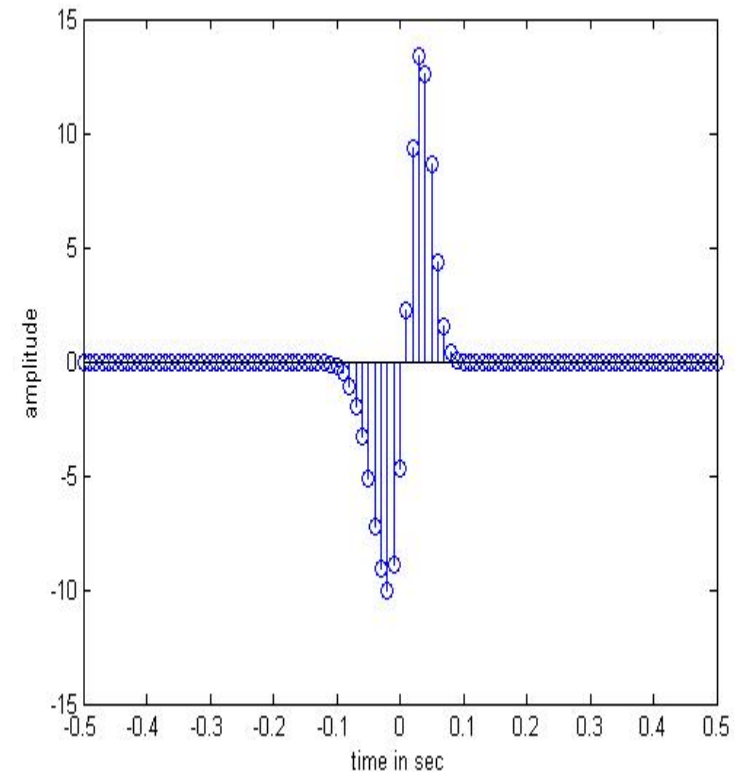
- As a bi-phasic function
  - involves multi-frame smoothing & differencing
  - is less susceptible to rapid local fluctuations



# Bi-phasic function as smoothed differentiator



- Hermes[15] suggested the parameters of bi-phasic function as
  - $\tau_1 = 0.015s$ ;  $\tau_2 = 0.02s$ ;  $d_1 = 0.025s$ ;  $d_2 = 0.05s$
  - to simulate short term adaptation characteristic of human ear
  - That emphasizes recent inputs while masking rapid modulations ( $\tau_2 > \tau_1$ )
- Parameters based on acoustics of tabla strokes
  - $\tau_1 = 0.02s$ ;  $\tau_2 = 0.0333s$ ;  $d_1 = 0.0289s$ ;  $d_2 = 0.0067s$
  - Stroke durations are ranging from 400ms for tonal strokes to 30ms for impulsive strokes.



$$P(t) = \frac{1}{\tau_1 \sqrt{2\pi}} \exp\left(-\frac{(t-d_1)^2}{2\tau_1^2}\right) - \frac{1}{\tau_2 \sqrt{2\pi}} \exp\left(-\frac{(t+d_2)^2}{2\tau_2^2}\right)$$



# Post -Processing



- Normalization
  - by subtracting the mean
    - so that the average will be zero
  - dividing by the max absolute deviation,
    - so that the function will be in the interval [-1,1].
- Fixed threshold method
  - peaks where the detection function exceed the threshold  $\delta(n)$ , a positive constant are considered as onsets
    - to separate the event-related and non event-related onsets
- *Adaptive threshold method*
  - based on the local mean is implemented as,

$$\tilde{\delta} [n] = \delta + \text{mean}(|d(n-M)|, \dots, |d(n+M)|)$$



# Audio Segmentation system





# Computing rhythmic representation



- Autocorrelation method avoids explicit extraction of onsets by comparing the novelty curve with the time shifted copies of it to analyse the periodicity in the rhythmic pattern
- Here, autocorrelation of the spectral flux feature is computed.
- **Rhythmic feature of block n can be expressed as,**

$$r_n(k) = \sum_{m=0}^{N-1-k} [sf(n+m)w'(m)][sf(n+m+k)w'(k+m)]$$

where, n is the block index, N is the block size and k is the lag.

- Only the information between zero and 3s lag is retained as rhythmic feature.



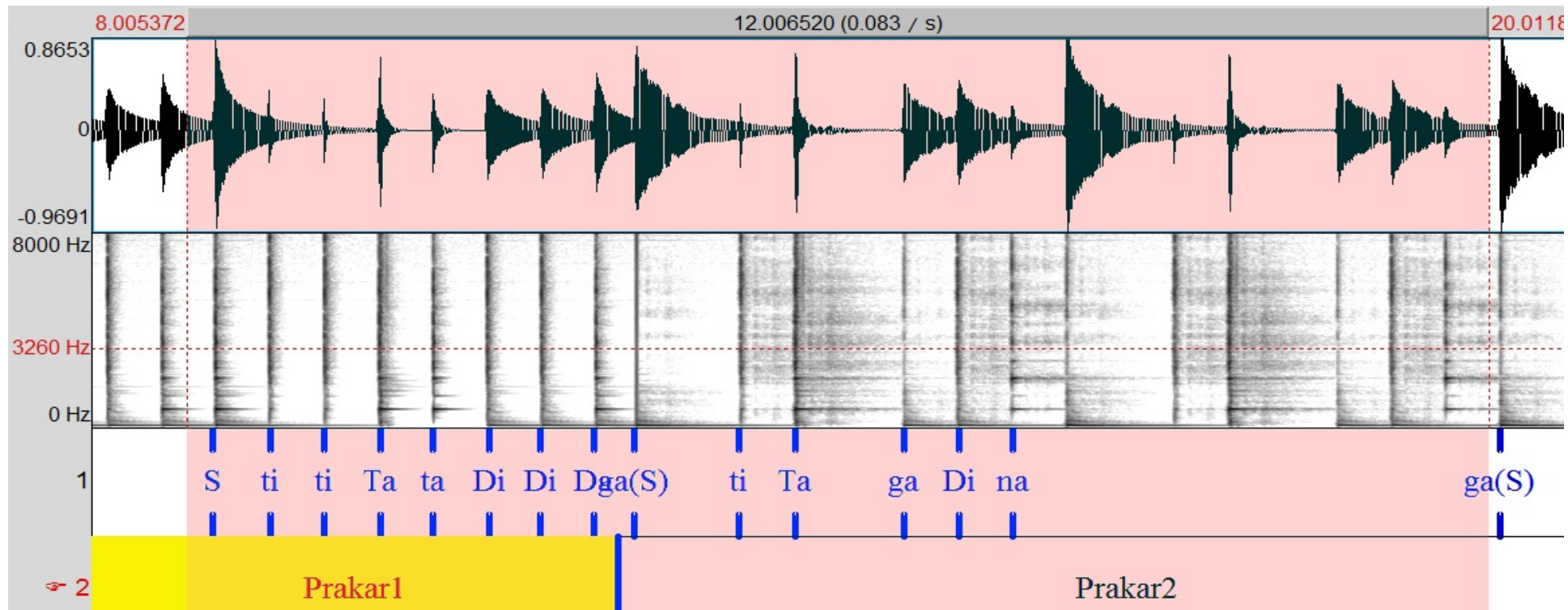
# Acoustic description of

## Segments with deviation from basic structure of tal



- Few cycles of 3 prakaras are concatenated and the resulting audio is analyzed.
- Here, prakaras are of nearly same tempo but of different rhythmic pattern.
- The cycle length of each prakara is 4.027s, 3.98s and 3.99s respectively.

### Prakar1 & Prakar2 boundary



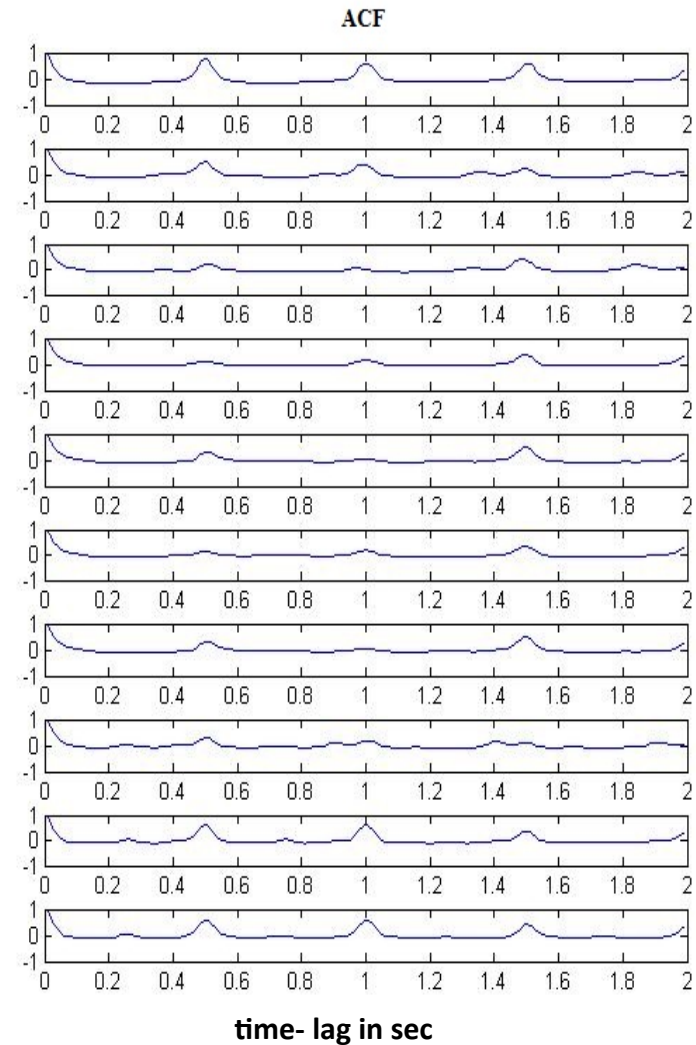
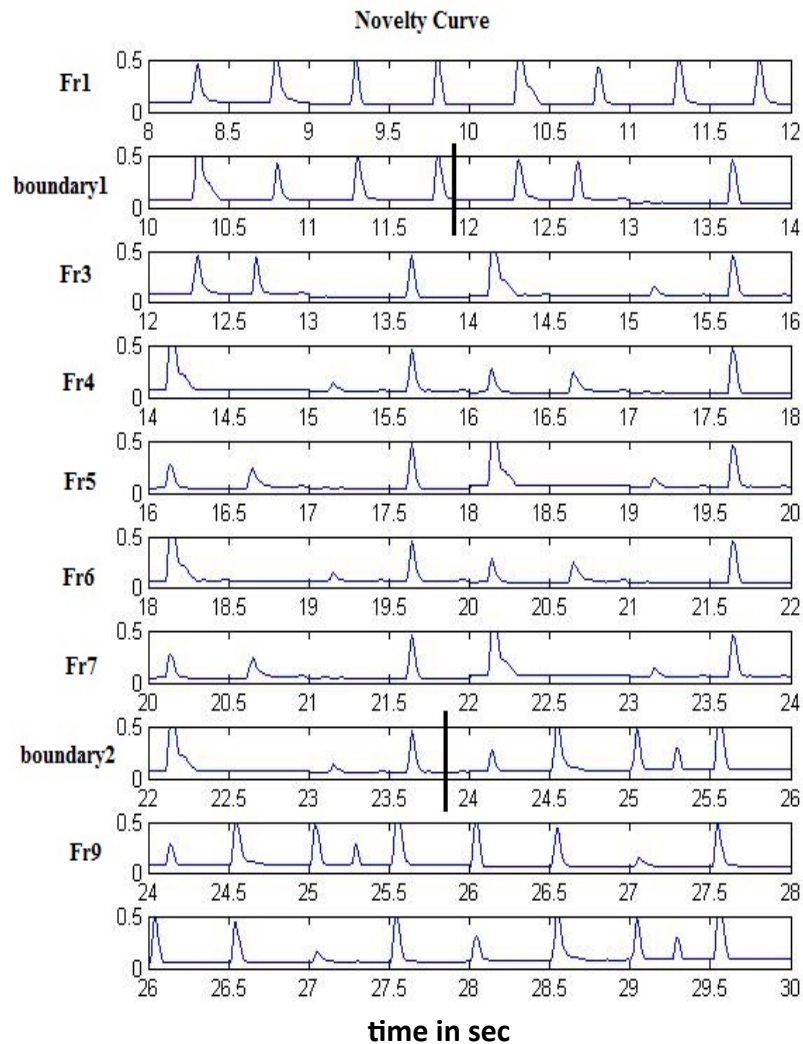


# Novelty function and ACF



Segments with deviation from basic structure of tal  
at Pr1 and Pr2 boundary & Pr2 and Pr3 boundary

4s of rhythmic window, hop 2s





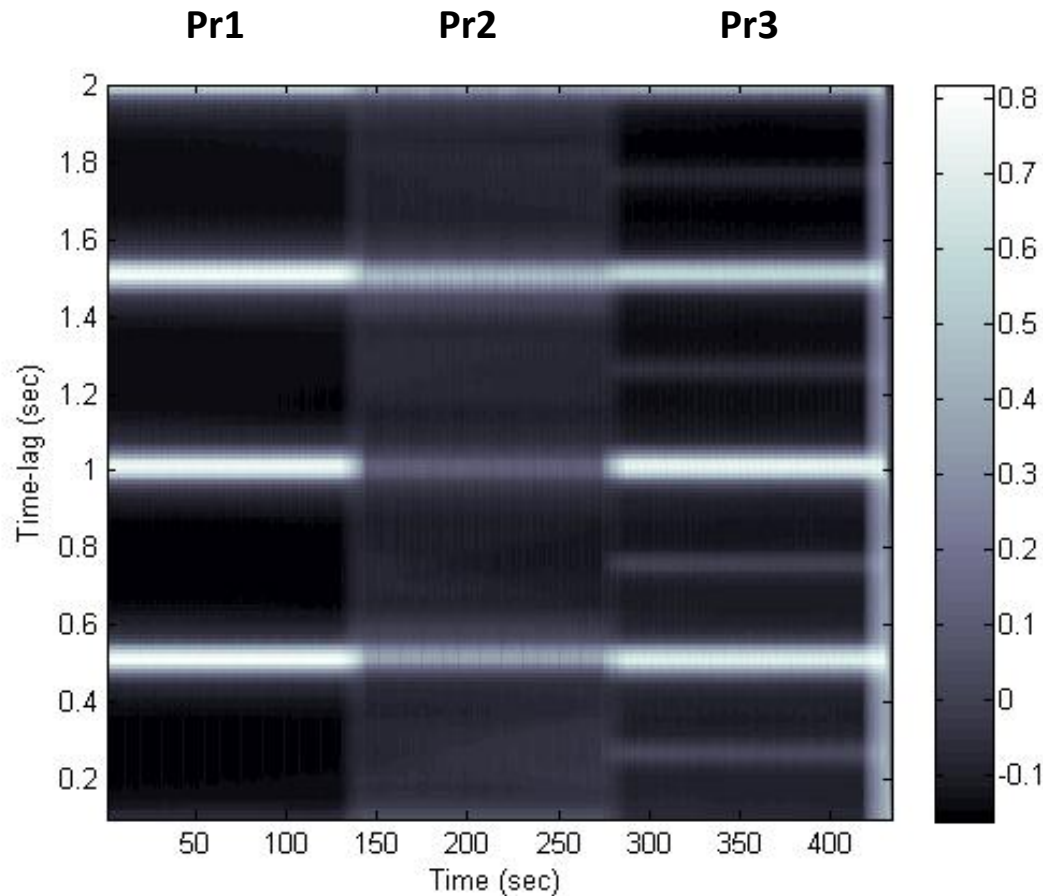
# Rhythmogram



## Prakar1 , Prakar2, Prakar3 concatenated

•A two dimensional time-pulse representation with lag-time on y- axis, time position on the x-axis and the autocorrelation values visualized as intensity, that displays the Progression of rhythm with time

16s of rhythmic window, hop 0.5s



Bright peaks are at the interval of 0.5s, representing the Inter Stroke Interval (ISI)

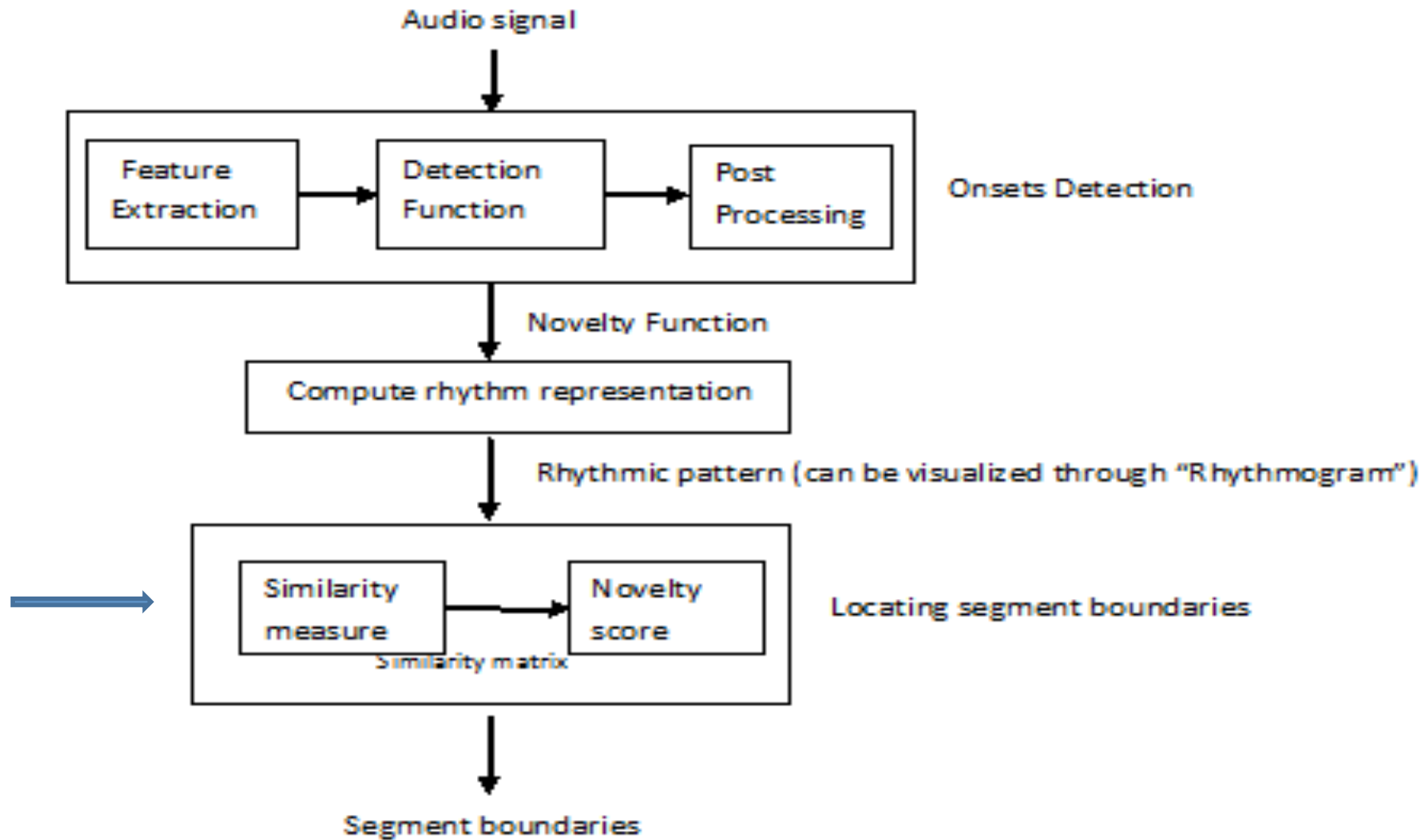
Pr1 , Pr2, Pr3 are of same tempo, hence the similarity in the interval between bright peaks

Due to pause bols, acf peaks of Pr2 are of reduced strength

Increased surface rhythm in prakar3 is due to the half beats at matra2 of the cycle. This has contributed to feeble peaks at the multiples of 0.25s in the rhythmogram



# Audio Segmentation system





# Locating boundaries

## Self similarity matrix



- Each rhythmic frame can be compared with all other frames in a pair wise fashion to compute the Self Distance matrix(SDM)
- $D(i,j)=d(x_i,x_j)$  for  $i,j \in \{1, 2,\dots,N\}$ , where the distance function  $d$  specifies the correlation distance between two feature vectors  $x_i$  and  $x_j$  represented as,

$$d_{cor} = 1 - \frac{(x_i - \bar{x}_i)(x_j - \bar{x}_j)}{\sqrt{(x_i - \bar{x}_i)(x_i - \bar{x}_i)} \sqrt{(x_j - \bar{x}_j)(x_j - \bar{x}_j)}}$$



# Self Distance matrix



at the boundary of pr1 - pr2 and pr2 - pr3

	Pr1	boundary1		Pr2		Boundary2		Pr3			
Frame	1	2	3	4	5	6	7	8	9	10	
Pr1	1	0.0000	0.0984	0.2622	0.1840	0.1461	0.1824	0.1469	0.2125	0.0373	0.0304
	2	0.0984	0.0000	0.1613	0.1414	0.1554	0.1407	0.1565	0.1022	0.0782	0.1005
Pr2	3	0.2622	0.1613	0.0000	0.1056	0.0834	0.1052	0.0834	0.1945	0.2857	0.2808
	4	0.1840	0.1414	0.1056	0.0000	0.0435	0.0000	0.0437	0.1191	0.1602	0.1549
	5	0.1461	0.1554	0.0834	0.0435	0.0000	0.0431	0.0000	0.1565	0.1707	0.1520
Pr3	6	0.1824	0.1407	0.1052	0.0000	0.0431	0.0000	0.0433	0.1190	0.1593	0.1542
	7	0.1469	0.1565	0.0834	0.0437	0.0000	0.0433	0.0000	0.1574	0.1719	0.1532
	8	0.2125	0.1022	0.1945	0.1191	0.1565	0.1190	0.1574	0.0000	0.1432	0.1746
	9	0.0373	0.0782	0.2857	0.1602	0.1707	0.1593	0.1719	0.1432	0.0000	0.0137
	10	0.0304	0.1005	0.2808	0.1549	0.1520	0.1542	0.1532	0.1746	0.0137	0.0000

•Boundary1 is appearing in fr2. Fr3 is having 2 pause bols and hence the distance between fr1 (which is in Pr1) and fr3 (Which is in Pr2) is more.

•Boundary2 is appearing in fr8.



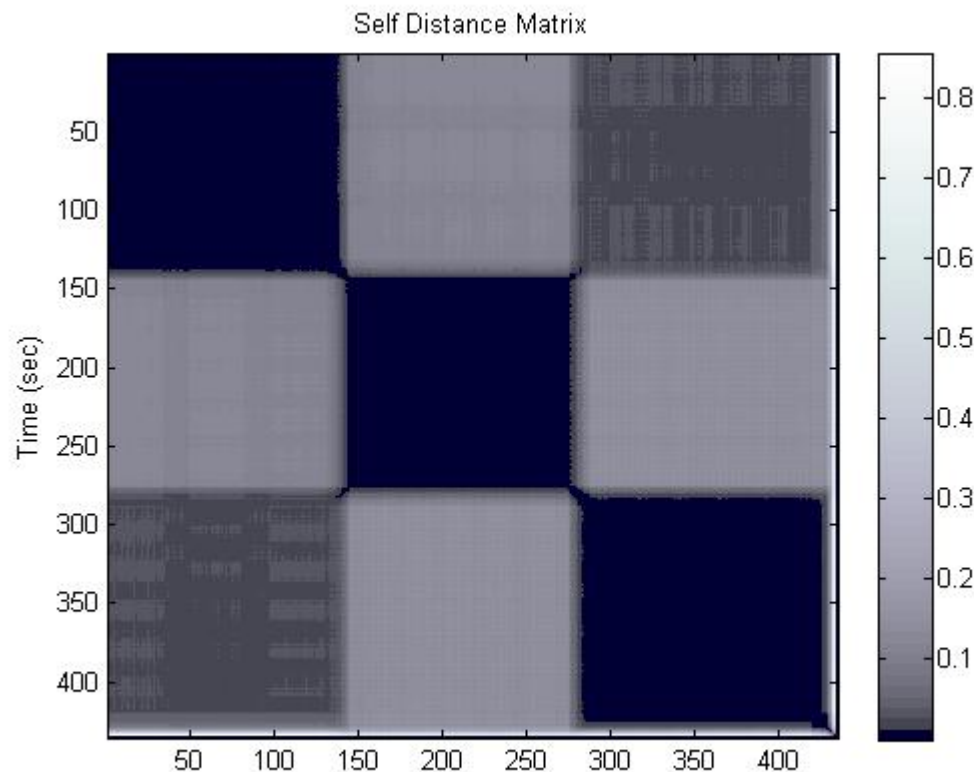
# Similarity Matrix



16s of rhythmic window, hop 0.5s

Audio of 36 cycles of Pr1, Pr2 and Pr3

Cycle length in each Prakar is about 4s



- Choice of window length : larger than the expected periodicity (suggested as 4 times the periodicity in [18])
  - to capture tactus (matra) periodicity
  - to capture measure (vibhag) periodicity
  - *to capture periodicity patterns in a tala cycle*

Similarity matrix



# Locating boundaries



## Novelty Score

- Structure of the Self Distance matrix,  $D$  is an indication of the novelty measure that will have high values at the segment boundaries
- ***Correlating the checker board kernel  $C$  along the diagonal of the matrix  $D$  will yield the measure of novelty, novelty score.***
- A simple 4×4 checker board pattern represented by the kernel matrix  $C_4$  with kernel length 4 as,

$$C_4 = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}$$

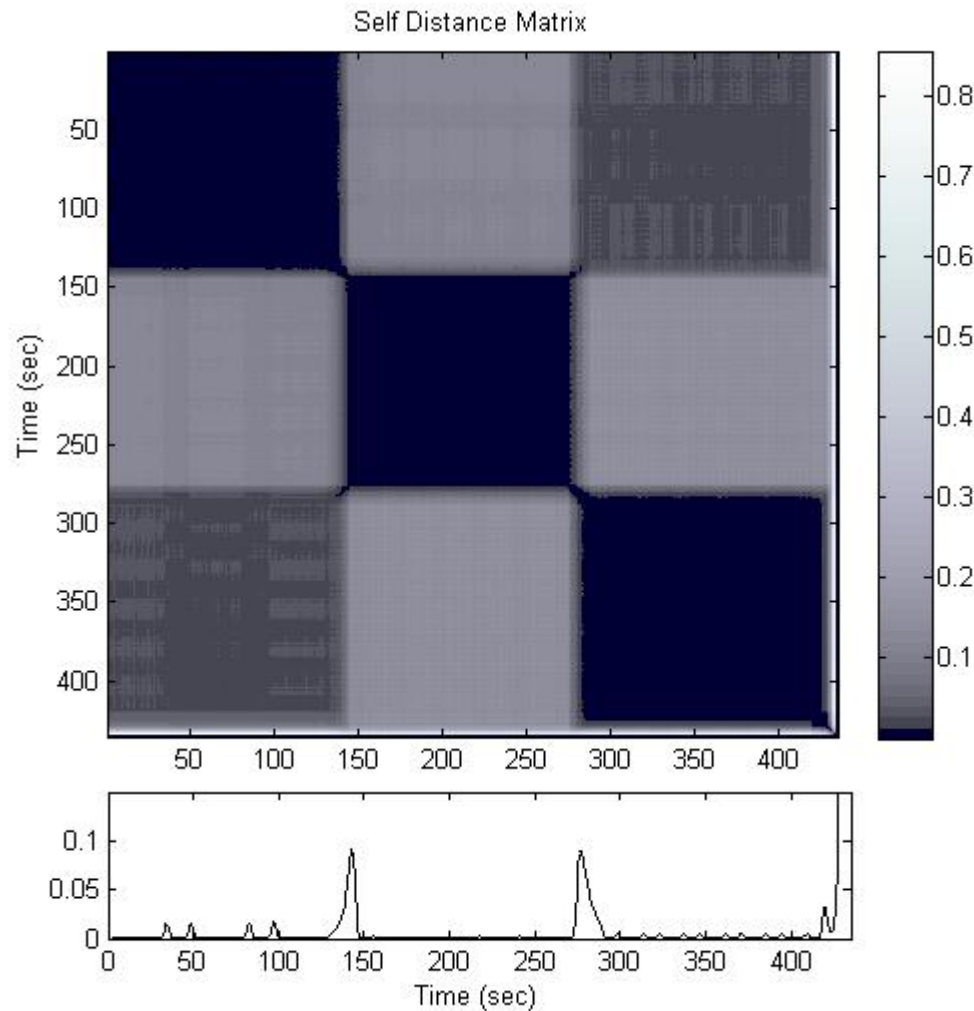
- Larger kernels are formed by kronecker product of  $C$  with the matrix of ones
- ***Extrema in the novelty score correspond to large changes in the rhythmic structure and thus indicate the boundaries***



# Locating the boundaries



Audio of 36 cycles of Pr1, Pr2 and Pr3



- Novelty Score indicating the boundaries
  - Got by correlating the Kernel of width 16 along the diagonal of the matrix
  - The width of the checker board kernel decides the scale of the novelty measure

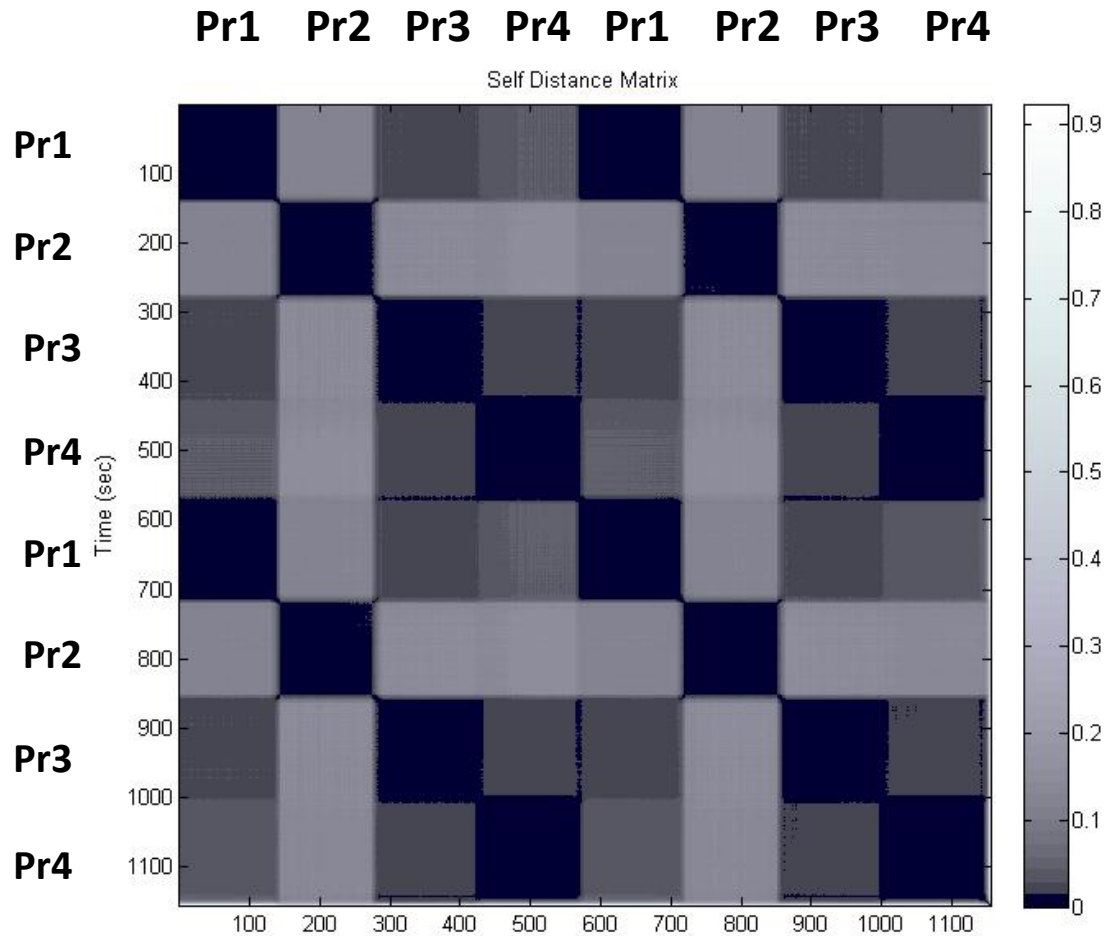
**Similarity matrix and Novelty score**



# Repetitive structure of Prakars



Audio of 36 cycles of Pr1, Pr2 and Pr3 repeated



Prakar compared with repeated Prakar is represented by dark region



# Tabla Solo performance



Tabla



Melodic  
accompaniment





# Segments in a tabla solo performance



- Segments usually appearing in a tabla solo performance are,
- Peshkar:
  - Introductory section like alap
- Kaida:
  - The overall structure of a kaida can be divided into three sections, an opening theme, a series of variations based on the opening theme and a concluding *tihai*
  - *In these sections the tempo will be maintained same even though rhythmic density will increase after the opening of kaida.*
- Gat:
  - very distinct compositional structure, which is pre-composed
  - *Tempo of the performance increases from this section*
  - ***From here, it is drut laya***
- Rela
  - is a fast and flowing compositional form
- Raun
  - is a fast compositional form characterized by droning effect.



# Acoustic descriptions of

## *Tabla solo by Zakir Hussain*



Segment (duration)			Duration of one cycle	Strokes per cycle	Remarks	
Peshkar (8min)	Slow		6.8min	26s to 24.6s		
	Fast		1.2min			
Kaidas	Kaida1 (4.1min)	S	2.7min	28.3s to 24.9s	128 strokes in slower section and doubled in faster section	
		F	1.4min			
	Kaida2 (2.15min)	S	0.83min	27.3s to 25.2s		
		F	1.72min			
	Kaida3 (2.23min)	S	1.37min	28.6s to 25.3s		
		F	0.86min			
	Kaida4 (3.6min)	S	2.2min	21.7s to 20.9s		96 strokes in slower section and increased in multiples of 1.5, 3 6 and 8 times in faster sections
		F	1.4min			
	Kaida5 (7.2min)		50.6s	16s to 15s		Tempo increased slightly
			6.3min			
Gat 1(2.2min)			4s		Tempo has increased much. Ati-drut section	
Rela (1.08min)			4s			
Gat2			3.7s			
Gat3			2.98s			



# Kaida : Theme & Variation



## ***Kaida Rules:***

- The main focus during a kaida is the thematic development that is achieved through a series of variations

## **Kaida1 theme**



- 1 2 3            Dha Ti Ta
- 1 2 3            Dha Ti Ta
- 1 2 3 4          Dha Dha Ti Ta
- 1 2 3 4 5 6      Dha Ge Ti Na Ge Na

## **Kaida1 variation**

- 1 2 3            Dha Ti Ta
- 1 2 3 4          Dha DhaTi Ta
- 1 2 3            Dha Ti Ta
- 1 2 3 4 5 6      Dha Ge Ti Na Ge Na

**One must use only the bols in the original theme of the kaida in variations**

## **Kaida2 theme**

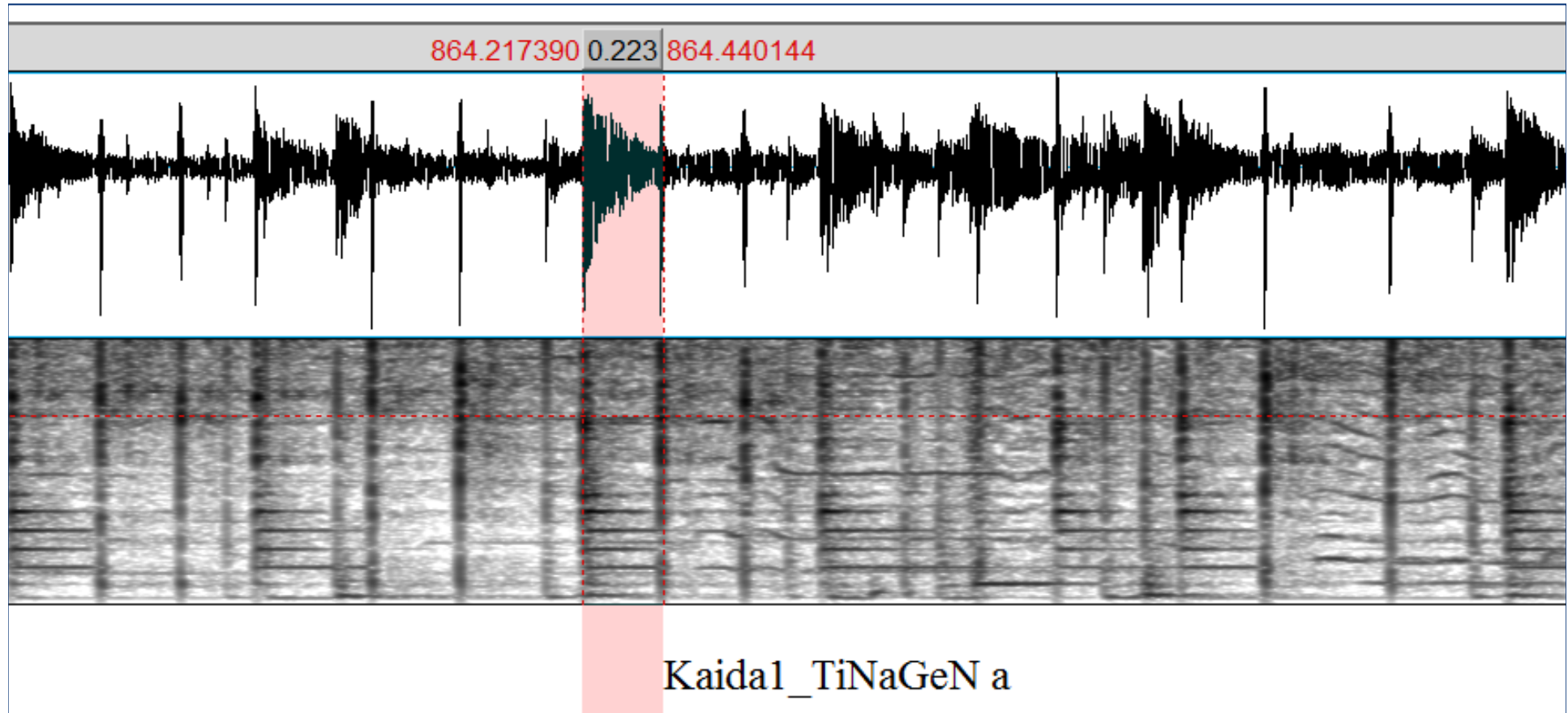


- 1 2            Dha Ti
- 1 2            Dha Ge
- 1 2            Na Dha
- 1 2            Tire kita
- 1 2            Dha Ti
- 1 2            Dha Ge
- 1 2            Ti Na
- 1 2            Ge Na

***Theme of Kaida2 constituted through bols and grouping of bols is different from that of kaida1***



# Spectrogram of Kaida1 slower section

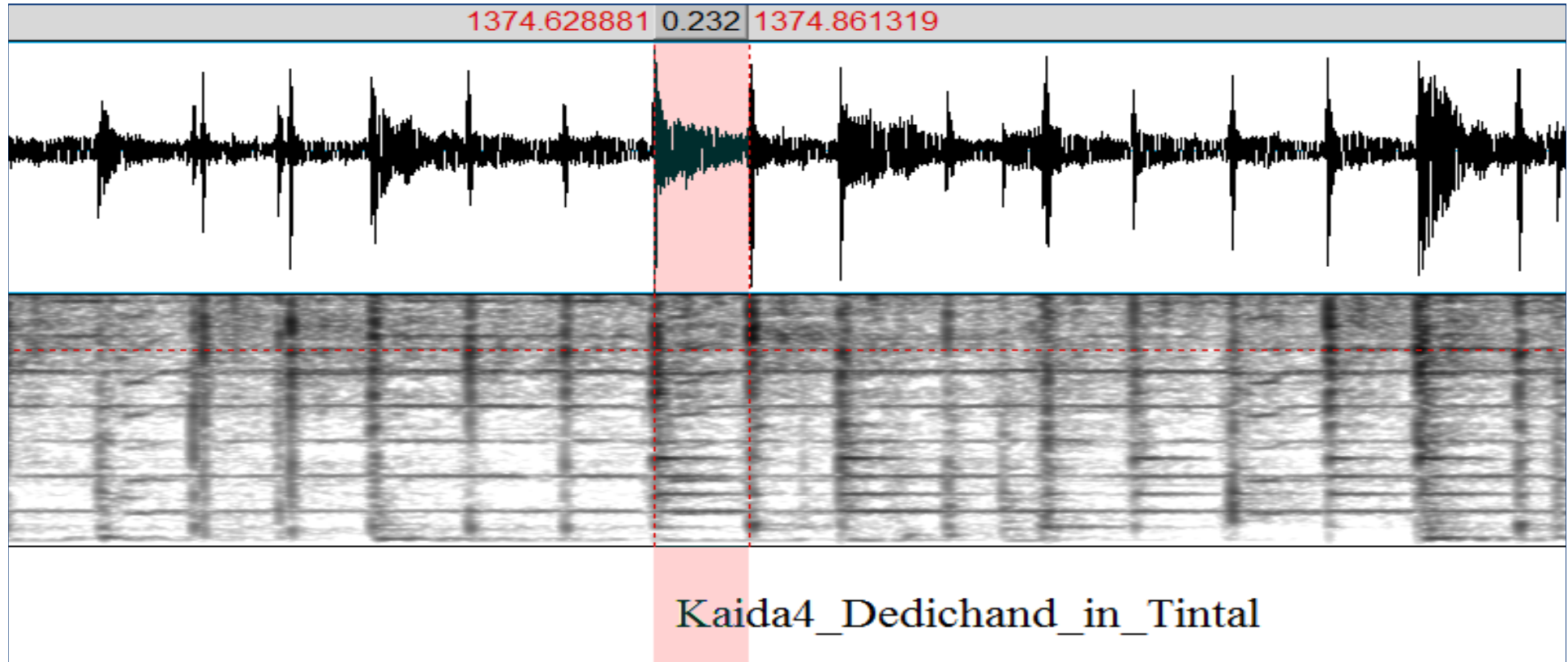


**In slower sections of kaida1, stroke density is 128 strokes per cycle with IOI of 0.22s**

**In faster sections of kaida1, kaida2 & kaida3, strokes density has doubled with IOI of 0.1s.**



# Spectrogram of Kaida4 slower section



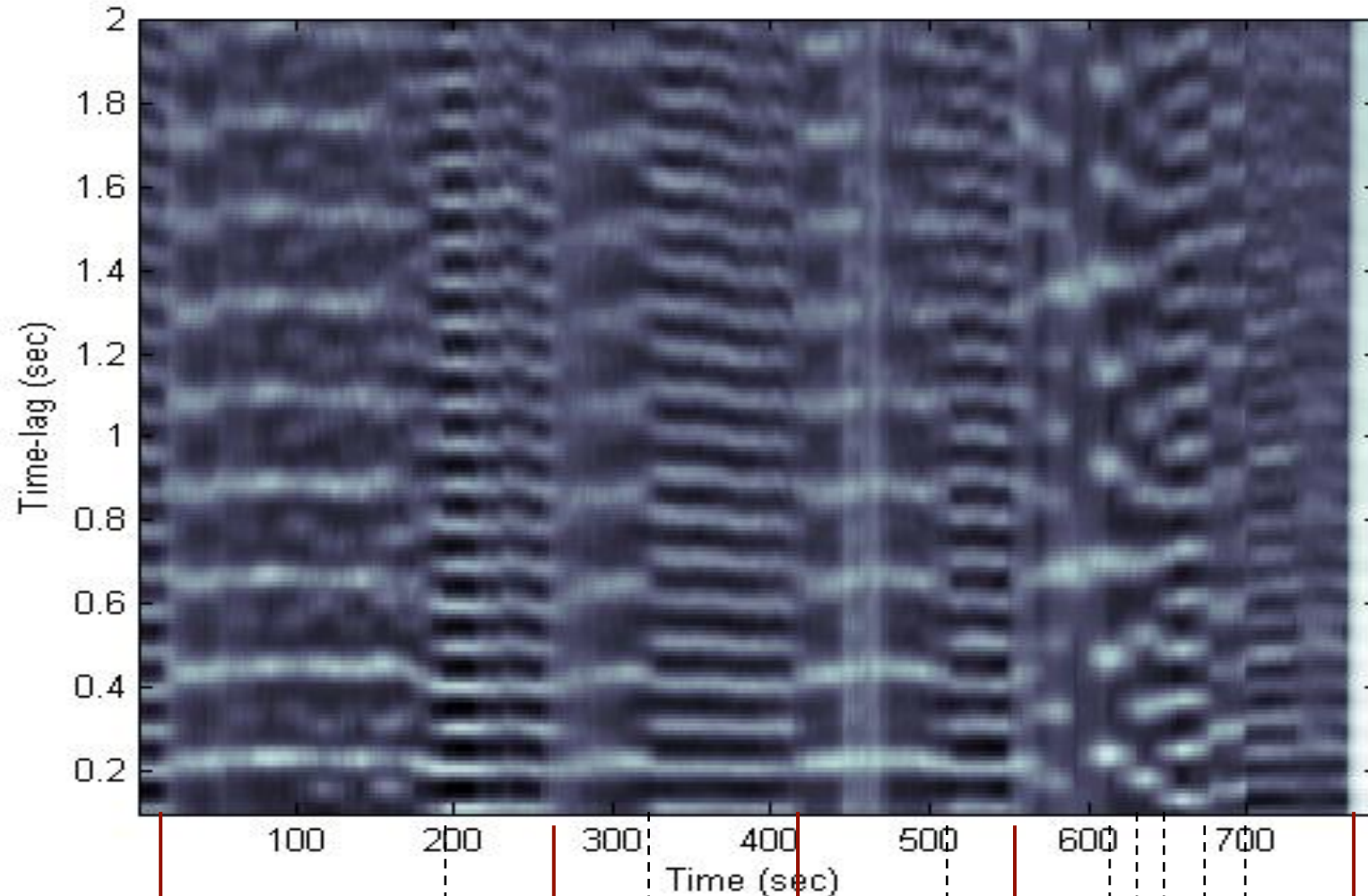
**Dedichand kaida- indicates 3 matras in 2 beats**

**Stroke density of 96 strokes/cycle in the slower section  
with the IOI of 0.23s**



# Rhythmogram

Tabla solo by ZH, Kolkatta (Kaidas section, 13 min)



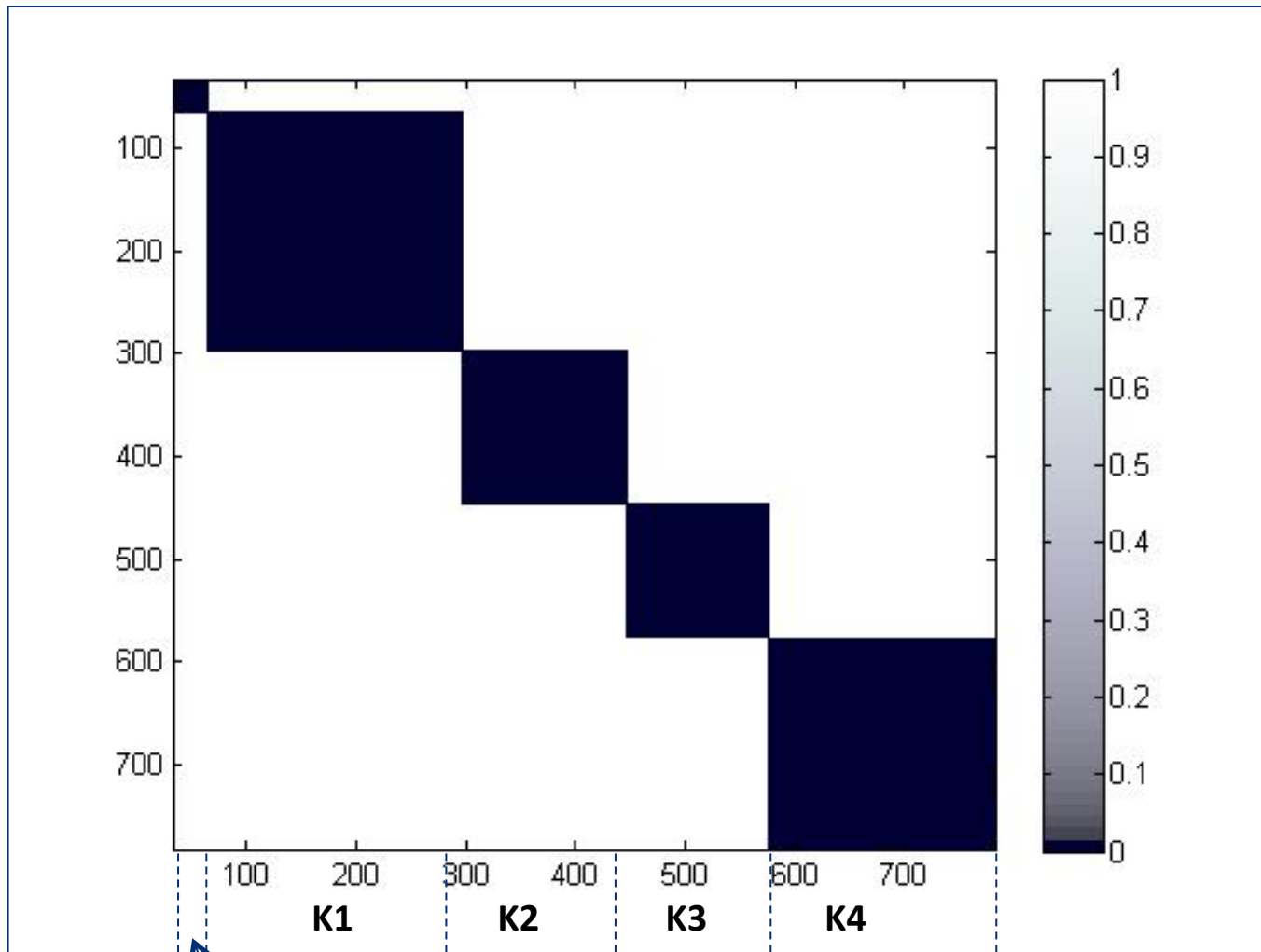
100 200 300 400 500 600 700  
Time (sec)  
**Kaida1** **Kaida2** **Kaida3** **Kaida4**

S	F	S	F	S	F	S	F	



# Ideal Similarity matrix

tabla solo by ZH,Kolkatta (Kaidas section, 13 min)



Peshkar

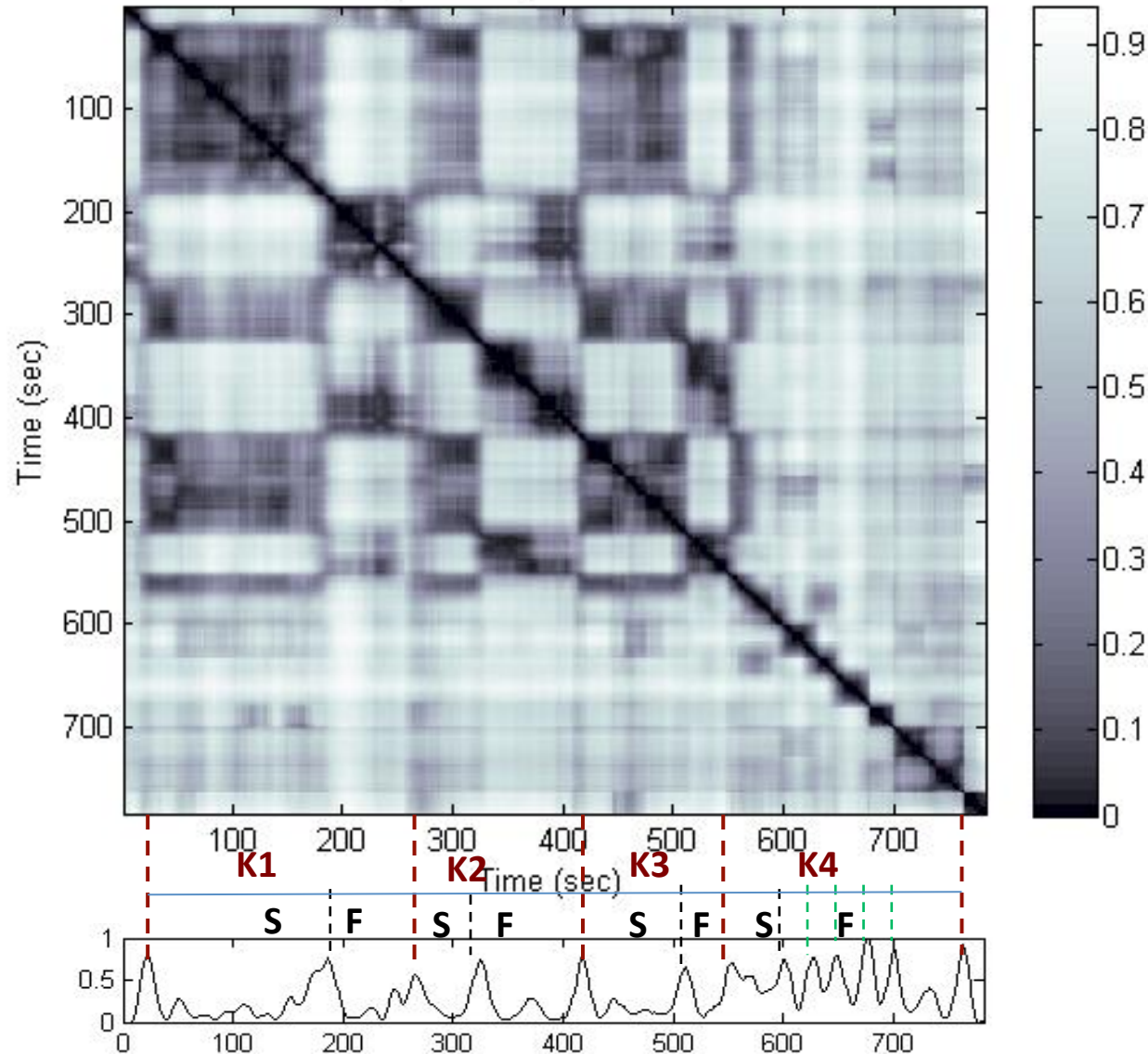


# Similarity Matrix & Novelty Score

Tabla solo by ZH,Kolkatta (Kaidas section, 13 min)



Novelty score by kernel of width 16



**Rhythmic window of 28s & hop of 0.5s is used**

**Dissimilarity of faster section of kaida with that of slower section, due to difference in surface rhythm (higher rhythmic density)**



# Segments in a Khyal Vocal Concert



- **Khyal vocal performance is a polyphonic audio with simultaneous presence of many instruments with voice**
- **Aalap:** *Unmetered melodic improvisation (no tabla).*
- **Bada Khyal :** Will be rendered either in vilambit or madhya laya.
  - **Sthayi:** rendered in lower octave
    - **Vistar**
      - Aakar Vistar
      - Bol- Vistar
  - **Antara:** Rendered in upper octave
  - **Bol-baat:** Importance is there for lyrics
  - **Sargam:** Swaras are used, matras in lyrics will increase. ***So, fillers appear in percussion***
  - **Tan:** ***Basic theka will be played by tabla.*** Dynamic in melodic aspect
    - **Aakar-tan:** Intensity will display great regularity
    - **Bol-tan:** Lyrics of composition will be used in improvisation.
  - **Miscellaneous**
- **Chota Khyal :** Drut laya or ati-drut laya.
  - *Sometimes this may be in different tal.*
  - *Will have subsections of sthayi, antara, vistar, tan and sargam*



# Database



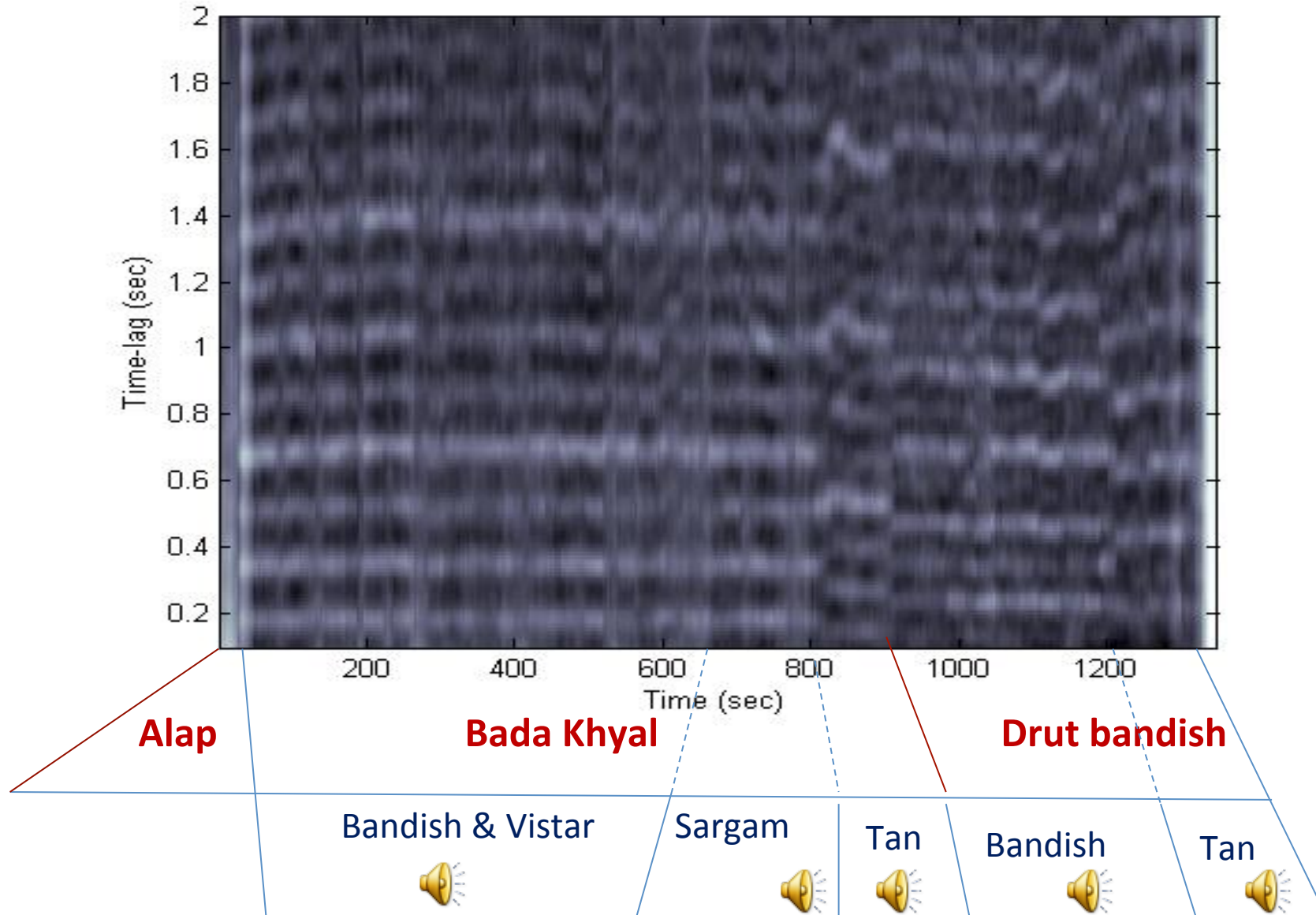
Raga	Vocalist	Bada khyal			Chota Khyal		
		Bandish	Tal	Laya (Dur of 1 cycle)	Bandish	Tal	Laya (Dur of 1 cycle)
Bhupali	Rashid Khan (RK)	Karo ge tum paar	<b>Tintal</b>	<b>Madhya (10.2s)</b>	Tu kari	<b>Tintal</b>	<b>Drut (3.67s)</b>
Deshkar	Kishori Amonkar (KA)	PiyaJaag	<b>Tintal</b>	<b>Vilambit (23s)</b>	HotoTore	<b>Tintal</b>	<b>Drut (6.8s)</b>
Miyan ki Todi	Prabha Atre (PA)	Mana Panchi	<b>Ektal</b>	<b>Vilambit (50.3s)</b>	Ja..re Parigama	<b>Ektal</b>	<b>Drut (3.01s)</b>
Miyan ki Todi	Kaivalya Kumar (KK)	Sa..i..	<b>Ektal</b>	<b>Vilambit (67s)</b>	Ab more	<b>Ektal</b>	<b>Drut (3.98s)</b>
Marwa	Rashid Khan (RK)	Piya more	<b>Ektal</b>	<b>Vilambit (72.4s)</b>	Kavu ki rit	<b>Tintal</b>	<b>Drut (4.58s)</b>



# Rhythmogram of Bhoopali, RK



Bada Khyal in Madhya laya, Tintal  
Chota khyal in Drut laya, Tintal



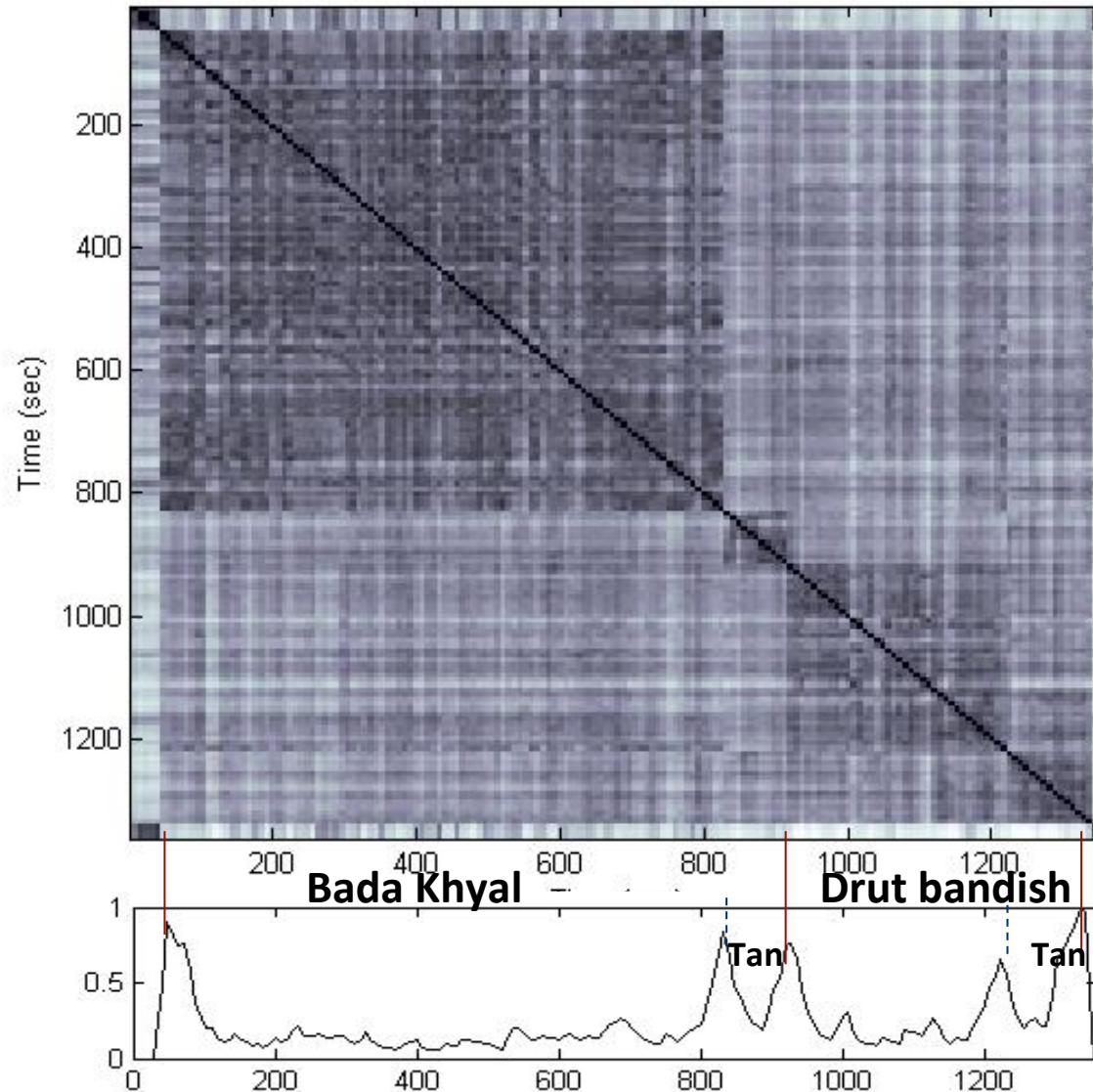


# Similarity Matrix of Bhoopali, RK



Bada Khyal in Madhya laya, Tintal  
Chota khyal in Drut laya, Tintal

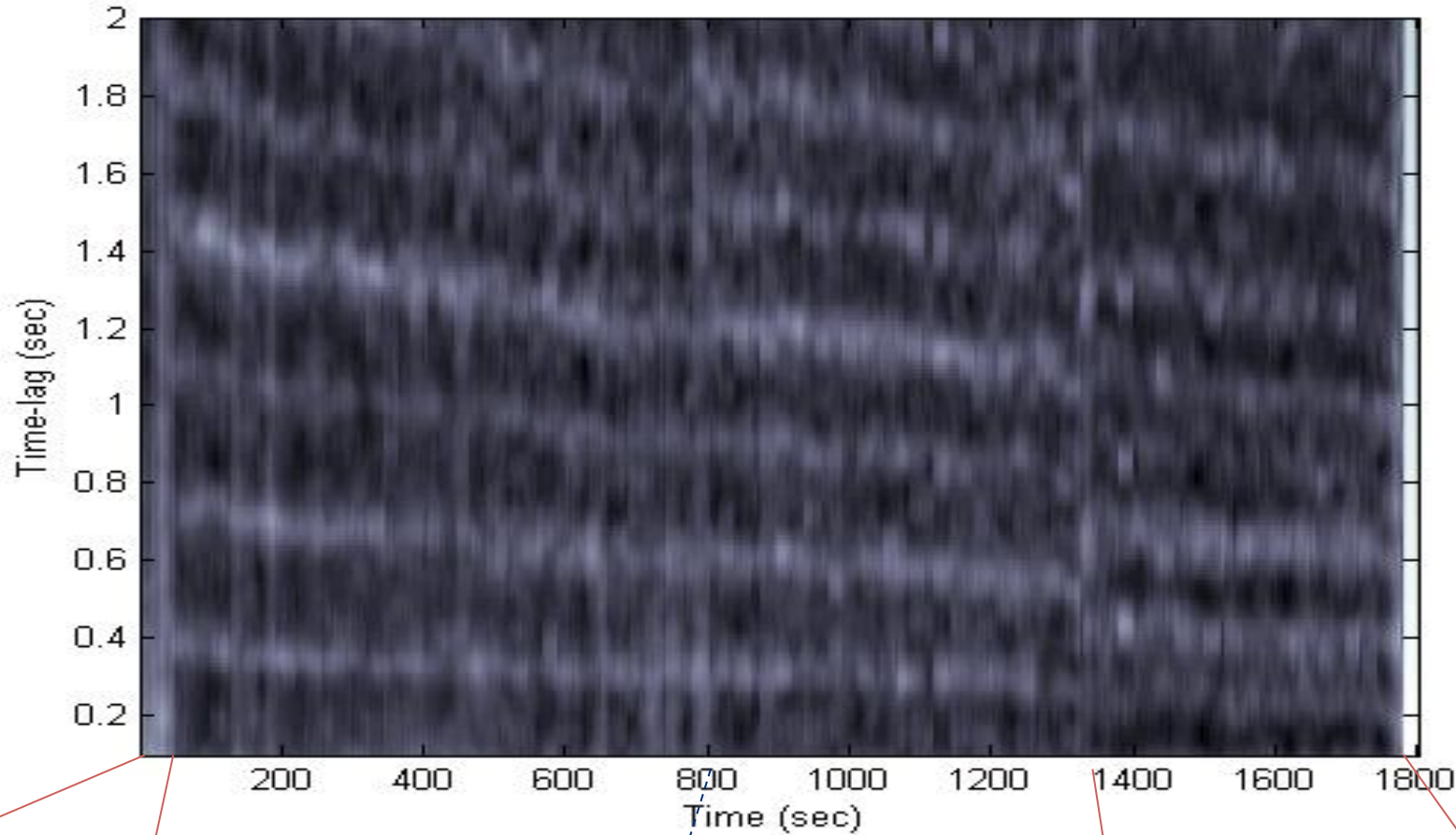
Novelty score by kernel of width 16





# Rhythmogram of Raag Deshkar by Kishori Amonkar

Bada Khyal in Vilambit laya, Tintal & Chota khyal in Drut laya, Tintal



**Alaap**

**Bada Khyal**

**Drut bandish**

**Bandish & Aakar  
Vistar**

**Bol-Vistaar**





# Conclusions



- Spectral features combined with the auditory processing motivated bi-phasic function achieved good time localization of extracted onsets
- ACF has proved to be tough enough to reveal the inherent periodicities and strengths of accents of complicate polyphonic music
- Audio segmentation algorithm has returned the boundaries between major sections of alap, bada khyal and chota khyal sections in vocal concerts and also the boundaries of sections like tan, sargam even within drut segment of the performance
- Boundaries of the segments in the intricate rhythmic repertoire like tabla solo has also been identified



# Future work

- In similarity matrix, Kaida1 slower section is showing more dis-similarity with faster section of kaida1, compared to Kaida2 slower section. ***Tempo due to surface rhythm has become a prominent criteria in this analysis.***
- Novelty function has to be refined to capture the stroke level changes ie., when a bol of basic theka is replaced by other bol. For this, timbral dis-similarity of strokes has to be captured.
- *Along with the percussive onsets, non-percussive onsets have to be extracted in the novelty detection stage so as to capture the complete rhythmic structure*
- *Efficiency of autocorrelation approach of rhythmic pattern representation has to be compared with other prevalent approaches like combination of DFT and frequency mapped autocorrelation method, that relies on the dominant metrical level of the segment*
- *Other aspects of music like timbre and melody have to be combined with rhythm to extract the boundaries of segments within bada khyal other than tan and sargam*



# REFERENCES

- [1] M. Cooper and J. Foote, “Summarizing popular music via structural similarity analysis,” in Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA '03), pp. 127–130, New Paltz, NY, USA, October 2003.
- [2] J. Foote, “Automatic audio segmentation using a measure of audio novelty,” in Proceedings of IEEE International Conference on Multimedia and Expo (ICME '00), vol. 1, pp. 452–455, New York, NY, USA, July-August 2000.
- [3] K. Jensen, J. Xu, and M. Zachariasen, “Rhythm-based segmentation of popular chinese music,” in Proc. Int. Conf. Music Inf. Retrieval (ISMIR), London, U.K., 2005.
- [4] M. Miron, ‘ Automatic Detection of Hindustani Talas’, Master’s thesis, Universitat Pompeu Fabra, Barcelona, Spain, 2011.
- [5] M. Clayton, Time in Indian Music : “Rhythm , Metre and Form in North Indian Rag Performance”, Oxford University Press, 2000.
- [6] K. Jensen, ‘Multiple Scale Music Segmentation Using Rhythm, Timbre, and Harmony’, EURASIP Journal on Advances in Signal Processing, Aug 2006.
- [7] J. P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M. B. Sandler: “A Tutorial on Onset Detection in Music Signals”, IEEE transactions on Speech and Audio, vol. 13, no. 5, pp. 1035–1047, 2005.
- [8] S. Dixon : “ Onset Detection Revisited”, In Proceedings of the International Conference on Digital Audio Effects (DAFx'06), Montreal, Canada. 2006
- [9] P. Grosche and M. Müller, “Extracting predominant local pulse information from music recordings”, IEEE Transactions on Audio, Speech, and Language Processing, 19(6): 1688–1701, 2011.



# REFERENCES Contd.

- [10] S. Gulati and P. Rao, "Rhythm pattern representation for tempo detection in music", Proc. of the First International Conference on Intelligent Interactive Technologies and Multimedia, Dec, 2010, Allahabad, India.
- [11] V.Rao, P.Rao, "Vocal melody detection in the presence of pitched accompaniment using harmonic matching methods", Proc. of the 11th Int. Conference on Digital Audio Effects (DAFx-08), Espoo, Finland, September 1-4, 2008
- [12] A.Bapat and P.Rao, "Pitch tracking of voice in tabla background by the two-way mismatch method", Proc. of the 13th Int. Conf. on Advanced Computing and Communications, 2005, Coimbatore, India.
- [13] J. Cheri Ross, Vinutha T.P. and P.Rao, "Detecting melodic motifs from audio for Hindustani classical music", Proc. of ISMIR 2012.
- [14] P. Pradeep Kumar., P. Rao, and S. D. Roy, "Note onset detection in natural humming", Conference on Computational Intelligence and Multimedia Applications, 2007, International Conference on. Vol. 4. IEEE, 2007.
- [15] D. J. Hermes, "Vowel onset detection," *J. Acoust. Soc. of America*, 87(2), pp.866-873, 1990.
- [16] C. Ros~ao, R. Ribeiro, and D. Martins de Matos "Influence of Peak selection methods on Onset detection", 13th International Society for Music Information Retrieval Conference (ISMIR 2012)
- [17] T. Pr~atzlich, "A Cross-Version Approach for Novelty Detection in Music Recordings", Dec 2011, Mast.Thesis, Saraland Univ.
- [18] G.Peeters, "Template based Estimation of Time-varying Tempo"



**Thank You!!**