A Data-driven Bayesian Approach to Automatic Rhythm Analysis of Indian Art Music

17 Nov 2016, PhD defense

Ajay Srinivasamurthy

Dept. of Information and Communication Technologies

Universitat Pompeu Fabra

Thesis Supervisor

Dr. Xavier Serra

Music Technology Group

Universitat Pompeu Fabra

Thesis Committee:

Dr. Simon Dixon (QMUL)

Dr. Geoffroy Peeters (IRCAM)

Dr. Juan Pablo Bello (NYU)

compmusic

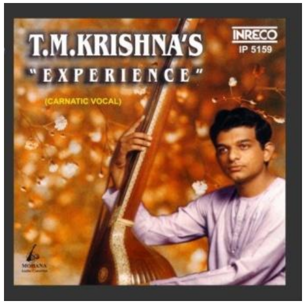
upf. Music Technology Group

Acknowledgements

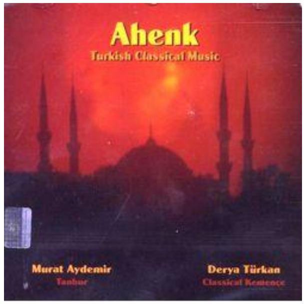
Context – CompMusic

- Information technologies in a multicultural world
- Culture aware Music Information Research (MIR)
- Focus on rhythm

Hindustani



Carnatic



Turkish-makam

Arab-Andalusian

Beijing Opera

http://compmusic.upf.edu/

Motivation – Automatic rhythm analysis

compmusic

- Rhythm is a fundamental dimension of music
- Rhythmic structures and patterns
  - **Rhythm similarity**
- Applications
  - Enriched listening
  - Meaningful interaction with large collections of music

Scope and Objectives – Long title

Culture-aware and data-driven

signal processing and machine learning approaches for

automatic analysis, description and discovery of

rhythmic structures and patterns in

audio music collections of Indian art music

… automatic analysis, description and discovery …

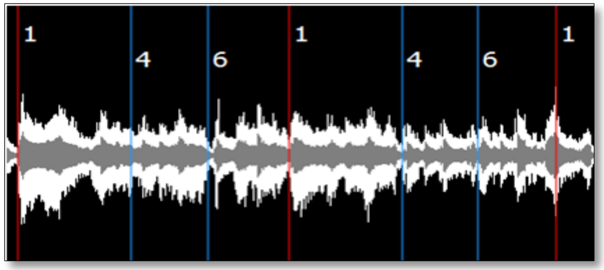
compmusic

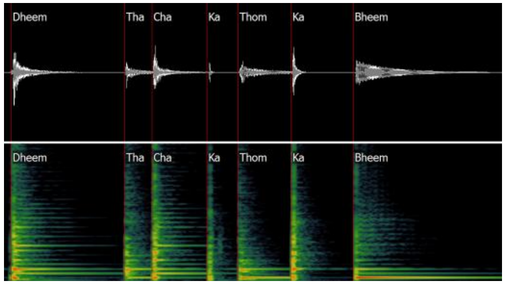
- Automatic rhythm analysis and description
- No focus on synthesis or composition

Culture-aware and data-driven signal processing and machine learning approaches for automatic analysis, description and discovery of rhythmic structures and patterns in audio music collections of Indian art music

. . . of rhythmic structures and patterns . . .

- Analysis of **metrical structures**
- Discovery of rhythm and **percussion patterns**





Culture-aware and data-driven signal processing and machine learning approaches for automatic analysis, description and discovery of rhythmic structures and patterns in audio music collections of Indian art music

Culture-aware and data-driven . . .

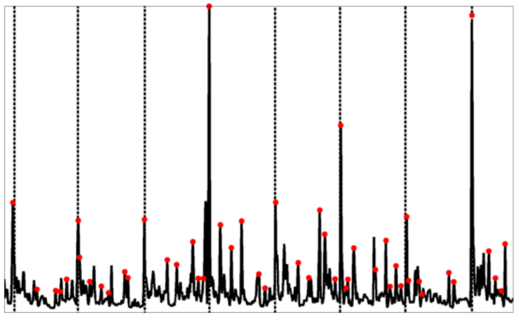
- Culture-aware
  - Work within CompMusic and its community
  - Bring music knowledge into methods
  - Musical concepts to engineering formulations
- Data-driven engineering approaches
  - Emphasis on **data and methods**

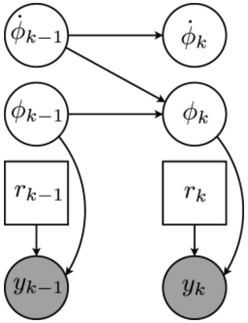
Culture-aware and data-driven signal processing and machine learning approaches for automatic analysis, description and discovery of rhythmic structures and patterns in audio music collections of Indian art music

. . . signal processing and machine learning approaches . . .

compmusic

- Informed signal processing
- **Bayesian machine learning methods**
  - Generative models – explicitly include music knowledge
- Novel methodologies for analysis and discovery

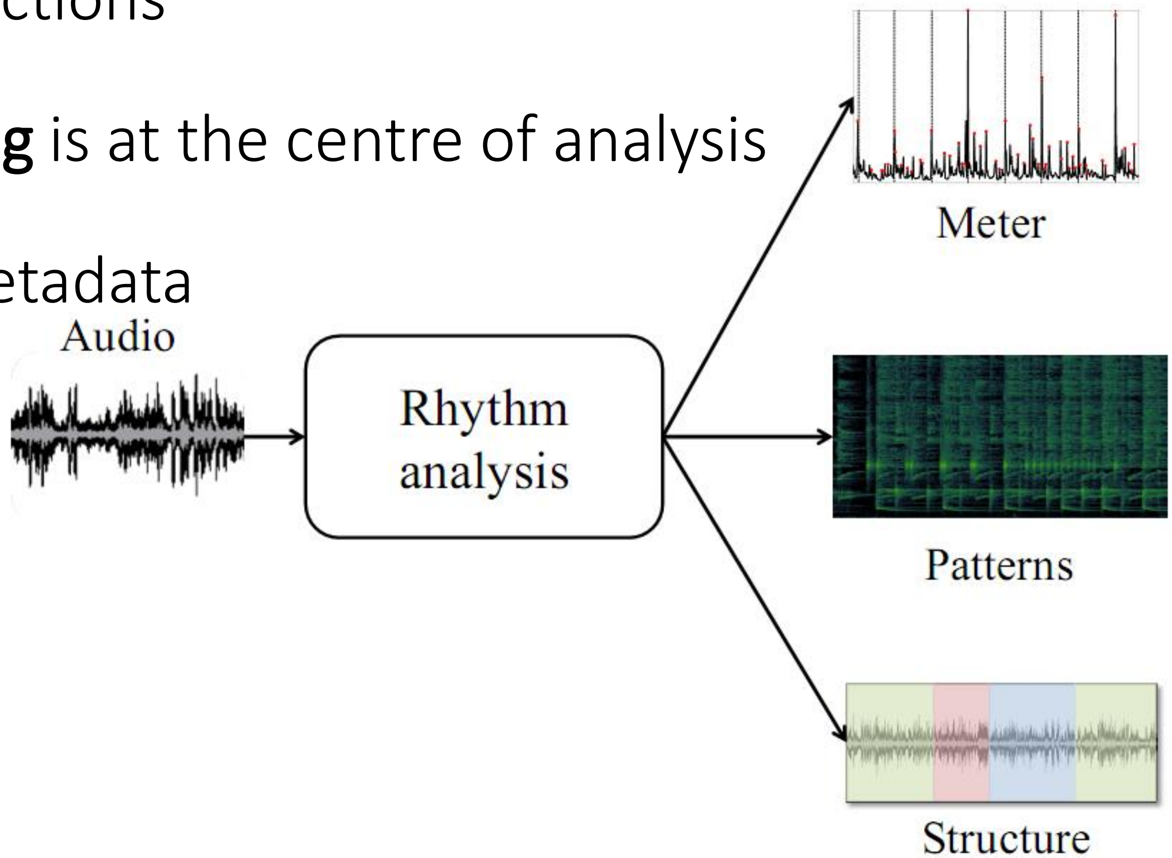




Culture-aware and data-driven **signal processing and machine learning approaches** for automatic analysis, description and discovery of rhythmic structures and patterns in audio music collections of Indian art music

. . . audio music collections . . .

- Audio music collections
- An **audio recording** is at the centre of analysis
- Use additional metadata



Culture-aware and data-driven signal processing and machine learning approaches for automatic analysis, description and discovery of rhythmic structures and patterns in audio music collections of Indian art music

. . . Indian art music

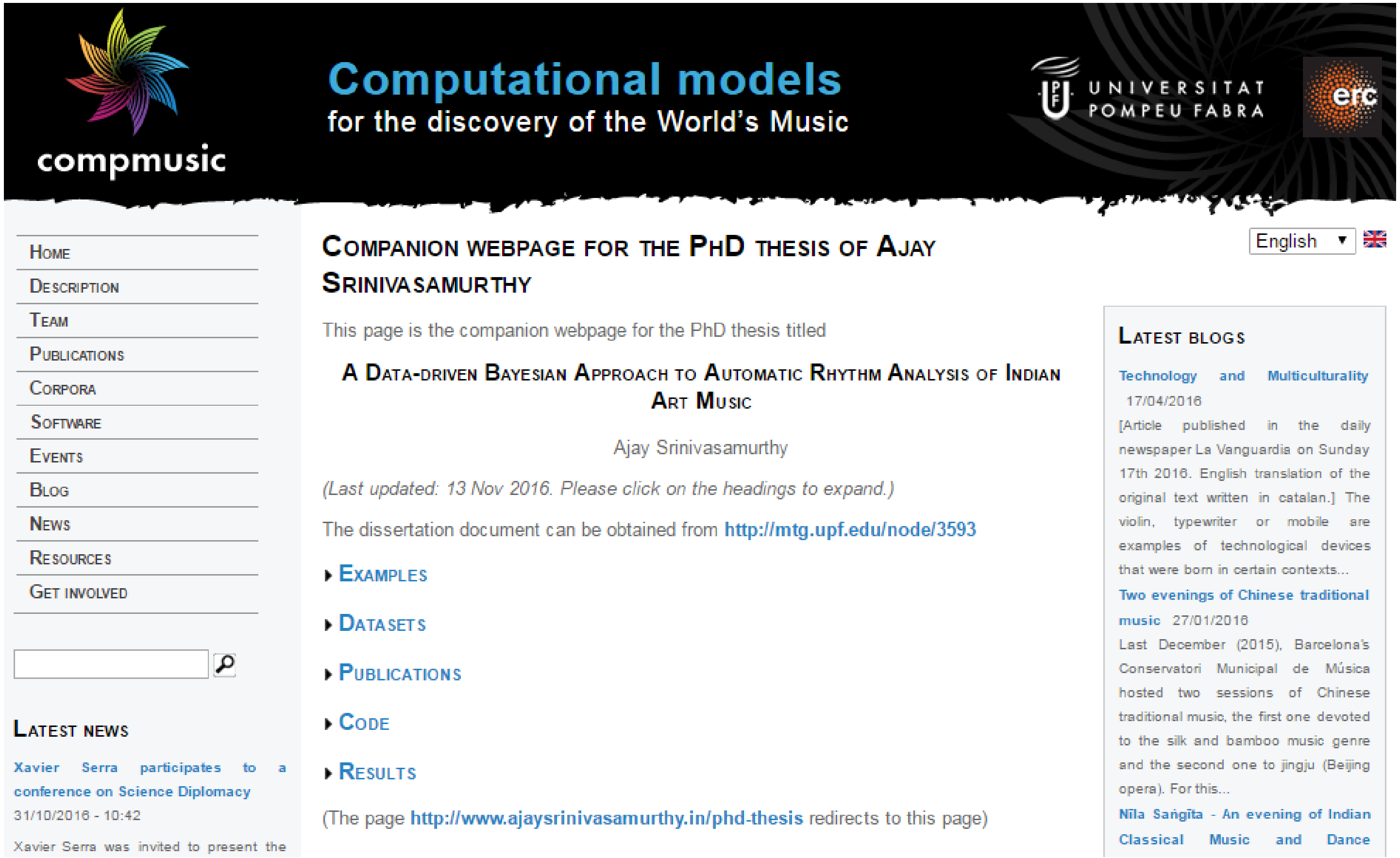
compmusic

- The focus is on Indian art music
  - Well-studied with significant musicological literature
  - A large audience
  - Unique challenges
- Identifying challenges and opportunities
- Extensions to other music cultures

Culture-aware and data-driven signal processing and machine learning approaches for automatic analysis, description and discovery of rhythmic structures and patterns in audio music collections of Indian art music

Additionally . . .

- Emphasis on reproducible and open research
- Code – Open source, Data – Easily accessible



Primary research questions

compmusic

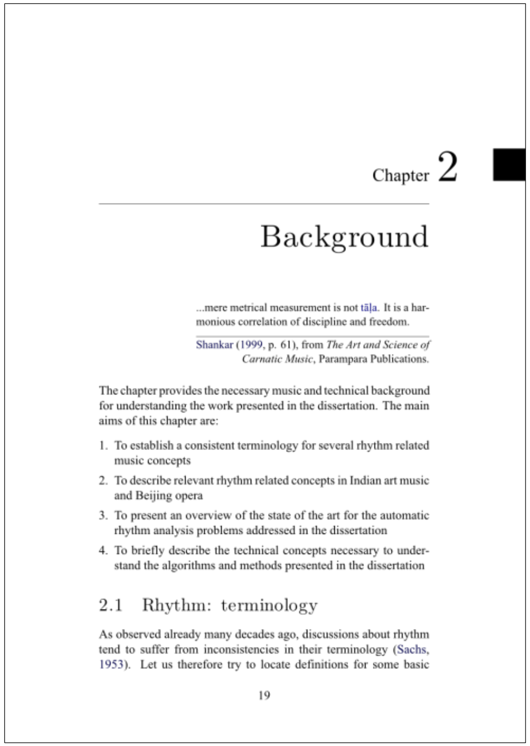
- Data corpora for MIR
  - Musicological insights on music performance
- Culture-aware specialized approaches
  - Utility of the explicit use of music knowledge

Primary research questions

- Data corpora for MIR
  - Musicological insights on music performance
- Culture-aware specialized approaches
  - Utility of the explicit use of music knowledge

Topics

- Background
- Main contributions
  - Data corpora for rhythm analysis
  - Meter inference and tracking
  - Percussion pattern discovery
- Summary



Chapter 2

Background

...mere metrical measurement is not tāḷa. It is a harmonious correlation of discipline and freedom.

Shankar (1999, p. 61), from *The Art and Science of Carnatic Music*, Parampara Publications.

The chapter provides the necessary music and technical background for understanding the work presented in the dissertation. The main aims of this chapter are:

1. To establish a consistent terminology for several rhythm related music concepts
2. To describe relevant rhythm related concepts in Indian art music and Beijing opera
3. To present an overview of the state of the art for the automatic rhythm analysis problems addressed in the dissertation
4. To briefly describe the technical concepts necessary to understand the algorithms and methods presented in the dissertation

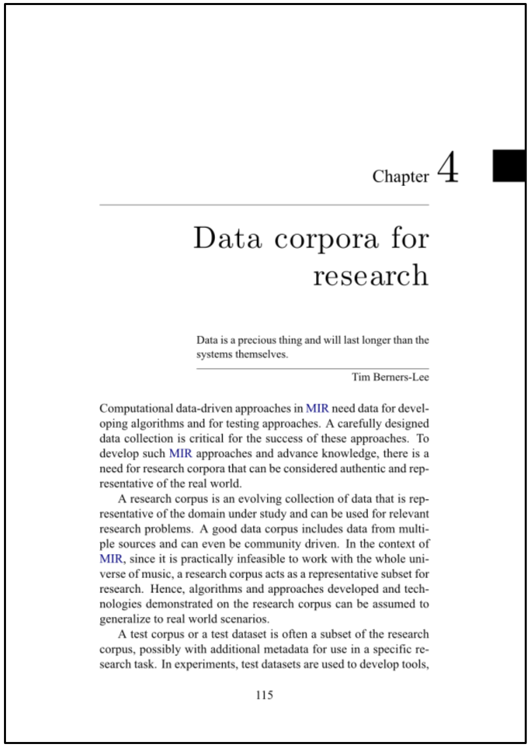
2.1 Rhythm: terminology

As observed already many decades ago, discussions about rhythm tend to suffer from inconsistencies in their terminology (Sachs, 1953). Let us therefore try to locate definitions for some basic

19

Topics

- Background
- Main contributions
  - Data corpora for rhythm analysis
  - Meter inference and tracking
  - Percussion pattern discovery
- Summary



Chapter 4

Data corpora for research

Data is a precious thing and will last longer than the systems themselves.

Tim Berners-Lee

Computational data-driven approaches in MIR need data for developing algorithms and for testing approaches. A carefully designed data collection is critical for the success of these approaches. To develop such MIR approaches and advance knowledge, there is a need for research corpora that can be considered authentic and representative of the real world.

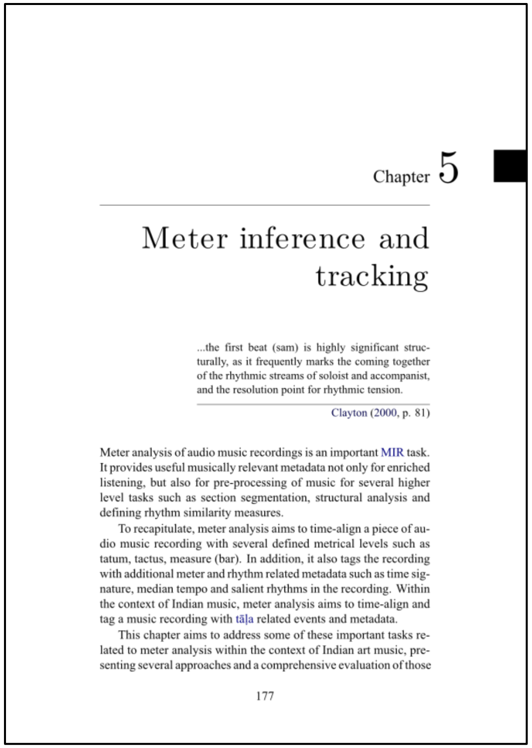
A research corpus is an evolving collection of data that is representative of the domain under study and can be used for relevant research problems. A good data corpus includes data from multiple sources and can even be community driven. In the context of MIR, since it is practically infeasible to work with the whole universe of music, a research corpus acts as a representative subset for research. Hence, algorithms and approaches developed and technologies demonstrated on the research corpus can be assumed to generalize to real world scenarios.

A test corpus or a test dataset is often a subset of the research corpus, possibly with additional metadata for use in a specific research task. In experiments, test datasets are used to develop tools,

115

Topics

- Background
- Main contributions
  - Data corpora for rhythm analysis
  - Meter inference and tracking
  - Percussion pattern discovery
- Summary



Chapter 5

Meter inference and tracking

...the first beat (sam) is highly significant structurally, as it frequently marks the coming together of the rhythmic streams of soloist and accompanist, and the resolution point for rhythmic tension.

Clayton (2000, p. 81)

Meter analysis of audio music recordings is an important MIR task. It provides useful musically relevant metadata not only for enriched listening, but also for pre-processing of music for several higher level tasks such as section segmentation, structural analysis and defining rhythm similarity measures.

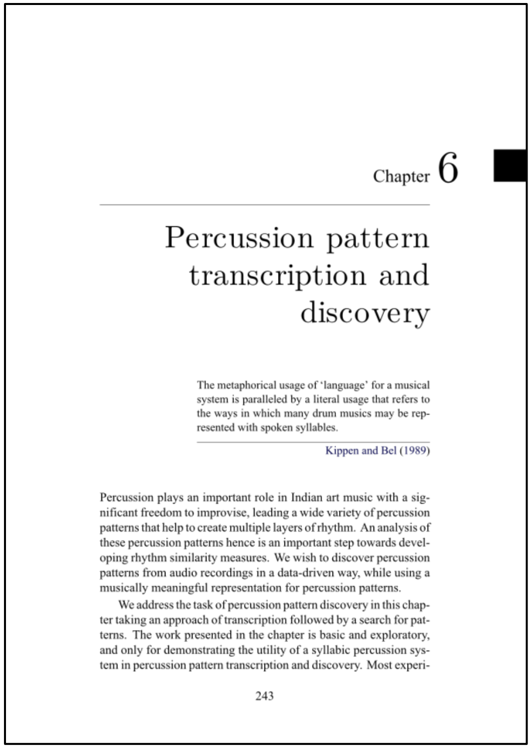
To recapitulate, meter analysis aims to time-align a piece of audio music recording with several defined metrical levels such as tatum, tactus, measure (bar). In addition, it also tags the recording with additional meter and rhythm related metadata such as time signature, median tempo and salient rhythms in the recording. Within the context of Indian music, meter analysis aims to time-align and tag a music recording with tāḷa related events and metadata.

This chapter aims to address some of these important tasks related to meter analysis within the context of Indian art music, presenting several approaches and a comprehensive evaluation of those

177

Topics

- Background
- Main contributions
  - Data corpora for rhythm analysis
  - Meter inference and tracking
  - Percussion pattern discovery
- Summary



Chapter 6

Percussion pattern transcription and discovery

The metaphorical usage of 'language' for a musical system is paralleled by a literal usage that refers to the ways in which many drum musics may be represented with spoken syllables.

Kippen and Bel (1989)

Percussion plays an important role in Indian art music with a significant freedom to improvise, leading a wide variety of percussion patterns that help to create multiple layers of rhythm. An analysis of these percussion patterns hence is an important step towards developing rhythm similarity measures. We wish to discover percussion patterns from audio recordings in a data-driven way, while using a musically meaningful representation for percussion patterns.

We address the task of percussion pattern discovery in this chapter taking an approach of transcription followed by a search for patterns. The work presented in the chapter is basic and exploratory, and only for demonstrating the utility of a syllabic percussion system in percussion pattern transcription and discovery. Most experi-

243

Topics

- Background
- Main contributions
  - Data corpora for rhythm analysis
  - Meter inference and tracking
  - Percussion pattern discovery
- Summary

Chapter 7

Applications, Summary and Conclusions

The outcome of any serious research can only be to make two questions grow where only one grew before.

Thorstein Veblen (1908)

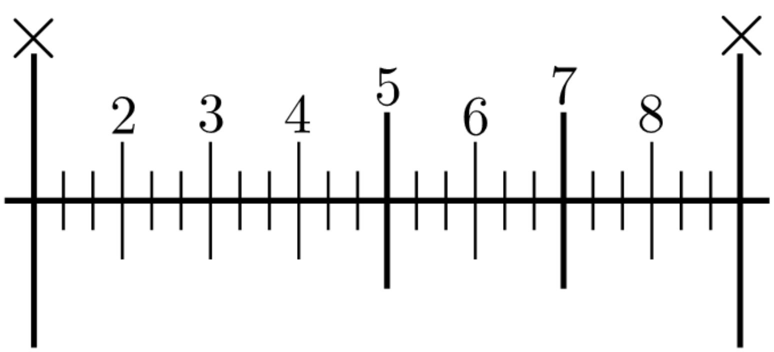
The concluding chapter of the dissertation aims to present some concrete applications of the rhythm analysis approaches and results presented in the previous chapters. It is followed by a summary of the work presented in the dissertation, along with some key results and conclusions. The thesis opens up a host of open problems - pointers and directions for future work based on the thesis form the last part of the chapter.

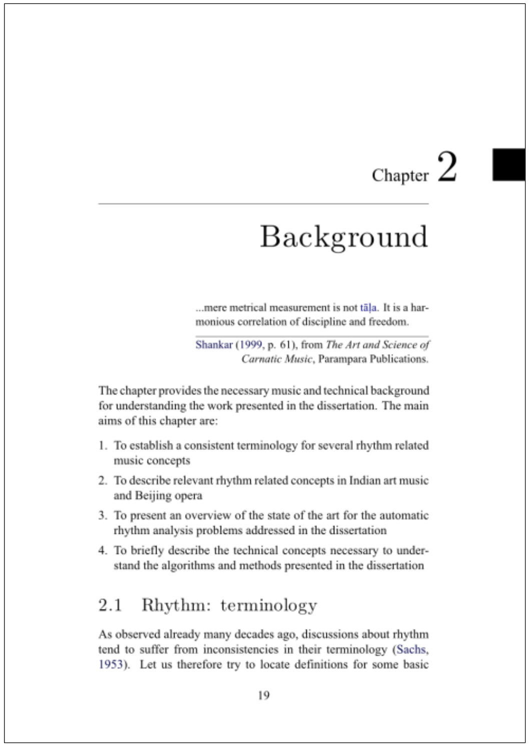
7.1 Applications

There are several applications for the research work presented in the dissertation. Some of these applications have already been identified in Chapter 1 and Chapter 3. The goal of this section is to present concrete examples of such applications, and further suggest other applications that might be built or get benefited from the work presented here. The section describes some of the applications that

277

Background





Chapter 2

Background

...mere metrical measurement is not tāḷa. It is a harmonious correlation of discipline and freedom.

Shankar (1999, p. 61), from *The Art and Science of Carnatic Music*, Parampara Publications.

The chapter provides the necessary music and technical background for understanding the work presented in the dissertation. The main aims of this chapter are:

1. To establish a consistent terminology for several rhythm related music concepts
2. To describe relevant rhythm related concepts in Indian art music and Beijing opera
3. To present an overview of the state of the art for the automatic rhythm analysis problems addressed in the dissertation
4. To briefly describe the technical concepts necessary to understand the algorithms and methods presented in the dissertation

2.1 Rhythm: terminology

As observed already many decades ago, discussions about rhythm tend to suffer from inconsistencies in their terminology (Sachs, 1953). Let us therefore try to locate definitions for some basic

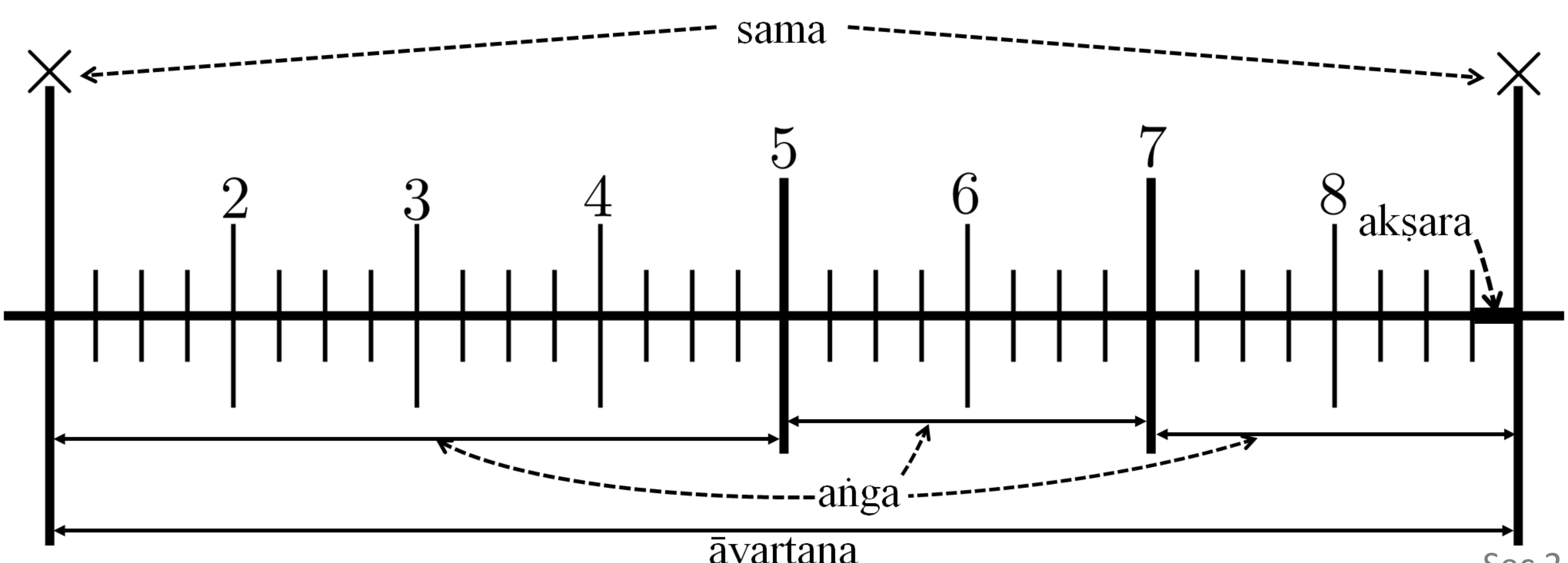
19

Rhythm in Carnatic music

Vignesh Ishwar in concert at Arkay Convention Center, Chennai, India

Tāla in Carnatic music

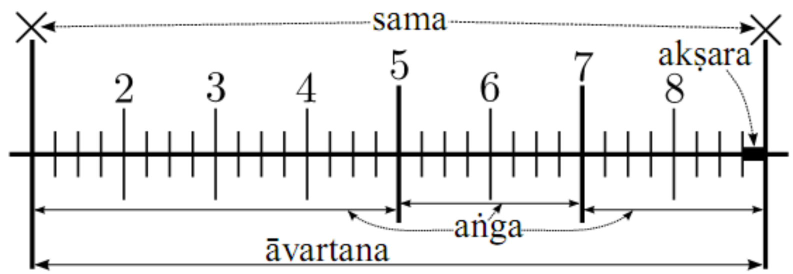
- Time cycles
  - Broad structure for rendition and repetition of melodic and rhythmic phrases, motifs, and improvisations
  - **Akṣara**, “beats”, **sama** (downbeat), **aṅga** (section)



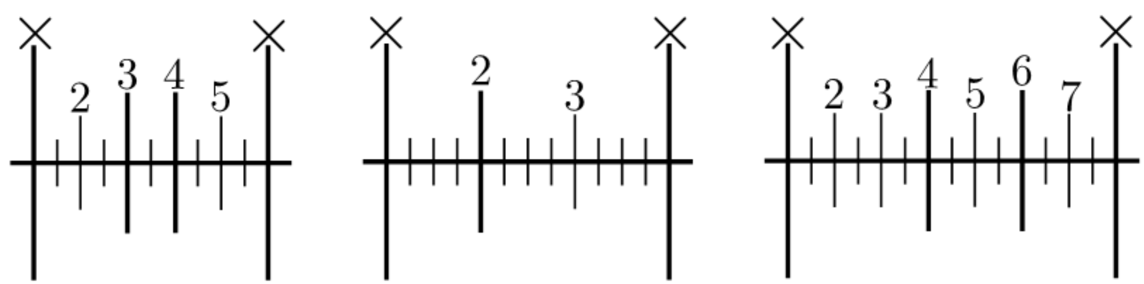
Popular tālas

compmusic

Tāḷa	# beats	naḍe	# Akṣara
Ādi	8	4	32
Rūpaka	3	4	12
Miśra chāpu	7	2	14
Khaṇḍa chāpu	5	2	10

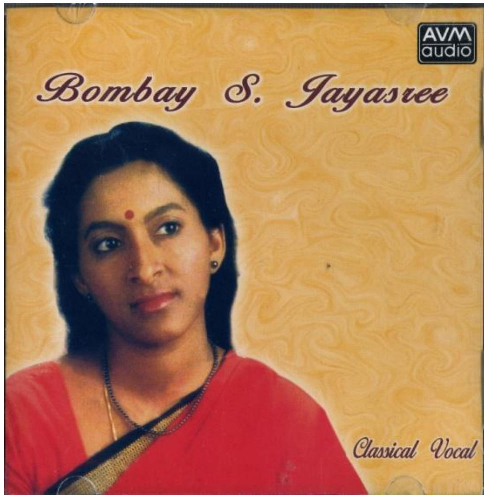


(a) Ādi tāḷa, illustrated

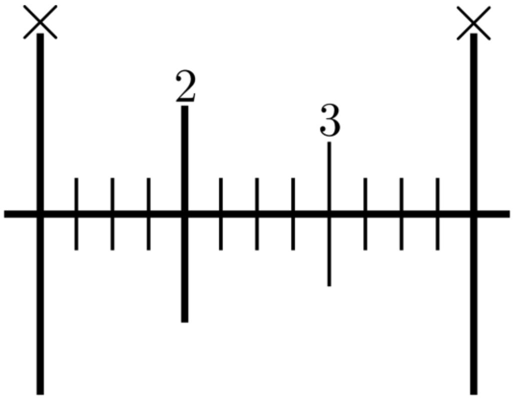


(b) Khaṇḍa chāpu tāḷa **(c)** Rūpaka tāḷa **(d)** Miśra chāpu tāḷa

An example



Artist: Bombay Jayasree (vocal)
Release: Classical Vocal
Composition: Śaṅkari nīvē
Composer: Subbaraya Sastry
Rāga: Bēgaḍa
Tāla: Rūpaka (Cycle of 12 akṣara)



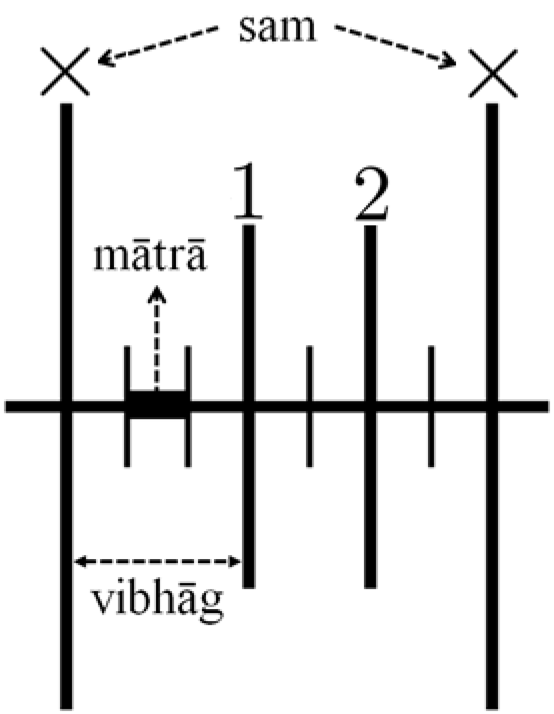
http://musicbrainz.org/recording/fe63aa10-bfa0-4dc6-9845-e57811c2feaa

Rhythm in Hindustani music

https://musicboxnews.files.wordpress.com/2011/09/darbar-festival-20081.jpg

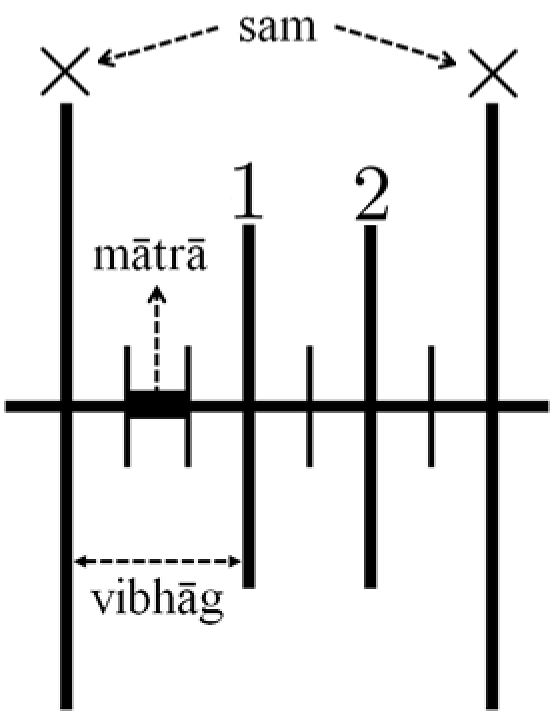
Tāl in Hindustani music

- Metrical time cycles
  - Broad structure for rendition and repetition of melodic and rhythmic phrases, motifs, and improvisations
  - **mātrā** (beat), **sam** (downbeat), **vibhāg** (section)
- Tempo classes (lay)
  - Wide range of tempi
  - Slow (vilaṁbit): 10-60 mātrā per minute (MPM)
  - Medium (madhya): 60-150 MPM
  - Fast (drut): >150 MPM



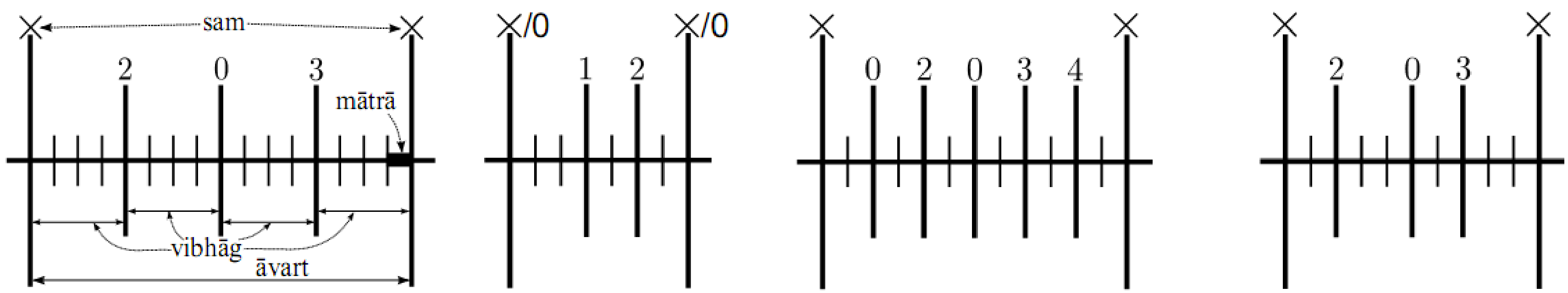
Tāl in Hindustani music

- Metrical time cycles
  - Broad structure for rendition and repetition of melodic and rhythmic phrases, motifs, and improvisations
  - **mātrā** (beat), **sam** (downbeat), **vibhāg** (section)
- Tempo classes (lay)
  - Wide range of tempi
  - Slow (vilaṁbit): 10-60 mātrā per minute (MPM)
  - Medium (madhya): 60-150 MPM
  - Fast (drut): >150 MPM



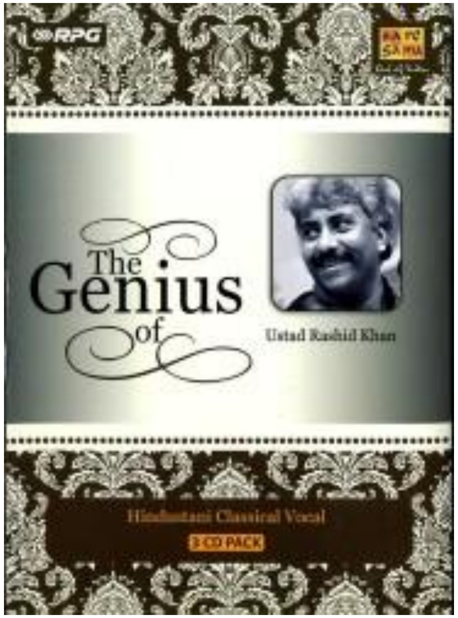
Popular tāls

Tāl	# vibhāg	# mātrās	mātrā grouping
Tīntāl	4	16	4,4,4,4
Ēktāl	6	12	2,2,2,2,2,2
Jhaptāl	4	10	2,3,2,3
Rūpak tāl	3	7	3,2,2

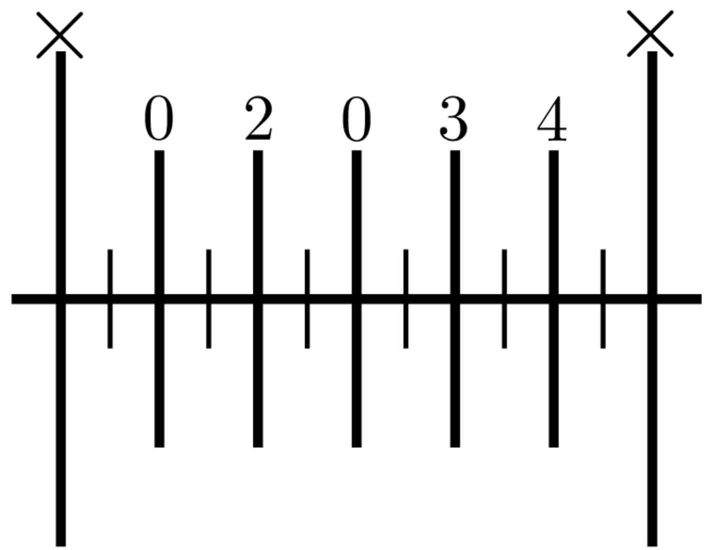


(a) Tīntāl, illustrated (b) Rūpak tāl (c) Ēktāl (d) Jhaptāl

An example



Artist: Rashid Khan (vocal)
Release: The Genius Of Ustad Rashid Khan
Composition: Rasiya Maara Ama Laara
Rāga: Ahir Bhairav
Tāla: Ektāl (Cycle of 12 mātrās)



https://musicbrainz.org/recording/def59da7-8d7b-4047-b16c-eeee65482e87

Some observations

compmusic

- Phrasal, melodic and rhythmic changes closely tied to tāl/a
- State of the art – concepts do not extend
  - E.g. Possibly non-isochronous beats/sections
- Underlying rhythmic structures
  - Not explicit in the signal
- Musical definitions → Engineering formulations

A. Srinivasamurthy, A. Holzapfel, and X. Serra (2014). In Search of Automatic Rhythm Analysis Methods for Turkish and Indian Art Music. Journal of New Music Research, 43(1):97–117.

Syllabic percussion systems

compmusic

- Mnemonics for oral transmission – **syllables**
  - Possibly **onomatopoeic**
- Define and describe percussion patterns
  - Also an important role in many aspects of rhythm
- **A language for percussion**
  - Clear analogy to speech
  - Musically meaningful representation
  - Subtleties of articulation, timbre and dynamics
- Vocal percussion – art form

Clayton, Martin (2000). Time in Indian music : rhythm, metre and form in North Indian rag performance, p. 58, Oxford University Press.

Keçecioğlu, M. U. (2010). Usul-Velvele Editor. `http://notist.org/usul_velvele%20editor_ing.html`

Bōl in Hindustani music

- Strokes of the tabla
- Timbrally grouped for automatic analysis

Bōls	Symbol	Description
DHET	DHET	A combined stroke played with a closed stroke on dāyān with GE
DHI	DHI	A closed combined stroke with GE and a soft resonant stroke on dāyān

Bōls	Symbol	Description
D, DA, DAA	DA	A closed stroke on the dāyān (right drum)
N, NA, TAA	NA	A ringing stroke on the dāyān

Bōl in Hindustani music

- Strokes of the tabla
- Timbrally grouped for automatic analysis

Bōls	Symbol	Description
DHET	DHET	A combined stroke played with a closed stroke on dāyān with GE
DHI	DHI	A closed combined stroke with GE and a soft resonant stroke on dāyān

Bōls	Symbol	Description
D, DA, DAA	DA	A closed stroke on the dāyān (right drum)
N, NA, TAA	NA	A ringing stroke on the dāyān

Canonical patterns of a tāl (ṭhēkā)

×				**2**			
×	2	3	4	5	6	7	8
DHA	DHIN	DHIN	DHA	DHA	DHIN	DHIN	DHA
0				**3**			
9	10	11	12	13	14	15	16
DHA	TIN	TIN	NA	NA	DHIN	DHIN	DHA

(a) Tīntāl

×		**0**		**2**	
×	2	3	4	5	6
DHIN	DHIN	DHA GE	TIRAKITA	TUN	NA
0		**3**		**4**	
7	8	9	10	11	12
KAT	TA	DHA GE	TIRAKITA	DHIN	NA

(b) Ēktāl

Carnatic music – Solkattu

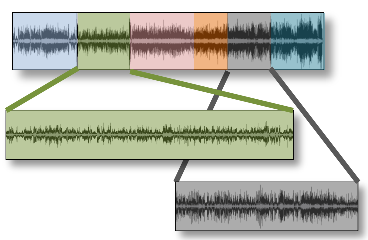
- Strokes of the mridangam
- Compound strokes grouped together

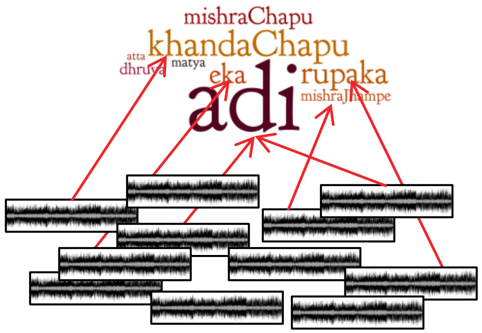
Syllable	Description
AC	A semi ringing stroke on the right head
ACT	AC with TH/TM
CH	A ringing stroke on the right head
CHT	CH with TH/TM
DM	A strong ringing stroke on the right head

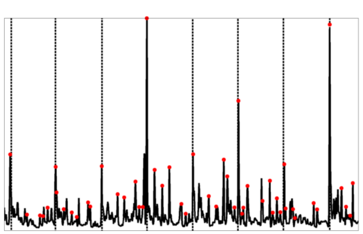
Summary – Indian art music

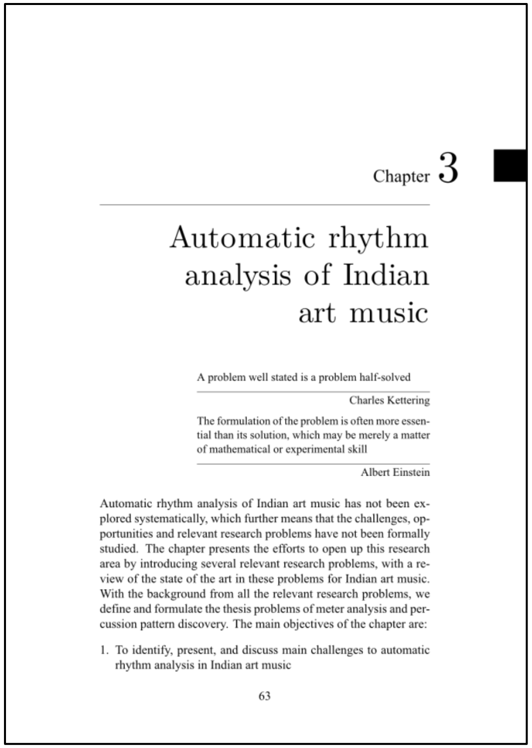
- Oral traditions, improvised
  - Tāla is the basic skeletal rhythmic framework
- No metronome, flexible timing
  - Wide tempo range in HM, smaller in CM
- Surface rhythm – indicators to underlying structures
- Well defined syllabic percussion systems
- Have unmetered forms
  - Beyond the scope of automatic rhythm analysis

Automatic rhythm analysis problems








Chapter 3

Automatic rhythm analysis of Indian art music

A problem well stated is a problem half-solved

Charles Kettering

The formulation of the problem is often more essential than its solution, which may be merely a matter of mathematical or experimental skill

Albert Einstein

Automatic rhythm analysis of Indian art music has not been explored systematically, which further means that the challenges, opportunities and relevant research problems have not been formally studied. The chapter presents the efforts to open up this research area by introducing several relevant research problems, with a review of the state of the art in these problems for Indian art music. With the background from all the relevant research problems, we define and formulate the thesis problems of meter analysis and percussion pattern discovery. The main objectives of the chapter are:

1. To identify, present, and discuss main challenges to automatic rhythm analysis in Indian art music

63

Automatic rhythm analysis in Indian art music

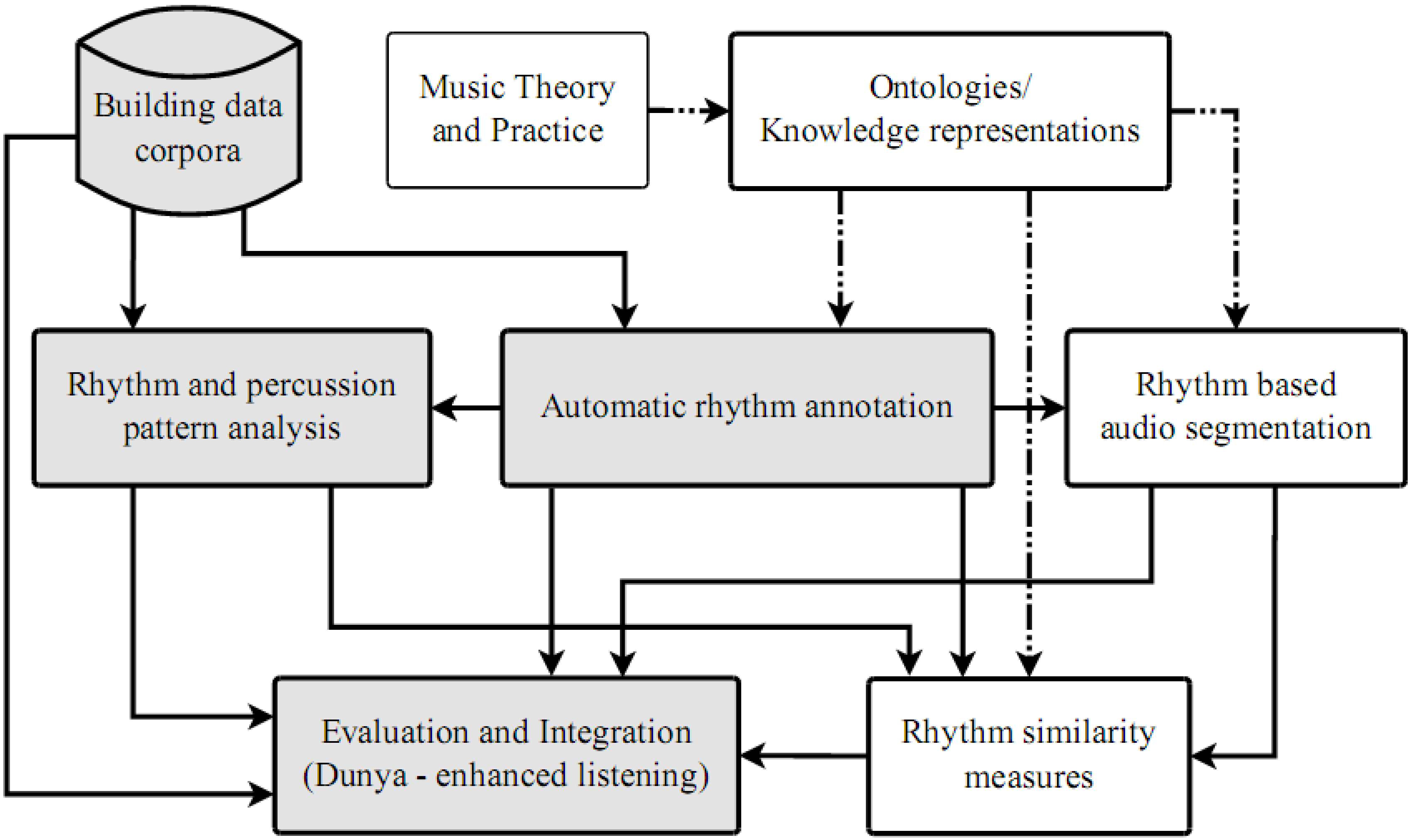


Fig 3.2

Automatic rhythm analysis in Indian art music

compmusic

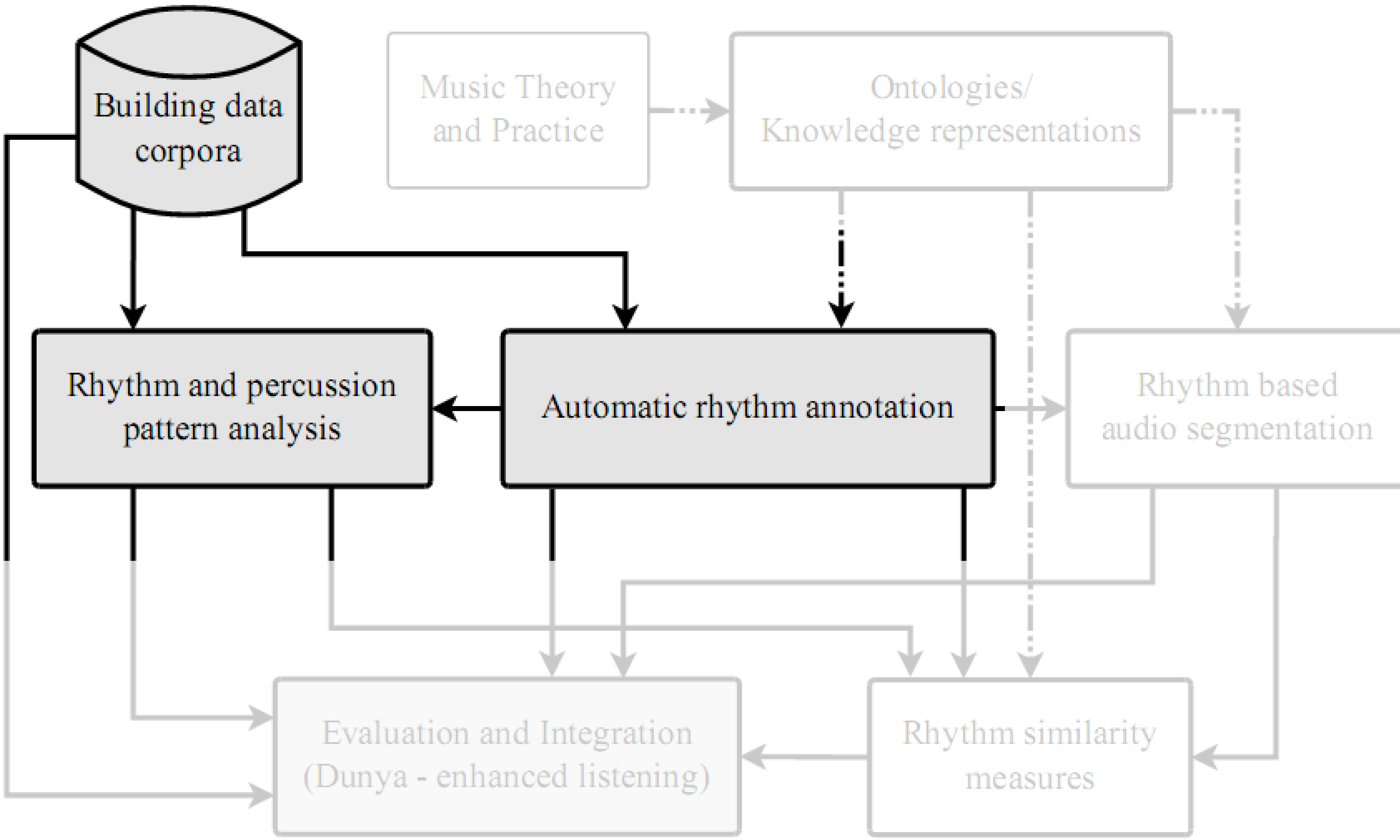
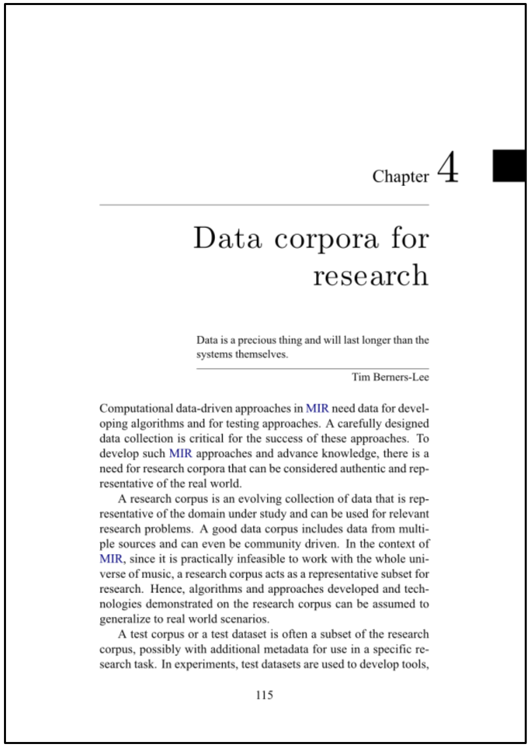


Fig 3.2

Data corpora for research



Chapter 4

Data corpora for research

Data is a precious thing and will last longer than the systems themselves.

Tim Berners-Lee

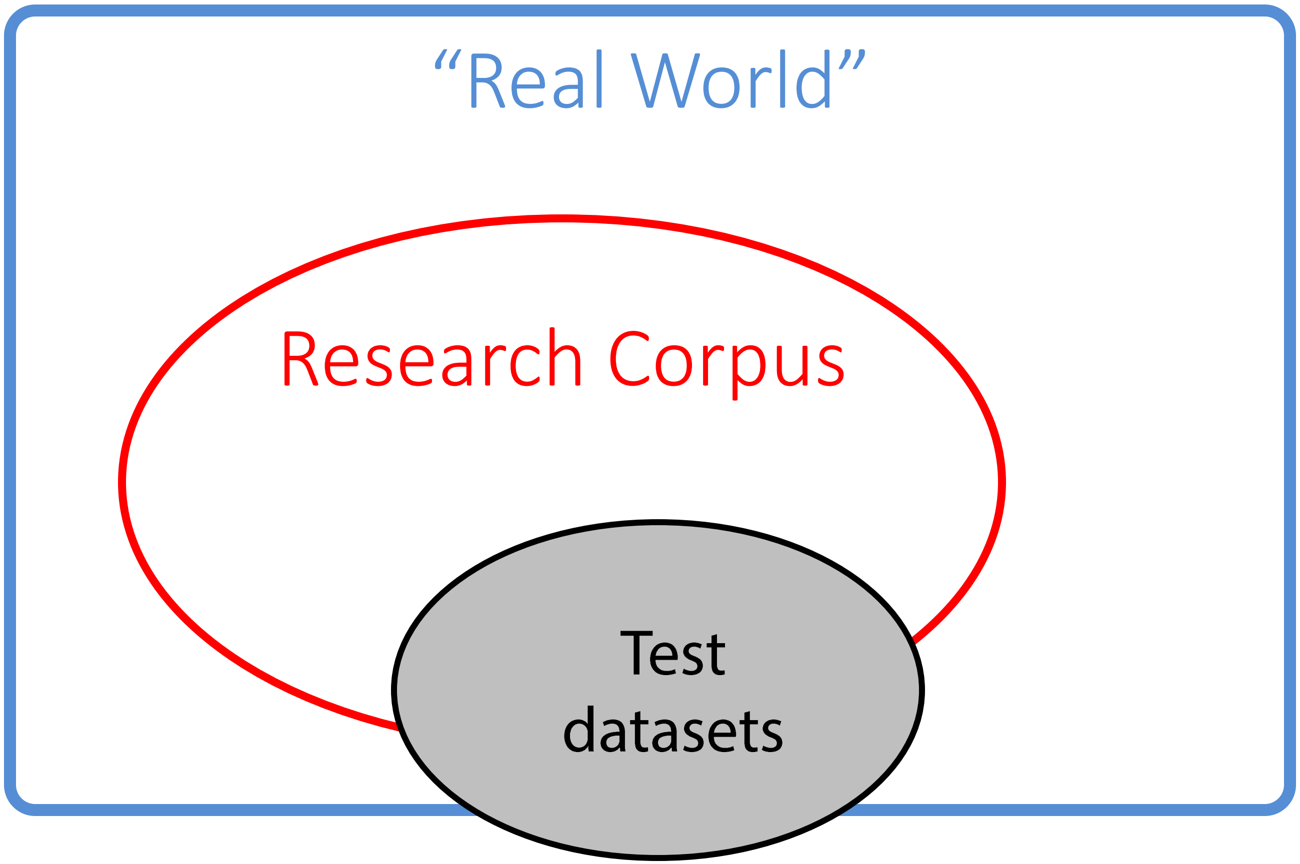
Computational data-driven approaches in MIR need data for developing algorithms and for testing approaches. A carefully designed data collection is critical for the success of these approaches. To develop such MIR approaches and advance knowledge, there is a need for research corpora that can be considered authentic and representative of the real world.

A research corpus is an evolving collection of data that is representative of the domain under study and can be used for relevant research problems. A good data corpus includes data from multiple sources and can even be community driven. In the context of MIR, since it is practically infeasible to work with the whole universe of music, a research corpus acts as a representative subset for research. Hence, algorithms and approaches developed and technologies demonstrated on the research corpus can be assumed to generalize to real world scenarios.

A test corpus or a test dataset is often a subset of the research corpus, possibly with additional metadata for use in a specific research task. In experiments, test datasets are used to develop tools,

115

Data corpora for research



Serra, X. (2014, January). Creating Research Corpora for the Computational Study of Music: the case of the CompMusic Project. In Proceedings of the 53rd AES International Conference on Semantic Audio. London.

Building research corpora

- Building research corpora and test datasets
  - A research problem in itself !
- Systematic approach to building MIR corpora
  - Criteria: Purpose, Coverage, Completeness, Quality and Reusability [Serra, 2014]
  - Based on reliable references
- Corpora for Indian art music
  - An analysis of the corpus
- **Test datasets useful for different MIR tasks**

Srinivasamurthy, A., Koduri, G. K., Gulati, S., Ishwar, V., & Serra, X. (2014, September). Corpora for Music Information Research in Indian Art Music. In Proceedings of Joint 42nd International ComputerMusic Conference/13th Sound and Music Computing Conference (pp. 1029–1036).

Datasets for rhythm analysis

compmusic

- Carnatic music rhythm dataset (CMR dataset)
- Hindustani music rhythm dataset (HMR dataset)

- Mulgaonkar Tabla Solo dataset (MTS dataset)
- UKS Mridangam Solo dataset (UMS dataset)

Rhythm annotated datasets

Datasets for rhythm analysis

compmusic

- Carnatic music rhythm dataset (CMR dataset)
- Hindustani music rhythm dataset (HMR dataset)

- Mulgaonkar Tabla Solo dataset (MTS dataset)
- UKS Mridangam Solo dataset (UMS dataset)

Datasets for percussion analysis

Datasets for rhythm analysis

- Carnatic music rhythm dataset (CMR dataset)
- Hindustani music rhythm dataset (HMR dataset)

- Mulgaonkar Tabla Solo dataset (MTS dataset)
- UKS Mridangam Solo dataset (UMS dataset)

Annotations are public. Audio is commercially available and easily accessible.

Carnatic music rhythm dataset

compmusic

Tāḷa	#Pieces	Total Duration hours (min)	$\overline{T_f}$	#Ann.	#Sama
Ādi	50	4.21 (252.78)	4m51s	22793	2882
Rūpaka	50	4.45 (267.45)	4m37s	22668	7582
Miśra chāpu	48	5.70 (342.13)	6m35s	54309	7795
Khaṇḍa chāpu	28	2.24 (134.62)	4m25s	21382	4387
Total	176	16.61 (996.98)	5m4s	121602	22646

CMR_f

Srinivasamurthy, A., & Serra, X. (2014, May). A Supervised Approach to Hierarchical Metrical Cycle Tracking from Audio Music Recordings. In Proc. of the 39th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2014) (pp. 5237–5241). Florence, Italy.

Carnatic music rhythm dataset

compmusic

Tāḷa	#Pieces	Total Duration hours (min)	$\overline{T_f}$	#Ann.	#Sama
Ādi	50	4.21 (252.78)	4m51s	22793	2882
Rūpaka	50	4.45 (267.45)	4m37s	22668	7582
Miśra chāpu	48	5.70 (342.13)	6m35s	54309	7795
Khaṇḍa chāpu	28	2.24 (134.62)	4m25s	21382	4387
Total	176	16.61 (996.98)	5m4s	121602	22646

CMR_f

Tāḷa	# Pieces	Total Duration hours (min)	# Ann.	# Sama
Ādi	30	0.98 (58.87)	5452	696
Rūpaka	30	1.00 (60.00)	5148	1725
Miśra chāpu	30	1.00 (60.00)	8992	1299
Khaṇḍa chāpu	28	0.93 (55.93)	9133	1840
Total	118	3.91 (234.80)	28725	5560

CMR

Two minute excerpts

http://compmusic.upf.edu/carnatic-rhythm-dataset

Carnatic music rhythm dataset

compmusic

CMR$_f$

Tāḷa	#Pieces	Total Duration hours (min)	$\overline{T_f}$	#Ann.	#Sama
Ādi	50	4.21 (252.78)	4m51s	22793	2882
Rūpaka	50	4.45 (267.45)	4m37s	22668	7582
Miśra chāpu	48	5.70 (342.13)	6m35s	54309	7795
Khaṇḍa chāpu	28	2.24 (134.62)	4m25s	21382	4387
Total	176	16.61 (996.98)	5m4s	121602	22646

CMR

Tāḷa	# Pieces	Total Duration hours (min)	# Ann.	# Sama
Ādi	30	0.98 (58.87)	5452	696
Rūpaka	30	1.00 (60.00)	5148	1725
Miśra chāpu	30	1.00 (60.00)	8992	1299
Khaṇḍa chāpu	28	0.93 (55.93)	9133	1840
Total	118	3.91 (234.80)	28725	5560

Hindustani music rhythm dataset

compmusic

- HMR_f dataset, two minute excerpts from full pieces

Tāl	# Pieces	Total Duration hours (min)	# Ann.	# Sam
Tīntāl	54	1.80 (108)	17142	1081
Ēktāl	58	1.93 (116)	12999	1087
Jhaptāl	19	0.63 (38)	3029	302
Rūpak tāl	20	0.67 (40)	2841	406
Total	151	5.03 (302)	36011	2876

Tāl	$\overline{\tau_s} \pm \sigma_s$	$\overline{\tau_b} \pm \sigma_b$	$[\tau_{s,\min}, \tau_{s,\max}]$
Tīntāl	10.36 ± 9.875	0.65 ± 0.617	[2.32, 44.14]
Ēktāl	30.20 ± 26.258	2.52 ± 2.188	[2.23, 69.73]
Jhaptāl	8.51 ± 3.149	0.85 ± 0.315	[4.06, 16.23]
Rūpak tāl	7.11 ± 3.360	1.02 ± 0.480	[2.82, 16.09]

Srinivasamurthy, A., Holzapfel, A., Cemgil, A. T., & Serra, X. (2016, March). A generalized Bayesian model for tracking long metrical cycles in acoustic music signals. In Proceedings of the 41st IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2016) (pp. 76–80). Shanghai, China.

Hindustani music rhythm dataset

compmusic

- HMR$_f$ dataset

Tāl	# Pieces	Total Duration hours (min)	# Ann.	# Sam
Tīntāl	54	1.80 (108)	17142	1081
Ēktāl	58	1.93 (116)	12999	1087
Jhaptāl	19	0.63 (38)	3029	302
Rūpak tāl	20	0.67 (40)	2841	406
Total	151	5.03 (302)	36011	2876

Tāl	$\overline{\tau_s} \pm \sigma_s$	$\overline{\tau_b} \pm \sigma_b$	$[\tau_{s,\min}\ ,\ \tau_{s,\max}]$
Tīntāl	10.36 ± 9.875	0.65 ± 0.617	[2.32, 44.14]
Ēktāl	30.20 ± 26.258	2.52 ± 2.188	[2.23, 69.73]
Jhaptāl	8.51 ± 3.149	0.85 ± 0.315	[4.06, 16.23]
Rūpak tāl	7.11 ± 3.360	1.02 ± 0.480	[2.82, 16.09]

http://compmusic.upf.edu/hindustani-rhythm-dataset

Long cycle (duration) subset

- HMR_l dataset (< 60 MPM, vilaṁbit)

Tāl	# Pieces	Total Duration hours (min)	# Ann.	# Sam
Tīntāl	13	0.43 (26)	1020	65
Ēktāl	32	1.07 (64)	967	79
Jhaptāl	6	0.2 (12)	592	59
Rūpak tāl	8	0.27 (16)	701	101
Total	59	1.97 (118)	3280	304

Tāl	$\overline{\tau_s} \pm \sigma_s$	$\overline{\tau_b} \pm \sigma_b$	$[\tau_{s,\min}, \tau_{s,\max}]$
Tīntāl	26.16 ± 7.963	1.63 ± 0.498	[18.57, 44.14]
Ēktāl	52.16 ± 12.531	4.35 ± 1.044	[14.43, 69.73]
Jhaptāl	12.30 ± 1.935	1.23 ± 0.194	[10.20, 16.23]
Rūpak tāl	10.28 ± 3.050	1.47 ± 0.436	[6.95, 16.09]

Short cycle (duration) subset

- HMR_s dataset (> 60 MPM, madhya and drut)

Tāl	# Pieces	Total Duration hours (min)	# Ann.	# Sam
Tīntāl	41	1.37 (82)	16122	1016
Ēktāl	26	0.87 (52)	12032	1008
Jhaptāl	13	0.43 (26)	2437	243
Rūpak tāl	12	0.40 (24)	2140	305
Total	92	3.07 (184)	32731	2572

Tāl	$\overline{\tau_s} \pm \sigma_s$	$\overline{\tau_b} \pm \sigma_b$	$[\tau_{s,\min}, \tau_{s,\max}]$
Tīntāl	5.35 ± 1.823	0.33 ± 0.114	[2.32, 9.89]
Ēktāl	3.17 ± 0.471	0.26 ± 0.039	[2.23, 4.11]
Jhaptāl	6.77 ± 1.688	0.68 ± 0.169	[4.06, 9.97]
Rūpak tāl	5.00 ± 1.191	0.71 ± 0.170	[2.82, 6.68]

Corpus level analysis

- Tempo range and distribution
- **Analysis of cycle length rhythm patterns**
- Intra-cycle tempo variations

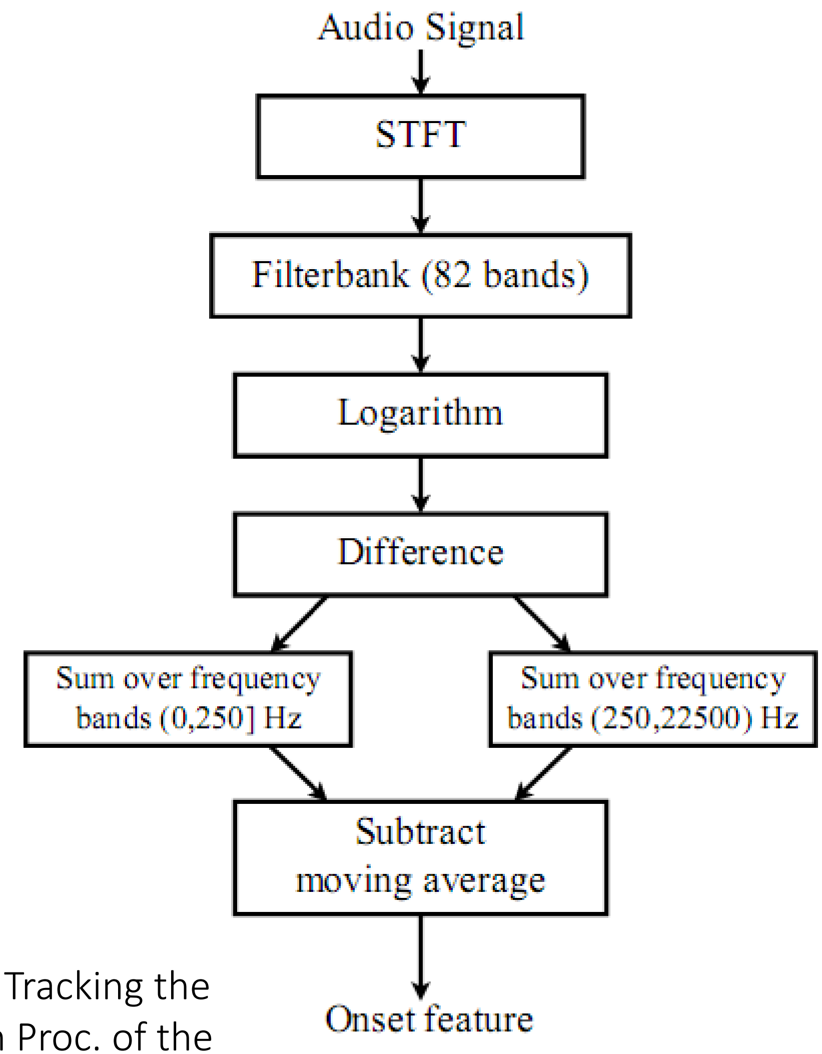
Srinivasamurthy, A., Holzapfel, A., Ganguli, K. K., & Serra, X. Computational Rhythm analysis of a Hindustani Music Corpus (under preparation)

Cycle length rhythm patterns

- Musically relevant
  - Useful for meter analysis
- Spectral flux feature
  - Indicative of onsets
- Average patterns
  - Over the complete dataset

Cycle length rhythm patterns

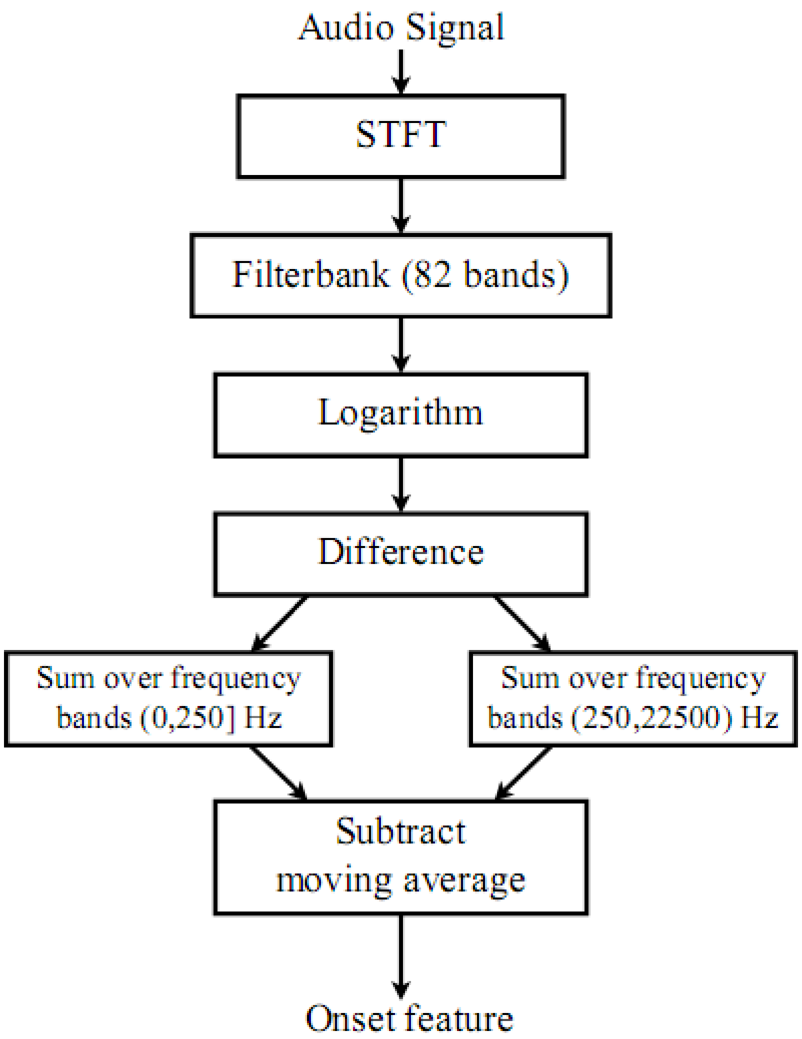
- Musically relevant
  - Useful for meter analysis
- Spectral flux feature
  - Indicative of onsets
- Average patterns
  - Over the complete dataset



Holzapfel, A., Krebs, F., & **Srinivasamurthy, A.** (2014, October). Tracking the "odd": Meter inference in a culturally diverse music corpus. In Proc. of the 15th International Society forMusic Information Retrieval Conference (ISMIR 2014) (pp. 425–430). Taipei, Taiwan.

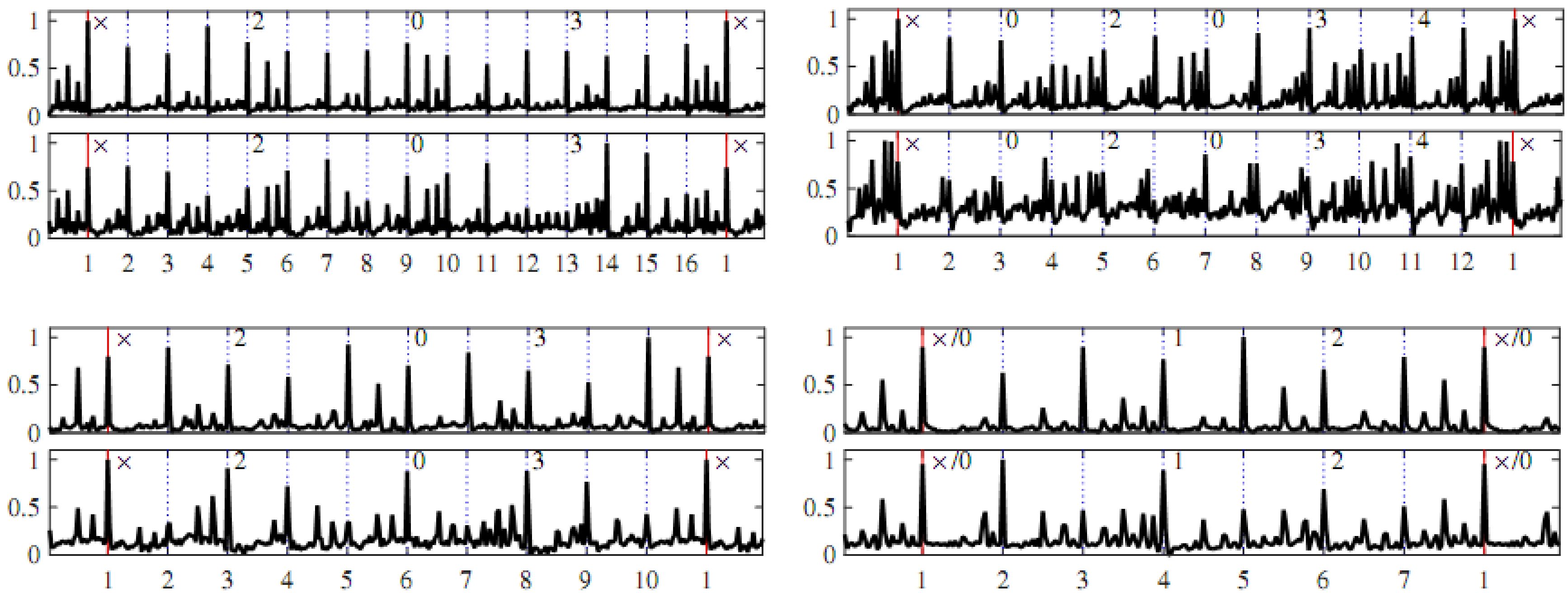
Cycle length rhythm patterns

- Musically relevant
  - Useful for meter analysis
- Spectral flux feature
  - Indicative of onsets
- Average patterns
  - Over the complete dataset



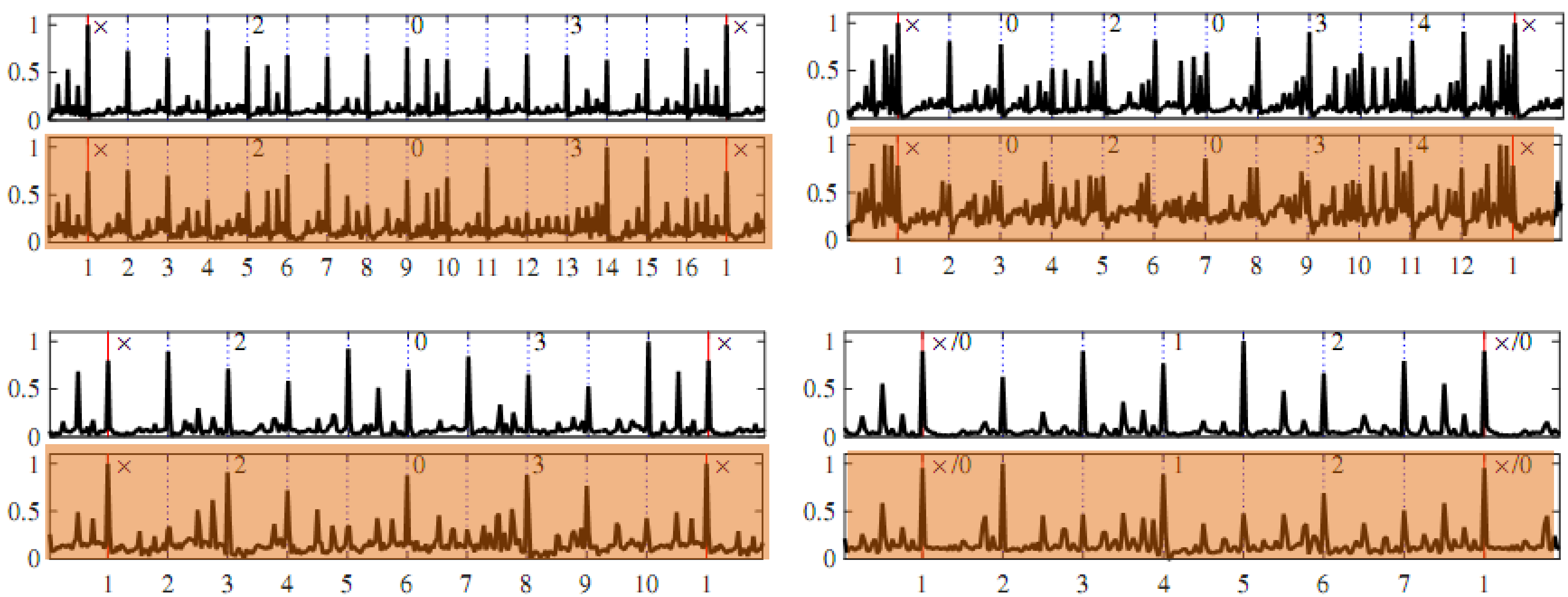
Rhythm patterns in HMR_l dataset

- Rhythm patterns on long and short cycle subsets



Rhythm patterns in HMR$_l$ dataset

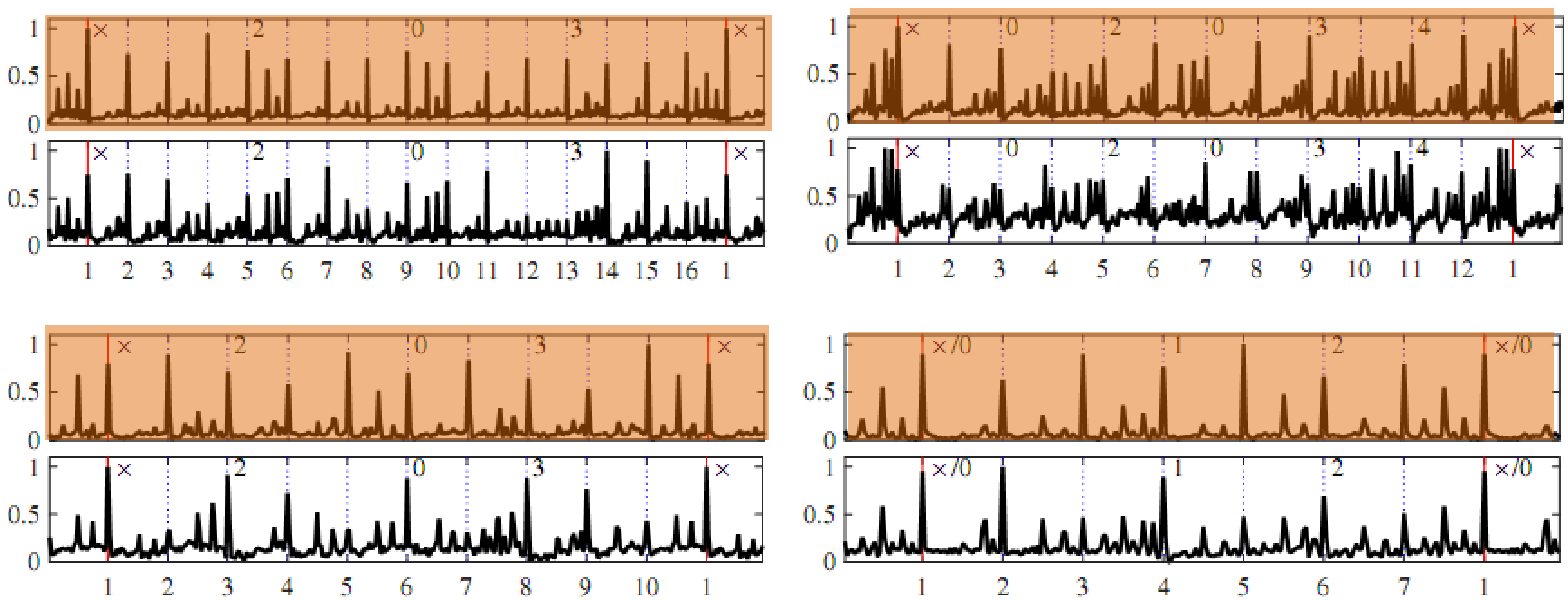
- Low frequency band, left bass strokes of the tabla



Rhythm patterns in HMR_l dataset

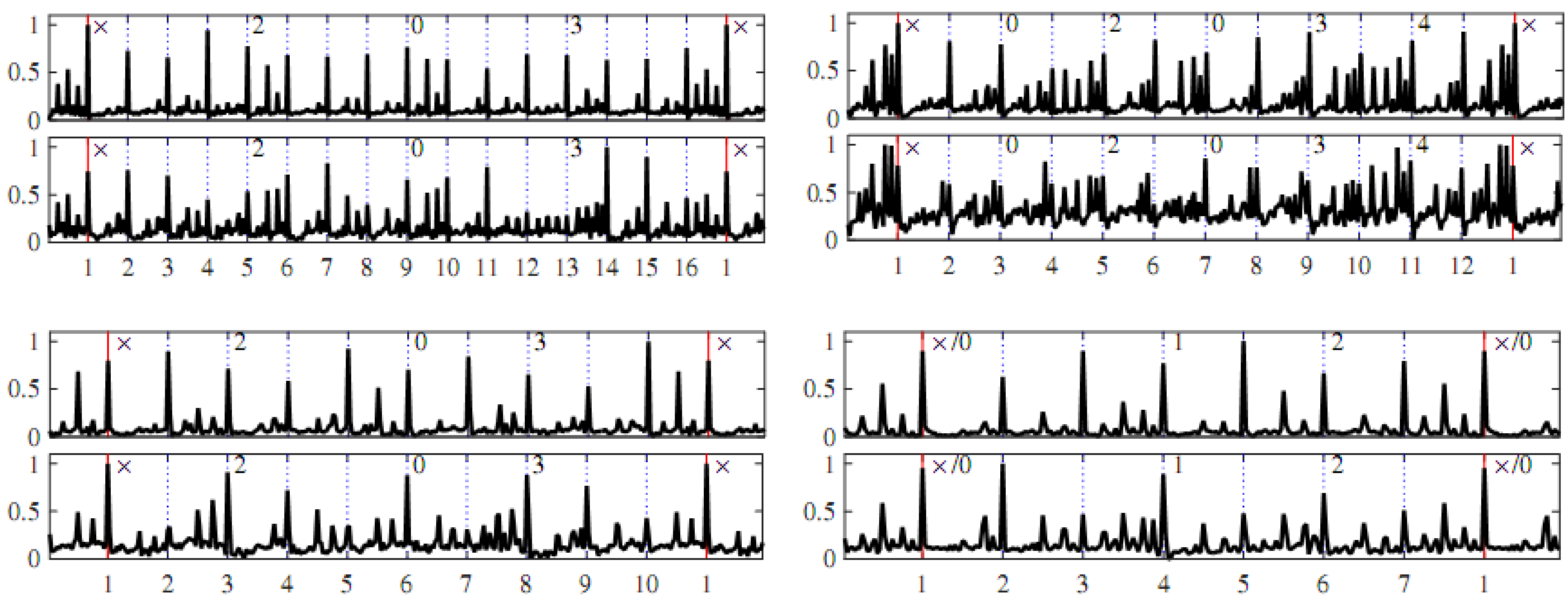
compmusic

- High frequency band, right strokes of tabla



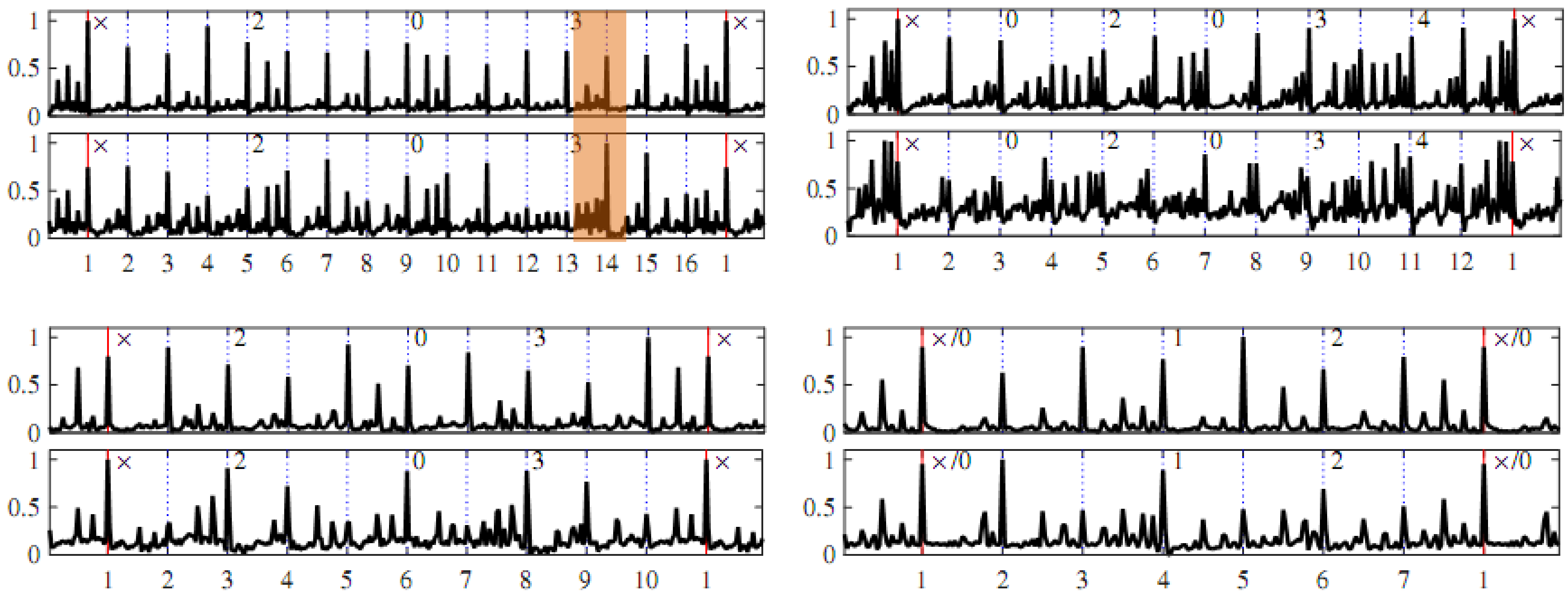
Rhythm patterns in HMR_l dataset

- Stronger accents on the mātrā – mātrās are anchors
- Fillers between mātrās – vilaṁbit



Rhythm patterns in HMR$_l$ dataset

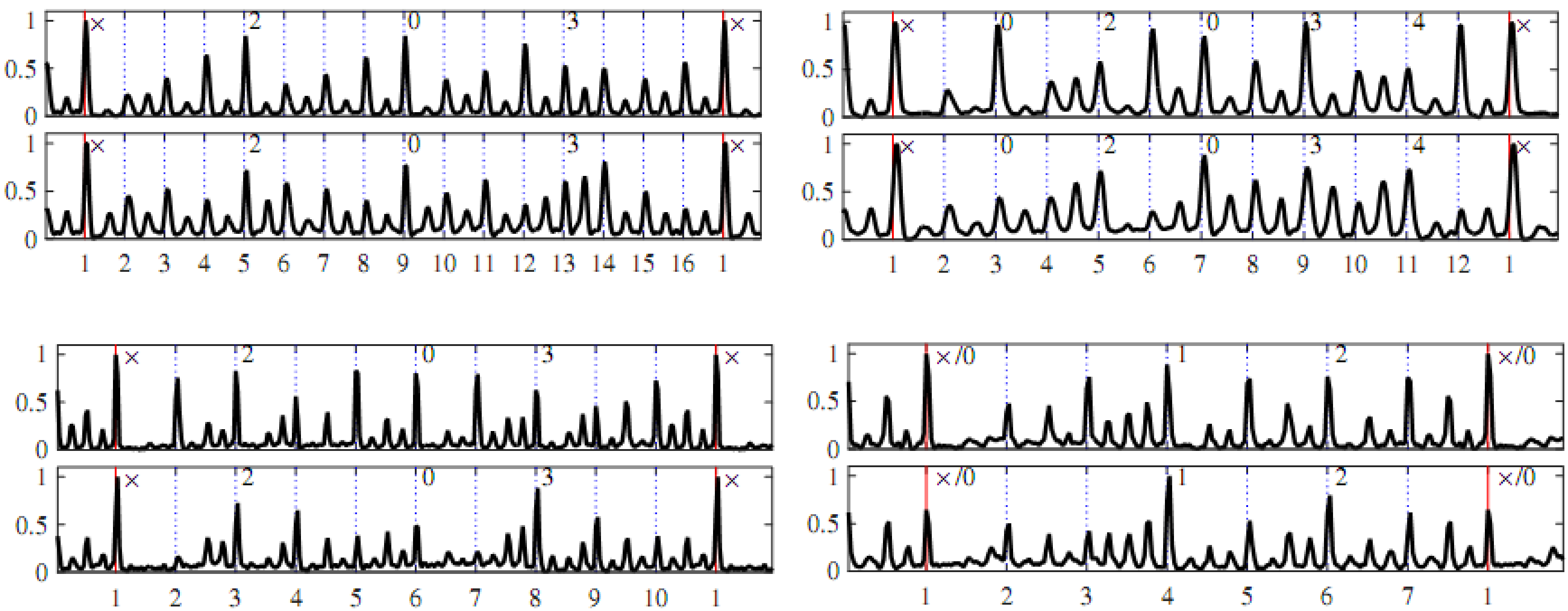
- Tīntāl – 14th mātra strong left accent
  - Indicating the arrival of sam (āmad)



Rhythm patterns in HMR$_s$ dataset

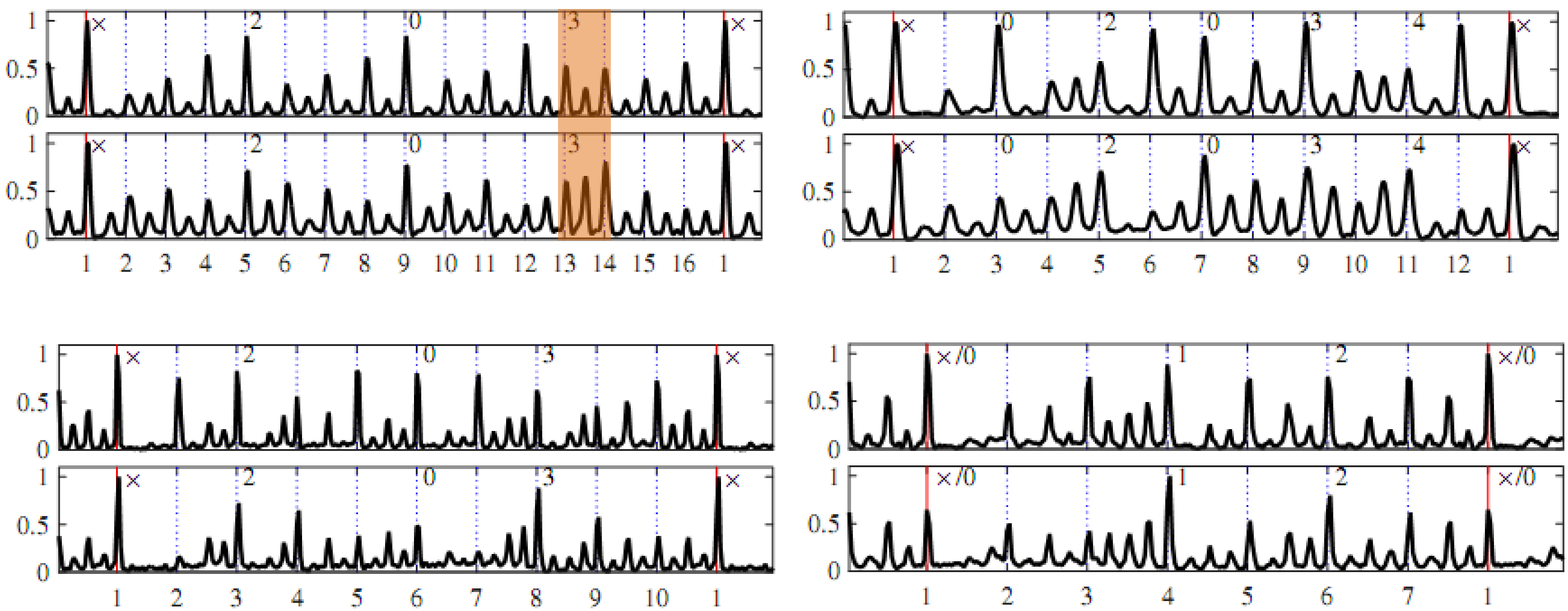
compmusic

- Less fillers, shorter cycles
- Vibhāg are anchors



Rhythm patterns in HMR$_s$ dataset

- Tīntāl – start of last vibhāg, more fillers
  - Indication of sam coming up



Mulgaonkar Tabla Solo (MTS) dataset

compmusic

- 8245 syllables, 38 compositions, tīntāl, 17 min, time-aligned

ID	Syllable	#Inst.	ID	Syllable	#Inst.
1	DA	132	10	KI	1482
2	DHA	582	11	NA	1308
3	DHE	277	12	RE	294
4	DHET	67	13	TA	2375
5	DHI	156	14	TE	18
6	DHIN	149	15	TII	64
7	DIN	117	16	TIN	61
8	GE	961	17	TIT	43
9	KDA	95	18	TRA	64

http://compmusic.upf.edu/tabla-solo-dataset

Gupta, S., Srinivasamurthy, A., Kumar, M., Murthy, H., & Serra, X. (2015, October). Discovery of Syllabic Percussion Patterns in Tabla Solo Recordings. In Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR 2015) (pp. 385–391). Malaga, Spain.

UKS Mridangam Solo (UMS) dataset

- 8877 syllables, 24 min, ādi and rūpaka tāla, not time-aligned

ID	Syllable	#Inst.	ID	Syllable	#Inst.
1	AC	119	12	DNT	922
2	ACT	50	13	LF	467
3	CH	114	14	LFT	12
4	CHT	112	15	NM	850
5	DM	14	16	NMT	632
6	DH3	1266	17	TH	776
7	DH3T	23	18	TA	754
8	DH3M	602	19	TAT	13
9	DH4	367	20	TM	913
10	DH4T	12	21	TG	30
11	DN	829	-	-	-

http://compmusic.upf.edu/mridangam-tani-dataset

Kuriakose, J., Kumar, J. C., Sarala, P., Murthy, H. A., Sivaraman,U.K. (2015, February). Akshara Transcription of Mrudangam Strokes in CarnaticMusic. In Proc. 21st National Conference on Communication. Mumbai, India.

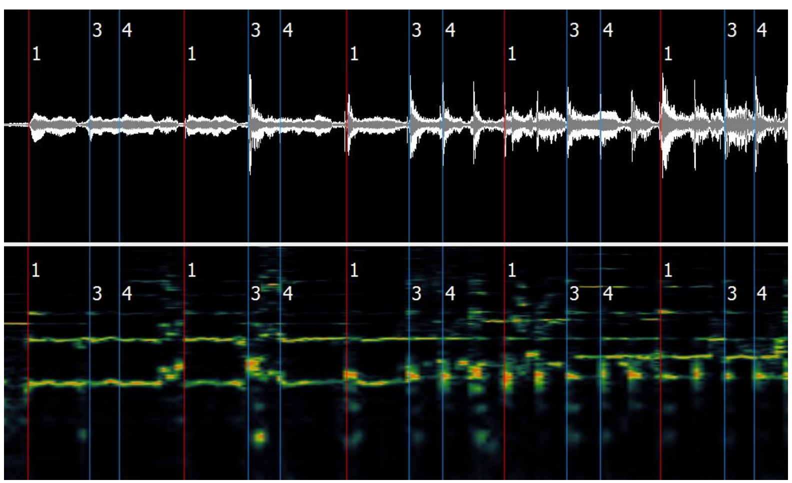
Ballroom dataset

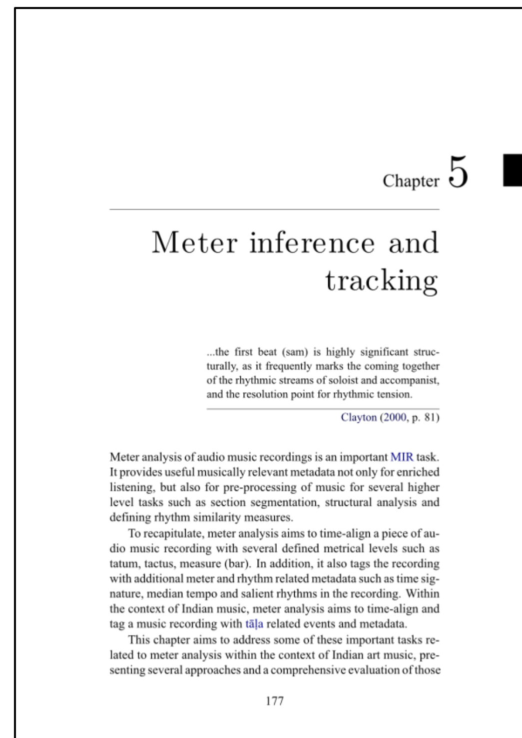
compmusic

- Several dance styles from BallroomDancers.com
- 697 thirty second clips
- Beat and downbeat annotated
- Time signatures 3/4 and 4/4
  - fixed through the piece

Gouyon, F., Klapuri, A., Dixon, S., Alonso, M., Tzanetakis, G., Uhle, C., & Cano, P. (2006). An experimental comparison of audio tempo induction algorithms. IEEE Transactions on Audio, Speech and Language Processing, 14(5), 1832–1844

Meter inference and tracking





Chapter 5

Meter inference and tracking

> ...the first beat (sam) is highly significant structurally, as it frequently marks the coming together of the rhythmic streams of soloist and accompanist, and the resolution point for rhythmic tension.
>
> Clayton (2000, p. 81)

Meter analysis of audio music recordings is an important MIR task. It provides useful musically relevant metadata not only for enriched listening, but also for pre-processing of music for several higher level tasks such as section segmentation, structural analysis and defining rhythm similarity measures.

To recapitulate, meter analysis aims to time-align a piece of audio music recording with several defined metrical levels such as tatum, tactus, measure (bar). In addition, it also tags the recording with additional meter and rhythm related metadata such as time signature, median tempo and salient rhythms in the recording. Within the context of Indian music, meter analysis aims to time-align and tag a music recording with tāḷa related events and metadata.

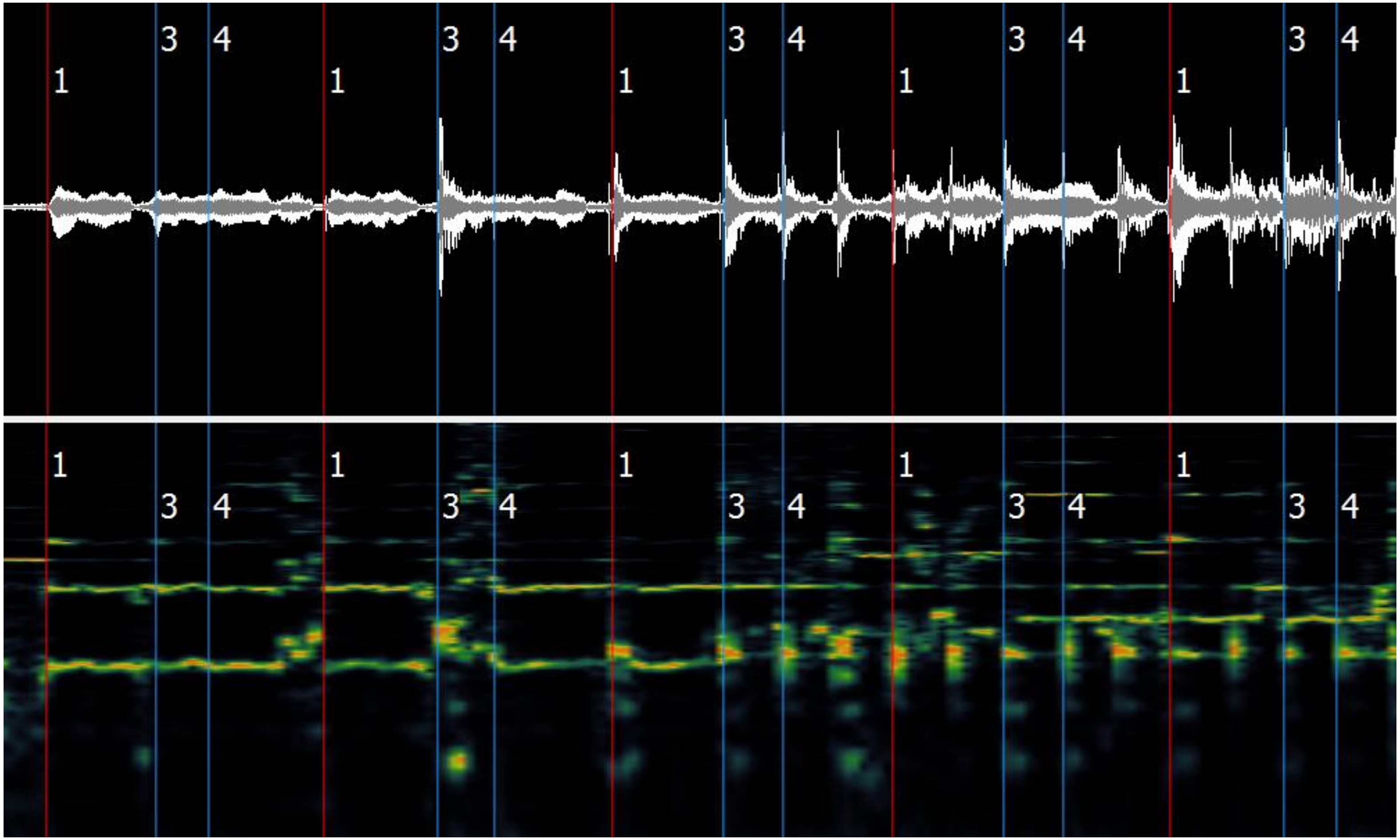
This chapter aims to address some of these important tasks related to meter analysis within the context of Indian art music, presenting several approaches and a comprehensive evaluation of those

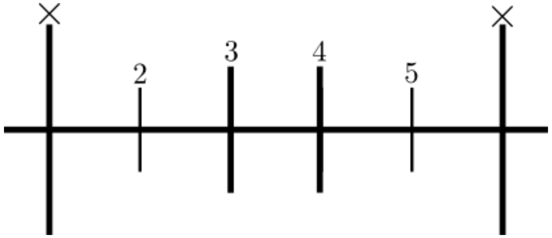
177

Meter analysis

- Analysis of different aspects of meter
  - Events such as beats, downbeats
  - Tempo, rhythm class (tāl/a), patterns

The goal !





Meter analysis tasks

compmusic

- Meter inference
  - Rhythm class + tempo + beats + downbeats
- Meter tracking
  - Rhythm class/metrical structure known
  - Tempo + beats + downbeats
- Informed meter tracking
  - Tempo informed meter tracking – tight tempo range known/given
  - Tempo + downbeat (sama) informed meter tracking
    - A tight tempo range known
    - A few downbeat instances known/given

Meter analysis tasks

compmusic

- Meter inference
  - Rhythm class + tempo + beats + downbeats
- Meter tracking
  - Rhythm class/metrical structure known
  - Tempo + beats + downbeats
- Informed meter tracking
  - Tempo informed meter tracking – tight tempo range known/given
  - Tempo + downbeat (sama) informed meter tracking
    - A tight tempo range known
    - A few downbeat instances known/given

Meter analysis tasks

compmusic

- Meter inference
  - Rhythm class + tempo + beats + downbeats
- Meter tracking
  - Rhythm class/metrical structure known
  - Tempo + beats + downbeats
- Informed meter tracking
  - Tempo informed meter tracking – tight tempo range known/given
  - Tempo + downbeat (sama) informed meter tracking
    - A tight tempo range known
    - A few downbeat instances known/given

Meter analysis tasks

- Meter inference
  - Rhythm class + tempo + beats + downbeats
- Meter tracking
  - Rhythm class/metrical structure known
  - Tempo + beats + downbeats
- Informed meter tracking
  - Tempo informed meter tracking – tight tempo range known/given
  - Tempo + downbeat (sama) informed meter tracking
    - A tight tempo range known
    - A few downbeat instances known/given

State of the art

- Tempo estimation
  - Well explored, utilizes periodicity in onsets
- Beat tracking
  - “Foot tapping times” in a piece of music
  - Several approaches, but assume isochronicity
- Time signature estimation
  - Estimating the metrical structure
- Downbeat (sama) tracking
  - Recently addressed
- Integrated approaches – recent efforts

State of the art

- Tempo estimation
  - Well explored, utilizes periodicity in onsets
- Beat tracking
  - “Foot tapping times” in a piece of music
  - Several approaches, but assume isochronicity
- Time signature estimation
  - Estimating the metrical structure
- Downbeat (sama) tracking
  - Recently addressed
- Integrated approaches – recent efforts

Evaluation of meter analysis tasks

compmusic

- Equivalent definitions for beats and downbeats
- Metrical ambiguity
  - CML and AML
- Precision ($\mathfrak{p}$)
- Recall ($\mathfrak{r}$)
- f-measure ($\mathfrak{f}$)
- Information Gain ($\mathfrak{I}$)
- Continuity based measure: AML_t

Preliminary check – MIR state of the art

compmusic

Method	ādi (8)	rūpaka (3)
DAV	21.7	41.2
HOC-SVM	22.9	42.1
HOC	49.9	64.4

"Downbeat" estimation

- Poor, insufficient performance
- Lack of engineering definitions of musical concepts
- Need for methods that explicitly use the knowledge of metrical structures

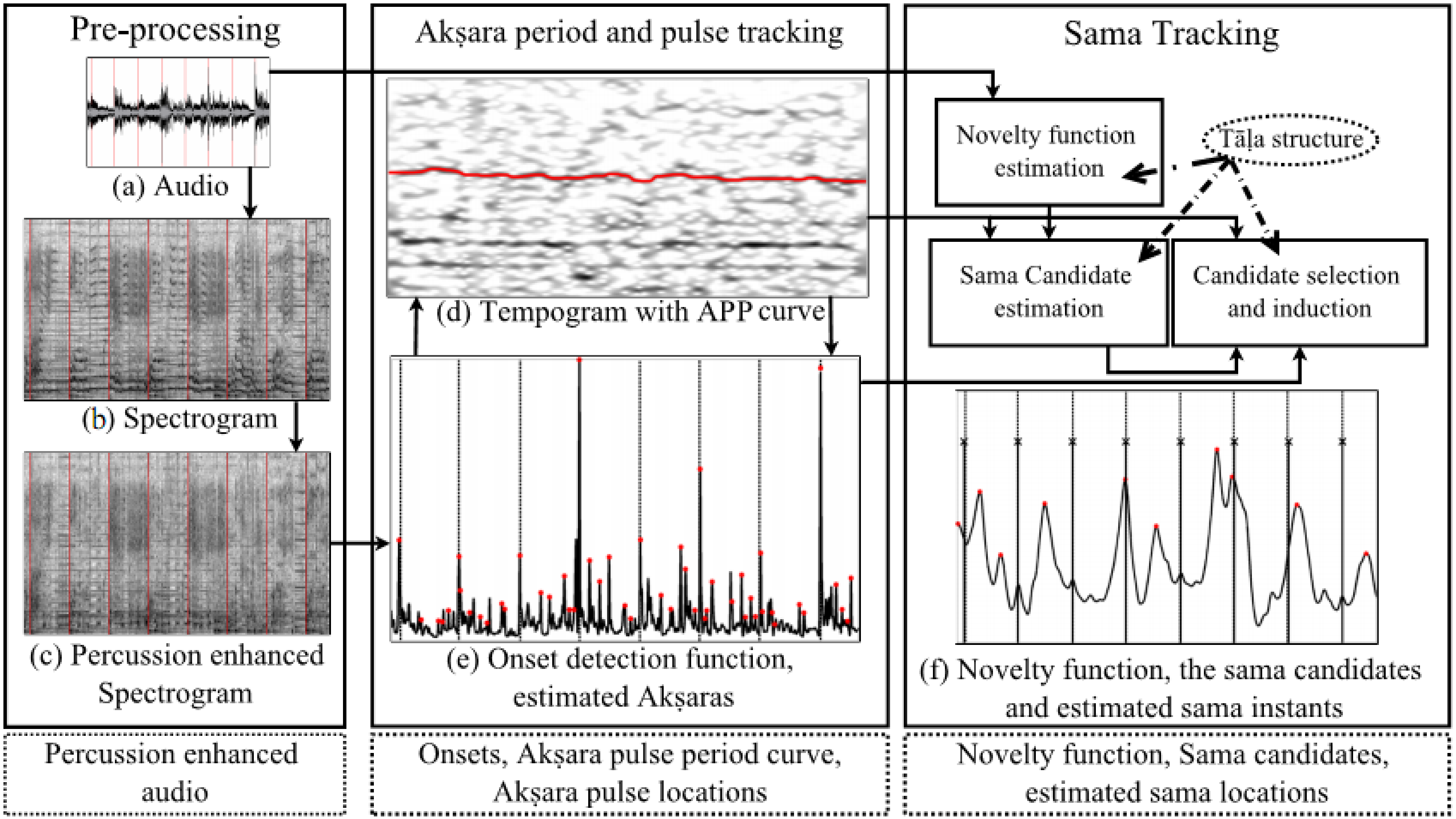
	CML (%)			AML (%)		
Dataset	L_{cb}	L_{ca}	L_{ba}	L_{cb}	L_{ca}	L_{ba}
Carnatic	11.06	8.76	4.15	34.10	45.16	25.81
Hindustani	0.00	25.40	-	45.22	46.50	-

Dataset	OP (%)	STM (%)
Carnatic	41.0	42.2
Hindustani	47.8	51.6

"Cycle length" estimation

Srinivasamurthy, A., Holzapfel, A., & Serra, X. (2014). In Search of Automatic Rhythm Analysis Methods for Turkish and Indian Art Music. Journal of New Music Research. 43(1),97–117

Preliminary approach – dynamic programming



Srinivasamurthy, A., & Serra, X. (2014, May). A Supervised Approach to Hierarchical Metrical Cycle Tracking from Audio Music Recordings. In Proc. of the 39th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2014) (pp. 5237–5241). Florence, Italy.

Preliminary results

- Evaluated on CMR_f dataset
- Accurate tempo tracking
- Poor sama tracking
- **Adhoc approach** – cannot be extended or improved easily

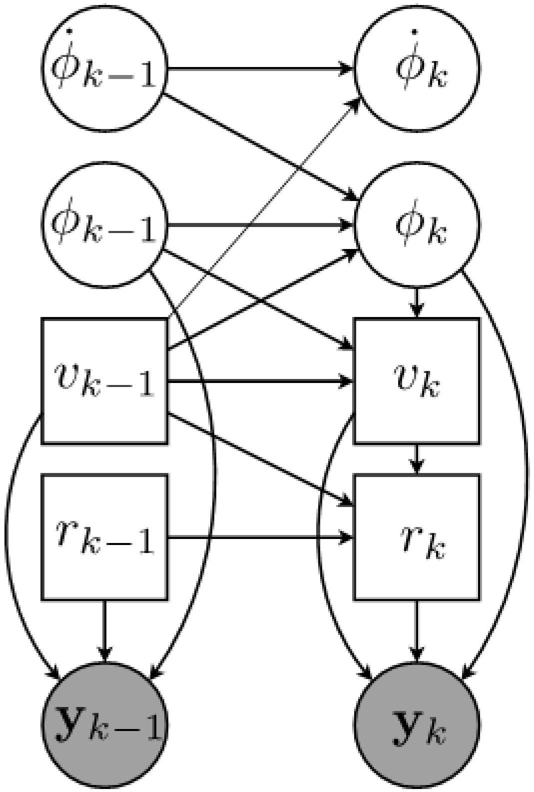
Measure	CML	AML
$\overline{\tau_o}$ estimation	81.2	98.9
τ_o tracking	80.4	96.3

Tempo estimation and tracking

Variant	$\mathfrak{p}$	$\mathfrak{r}$	$\mathfrak{f}$	$\mathfrak{I}$ (bits)	Cand. Accu. (%)
Method-A	0.290	0.190	0.216	1.17	20.46
Method-B	0.246	0.202	0.215	1.25	27.85
RB-1	0.155	0.175	0.137	0.40	-
RB-2	0.228	0.200	0.206	1.11	15.3

Sama tracking

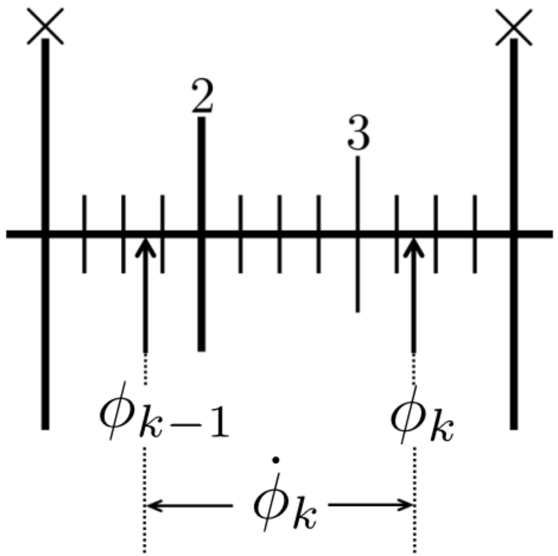
Bayesian models for meter analysis



Bayesian models for meter analysis

- Bayesian models for joint estimation of tempo, beats and downbeats
  - **The bar pointer model**
- Bar/beat length rhythmic patterns
  - Often learnt from data
- Limitation: Restricted to short bar/cycle durations
  - Problems with long metrical cycles

Bar pointer model – idea

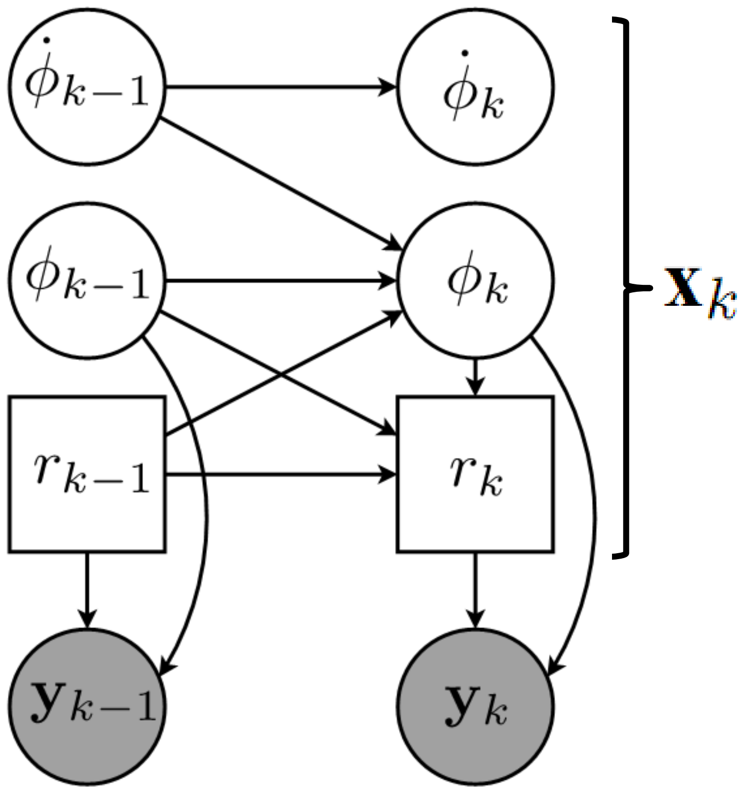


Exaggerated for illustration

- ϕ_k - Position in cycle
- $\dot{\phi}_k$ - Instantaneous tempo

Bar pointer model (BP-model)

- Dynamic Bayesian Network (DBN)

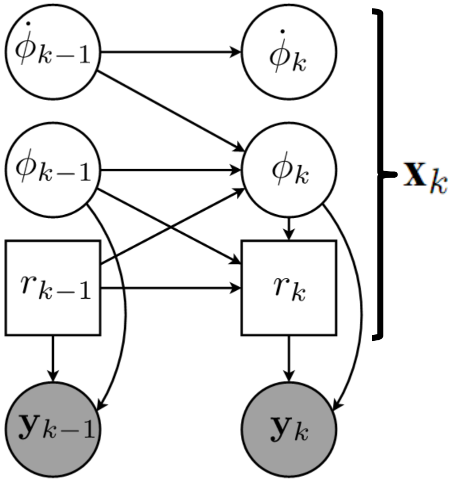


N. Whiteley, A. T. Cemgil, and S. Godsill (2007). Sequential inference of rhythmic structure in musical audio. In Proc. of the 33rd ICASSP, vol. 4, pp. 1321–1325.

Model structure

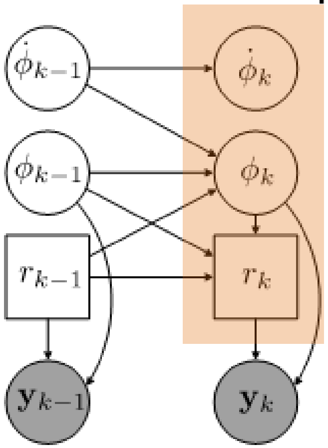
$$P(\mathbf{y}_{1:K}, \mathbf{x}_{0:K}) = P(\mathbf{x}_0) \cdot \prod_{k=1}^{K} P(\mathbf{x}_k|\mathbf{x}_{k-1})\, P(\mathbf{y}_k|\mathbf{x}_k)$$

- Observed feature sequence: $\mathbf{y}_{1:K} = \{\mathbf{y}_1, \ldots, \mathbf{y}_K\}$
- Latent variable sequence: $\mathbf{x}_{1:K} = \{\mathbf{x}_1, \ldots, \mathbf{x}_K\}$
- $P(\mathbf{x}_0)$ - Initial state distribution (uniform)
- $P(\mathbf{x}_k|\mathbf{x}_{k-1})$ - Transition model
- $P(\mathbf{y}_k|\mathbf{x}_k)$- Observation model



F. Krebs, A. Holzapfel, A. T. Cemgil, and G. Widmer (2015). Inferring metrical structure in music using particle filters. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 23(5):817–827.

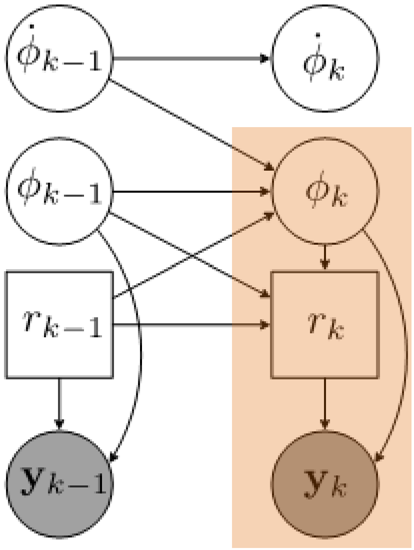
Latent variables



- State of a hypothetical bar pointer
- Position in cycle: $\phi \in [0, M)$
  - M is the length of cycle
- Rhythmic pattern indicator: $r \in \{1, \ldots, R\}$
  - Choose the observation model corresponding to the rhythmic pattern
- Instantaneous tempo: $\dot{\phi}_k \in [\dot{\phi}_{\text{min}}, \dot{\phi}_{\text{max}}]$
  - The rate at which the bar pointer progresses
  - Related to the length of the cycle ($\Delta \cdot M/\dot{\phi}_k$)

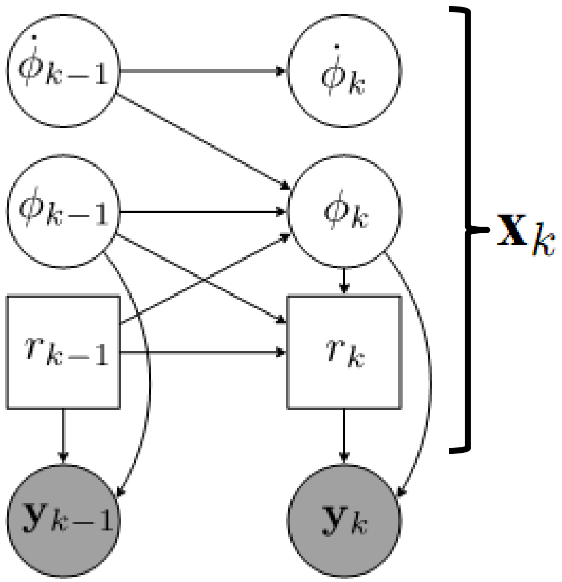
Observation model

- Spectral flux
  - Two frequency bands
- Two component GMM
  - For each position and pattern
  - ML estimates of parameters using annotated data
- State tying
  - 1/64 note, half akṣara



$$P(\mathbf{y} \mid \mathbf{x}) = P(\mathbf{y} \mid \phi, r) = \sum_{i=1}^{2} \pi_{\phi,r,i} \mathcal{N}(\mathbf{y}; \boldsymbol{\mu}_{\phi,r,i}, \boldsymbol{\Sigma}_{\phi,r,i})$$

Transition model



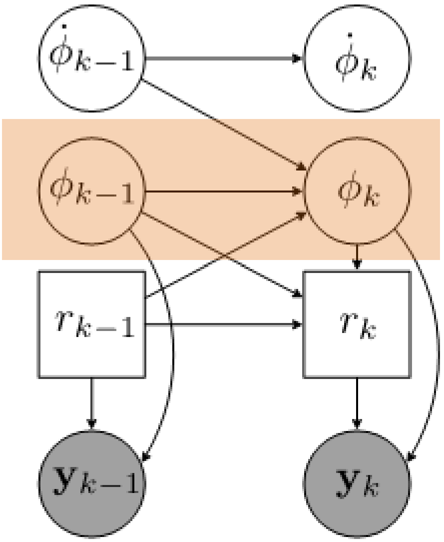
$$P(\mathbf{x}_k|\mathbf{x}_{k-1}) = P(\phi_k|\phi_{k-1}, \dot{\phi}_{k-1}, r_{k-1})P(\dot{\phi}_k|\dot{\phi}_{k-1}) \times P(r_k|r_{k-1}, \phi_k, \phi_{k-1})$$

Transition model

- Cyclic position pointer

$$P(\phi_k|\phi_{k-1}, \dot{\phi}_{k-1}, r_{k-1}) = \mathbb{1}_\phi$$

$$\phi_k = (\phi_{k-1} + \dot{\phi}_{k-1}) \bmod (M_{r_{k-1}})$$



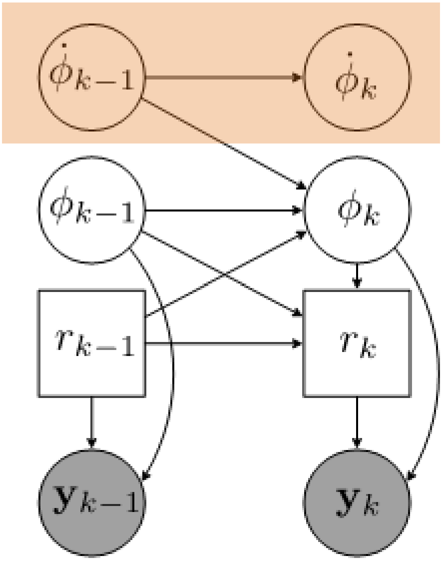
- Tracking: $M(r_k) = M$

$$P(\mathbf{x}_k|\mathbf{x}_{k-1}) = P(\phi_k|\phi_{k-1}, \dot{\phi}_{k-1}, r_{k-1}) P(\dot{\phi}_k|\dot{\phi}_{k-1}) P(r_k|r_{k-1}, \phi_k, \phi_{k-1})$$

Transition model

- Allow tempo changes

$$P(\dot{\phi}_k|\dot{\phi}_{k-1}) \propto \mathcal{N}(\dot{\phi}_{k-1}, \sigma^2_{\dot{\phi}}) \times \mathbb{1}_{\dot{\phi}}$$

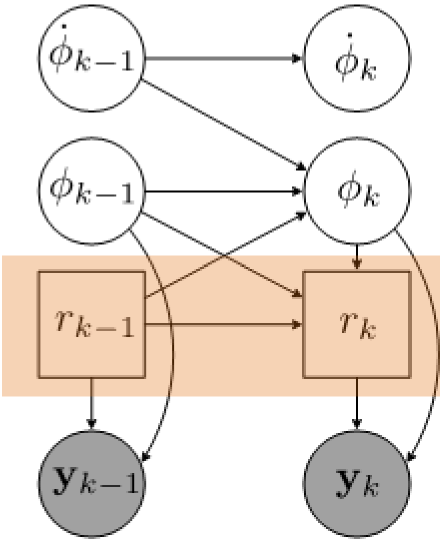


- Normal distribution
  - Within allowed range of tempo: $[\dot{\phi}_{\text{min}}, \dot{\phi}_{\text{max}}]$

$$P(\mathbf{x}_k|\mathbf{x}_{k-1}) = P(\phi_k|\phi_{k-1}, \dot{\phi}_{k-1}, r_{k-1}) P(\dot{\phi}_k|\dot{\phi}_{k-1}) P(r_k|r_{k-1}, \phi_k, \phi_{k-1})$$

Transition model

- Cycle length patterns
  - Capture diversity of patterns
  - Learnt from data
- Pattern transitions
  - Allow transitions only at end of cycle
- Time homogeneity in pattern transitions



$$P(r_k \mid r_{k-1}, \phi_k, \phi_{k-1}) = \begin{cases} \mathbb{A}(r_{k-1}, r_k) & \text{if } \phi_k < \phi_{k-1} \\ \mathbb{1}_r & \text{else} \end{cases}$$

$$P(\mathbf{x}_k|\mathbf{x}_{k-1}) = P(\phi_k|\phi_{k-1}, \dot{\phi}_{k-1}, r_{k-1}) P(\dot{\phi}_k|\dot{\phi}_{k-1}) P(r_k|r_{k-1}, \phi_k, \phi_{k-1})$$

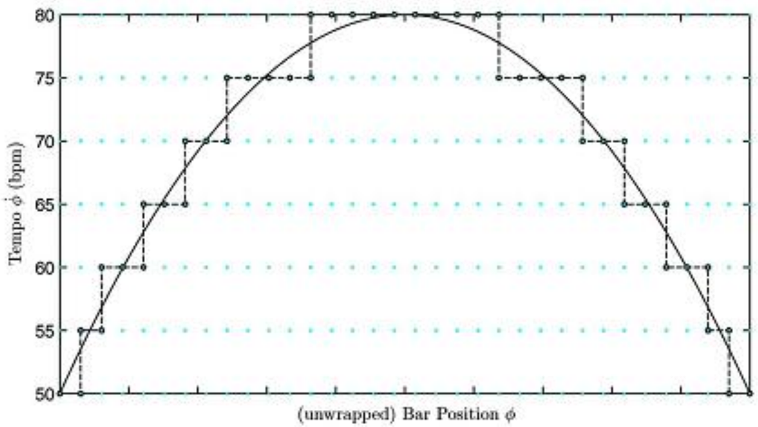
Inference

- A sequence $\mathbf{x}^*_{1:K}$ that maximizes the posterior $P(\mathbf{x}_{1:K}|\mathbf{y}_{1:K})$
- $\mathbf{x}^*_{1:K}$ can then be translated into
  - Sama instants $\mathcal{S}_z = \{t^*_k \mid \phi^*_k = 0\}$
  - Beat instants $\mathcal{B}_z = \{t^*_k \mid \phi^*_k = i \cdot {}^{M_{r^*}}/{}_{B_{r^*}}, i = 1, \ldots, B_r\}$
  - Instantaneous tempo track $(\dot{\phi}^*_k)$
- Two inference schemes
  - Viterbi algorithm
  - Particle filters

Hidden Markov model

- Discrete state space approximation of the model
- $\dot{\phi} : n \in \{n_{\min}, n_{\min}+1, n_{\min}+2, \cdots, N-1, N\}$
- $\phi : m \in \{1, 2, \ldots, \lceil M_r \rceil\}$

$$P(n_k \mid n_{k-1}) = \begin{cases} 1 - p_n & \text{if } n_k = n_{k-1} \\ \frac{p_n}{2} & \text{if } n_k = n_{k-1} \pm 1 \\ 0 & \text{otherwise} \end{cases}$$



- Viterbi algorithm to infer the sequence $\mathbf{x}^*_{1:K}$
- Good performance needs a fine grid
  - State space size: M.N.R; Large state spaces: scaling issues

Image from F. Krebs, A. Holzapfel, A. T. Cemgil, and G. Widmer (2015). Inferring metrical structure in music using particle filters. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 23(5):817–827.

Particle filters

- Computing $P(\mathbf{x}_{1:K}|\mathbf{y}_{1:K})$ is often intractable, but can be estimated point-wise
- Approximate using a weighted set of points (particles) in the state space

$$P(\mathbf{x}_{1:K}|\mathbf{y}_{1:K}) \approx \sum_{i=1}^{N_p} w_K^{(i)} \delta(\mathbf{x}_{1:K} - \mathbf{x}_{1:K}^{(i)})$$

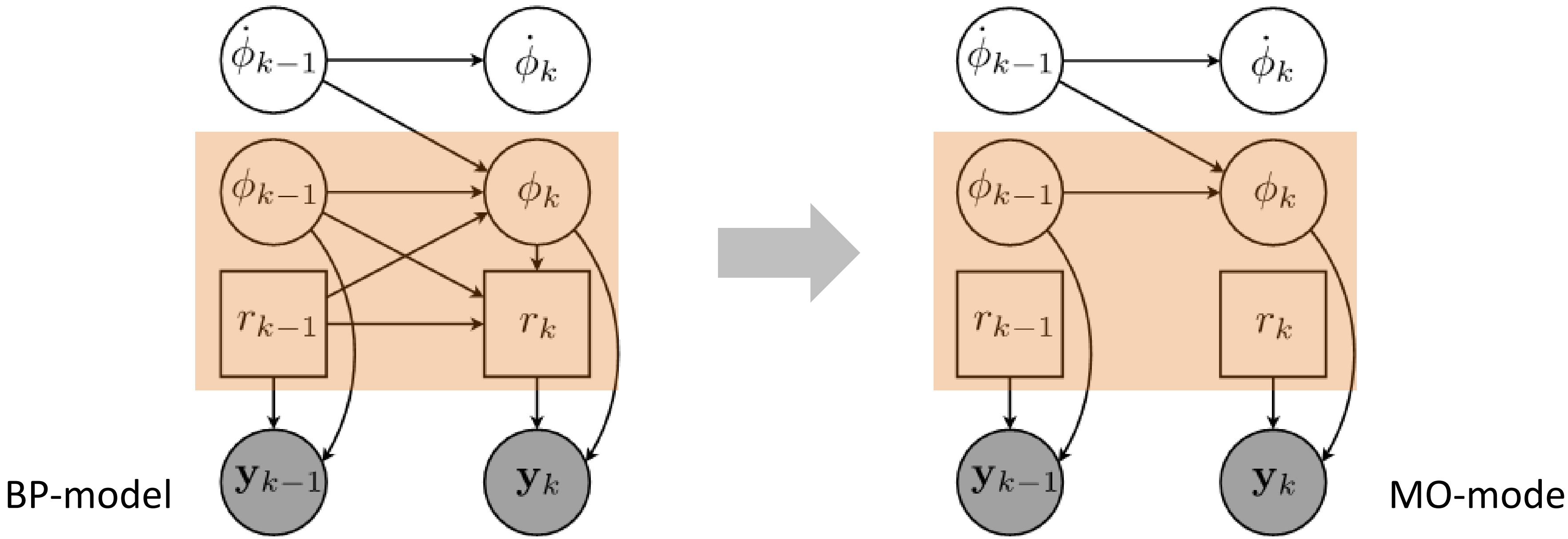
- Sample $\mathbf{x}_k^{(i)}$ and compute weights $w_k^{(i)}$ recursively at each time step
- **Auxiliary Mixture Particle Filter (AMPF)**

BP-model: Novel contributions

- Extended and evaluated on Indian art music
- Multiple rhythm patterns evaluated
- Novel extensions - both at model and inference level
  - Mixture Observation model (MO-model)
  - Section Pointer model (SP-model)
  - Inference extensions

Model extensions: MO-model

- Capture diversity of patterns + faster meter tracking
- Mixture observation model: marginalize over all patterns



$$P(\mathbf{x}_k|\mathbf{x}_{k-1}) = P(\phi_k|\phi_{k-1},\dot{\phi}_{k-1})P(\dot{\phi}_k|\dot{\phi}_{k-1})$$

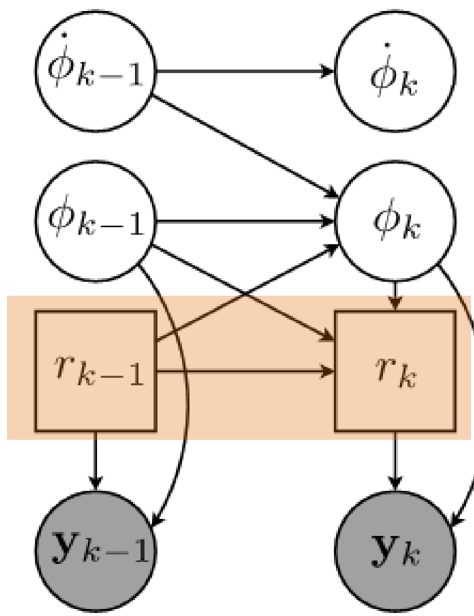
$$P(\mathbf{y} \mid \mathbf{x}) \propto \sum_{j=1}^{R} P(\mathbf{y} \mid \phi, r = j)$$

Srinivasamurthy, A., Holzapfel, A., Cemgil, A. T., & Serra, X.(2015, October). Particle Filters for Efficient Meter Tracking with Dynamic Bayesian Networks. In Proc. of the 16th International Society for Music Information Retrieval Conference (ISMIR 2015) (pp. 197–203). Malaga, Spain.

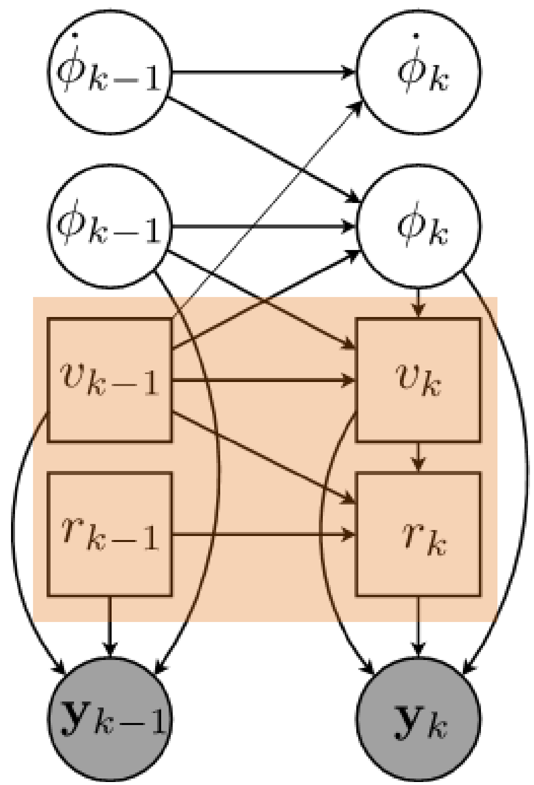
Model extensions: SP-model

compmusic

- Section Pointer model – Generalization of BP-model
  - Track meaningful and shorter sections
  - Section length patterns



BP-model



SP-model

Srinivasamurthy, A., Holzapfel, A., Cemgil, A. T., & Serra, X. (2016, March). A generalized Bayesian model for tracking long metrical cycles in acoustic music signals. In Proceedings of the 41st IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2016) (pp. 76–80). Shanghai, China.

Inference extensions to BP-model

- End-of-bar pattern sampling ($AMPF_e$)
  - Defer decision of sampling rhythm pattern to the end of cycle
- Faster Inference
  - Peak Hop inference ($AMPF_p$) – up to 10x faster
  - Onset gated weight update ($AMPF_g$)

Inference extensions to BP-model

- End-of-bar pattern sampling ($AMPF_e$)
  - Defer decision of sampling rhythm pattern to the end of cycle
- Faster Inference
  - Peak Hop inference ($AMPF_p$) – up to 10x faster
  - Onset gated weight update ($AMPF_g$)

Summary of models

Acronym	Model	Inference algorithm	Meter Analysis Inference	Meter Analysis Tracking
§HMM_0	BP-model†	Viterbi algorithm	✓	✓
§AMPF_0	BP-model	AMPF	✓	✓
*HMM_m	MO-model†	Viterbi algorithm	×	✓
*AMPF_m	MO-model	AMPF	×	✓
*HMM_s	SP-model†	Viterbi algorithm	✓	✓
*AMPF_s	SP-model	AMPF	✓	✓
*AMPF_e	BP-model	AMPF with end-of-bar pattern sampling	×	✓
*AMPF_p	BP-model	Peak hop inference in AMPF	✓	✓
*AMPF_g	BP-model	Onset gated weight update in AMPF	✓	✓

compmusic

Summary of models

Acronym	Model	Inference algorithm	Meter Analysis Inference	Meter Analysis Tracking
§HMM_0	BP-model†	Viterbi algorithm	✓	✓
§AMPF_0	BP-model	AMPF	✓	✓
*HMM_m	MO-model†	Viterbi algorithm	×	✓
*AMPF_m	MO-model	AMPF	×	✓
*HMM_s	SP-model†	Viterbi algorithm	✓	✓
*AMPF_s	SP-model	AMPF	✓	✓
*AMPF_e	BP-model	AMPF with end-of-bar pattern sampling	×	✓
*AMPF_p	BP-model	Peak hop inference in AMPF	✓	✓
*AMPF_g	BP-model	Onset gated weight update in AMPF	✓	✓

Summary of models

Acronym	Model	Inference algorithm	Meter Analysis Inference	Meter Analysis Tracking
§HMM_0	BP-model†	Viterbi algorithm	✓	✓
§AMPF_0	BP-model	AMPF	✓	✓
*HMM_m	MO-model†	Viterbi algorithm	×	✓
*AMPF_m	MO-model	AMPF	×	✓
*HMM_s	SP-model†	Viterbi algorithm	✓	✓
*AMPF_s	SP-model	AMPF	✓	✓
*AMPF_e	BP-model	AMPF with end-of-bar pattern sampling	×	✓
*AMPF_p	BP-model	Peak hop inference in AMPF	✓	✓
*AMPF_g	BP-model	Onset gated weight update in AMPF	✓	✓

Summary of models

Acronym	Model	Inference algorithm	Meter Analysis Inference	Meter Analysis Tracking
§HMM_0	BP-model†	Viterbi algorithm	✓	✓
§AMPF_0	BP-model	AMPF	✓	✓
*HMM_m	MO-model†	Viterbi algorithm	×	✓
*AMPF_m	MO-model	AMPF	×	✓
*HMM_s	SP-model†	Viterbi algorithm	✓	✓
*AMPF_s	SP-model	AMPF	✓	✓
*AMPF_e	BP-model	AMPF with end-of-bar pattern sampling	×	✓
*AMPF_p	BP-model	Peak hop inference in AMPF	✓	✓
*AMPF_g	BP-model	Onset gated weight update in AMPF	✓	✓

Experiments and results

Experiments: Goals

- Meter Inference, tracking and informed tracking
  - For both Carnatic and Hindustani music
- Compare performance across models for Indian music datasets
- Identify limitations of the models and challenges to meter analysis with Bayesian models

Experimental setup (1)

- Music culture known
  - Analysis done separately on each culture
- Results reported only on CMR, HMR_l, HMR_s datasets
  - Hindustani music: also lay informed
- f-measure, InfoGain, AML_t for beats, f-measure for sama/downbeats
  - 70 ms error margin
  - 6.25% of median IAI: margin for HMR_l dataset sama tracking
- Tracking and inference – tempo tracking
  - 5% margin, CML and AML

Experimental setup (2)

- Results reported for single pattern per cycle/section case
  - Average performance over all pieces over three trials
- Tempo range fixed for each dataset *a priori*
- AMPF
  - 1500 particles per rhythm pattern

Meter inference

	Algo.	$\mathfrak{f}_b$	$\text{AML}_{t,b}$	$\mathfrak{I}_b$	$\mathfrak{f}_s$	Tempo		Tāḷa
				Bits		CML	AML	%
CMR	HMM_0	0.718	0.722	1.44	0.440	0.718	0.938	64
	AMPF_0	0.825	0.906	2.17	0.574	0.802	1.000	68
HMR_s	HMM_0	0.759	0.698	1.21	0.551	0.533	0.721	60
	AMPF_0	0.828	0.834	1.54	0.569	0.714	0.946	63
HMR_l	HMM_0	0.338	0.225	0.77	0.280	0.119	0.350	37
	AMPF_0	0.390	0.427	1.35	0.268	0.350	0.740	27
Blrm.	HMM_0	0.853	0.910	2.52	0.666	0.755	0.988	91
	AMPF_0	0.813	0.850	2.15	0.529	0.709	0.957	89

Meter inference

compmusic

Beat f-measure

	Algo.	$\mathfrak{f}_b$	$\text{AML}_{t,b}$	$\mathfrak{I}_b$ Bits	$\mathfrak{f}_s$	Tempo CML	Tempo AML	Tāḷa %
CMR	HMM_0	0.718	0.722	1.44	0.440	0.718	0.938	64
	AMPF_0	0.825	0.906	2.17	0.574	0.802	1.000	68
HMR_s	HMM_0	0.759	0.698	1.21	0.551	0.533	0.721	60
	AMPF_0	0.828	0.834	1.54	0.569	0.714	0.946	63
HMR_l	HMM_0	0.338	0.225	0.77	0.280	0.119	0.350	37
	AMPF_0	0.390	0.427	1.35	0.268	0.350	0.740	27
Blrm.	HMM_0	0.853	0.910	2.52	0.666	0.755	0.988	91
	AMPF_0	0.813	0.850	2.15	0.529	0.709	0.957	89

Meter inference

compmusic

Beat AML_t

	Algo.	f_b	$AML_{t,b}$	$\mathfrak{I}_b$ Bits	f_s	Tempo CML	Tempo AML	Tāḷa %
CMR	HMM_0	0.718	0.722	1.44	0.440	0.718	0.938	64
	$AMPF_0$	0.825	0.906	2.17	0.574	0.802	1.000	68
HMR_s	HMM_0	0.759	0.698	1.21	0.551	0.533	0.721	60
	$AMPF_0$	0.828	0.834	1.54	0.569	0.714	0.946	63
HMR_l	HMM_0	0.338	0.225	0.77	0.280	0.119	0.350	37
	$AMPF_0$	0.390	0.427	1.35	0.268	0.350	0.740	27
Blrm.	HMM_0	0.853	0.910	2.52	0.666	0.755	0.988	91
	$AMPF_0$	0.813	0.850	2.15	0.529	0.709	0.957	89

Sec 5.4.2

Meter inference

compmusic

Beat infoGain

	Algo.	f_b	$\mathrm{AML}_{t,b}$	$\mathfrak{I}_b$ Bits	f_s	Tempo CML	Tempo AML	Tāḷa %
CMR	HMM_0	0.718	0.722	1.44	0.440	0.718	0.938	64
CMR	AMPF_0	0.825	0.906	2.17	0.574	0.802	1.000	68
HMR_s	HMM_0	0.759	0.698	1.21	0.551	0.533	0.721	60
HMR_s	AMPF_0	0.828	0.834	1.54	0.569	0.714	0.946	63
HMR_l	HMM_0	0.338	0.225	0.77	0.280	0.119	0.350	37
HMR_l	AMPF_0	0.390	0.427	1.35	0.268	0.350	0.740	27
Blrm.	HMM_0	0.853	0.910	2.52	0.666	0.755	0.988	91
Blrm.	AMPF_0	0.813	0.850	2.15	0.529	0.709	0.957	89

Sec 5.4.2

Meter inference

compmusic

Downbeat/sama f-measure

	Algo.	f_b	$AML_{t,b}$	$\mathfrak{I}_b$ Bits	f_s	Tempo CML	Tempo AML	Tāḷa %
CMR	HMM_0	0.718	0.722	1.44	0.440	0.718	0.938	64
CMR	$AMPF_0$	0.825	0.906	2.17	0.574	0.802	1.000	68
HMR_s	HMM_0	0.759	0.698	1.21	0.551	0.533	0.721	60
HMR_s	$AMPF_0$	0.828	0.834	1.54	0.569	0.714	0.946	63
HMR_l	HMM_0	0.338	0.225	0.77	0.280	0.119	0.350	37
HMR_l	$AMPF_0$	0.390	0.427	1.35	0.268	0.350	0.740	27
Blrm.	HMM_0	0.853	0.910	2.52	0.666	0.755	0.988	91
Blrm.	$AMPF_0$	0.813	0.850	2.15	0.529	0.709	0.957	89

Meter inference

compmusic

- Tāla classification – better on short duration cycles

	Algo.	f_b	$AML_{t,b}$	$\mathfrak{I}_b$ Bits	f_s	Tempo CML	Tempo AML	Tāḷa %
CMR	HMM_0	0.718	0.722	1.44	0.440	0.718	0.938	64
CMR	$AMPF_0$	0.825	0.906	2.17	0.574	0.802	1.000	68
HMR_s	HMM_0	0.759	0.698	1.21	0.551	0.533	0.721	60
HMR_s	$AMPF_0$	0.828	0.834	1.54	0.569	0.714	0.946	63
HMR_l	HMM_0	0.338	0.225	0.77	0.280	0.119	0.350	37
HMR_l	$AMPF_0$	0.390	0.427	1.35	0.268	0.350	0.740	27
Blrm.	HMM_0	0.853	0.910	2.52	0.666	0.755	0.988	91
Blrm.	$AMPF_0$	0.813	0.850	2.15	0.529	0.709	0.957	89

Sec 5.4.2

Meter inference

- Poor performance on Hindustani long cycle subset

	Algo.	$\mathfrak{f}_b$	$\mathrm{AML}_{t,b}$	$\mathfrak{I}_b$ Bits	$\mathfrak{f}_s$	Tempo CML	Tempo AML	Tāḷa %
CMR	HMM_0	0.718	0.722	1.44	0.440	0.718	0.938	64
CMR	AMPF_0	0.825	0.906	2.17	0.574	0.802	1.000	68
HMR_s	HMM_0	0.759	0.698	1.21	0.551	0.533	0.721	60
HMR_s	AMPF_0	0.828	0.834	1.54	0.569	0.714	0.946	63
HMR_l	HMM_0	0.338	0.225	0.77	0.280	0.119	0.350	37
HMR_l	AMPF_0	0.390	0.427	1.35	0.268	0.350	0.740	27
Blrm.	HMM_0	0.853	0.910	2.52	0.666	0.755	0.988	91
Blrm.	AMPF_0	0.813	0.850	2.15	0.529	0.709	0.957	89

Meter inference

- Better performance on Ballroom dataset than Indian music dataset

	Algo.	$\mathfrak{f}_b$	$\text{AML}_{t,b}$	$\mathfrak{I}_b$ Bits	$\mathfrak{f}_s$	Tempo CML	Tempo AML	Tāḷa %
CMR	HMM_0	0.718	0.722	1.44	0.440	0.718	0.938	64
CMR	AMPF_0	0.825	0.906	2.17	0.574	0.802	1.000	68
HMR_s	HMM_0	0.759	0.698	1.21	0.551	0.533	0.721	60
HMR_s	AMPF_0	0.828	0.834	1.54	0.569	0.714	0.946	63
HMR_l	HMM_0	0.338	0.225	0.77	0.280	0.119	0.350	37
HMR_l	AMPF_0	0.390	0.427	1.35	0.268	0.350	0.740	27
Blrm.	HMM_0	0.853	0.910	2.52	0.666	0.755	0.988	91
Blrm.	AMPF_0	0.813	0.850	2.15	0.529	0.709	0.957	89

Meter inference

- Sama tracking poorer than beat tracking

	Algo.	$\mathfrak{f}_b$	$\mathrm{AML}_{t,b}$	$\mathfrak{I}_b$ Bits	$\mathfrak{f}_s$	Tempo CML	Tempo AML	Tāḷa %
CMR	HMM_0	0.718	0.722	1.44	0.440	0.718	0.938	64
	AMPF_0	0.825	0.906	2.17	0.574	0.802	1.000	68
HMR_s	HMM_0	0.759	0.698	1.21	0.551	0.533	0.721	60
	AMPF_0	0.828	0.834	1.54	0.569	0.714	0.946	63
HMR_l	HMM_0	0.338	0.225	0.77	0.280	0.119	0.350	37
	AMPF_0	0.390	0.427	1.35	0.268	0.350	0.740	27
Blrm.	HMM_0	0.853	0.910	2.52	0.666	0.755	0.988	91
	AMPF_0	0.813	0.850	2.15	0.529	0.709	0.957	89

Meter tracking: BP-model

- Baseline – tracking performance better than inference
- HMM and AMPF equivalent performance

	Algo.	$\mathfrak{f}_b$	$\mathrm{AML}_{t,b}$	$\mathfrak{I}_b$ Bits	$\mathfrak{f}_s$	Tempo CML	Tempo AML
CMR	HMM_0	0.784	0.771	1.59	0.624	0.890	0.915
	AMPF_0	0.827	0.840	1.97	0.671	0.955	0.997
HMR_s	HMM_0	0.835	0.796	1.39	0.733	0.663	0.830
	AMPF_0	0.884	0.858	1.64	0.772	0.844	0.964
HMR_l	HMM_0	0.353	0.305	0.86	0.429	0.294	0.435
	AMPF_0	0.374	0.513	1.40	0.396	0.390	0.610
Blrm.	HMM_0	0.929	0.921	2.78	0.821	0.987	0.989
	AMPF_0	0.909	0.895	2.56	0.735	0.98	0.98

Meter tracking: MO-model ($AMPF_m$)

- R=1, equivalent to $AMPF_0$
- Performance equivalent to BP-model

Dataset	$\mathfrak{f}_b$	$AML_{t,b}$	$\mathfrak{I}_b$ Bits	$\mathfrak{f}_s$	Tempo CML	Tempo AML
CMR	0.838	0.840	1.96	0.671	0.958	1.00
HMR_s	0.886	0.864	1.66	0.783	0.837	0.942
HMR_l	0.364	0.506	1.39	0.455	0.401	0.554
Ballroom	0.907	0.895	2.56	0.730	0.981	0.981

Meter tracking: SP-model ($AMPF_s$)

compmusic

- Ballroom dataset: No musically well defined sections
  - Evaluation only on Indian music datasets
- Sama tracking improves with SP-model
  - Beat tracking also improves, but to a lower extent

Dataset	$\mathfrak{f}_b$	$\mathrm{AML}_{t,b}$	$\mathfrak{I}_b$ Bits	$\mathfrak{f}_s$	Tempo CML	Tempo AML
CMR	0.868	0.879	2.16	0.717	0.958	1.00
$\mathrm{HMR_s}$	0.924	0.890	1.88	0.850	0.855	0.971
$\mathrm{HMR_l}$	0.414	0.590	1.63	0.509	0.458	0.644

Meter tracking: Inference extensions

compmusic

- No significant improvement over baseline

	Dataset	$\mathfrak{f}_b$	$\mathrm{AML}_{\mathrm{t},b}$	$\mathfrak{I}_b$ Bits	$\mathfrak{f}_s$	Tempo CML	Tempo AML
CMR	$\mathrm{AMPF_e}$	0.826	0.842	1.97	0.668	0.958	0.997
	$\mathrm{AMPF_p}$	0.519	0.561	0.67	0.213	0.927	0.969
	$\mathrm{AMPF_g}$	0.756	0.756	1.51	0.580	0.938	0.98
$\mathrm{HMR_s}$	$\mathrm{AMPF_e}$	0.882	0.858	1.64	0.777	0.833	0.935
	$\mathrm{AMPF_p}$	0.655	0.572	0.59	0.273	0.768	0.822
	$\mathrm{AMPF_g}$	0.821	0.653	1.25	0.653	0.743	0.895
Blrm.	$\mathrm{AMPF_e}$	0.908	0.895	2.56	0.734	0.98	0.98
	$\mathrm{AMPF_p}$	0.631	0.694	1.49	0.322	0.922	0.923
	$\mathrm{AMPF_g}$	0.831	0.815	2.13	0.579	0.939	0.943

Sec 5.4.3

Tempo informed meter tracking

- Tempo update restricted to 10% range around median tempo from ground truth
- Evaluated with BP-model and SP-model on Indian art music datasets

Dataset	Algo.	$\mathfrak{f}_b$	$\mathrm{AML}_{t,b}$	$\mathfrak{I}_b$	$\mathfrak{f}_s$
CMR	AMPF_0	0.899	0.952	2.35	0.792
	AMPF_s	0.898	0.950	2.38	0.814
HMR_s	AMPF_0	0.939	0.941	1.99	0.882
	AMPF_s	0.943	0.943	2.00	0.918
HMR_l	AMPF_0	0.425	0.959	2.76	0.786
	AMPF_s	0.439	0.979	2.83	0.848

Tempo-sama informed meter tracking

- Tempo range restricted + first instance of sama known
  - Semi-manual annotation of large corpora
- Improved performance over tempo informed tracking

Dataset	Algo.	$\mathfrak{f}_b$	$\mathrm{AML}_{\mathrm{t},b}$	$\mathfrak{I}_b$	$\mathfrak{f}_s$
CMR	AMPF_0	0.880	0.943	2.37	0.834
	AMPF_s	0.917	0.946	2.40	0.901
HMR_s	AMPF_0	0.959	0.920	2.02	0.911
	AMPF_s	0.958	0.915	2.01	0.933
HMR_l	AMPF_0	0.530	0.978	2.84	0.99
	AMPF_s	0.542	0.98	2.83	0.99

Summary of results: Indian art music

	Algo.	ID	$\mathfrak{f}_b$	$\text{AML}_{t,b}$	$\mathfrak{I}_b$ Bits	$\mathfrak{f}_s$	Tempo CML	Tempo AML
Inf.	HMM_0	1	0.648	0.605	1.21	0.443	0.51	0.72
	AMPF_0	2	0.730	0.776	1.77	0.505	0.67	0.92
Track	HMM_0	3	0.707	0.677	1.36	0.618	0.67	0.77
	AMPF_0	4	0.747	0.774	1.73	0.645	0.79	0.89
	AMPF_m	5	0.750	0.775	1.73	0.662	0.79	0.88
	AMPF_s	6	0.779	0.817	1.92	0.704	0.81	0.91
t-Tr.	AMPF_0	7	0.809	0.950	2.32	0.822	0.99	0.99
	AMPF_s	8	0.813	0.954	2.35	0.857	1.00	1.00
ts-Tr.	AMPF_0	9	0.830	0.943	2.35	0.896	1.00	1.00
	AMPF_s	10	0.849	0.943	2.36	0.931	1.00	1.00

Sec 5.4.5

Summary of results: Indian art music

compmusic

	Algo.	ID	$\mathfrak{f}_b$	$\mathrm{AML}_{t,b}$	$\mathfrak{I}_b$ Bits	$\mathfrak{f}_s$	Tempo CML	Tempo AML
Inf.	HMM_0	1	0.648	0.605	1.21	0.443	0.51	0.72
	AMPF_0	2	0.730	0.776	1.77	0.505	0.67	0.92
Track	HMM_0	3	0.707	0.677	1.36	0.618	0.67	0.77
	AMPF_0	4	0.747	0.774	1.73	0.645	0.79	0.89
	AMPF_m	5	0.750	0.775	1.73	0.662	0.79	0.88
	AMPF_s	6	0.779	0.817	1.92	0.704	0.81	0.91
t-Tr.	AMPF_0	7	0.809	0.950	2.32	0.822	0.99	0.99
	AMPF_s	8	0.813	0.954	2.35	0.857	1.00	1.00
ts-Tr.	AMPF_0	9	0.830	0.943	2.35	0.896	1.00	1.00
	AMPF_s	10	0.849	0.943	2.36	0.931	1.00	1.00

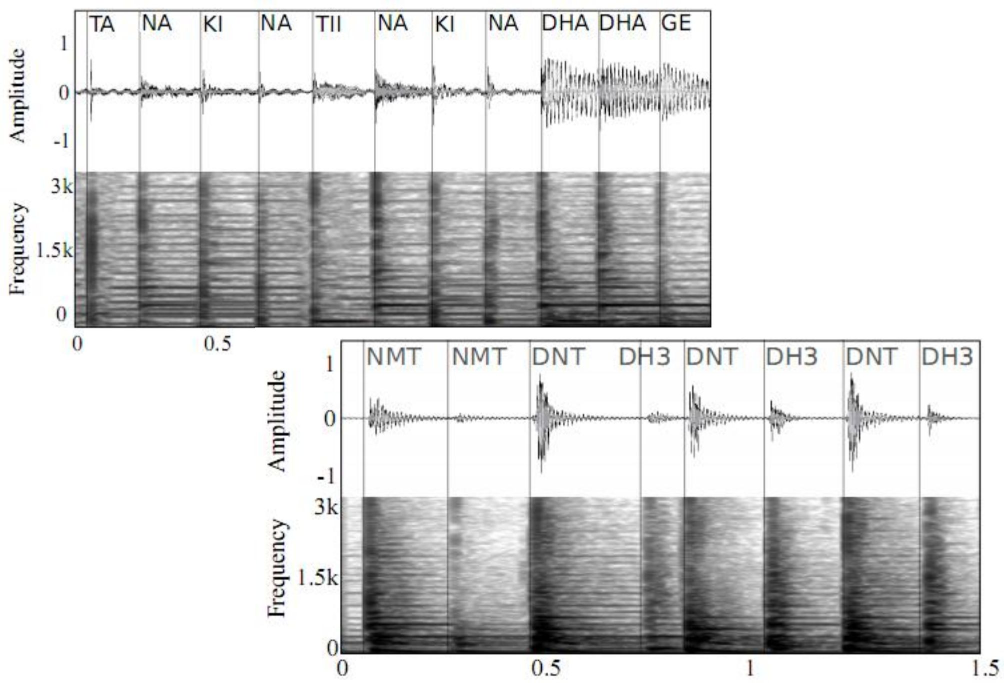
Some observations

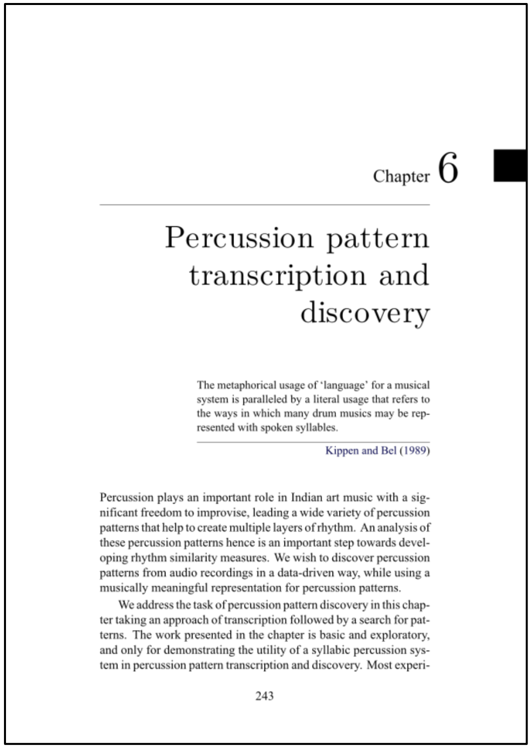
- **Extend to full pieces though tested on two min excerpts**
- Vilaṁbit pieces still difficult to track – need better formulation
  - Longer cycle – more scope for improvisation and difficult to track
- Multiple patterns – no significant improvement
  - need better features
- Inference extensions – no additional improvement
- **Section Pointer model** – generalization of all the state of the art models – most promising

Summary

- Bayesian models – explicitly consider metrical structures: **culture-aware data-driven algorithms**
  - Flexible and generalizable to similar metrical structures
  - Need training data
- Bayesian models can be used to impose additional structures
  - Tighter tempo, allow for human errors (skip a beat, e.g.)
- Models assume that tāla does not change
- Using melody based features in addition
  - Richer representations

Percussion pattern discovery





Chapter 6

Percussion pattern transcription and discovery

The metaphorical usage of 'language' for a musical system is paralleled by a literal usage that refers to the ways in which many drum musics may be represented with spoken syllables.

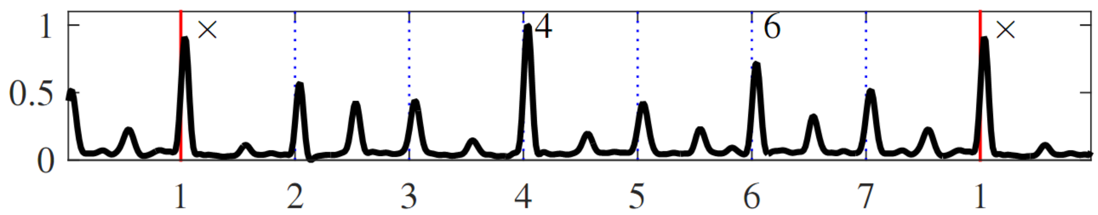
Kippen and Bel (1989)

Percussion plays an important role in Indian art music with a significant freedom to improvise, leading a wide variety of percussion patterns that help to create multiple layers of rhythm. An analysis of these percussion patterns hence is an important step towards developing rhythm similarity measures. We wish to discover percussion patterns from audio recordings in a data-driven way, while using a musically meaningful representation for percussion patterns.

We address the task of percussion pattern discovery in this chapter taking an approach of transcription followed by a search for patterns. The work presented in the chapter is basic and exploratory, and only for demonstrating the utility of a syllabic percussion system in percussion pattern transcription and discovery. Most experi-

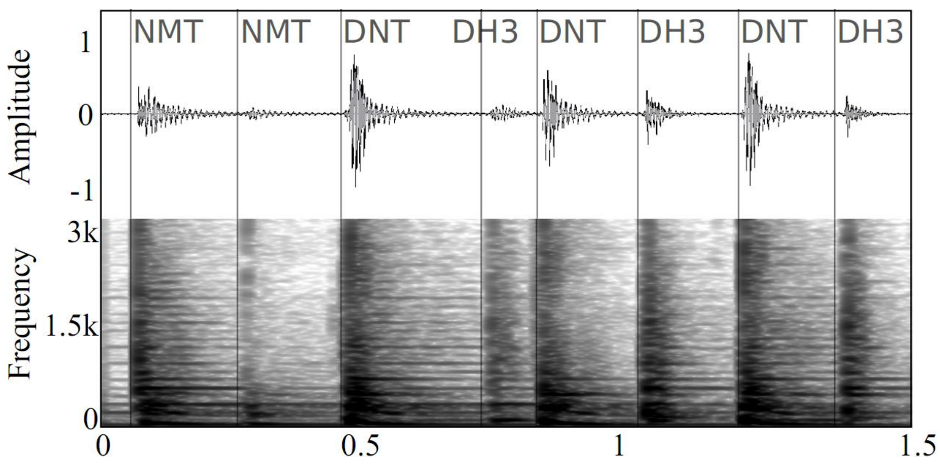
243

Rhythm and percussion patterns



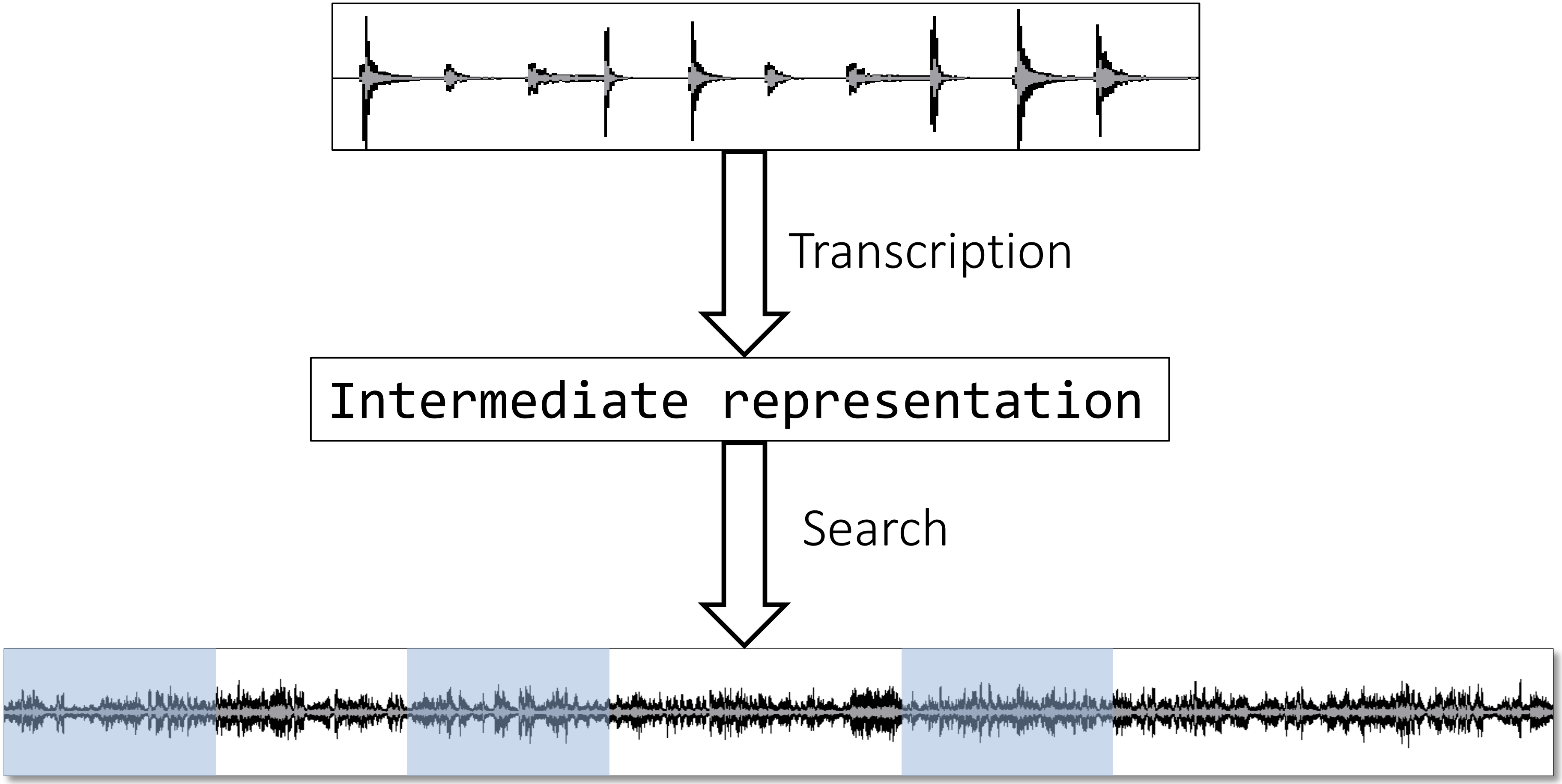
Rhythm pattern

Percussion pattern



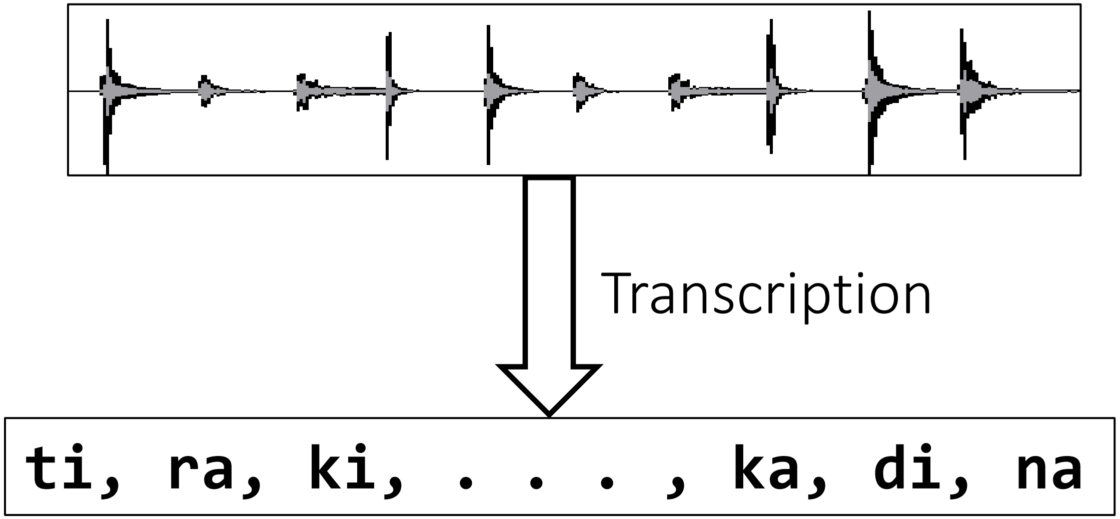
Percussion pattern discovery

- Pattern discovery: inter- and intra-piece



Transcription of percussion patterns

- Percussion patterns into constituent syllables
  - **Sparser but richer representation**
- Patterns – syllable sequences (simplistic!)
  - No timing or dynamics



The case of Indian art music

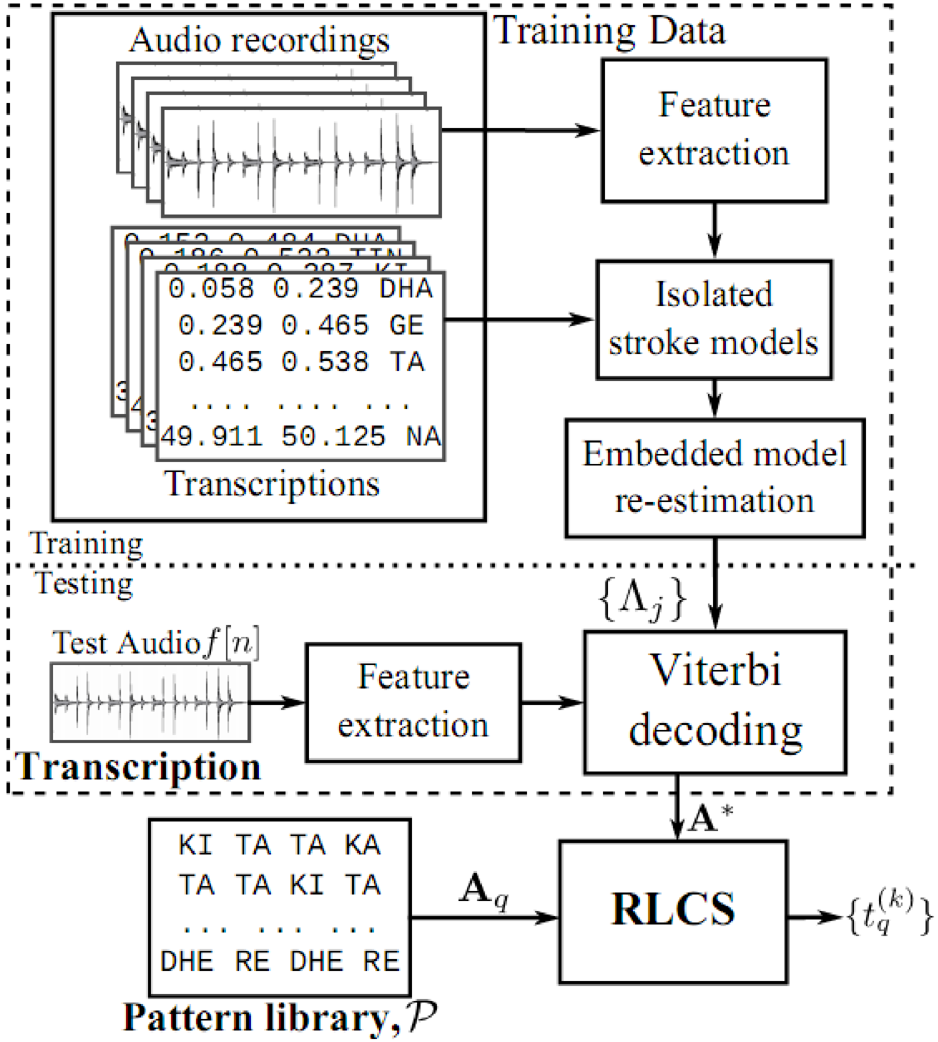
Percussion pattern discovery

compmusic

- Pattern library generation
  - Library of characteristic percussion patterns (query patterns)
  - Score corpus
- Automatic transcription
  - Percussion solo audio recording into a time aligned sequence of syllables
  - Syllable timbre models
- Approximate search
  - Search for the query patterns in the transcribed output syllable sequence
  - Robust to errors in transcription

Pattern discovery: Block diagram

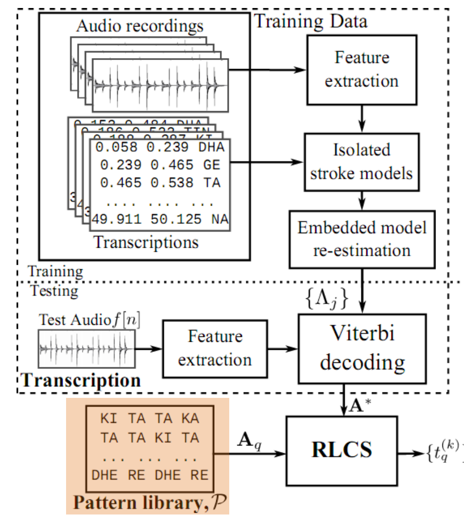
compmusic



Gupta, S., **Srinivasamurthy, A.,** Kumar, M., Murthy, H., & Serra, X. (2015, October). Discovery of Syllabic Percussion Patterns in Tabla Solo Recordings. In Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR 2015) (pp. 385–391). Malaga, Spain.

Pattern library

- Most occurring patterns
- Manual select musically representative patterns
- Lengths: 4, 6, 8, 16

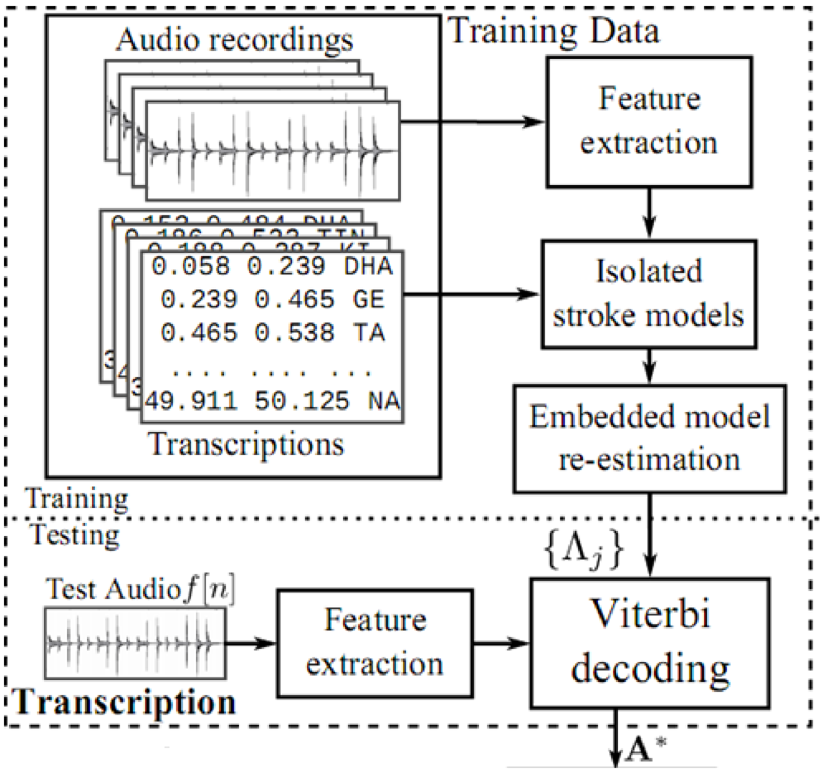


ID	Pattern	L	Count
$\mathbf{A}_1$	DHE, RE, DHE, RE, KI, TA, TA, KI, NA, TA, TA, KI, TA, TA, KI, NA	16	47
$\mathbf{A}_2$	TA, TA, KI, TA, TA, KI, TA, TA, KI, TA, TA, KI, TA, TA, KI, TA	16	10
$\mathbf{A}_3$	TA, KI, TA, TA, KI, TA, TA, KI	8	61
$\mathbf{A}_4$	TA, TA, KI, TA, TA, KI	6	214
$\mathbf{A}_5$	TA, TA, KI, TA	4	379
$\mathbf{A}_6$	KI, TA, TA, KI	4	450
$\mathbf{A}_7$	TA, TA, KI, NA	4	167
$\mathbf{A}_8$	DHA, GE, TA, TA	4	97

ID	Pattern	L	Count
$\mathbf{A}_1$	DH3, TA, DH3, TA, TH, DH3, TH, TA	8	70
$\mathbf{A}_2$	TA, DH3, TA, TH, DH3, TH, TA, TM	8	69
$\mathbf{A}_3$	DH3, TA, DH3, TA, TH, DH3	6	89
$\mathbf{A}_4$	DH3, TA, TH, DH3, TH, TA	6	70
$\mathbf{A}_5$	TA, TH, DH3, TH, TA, TM	6	69
$\mathbf{A}_6$	DH3, TA, TH, DH3	4	291
$\mathbf{A}_7$	DH3, TA, DH3, TA	4	114
$\mathbf{A}_8$	TH, DH3, TA, TH	4	102
$\mathbf{A}_9$	TA, TH, DH3, TH	4	102

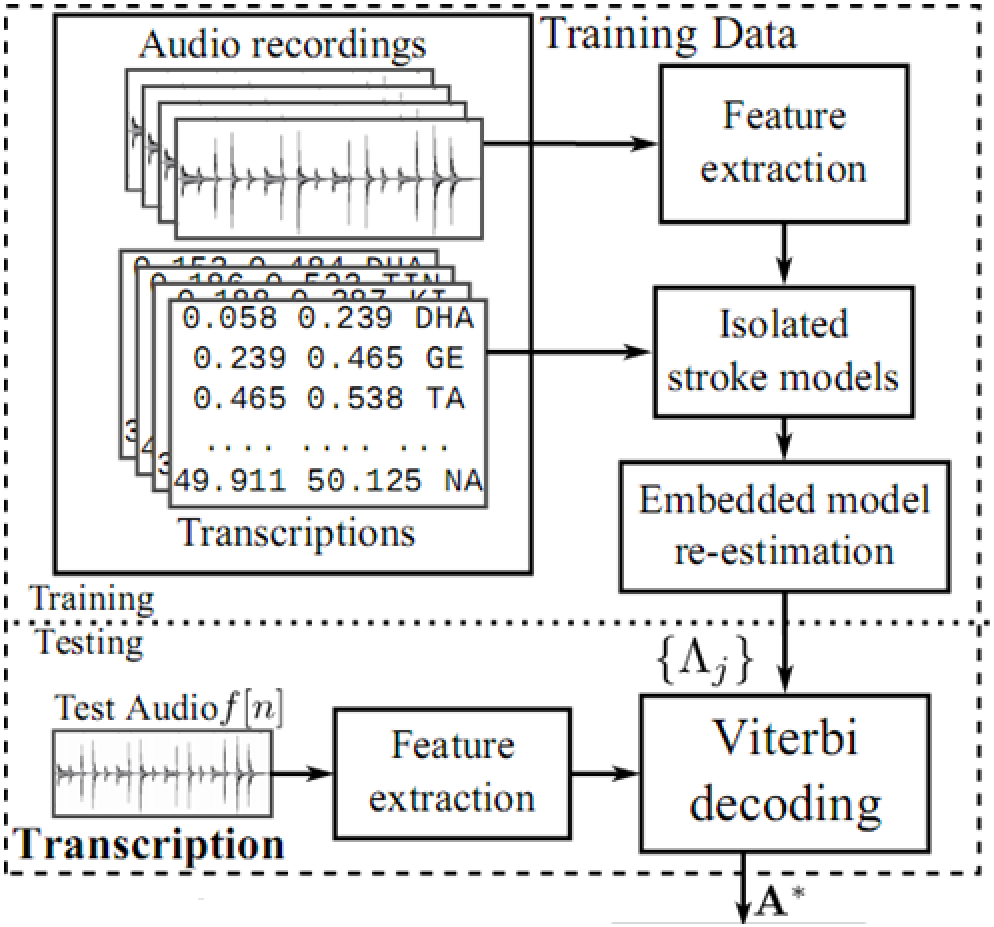
Automatic transcription

- Analogous to limited vocabulary continuous speech recognition with word models
- Syllable HMMs + language model



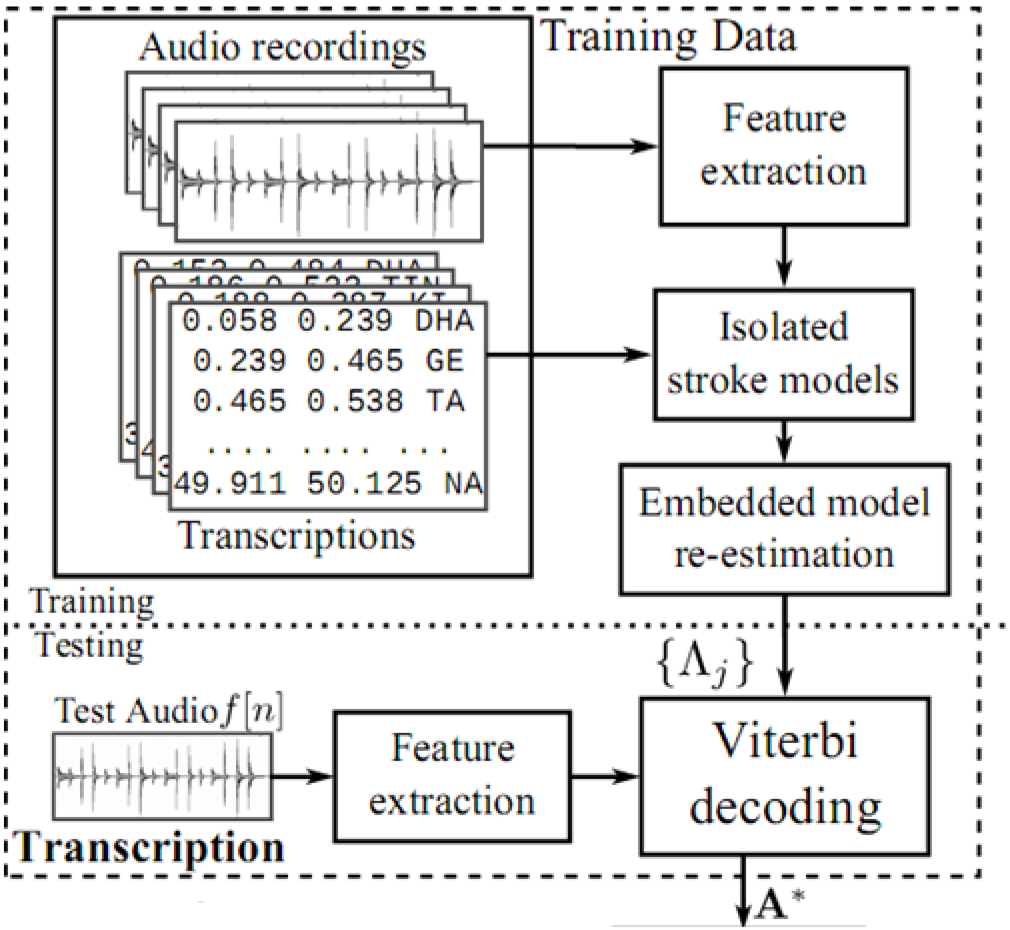
Automatic transcription

Audio + Transcriptions for training



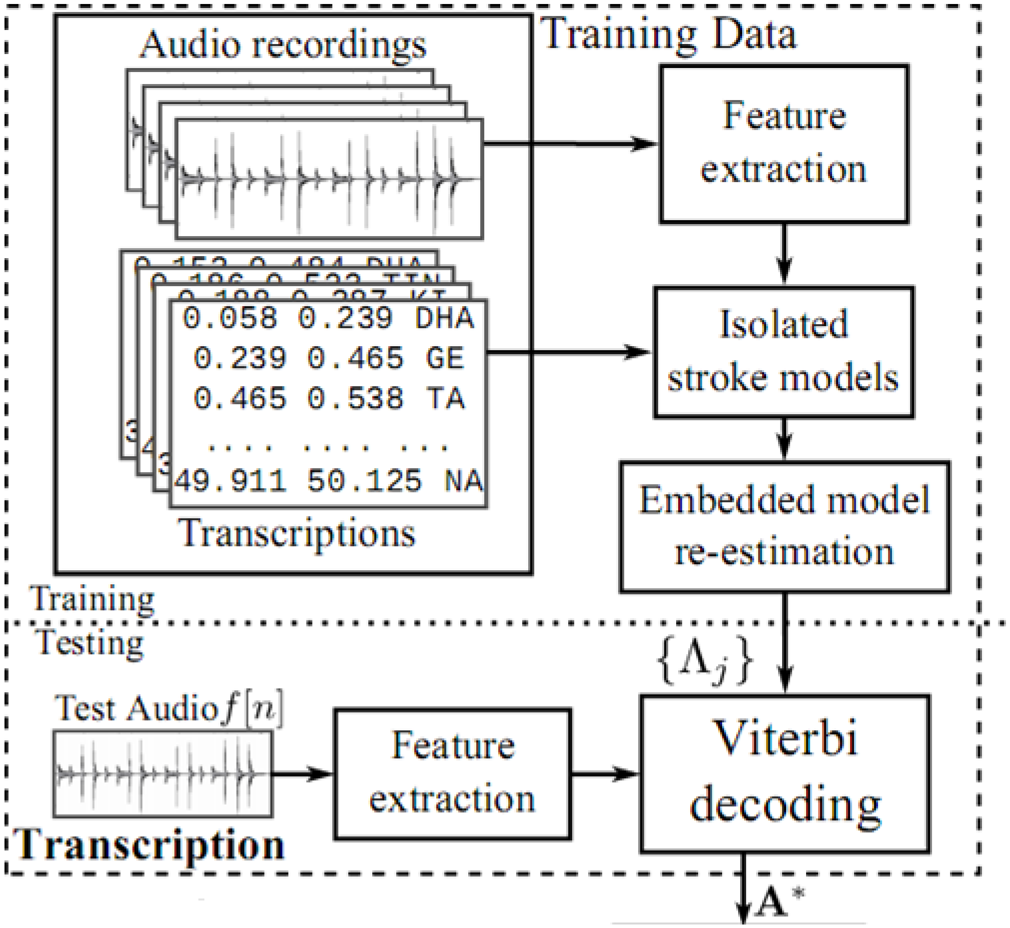
Automatic transcription

compmusic



MFCC_0_D_A
MFCC_D_A

Automatic transcription

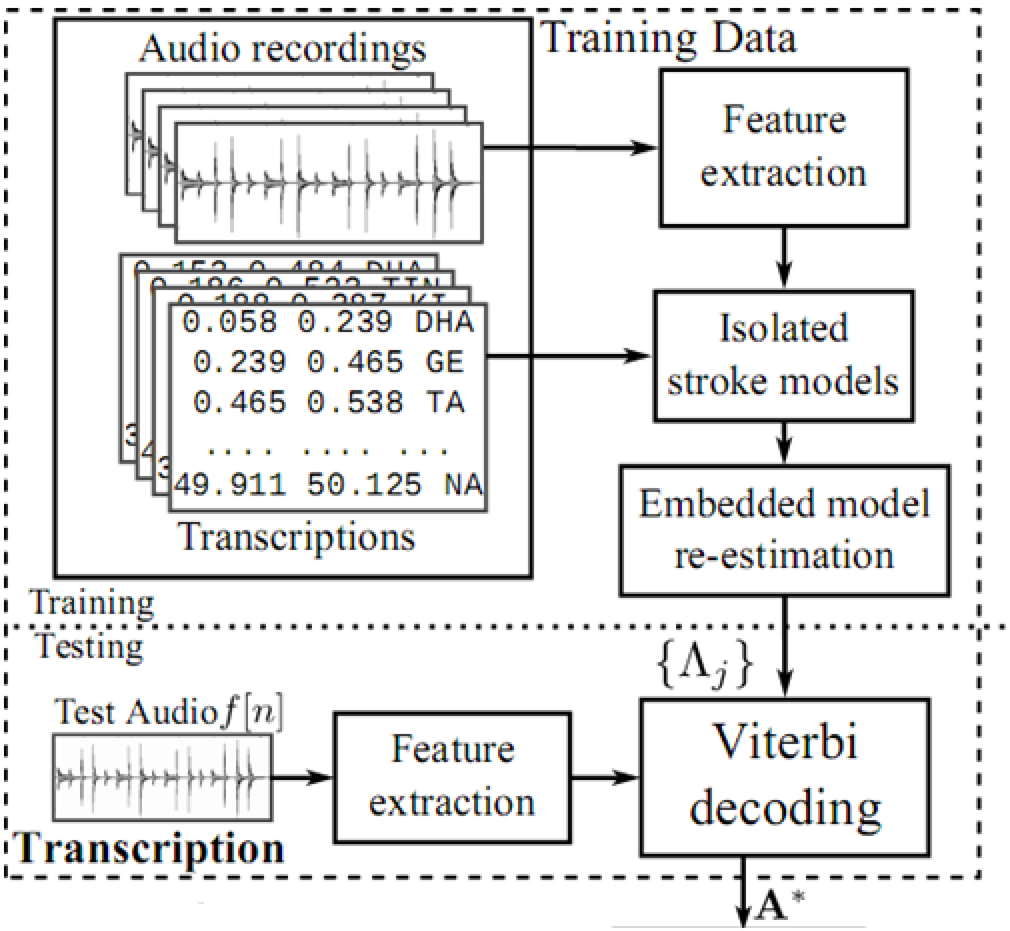


MFCC_0_D_A
MFCC_D_A

Time-aligned transcriptions to initialize models

7 state left-to-right HMM

Automatic transcription



MFCC_0_D_A
MFCC_D_A

Time-aligned transcriptions to initialize models

7 state left-to-right HMM

Flat and bigram LM

Transcription errors

compmusic

- Substitutions (S), Insertions (I), Deletions (D)
- Approximate string search, e.g. edit distance

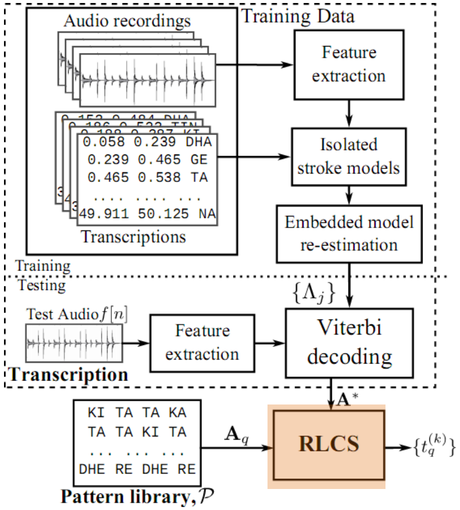
```
LAB:    ki ta ki na ki   dhin    ta ki    na ki        na dhin ta
REC: ki ki ta ki na na   dhin ki ta ki ta na ki ta dha na dhin ta

SENTENCE: @Correct=0.00 [H=0, S=1, N=1]
WORD:     @Corr=92.31, Acc=53.85 [H=12, D=0, S=1, I=5, N=13]
```

G. Navarro (2001). A Guided Tour to Approximate String Matching. ACM Computing Surveys, 33(1):31–88.

Approximate pattern search

- Rough Longest Common Subsequence (RLCS)
  - Robust to transcription errors
  - Multiple occurrences in a composition
- Dynamic Programming based method
  - Cost functions for rough matches
  - Similarity function between syllables
- Variants of RLCS
  - Binary syllable distance measure



Experiments

compmusic

- MTS and UMS datasets
  - Isolated stroke model initialization in MTS dataset
- Transcription evaluation measures
  - Correctness $\mathfrak{C} = \frac{L - \mathtt{D} - \mathtt{S}}{L}$
  - Accuracy $\mathfrak{A} = \frac{L - \mathtt{D} - \mathtt{S} - \mathtt{I}}{L}$
- Approximate search evaluation
  - Precision, recall, f-measure (70% overlap)
  - NO accurate boundary estimation

MTS dataset: Transcription

compmusic

LM	Feature	Training $\mathfrak{C}$	Training $\mathfrak{A}$	Test $\mathfrak{C}$	Test $\mathfrak{A}$
Flat	MFCC_D_A	67.82	46.05	64.21	37.94
	MFCC_0_D_A	70.63	51.78	66.30	43.86
Bigram	MFCC_D_A	**68.50**	**50.48**	**65.33**	**44.10**
	MFCC_0_D_A	69.33	46.72	64.49	39.48

Flat initialization

LM	Feature	Training $\mathfrak{C}$	Training $\mathfrak{A}$	Test $\mathfrak{C}$	Test $\mathfrak{A}$
Flat	MFCC_D_A	68.42	52.69	64.07	45.01
	MFCC_0_D_A	68.91	56.78	64.26	49.27
Bigram	MFCC_D_A	70.16	57.83	65.53	49.97
	MFCC_0_D_A	**70.71**	**60.77**	**66.23**	**53.13**

Time-aligned transcripts for stroke model initialization

Gupta, S., **Srinivasamurthy, A.**, Kumar, M., Murthy, H., & Serra, X. (2015, October). Discovery of Syllabic Percussion Patterns in Tabla Solo Recordings. In Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR 2015) (pp. 385–391). Malaga, Spain.

MTS dataset: Pattern search

- Baseline – exact string search
- RLCS – Improved recall with lower precision
  - Improved f-measure
- Different variants of RLCS tested
  - Needs further exploration

Variant	Parameter	Precision ($\mathfrak{p}$)	Recall ($\mathfrak{r}$)	f-measure ($\mathfrak{f}$)
Baseline	-	0.479	0.254	0.332
RLCS_0	$\delta = 0$	0.384	0.395	0.389

MTS dataset: Pattern search

compmusic

- Baseline – exact string search
- RLCS – Improved recall with lower precision
  - Improved f-measure
- Different variants of RLCS tested
  - Needs further exploration

Variant	Parameter	Precision ($\mathfrak{p}$)	Recall ($\mathfrak{r}$)	f-measure ($\mathfrak{f}$)
Baseline	-	0.479	0.254	0.332
RLCS_0	$\delta = 0$	0.384	0.395	0.389

MTS dataset: Pattern search

- Baseline – exact string search
- RLCS – Improved recall with lower precision
  - Improved f-measure
- Different variants of RLCS tested
  - Needs further exploration

Variant	Parameter	Precision ($\mathfrak{p}$)	Recall ($\mathfrak{r}$)	f-measure ($\mathfrak{f}$)
Baseline	-	0.479	0.254	0.332
$RLCS_0$	$\delta = 0$	0.384	0.395	0.389

Summary: MTS dataset

- Best transcription performance: a bigram language model
- RLCS improves over baseline
- RLCS variants did not add much value

UMS dataset: Transcription

compmusic

- Best performance: Flat language model + MFCC_0_D_A
- Better performance than on tabla dataset

LM	Feature	Training		Test	
		$\mathfrak{C}$	$\mathfrak{A}$	$\mathfrak{C}$	$\mathfrak{A}$
Flat	MFCC_D_A	76.66	59.43	74.08	55.64
	MFCC_0_D_A	**76.63**	**63.79**	**74.13**	**60.23**
Bigram	MFCC_D_A	78.12	57.69	75.90	54.02
	MFCC_0_D_A	78.78	60.54	76.50	57.38

Sec 6.3.4

UMS dataset: Pattern search

- Baseline – exact string search
- RLCS0-1: Best f-measure
  - High RLCS score threshold: equivalent to exact search
- RLCS0-2: Best recall

Variant	Parameter	Precision ($\mathfrak{p}$)	Recall ($\mathfrak{r}$)	f-measure ($\mathfrak{f}$)
Baseline	-	0.902	0.492	0.637
RLCS_0-1	$\delta = 0$	0.902	0.492	0.637
RLCS_0-2	$\delta = 0$	0.258	0.762	0.386

UMS dataset: Pattern search

- Baseline – exact string search
- RLCS0-1: Best f-measure
  - High RLCS score threshold: equivalent to exact search
- RLCS0-2: Best recall

Variant	Parameter	Precision ($\mathfrak{p}$)	Recall ($\mathfrak{r}$)	f-measure ($\mathfrak{f}$)
Baseline	-	0.902	0.492	0.637
$RLCS_0$-1	$\delta = 0$	0.902	0.492	0.637
$RLCS_0$-2	$\delta = 0$	0.258	0.762	0.386

UMS dataset: Pattern search

compmusic

- Baseline – exact string search
- RLCS0-1: Best f-measure
  - High RLCS score threshold: equivalent to exact search
- RLCS0-2: Best recall

Variant	Parameter	Precision ($\mathfrak{p}$)	Recall ($\mathfrak{r}$)	f-measure ($\mathfrak{f}$)
Baseline	-	0.902	0.492	0.637
RLCS_0-1	$\delta = 0$	0.902	0.492	0.637
RLCS_0-2	$\delta = 0$	0.258	0.762	0.386

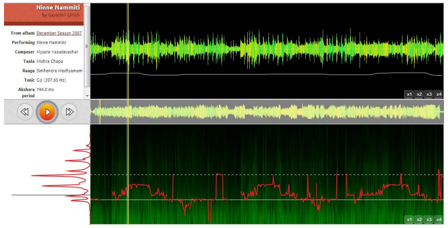
Summary

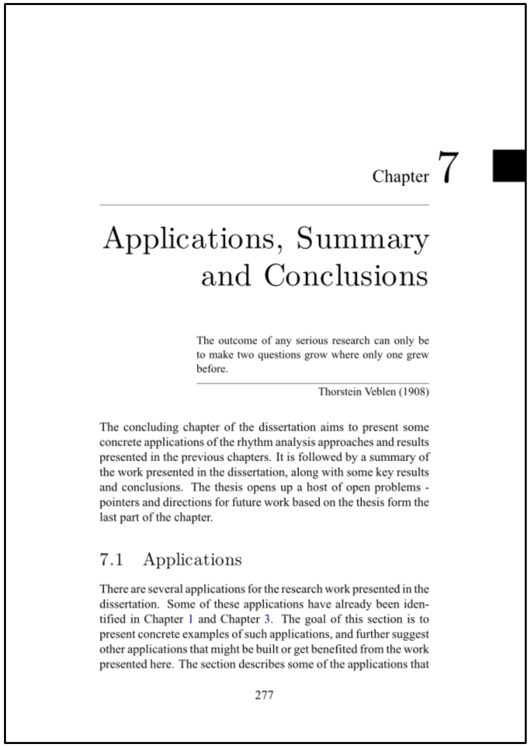
- Transcription performance better than that for tabla
  - Language model does not help much
- RLCS does not add much value here
- Nature of data
  - Pieces are very short
  - Pieces and patterns are of the same order of length

Percussion pattern discovery: Summary

- Preliminary experiments – need further exploration
  - More data to be added to MTS and UMS datasets
- Comprehensive pattern definition
  - Include dynamics and timing
- Improve pattern boundaries
  - Onset detection
- Better RLCS
  - Timbral syllable similarity measures

Summary and conclusions





Chapter 7

Applications, Summary and Conclusions

> The outcome of any serious research can only be to make two questions grow where only one grew before.
>
> Thorstein Veblen (1908)

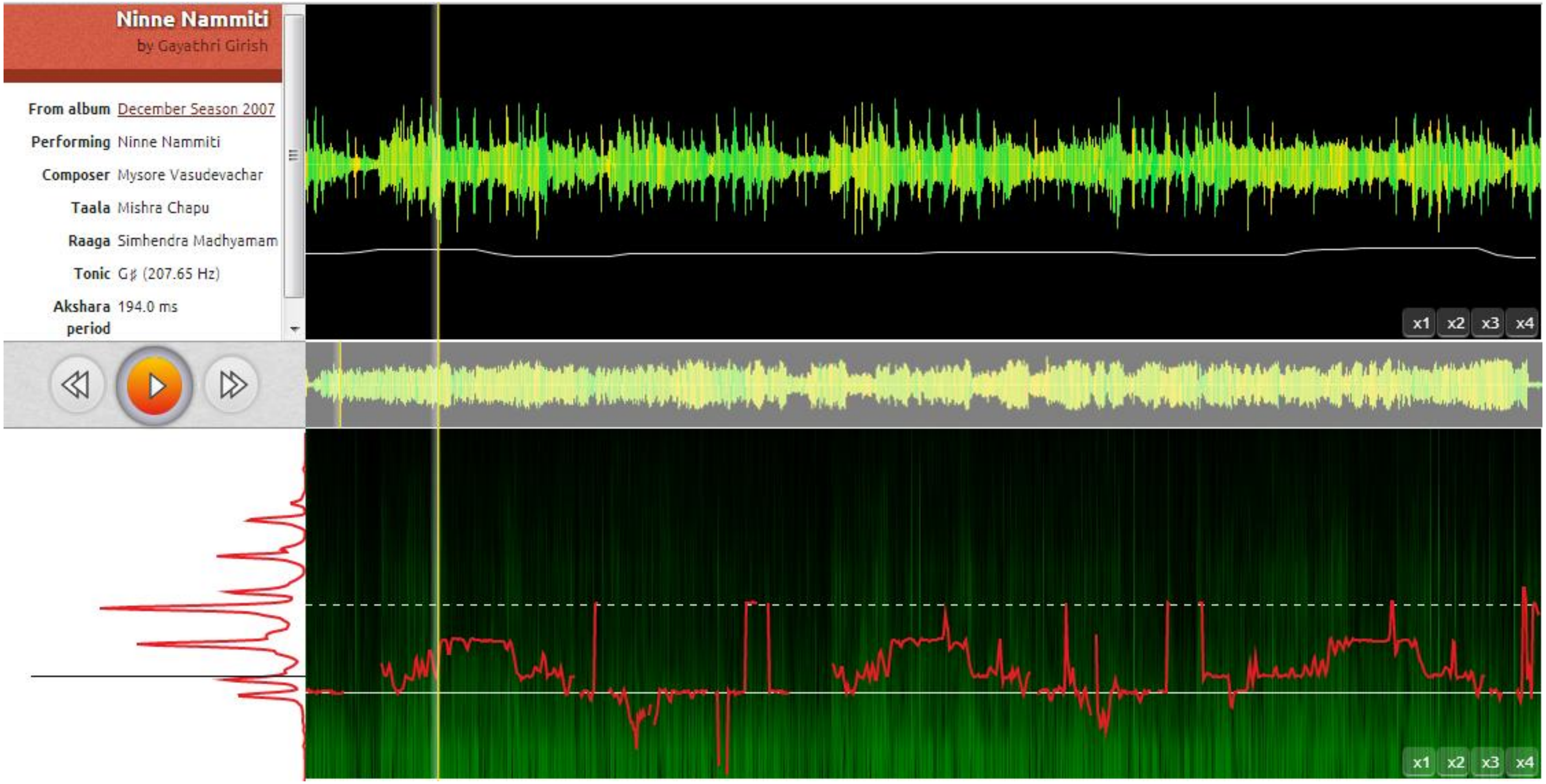
The concluding chapter of the dissertation aims to present some concrete applications of the rhythm analysis approaches and results presented in the previous chapters. It is followed by a summary of the work presented in the dissertation, along with some key results and conclusions. The thesis opens up a host of open problems - pointers and directions for future work based on the thesis form the last part of the chapter.

7.1 Applications

There are several applications for the research work presented in the dissertation. Some of these applications have already been identified in Chapter 1 and Chapter 3. The goal of this section is to present concrete examples of such applications, and further suggest other applications that might be built or get benefited from the work presented here. The section describes some of the applications that

277

Applications: Dunya



http://dunya.compmusic.upf.edu

Applications: Sarāga

http://musicmuni.com/

Contributions: Data corpora

- Sam/a and beat/mātra annotations in
  - Carnatic Music Rhythm (CMRf) dataset (16.6 hours)
  - Hindustani Music Rhythm (HMR) dataset (5 hours)
- Sam/a annotations in
  - **Carnatic Creative Commons music collection (41 hours)**
  - **Hindustani Creative Commons music collection (43 hours)**
- Percussion datasets
  - Mulgaonkar Tabla Solo (MTS) dataset (17 minutes)
  - Jingju percussion instrument dataset (3000 examples)
  - Jingju percussion patterns dataset (22 minutes)

Technical and scientific contributions (1)

- Identification of challenges and opportunities in automatic rhythm analysis of Indian art music
- Identification of relevant automatic rhythm analysis research problems in Indian art music
- An engineering formulation of meter analysis and percussion pattern discovery in Indian art music
- An illustrative evaluation of Carnatic and Hindustani research corpora
- Illustrative rhythm analysis of annotated rhythm datasets of Carnatic and Hindustani music

Technical and scientific contributions (2)

compmusic

- Bayesian methods for meter analysis in Indian art music
  - Novel extensions to the bar pointer model
- Percussion pattern discovery from percussion solos in Indian art music
  - Syllabic percussion systems, timbre mapping

Publications (1)

compmusic

Peer-reviewed journals

- **Srinivasamurthy, A.**, Holzapfel, A., & Serra, X. (2014). In Search of Automatic Rhythm Analysis Methods for Turkish and Indian Art Music. Journal of New Music Research, 43(1), 97–117.

Full articles in peer-reviewed conferences

- **Srinivasamurthy, A.**, Holzapfel, A., Cemgil, A. T., & Serra, X. (2016, March). A generalized Bayesian model for tracking long metrical cycles in acoustic music signals. In Proceedings of the 41st IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2016) (pp. 76–80). Shanghai, China.
- **Srinivasamurthy, A.**, Holzapfel, A., Cemgil, A. T., & Serra, X. (2015, October). Particle Filters for Efficient Meter Tracking with Dynamic Bayesian Networks. In Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR 2015) (pp. 197–203). Malaga, Spain.
- Gupta, S., **Srinivasamurthy, A.**, Kumar, M., Murthy, H., & Serra, X. (2015, October). Discovery of Syllabic Percussion Patterns in Tabla Solo Recordings. In Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR 2015) (pp. 385–391). Malaga, Spain.

Publications (2)

compmusic

- **Srinivasamurthy, A.**, Caro, R., Sundar, H., & Serra, X. (2014, October). Transcription and Recognition of Syllable based Percussion Patterns: The Case of Beijing Opera. In Proceedings of the 15th International Society for Music Information Retrieval Conference (ISMIR 2014) (pp. 431–436). Taipei, Taiwan.

- Holzapfel, A., Krebs, F., & **Srinivasamurthy, A.** (2014, October). Tracking the "odd": Meter inference in a culturally diverse music corpus. In Proceedings of the 15th International Society for Music Information Retrieval Conference (ISMIR 2014) (pp. 425–430). Taipei, Taiwan.

- **Srinivasamurthy, A.**, Koduri, G. K., Gulati, S., Ishwar, V., & Serra, X. (2014, September). Corpora for Music Information Research in Indian Art Music. In Proceedings of Joint International Computer Music Conference/Sound and Music Computing Conference. Athens, Greece.

- **Srinivasamurthy, A.**, & Serra, X. (2014, May). A Supervised Approach to Hierarchical Metrical Cycle Tracking from Audio Music Recordings. In Proceedings of the 39th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2014) (pp. 5237–5241). Florence, Italy.

- Tian, M., **Srinivasamurthy, A.**, Sandler, M., & Serra, X. (2014, May). A Study of Instrument-wise Onset Detection in Beijing Opera Percussion Ensembles. In Proceedings of the 39th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2014) (pp. 2174–2178). Florence, Italy.

- **Srinivasamurthy, A.**, Subramanian, S., Tronel, G., & Chordia, P. (2012, July). A Beat Tracking Approach to Complete Description of Rhythm in Indian Classical Music. In Proceedings of the 2nd CompMusic Workshop (pp. 72–78). Istanbul, Turkey.

Publications (3)

compmusic

Articles under preparation

- **Srinivasamurthy, A.**, Holzapfel, A., Ganguli, K. K., & Serra, X. Computational Rhythm analysis of a Hindustani Music Corpus (under preparation)

- **Srinivasamurthy, A.**, Holzapfel, A., & Serra, X. Bayesian models for meter analysis in Indian art music. (under preparation)

Resources

compmusic

Computational models for the discovery of the World's Music

UNIVERSITAT POMPEU FABRA

erc

- Home
- Description
- Team
- Publications
- Corpora
- Software
- Events
- Blog
- News
- Resources
- Get involved

English

Companion webpage for the PhD thesis of Ajay Srinivasamurthy

This page is the companion webpage for the PhD thesis titled

A Data-driven Bayesian Approach to Automatic Rhythm Analysis of Indian Art Music

Ajay Srinivasamurthy

(Last updated: 13 Nov 2016. Please click on the headings to expand.)

The dissertation document can be obtained from http://mtg.upf.edu/node/3593

- Examples
- Datasets
- Publications
- Code
- Results

(The page http://www.ajaysrinivasamurthy.in/phd-thesis redirects to this page)

Latest news

Xavier Serra participates to a conference on Science Diplomacy
31/10/2016 - 10:42
Xavier Serra was invited to present the

Latest blogs

Technology and Multiculturality
17/04/2016
[Article published in the daily newspaper La Vanguardia on Sunday 17th 2016. English translation of the original text written in catalan.] The violin, typewriter or mobile are examples of technological devices that were born in certain contexts...

Two evenings of Chinese traditional music 27/01/2016
Last December (2015), Barcelona's Conservatori Municipal de Música hosted two sessions of Chinese traditional music, the first one devoted to the silk and bamboo music genre and the second one to jingju (Beijing opera). For this...

Nīla Saṅgīta - An evening of Indian Classical Music and Dance

http://compmusic.upf.edu/phd-thesis-ajay

Examples, datasets and code

compmusic

Examples

- A resource page for Carnatic tāḷas
- A resource page for Hindustani tāls
- A resource page for Turkish usuls
- Percussion instruments used in Beijing opera
- Percussion patterns in Beijing opera

Datasets

All the datasets built within CompMusic can be found here: **http://compmusic.upf.edu/datasets**

Some of the datasets more extensively used in the thesis are linked below:

- Carnatic Music Rhythm (CMR) dataset
- Hindustani Music Rhythm (HMR) dataset
- Mulgaonkar Tabla Solo (MTS) dataset
- UKS Mridangam Solo (UMS) dataset
- Jingju Percussion Instrument (JPI) dataset
- Jingju Percussion Pattern (JPP) dataset

Code

- Essentia audio analysis library: **http://essentia.upf.edu/**
- Dunya API: **https://github.com/MTG/pycompmusic**
- Dunya server and back end: **https://github.com/MTG/dunya**
- A MATLAB package for meter analysis (maintained by Florian Krebs): **https://github.com/flokadillo/bayesbeat**
- A MATLAB/Python package for percussion pattern transcription using the **HTK** and RLCS: **https://github.com/swapnilgt/percPatternDiscovery**
- A MATLAB package for beat tracking evaluation (maintained by Matthew Davies): **https://code.soundsoftware.ac.uk/projects/beat-evaluation/**
- Rhythm analysis tools for jingju, from the tutorial in ISMIR 2014: **http://compmusic.upf.edu/jingju-tutorial**
- Sawaal-Jawaab Code and Demo: **http://compmusic.upf.edu/ismir-15-hacks**
- BeatStation, an interface to record beat tapping (first developed by Marius Miron, extension to Carnatic music by Ajay Srinivasamurthy): **https://github.com/ajaysmurthy/beatStation**

Results: Audio examples

compmusic

Results

Meter analysis

Some audio examples of meter analysis are here: **http://dunya.compmusic.upf.edu/ajay-thesis/examplesPage.htm**

Percussion pattern discovery

Some audio examples of percussion pattern discovery are here (examples generated with the help of Swapnil Gupta): **http://dunya.compmusic.upf.edu/ajay-thesis/tablaExamples.htm**

Audio Features

The examples below are a sonification of the spectral flux using a sawtooth signal of 800Hz with an amplitude envelope of the spectral flux feature used in the thesis. An audible click is also present at the true location of the sama (downbeat). The tracks are stereo: left and right channels have the spectral flux from low and high frequency bands, respectively.

Brochevarevarura in adi tala (**MusicBrainz**)

0:00 / 0:31

Sri Mahaganapati in mishra chapu tala (**MusicBrainz**)

0:00 / 0:18

Meter analysis results (audio examples)

This page is an accompanying page for the thesis **A Data-driven Bayesian Approach to Automatic Rhythm Analysis of Indian art Music** by Ajay Srinivasamurthy. Please see http://compmusic.upf.edu/phd-thesis-ajay for more details on the thesis.

The page lists some audio examples of automatic meter analysis in both Carnatic and Hindustani music. Click on each example to expand and listen to outputs of different meter analysis algorithms.

The audio examples have audible clicks marking the different events of the tala. Each track has a high frequency (2000 Hz) loud click at the sama, a lower frequency (1500 Hz) click at the anga/vibhaag boundaries, and a low frequency (1000 Hz) click at each beat of the tala.

The abbreviations correspond to the different algorithms presented in the thesis - Ground Truth: Manually annotated beats, Inf-AMPFo: Meter Inference with AMPF0, Track-AMPFo: Meter tracking with AMPF0, Track-AMPFm: Meter tracking with AMPFm, Track-AMPFs: Meter tracking with AMPFs, tInfTrack-AMPFo: Tempo-informed meter tracking with AMPF0, tInfTrack-AMPFs: Tempo-informed meter tracking with AMPFs, tsInfTrack-AMPFo: Tempo-sama-informed meter tracking with AMPF0, tsInfTrack-AMPFs: Tempo-sama-informed meter tracking with AMPFs.

The Carnatic examples are from CMR dataset

Carnatic Music

The Hindustani examples are from HMRs and HMRl datasets

Hindustani Music

Shankari Neeve in rupaka tala (MusicBrainz)

- Ground Truth

  0:00 / 0:30

- Inf-AMPFo

  0:00 / 0:30

- Track-AMPFo

  0:00 / 0:30

- Track-AMPFm

  0:00 / 0:30

- Track-AMPFs

  0:00 / 0:30

- tInfTrack-AMPFo

  0:00 / 0:30

- tInfTrack-AMPFs

  0:00 / 0:30

http://dunya.compmusic.upf.edu/ajay-thesis/examplesPage.htm

Percussion pattern discovery results (audio examples from tabla solo dataset)

This page is an accompanying page for the thesis **A Data-driven Bayesian Approach to Automatic Rhythm Analysis of Indian art Music** by Ajay Srinivasamurthy. Please see http://compmusic.upf.edu/phd-thesis-ajay for more details on the thesis.

The page lists some audio examples of percussion pattern discovery in tabla solo recordings. Click on each example to expand and listen to discovered patterns for each example. All the compositions are taken from Shades of Tabla by Pandit Arvind Mulagonkar, which was also used for experiments in the thesis.

For each composition, the complete audio recording of the composition is shown, along with some pattern examples discovered from the recordings using RLCS_0 algorithm. The numbers in parantheses for each example denotes the time location in the recording where the pattern was discovered.

Composition-1 (Dilli gharana, a quayda in tryasra jati teentaal, composed by Ustad Sidhar Khan Dhadi) (Link to MusicBrainz)

- Audio recording of the whole composition

  0:00 / 0:56

- Pattern DHA-GE-TA-TA Example-1 (9.77-10.41)

  0:00 / 0:00

- Pattern DHA-GE-TA-TA Example-2 (23.95 - 24.93)

  0:00 / 0:01

- Pattern DHA-GE-TA-TA Example-3 (21.63 - 22.63)

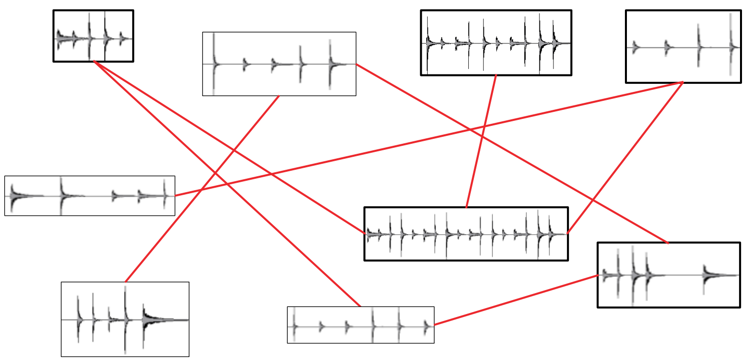
  0:00 / 0:01

http://dunya.compmusic.upf.edu/ajay-thesis/tablaExamples.htm

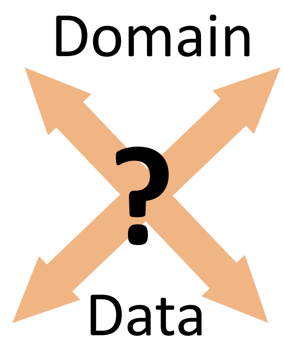
Impact and personal comments

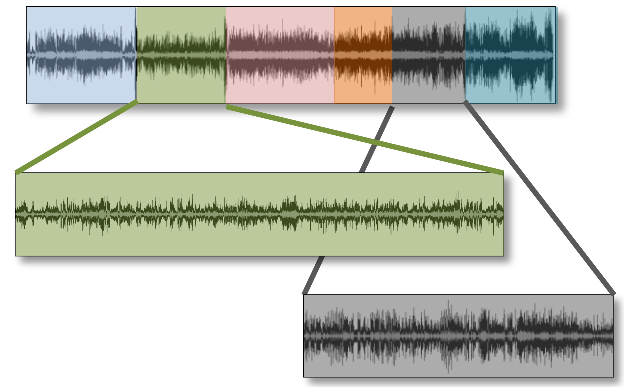
- Broad topics – first thesis!
  - Some experiments are exploratory
- New research problems identified in Indian art music
  - Good quality representative corpora for data-driven research
- CompMusic impact
  - Multiple collaborations with engineers, musicians, musicologists and music listeners

Many open ends!

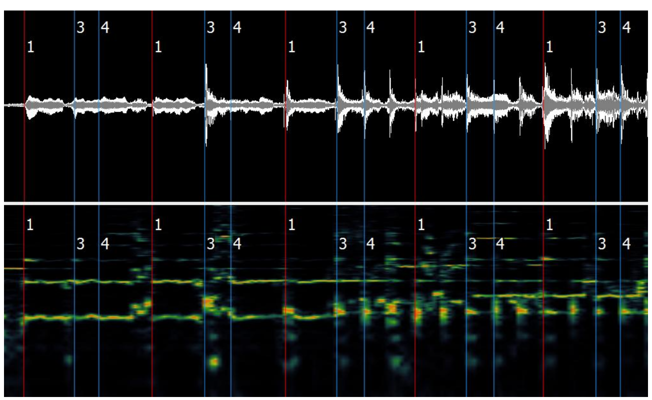


Rhythm similarity

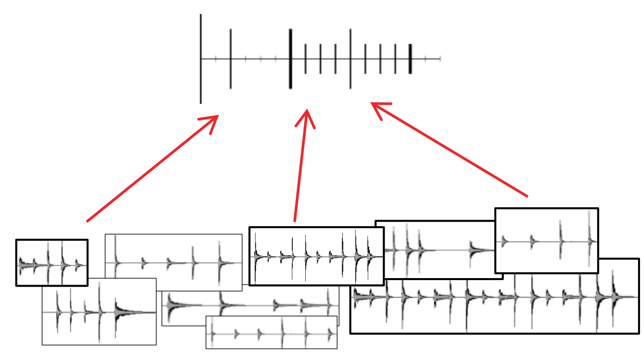




Structural analysis



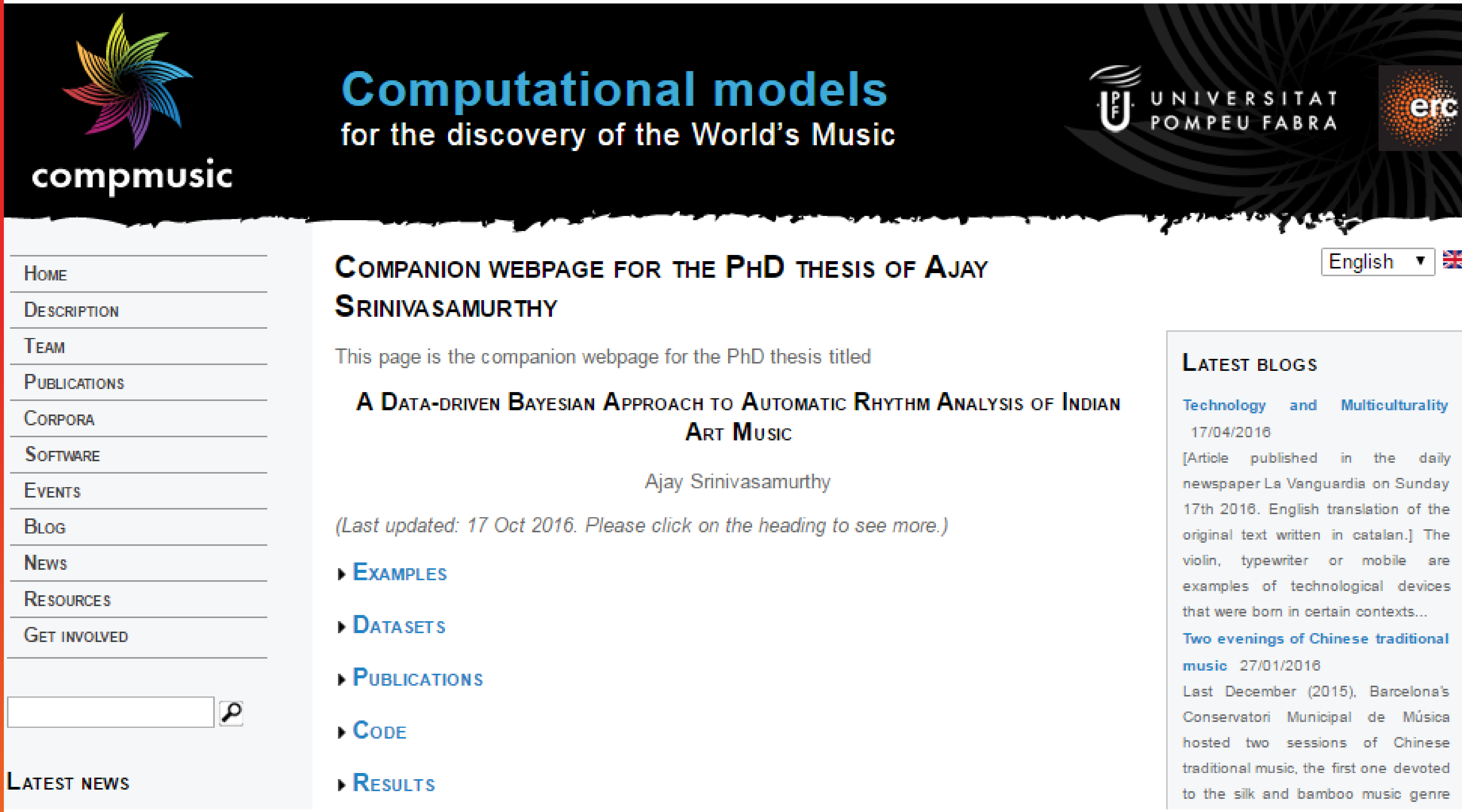
Meter analysis



Pattern Discovery

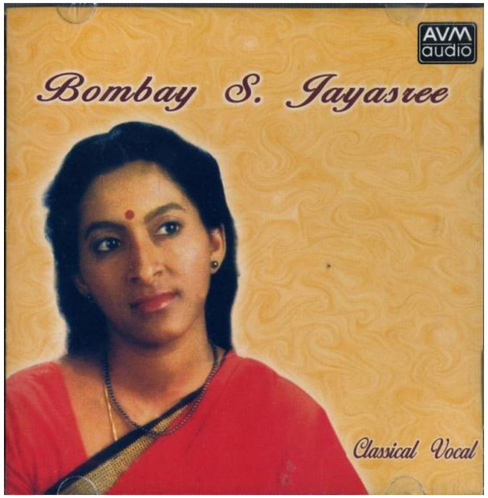
compmusic

Companion page

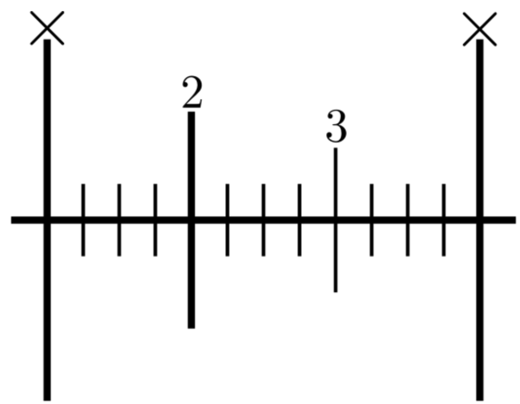


http://compmusic.upf.edu/phd-thesis-ajay

Meter analysis example: Carnatic

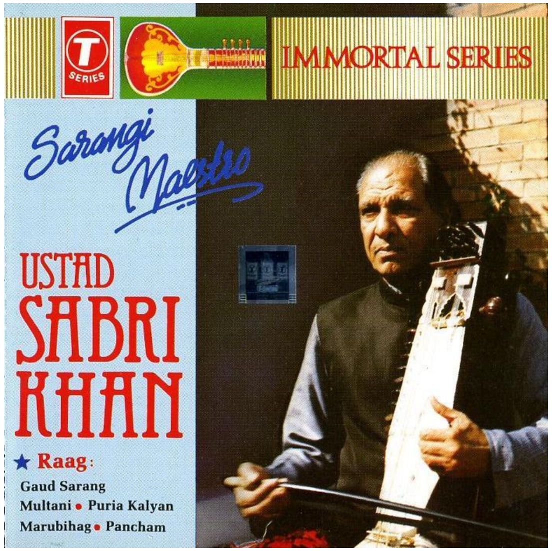


Artist:	Bombay Jayasree (vocal)
Release:	Classical Vocal
Composition:	Śaṅkari nīvē
Composer:	Subbaraya Sastry
Rāga:	Bēgaḍa
Tāla:	Rūpaka (Cycle of 12 akṣara)

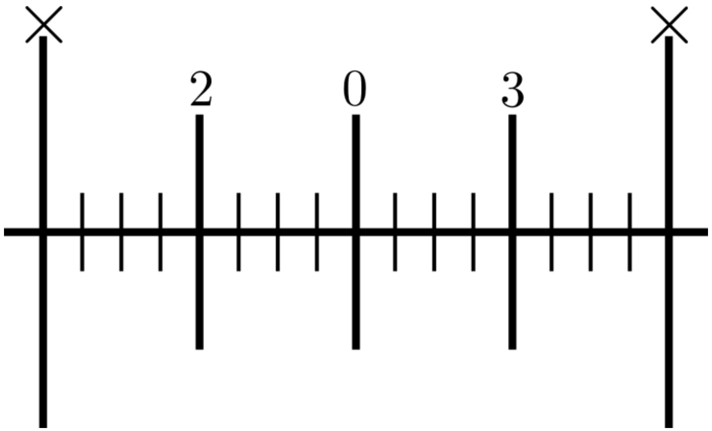


http://musicbrainz.org/recording/fe63aa10-bfa0-4dc6-9845-e57811c2feaa

Meter analysis example: Hindustani



Artist: Sabri Khan (sarangi)
Release: Immortal Series: Sarangi Maestro
Composition: Raga Pancham
Rāga: Pancham
Tāla: Tintal (Cycle of 16 matra)



https://musicbrainz.org/recording/0b29d7fc-9499-4e4f-87a1-869ec8da7ab0

A Data-driven Bayesian Approach to Automatic Rhythm Analysis of Indian Art Music

17 Nov 2016, PhD defense

Ajay Srinivasamurthy

Dept. of Information and Communication Technologies

Universitat Pompeu Fabra

Thesis Supervisor

Dr. Xavier Serra

Music Technology Group

Universitat Pompeu Fabra

Thesis Committee:

Dr. Simon Dixon (QMUL)

Dr. Geoffroy Peeters (IRCAM)

Dr. Juan Pablo Bello (NYU)

compmusic

upf. Music Technology Group