

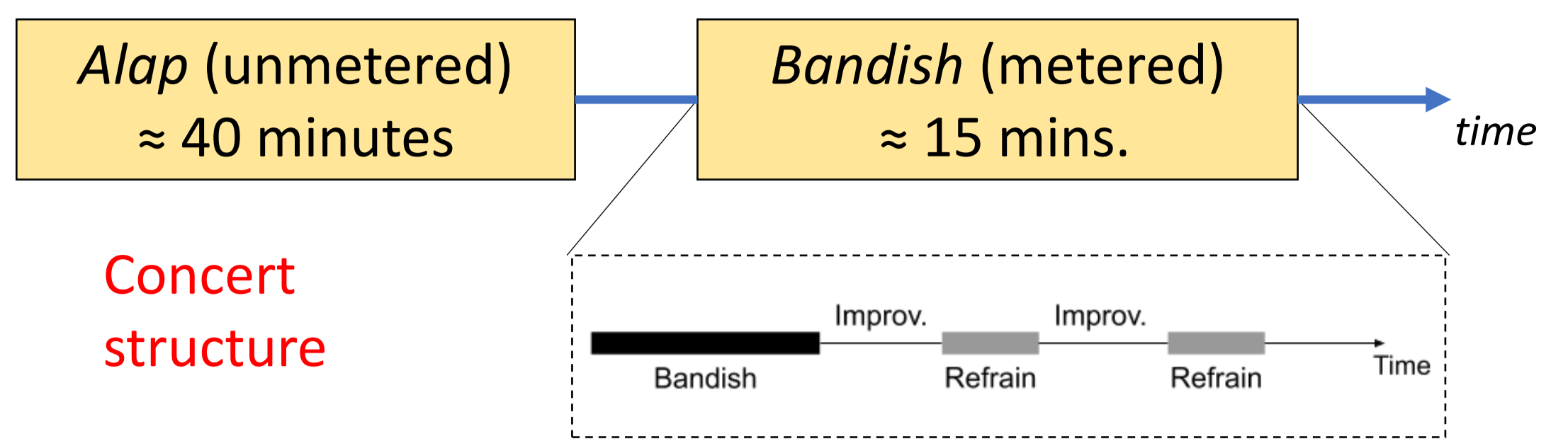
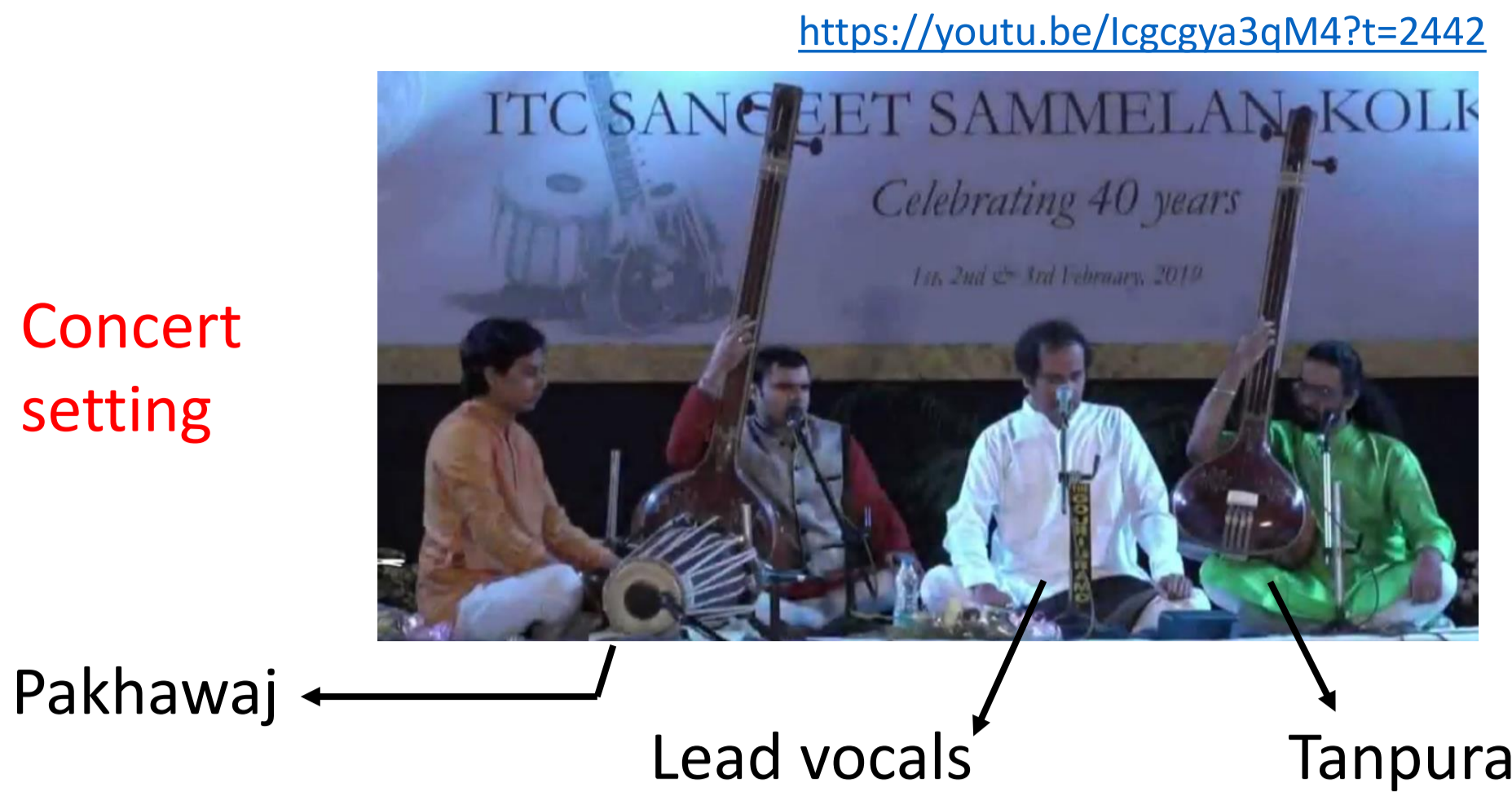
Structural Segmentation Of Dhrupad Vocal Bandish Audio Based On Tempo



Rohit M. A., Vinutha T. P., Preeti Rao
Dept. of Electrical Engineering, I.I.T. Bombay



Dhrupad Vocal Concert



Rhythmic Structure in the Bandish

- Metric tempo (m.t.) - Underlying tempo of composition
 - Range: 30 – 85 BPM
- Surface tempo (s.t.) - Rate of sung syllables or played strokes
 - Range: 30 – 960 BPM!
 - Generally an integer multiple of m.t. – 1, 2, ..., 8, 16
- Surface tempo multiple (s.t.m.) = $s.t. \div m.t.$

A section – during which surface tempo of neither instrument changes

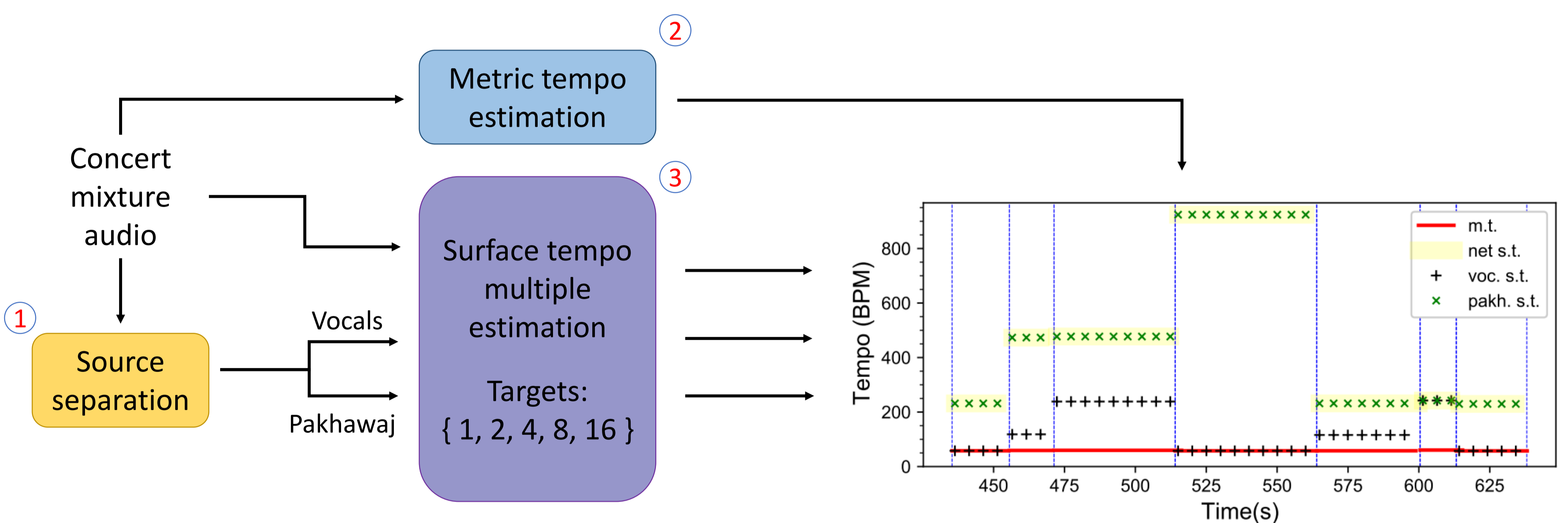


Click for audio examples

Tasks

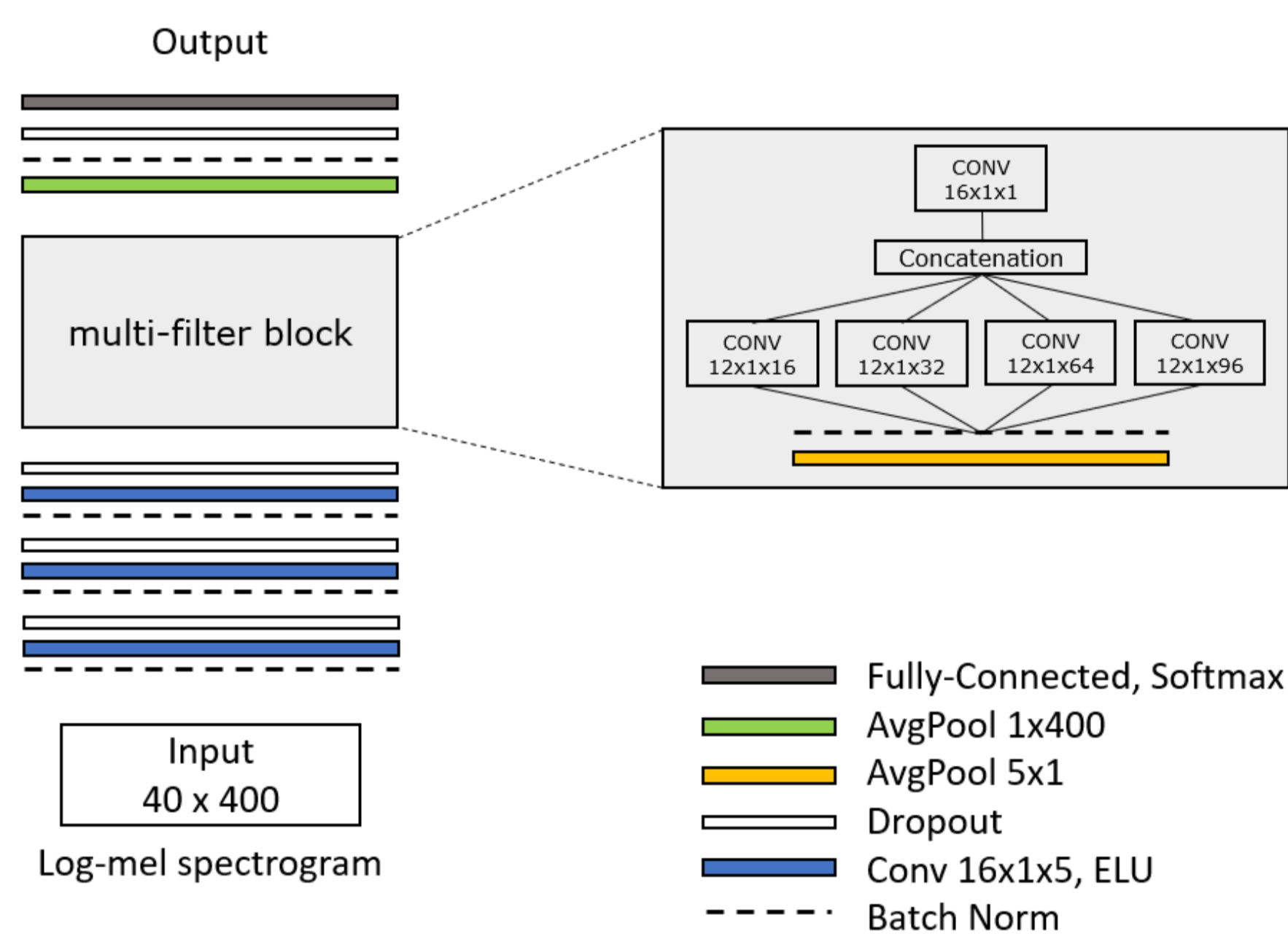
- Track metric tempo and overall surface tempo across a concert
- Track the surface tempi of source separated vocals and pakhawaj and obtain section boundaries

System Overview



Methods

- Source separation – Spleeter 2 stems model [2]
- Metric tempo estimation – tempo-cnn with octave-error correction [3]
- Surface tempo multiple estimation - **modified tempo-cnn**



Layer	Dimensions
Input	40 x 400
(BN, Conv, ELU, DO) x3	16 x 1 x 5
AvgPool	5 x 1
BN, MF Conv, DO	12x {1x16, 1x32, 1x64, 1x96}
Concat, Conv	16 x 1 x 1
AvgPool	1 x 400
BN, DO, FC, Softmax	# output classes

Modifications to tempo-cnn:

- Use of more dropout
- Fewer multi-filter blocks with shorter filters
- AvgPool across time at the end

Dataset

- 14 concerts (Source: Dunya corpus [4], YouTube)
- 634 sections → 1127 8-second chunks (training examples)
- Data augmentation using time-scaling and overlap between examples

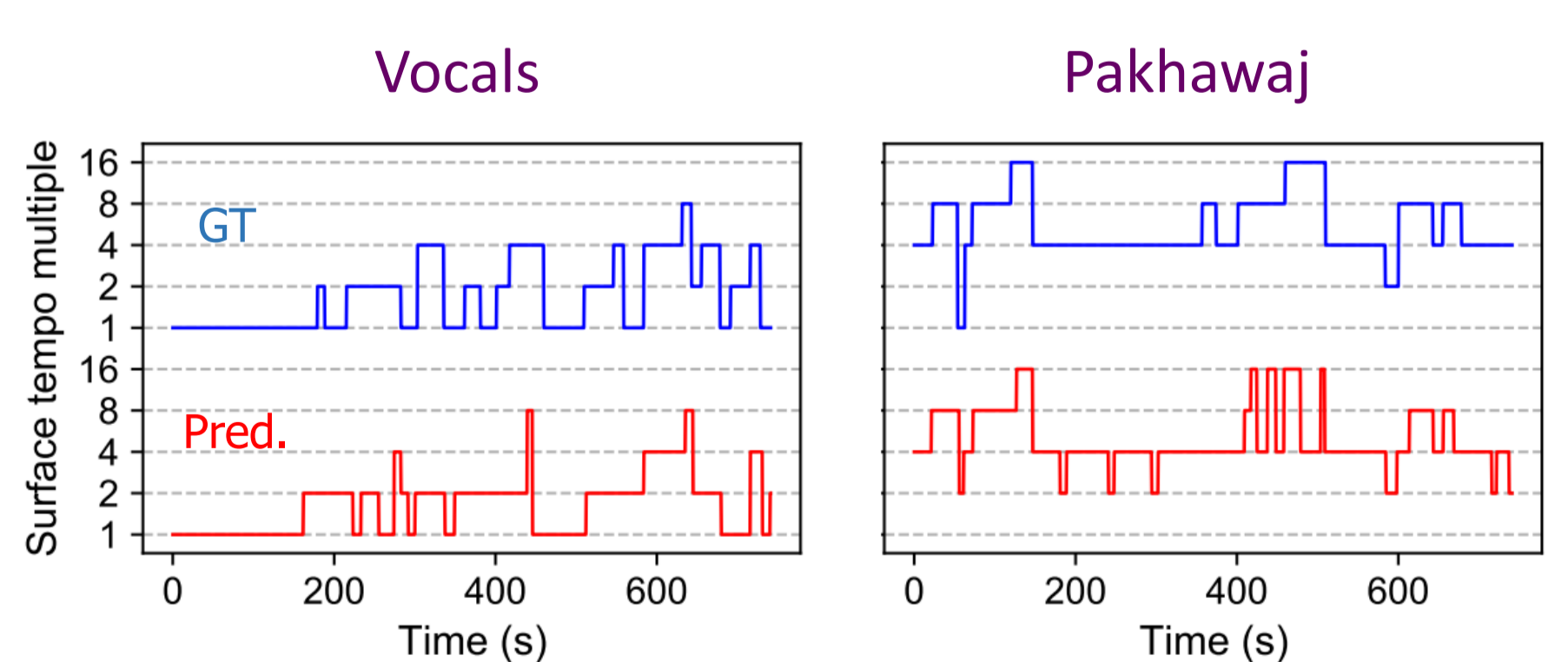
Results

M.t. estimation (% accuracy) S.t.m. estimation (% accuracy)

Acc. 1	Acc. 2	Vocals	Pakhawaj	Net
71.6	74.7	67.7	71.0	70.4

Boundary detection (± 3s tolerance)

Precision	Recall	F-score
0.40	0.53	0.45



Takeaways

- Better results observed on pakhawaj and mixture audios
- Imperfect source separation and melismatic singing are challenges in vocals
- Confusions in s.t.m. due to accents on alternate beats
- Using metric tempo as conditioning could help resolve confusions in s.t.m. prediction

References

- Clayton, M., Time in Indian Music, 2000
- <https://github.com/deezer/spleeter>
- <https://github.com/hendriks73/tempo-cnn>
- <https://dunya.compmusic.upf.edu/>