Adaptive Time–Frequency Scattering for Periodic Modulation Recognition in Music Signals

Changhong Wang[1], Emmanouil Benetos[1], Vincent Lostanlen[2], Elaine Chew[3]

[1]C4DM, Queen Mary University of London, UK; [2]MARL, New York University, USA; [3]CNRS-UMR9912/STMS IRCAM, France

changhong.wang@qmul.ac.uk

1. Introduction

Periodic modulations, such as vibratos, tremolos, trills, and flutter-tongue, are playing techniques elaborated on a stable pitch. The periodic nature of their **spectro-temporal patterns** motivates modeling and discriminating between these techniques.

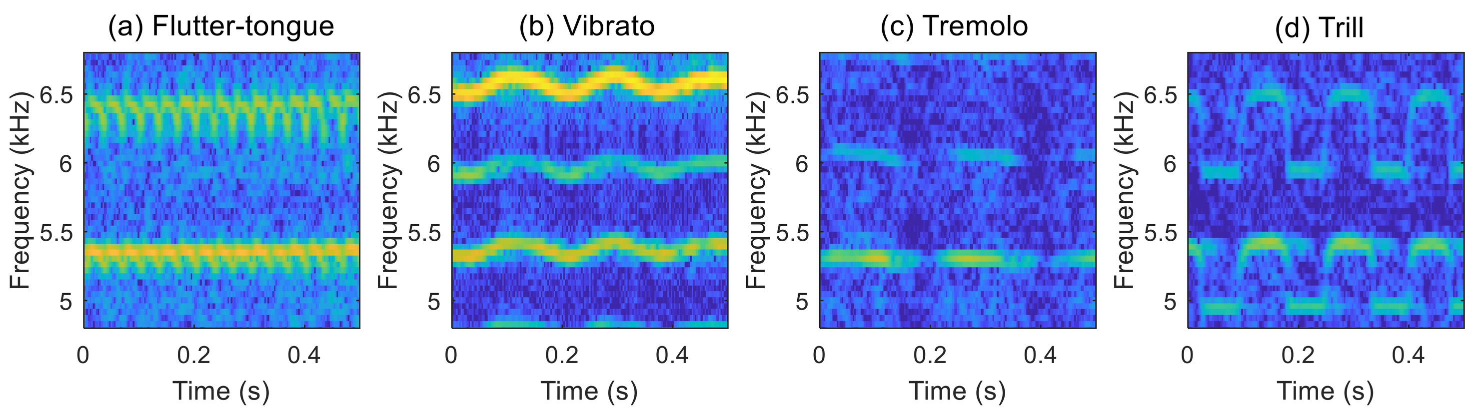
Contributions:

- **Representation**: adaptive time–frequency scattering transform for representing playing techniques.
- **Application**: fine discrimination between periodic modulations.
- **Dataset**: CBFdataset, an annotated dataset of Chinese bamboo flute performances for playing technique analysis.

2. Adaptive Time–Frequency Scattering

➢ Characteristics of Periodic Modulations

- Common feature: **parallel** spectro-temporal patterns.



Type	Rate (Hz)	Extent	Shape
Flutter-tongue	25-50	< 1 semitone	Sawtooth-like
Vibrato	3-10	< 1 semitone	Sinusoidal (FM)
Tremolo	3-8	≈ 0 semitone	Sinusoidal (AM)
Trill	3-10	Note level	Square-like

Figure 1: Characteristics of periodic modulations

➢ Adaptive Time–Frequency Scattering

- **Scattering transform**: a locally translation invariant representation [1].
- **Adaptivity**: transform is calculated **adaptively around the dominant frequency band** (frequency band with maximum acoustic energy).

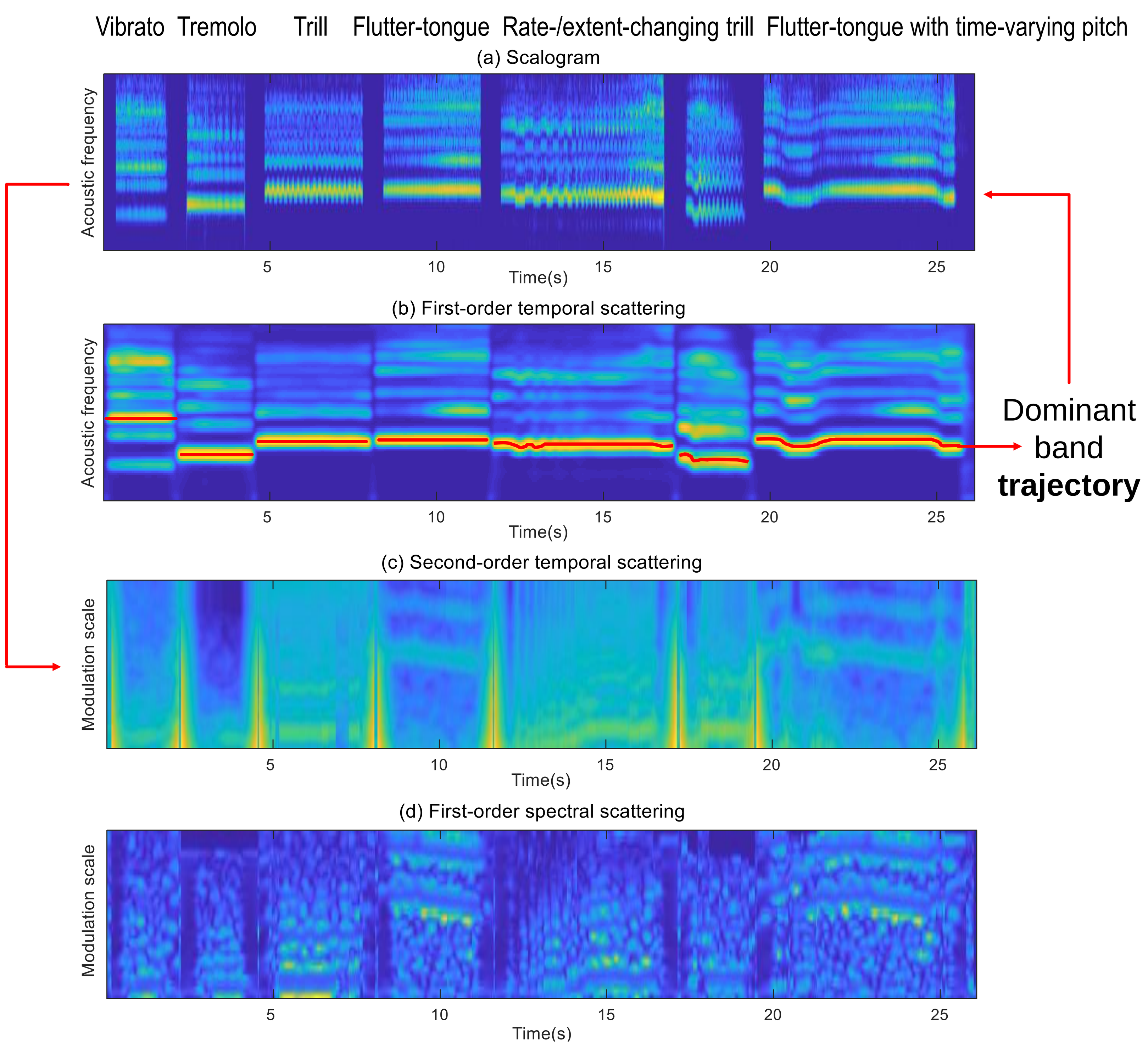


Figure 2: Visualisation of adaptive time–frequency scattering at each stage.

(a) Obtain scalogram (time–frequency image) from audio waveform.

(b) Extract the trajectory of framewise dominant frequency bands.

(c) Decompose the trajectory in (a); obtain modulation representation.

(d) Remove high energy boundaries by applying a spectral scattering along log-frequency axis.

3. Periodic Modulation Recognition

➢ Feature Extraction

- **Band-expanding technique** for discriminating between vibrato and tremolo: **N-band tolerance** centred at the dominant band trajectory.

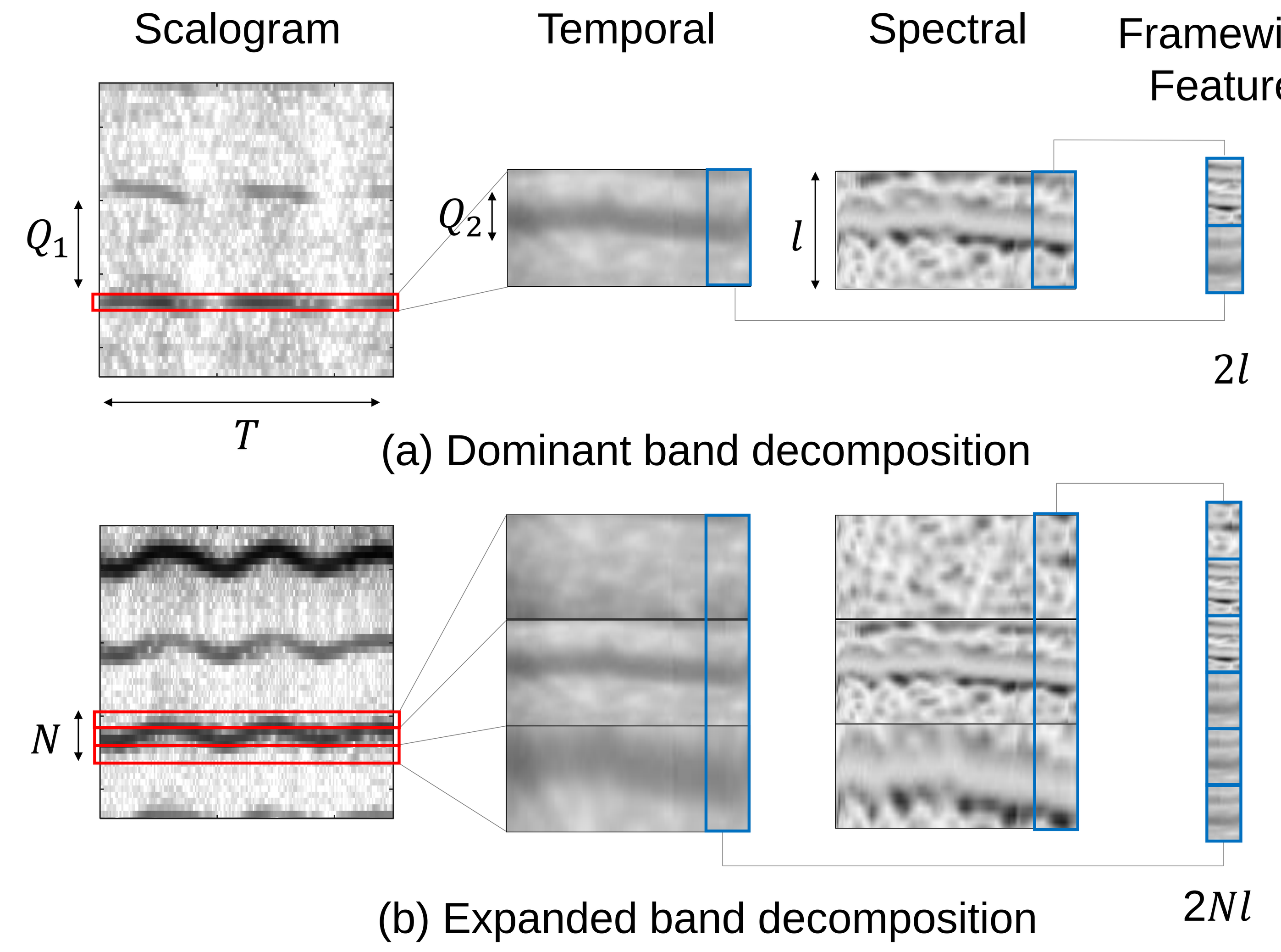


Figure 3: Feature extraction process

➢ Recognition System

- Binary SVM classifier for each playing technique.
- Framewise input with frame size = 46ms, feature dimension = 224.
- Discriminative information can be observed from the fundamental modulation rate and the richness of the harmonics in Figure 2(d).

4. Evaluation

➢ Dataset

- **CBF-periDB:** A Chinese bamboo flute dataset for **peri**odic modulation analysis [2], with full-length performances recorded and annotated by professional players.
- Download (recording + annotations): zenodo.org/record/3250223.

➢ Results

- **Baseline:** filter diagonalisation method for vibrato detection [3]. The best F-measure obtained is 45%.

Type	Dominant band		Expanded band	
	Temporal	Temporal+spectral	Temporal	Temporal+spectral
Flutter-tongue	97.9	98.0	98.1	**98.7**
Trill	75.1	76.2	80.4	**82.3**
Vibrato	26.4	45.3	66.5	**69.3**
Tremolo	2.2	10.6	49.1	**50.7**

Table 2: Performance comparison in framewise F-measures (%).

5. Conclusions

- Adaptive scattering transform presents a versatile and interpretable representation for analysing periodic modulations in performed music.
- Dominant band decomposition is sufficient to detect high-rate and large-extent modulations; expanded band decomposition captures subtle ones.
- Future work: compare adaptive scattering with equivalent representations and further verify the proposed method on other datasets.

[1] J. Andén, and S. Mallat. Deep scattering spectrum. *IEEE Transactions on Signal Processing,* 62(16): 4114–4128, 2014.

[2] C. Wang, E. Benetos, and E. Chew. CBF-periDB: A Chinese bamboo flute dataset for periodic modulation analysis, *ISMIR LBD Session*, Delft, Nov 2019.

[3] L. Yang K. Z. Rajab, and E. Chew. The filter diagonalisation method for music signal analysis: frame-wise vibrato detection and estimation. *Journal of Mathematics and Music,* 11(1): 42–60, 2017.

Queen Mary University of London
centre for digital music
MARL NYU Music and Audio Research Laboratory
ircam Centre Pompidou
erc European Research Council