Joint analysis of mode and playing technique in guqin performance with machine learning

Yu-Fen Huang [1], Jeng-I Liang [2], I-Chieh Wei [1], & Li Su [1]

[1] Music & Cultural Technology Lab, Institute of Information Science, Academia Sinica, Taiwan
[2] Department of Traditional Music, Taipei National University of the Arts, Taiwan

Introduction

* ## Research aims:

  - To bridge the knowledge gap between the music theory and empirical observations

  - To highlight the connection between the high-level structure and local elements in guqin music

* ## Research methods:

  - Compile a gun performance dataset, GQ 39

  - Perform pattern matching for the mode detection

  - Construct Convolutional neural network for the playing technique classification

GQ39 dataset

* ## Guqin (古琴):

  - Plucked seven-string musical instrument

  - UNESCO World Cultural Heritage

* ## Pentatonic scale:

Note degree	$\hat{1}$	$\hat{2}$	$\hat{3}$	$\hat{5}$	$\hat{6}$
Note name	宫(Gong)	商(Shang)	角(Jue)	徵(Zhi)	羽(Yu)

| Rank	Mode 1	Mode 2	Mode 3	Mode 5	Mode 6
1st	1	2	3	5	6
2nd	5	6	-	2	3
3rd	2	3	-	6	-

* ## GQ39 dataset:

  - 39 audio clips from 10 gun solo compositions

  - Historical recordings from 1960 to 1990

  - Note-by-note annotations by professional musician

Mode detection

* The relationship between mode & degree
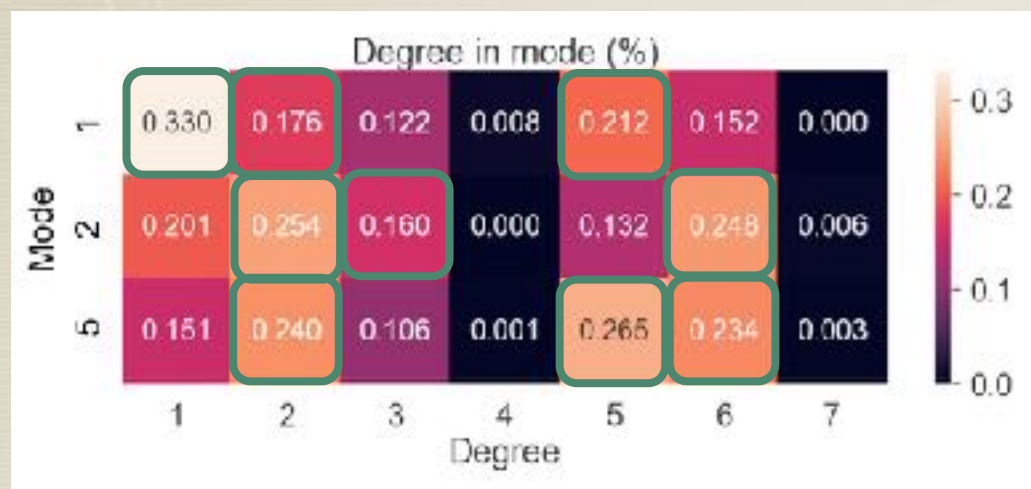
* Mode detection with uniform template (left) &
  Ordinal template with structural information (right)
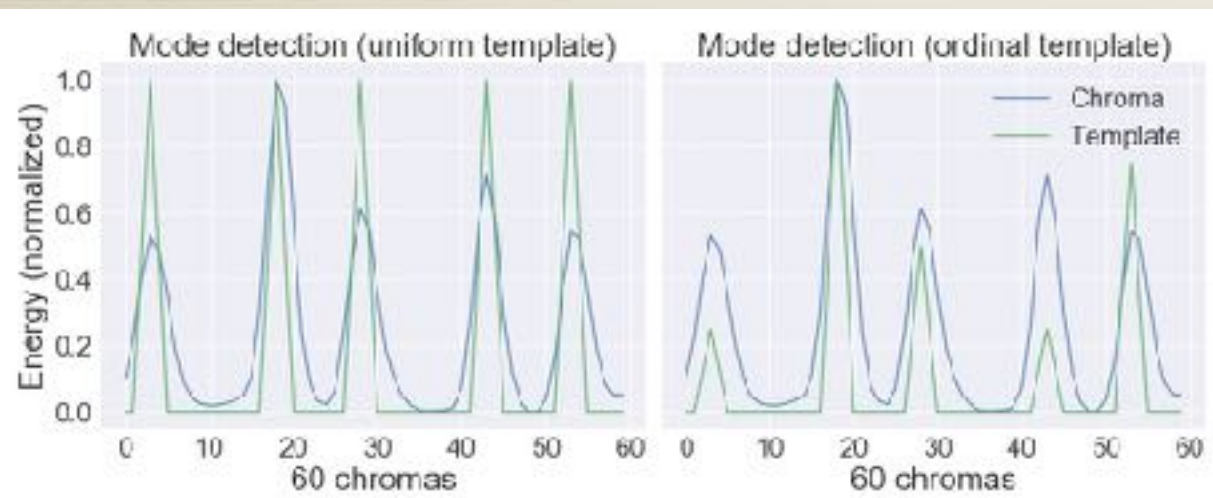
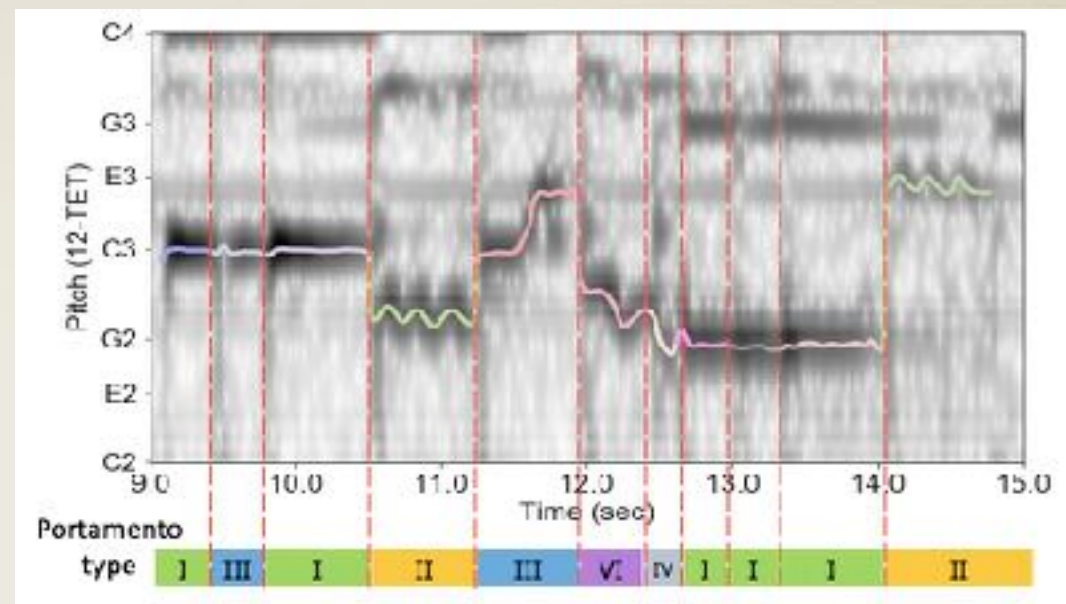| Type	CQT		Salience		Contour	
Result	uni	ord	uni	ord	uni	ord
Correct #	14	**34**	19	29	17	30
Fifth #	4	**4**	5	8	6	8
Miss #	21	**1**	15	2	16	1
Accuracy	0.41	**0.92**	0.55	0.84	0.51	0.87

Playing technique classification

* Data extraction

- Constant-Q transform (CQT)
- Pitch salience function: Crepe
- Estimated pitch contour
  Estimate pitch contour $\mathcal{S} := \{s_i\}_{i=1}^N$
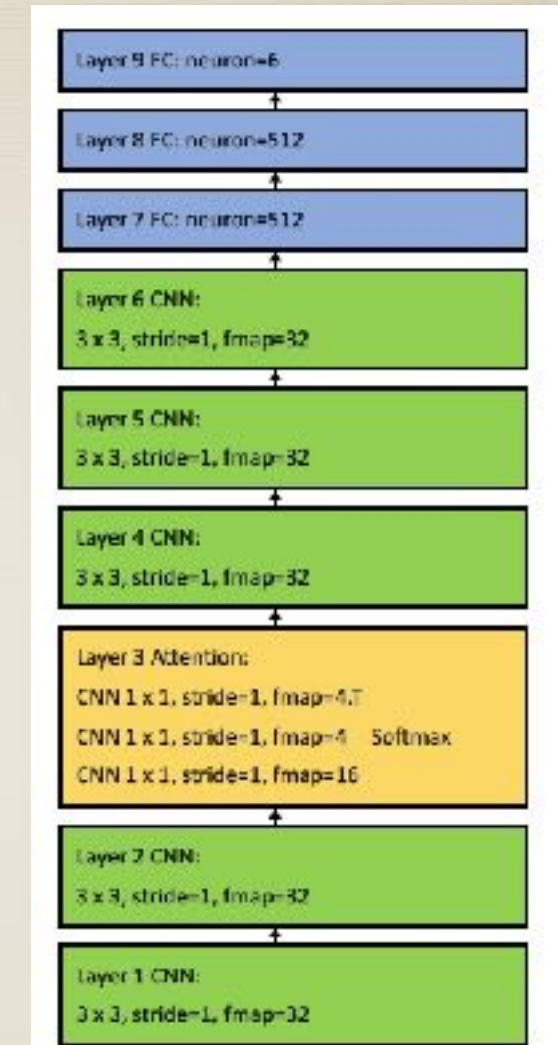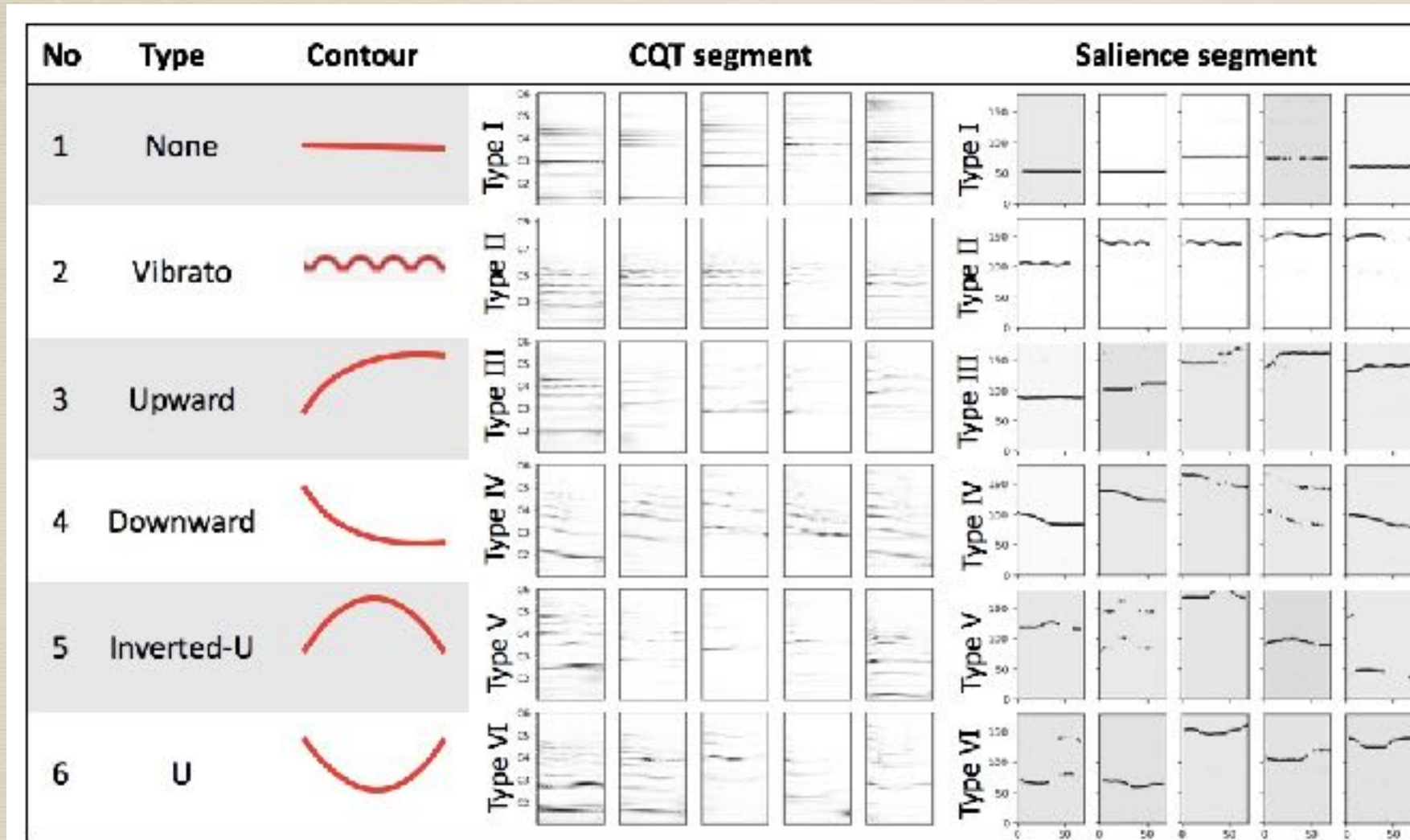 From Crepe $X \in \mathbb{R}^{K \times N}$

$$S^* = \arg\max_{\mathcal{S}} \sum_{i=1}^N X[:, s_i] - \lambda \sum_{i=2}^N |s_{i+1} - s_i|.$$

Classification for 6 types of left-hand playing techniques
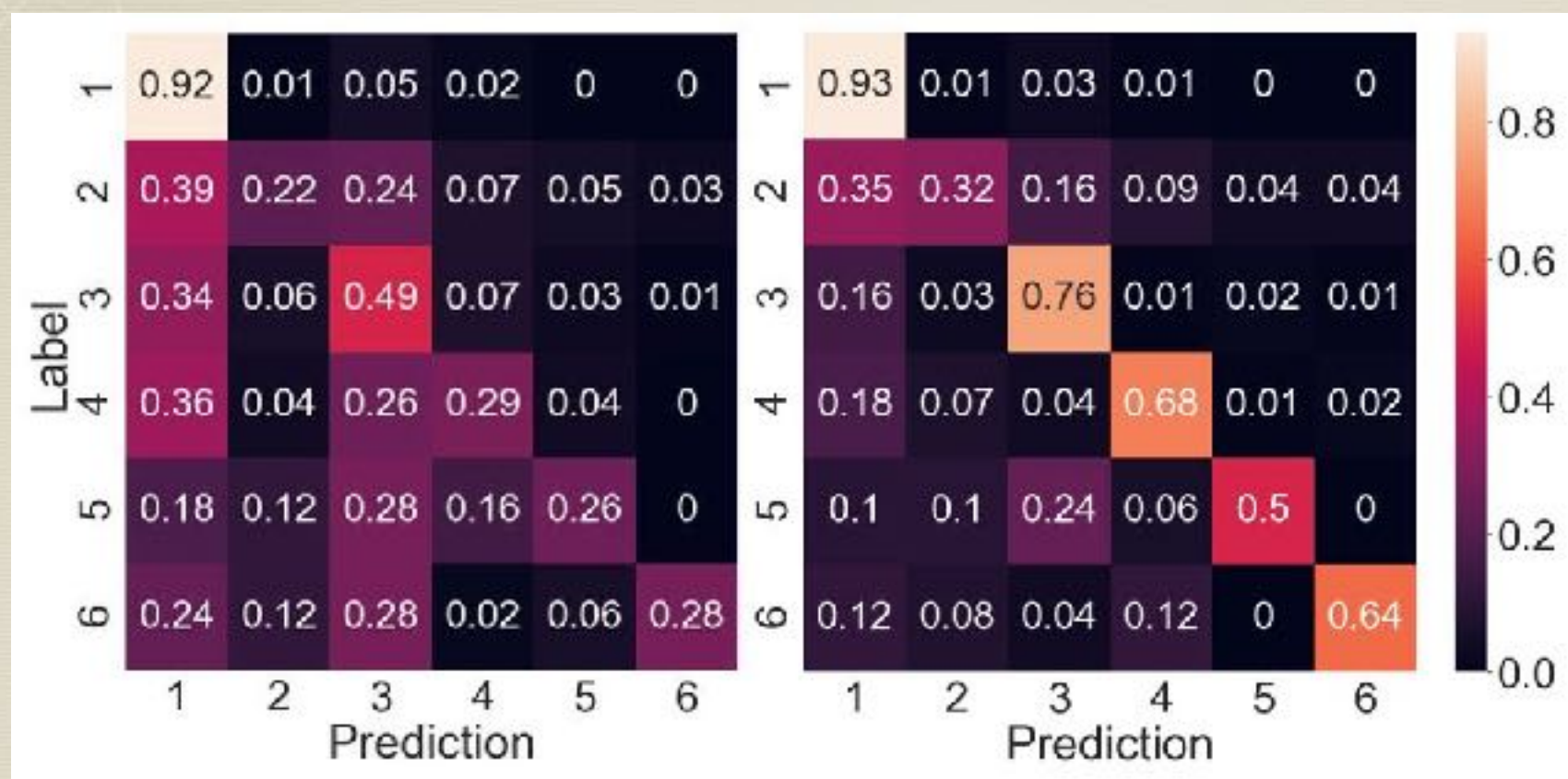
6 types of left-hand portamento techniques

CNN architecture

Classification results

Frame-level features only   With mid-, high-level features

Data type	CQT	Salience	Contour
Frame-level only	0.743	0.845	0.842
With mid-, high-level	0.840	0.839	0.842

Conclusion

The experimental results:

- Identify crucial components contributing to the mode detection and playing technique classification tasks

- Highlight the connection between the high-level structure and local elements in guqin music

For more information regarding GQ39 dataset, please visit out GitHub website: