

CHORD JAZZIFICATION

LEARNING JAZZ INTERPRETATIONS OF CHORD SYMBOLS

Tsung-Ping Chen¹, Satoru Fukayama², Masataka Goto³, and Li Su¹

¹Music and Culture Technology Lab, Institute of Information Science, Academia Sinica, Taiwan
²National Institute of Advanced Industrial Science and Technology (AIST), Japan



INTRODUCTION

- ◆ Chord symbols provide basic guides to musical harmony
- ◆ There are various ways to interpret chord symbols



RESEARCH AIM

- ◆ Generating jazz voicings from chord symbols (**triads only**)

METHODS

- ◆ Compiling chord voicings in pop-jazz sheet music
- ◆ Two-stage jazzification (**coloring** and **voicing**) using supervised learning

THE NEW DATASET

- ◆ 50 pop-jazz piano solos
- ◆ 6700 chords (796 phrases)

Sheet music
 Coloring
 Roman numeral analysis
 Voicing



THE ANNOTATIONS



Chord Symbol	Db:M	D#m	Gb:M7	D#M	Bb:m7	Ebm:7	Ab:7	G:dim	Fm:7
Time: Onset	55	57	59	62	63	64	65	66	67
Time: Duration	2	2	3	1	1	1	1	1	2
Voicing	(Db2,Db3,F4,Ab4,Eb5)	(D#2,D#3,G#3,Bb4,F5)	(Gb2,Db3,Ab3,Bb4,F5)	(D#2,D4,F#4,Bb4)	(Bb2,Ab4,C5,D#5,F5)	(Ab2,Gb3,Bb4,Db5)	(F2,Bb3,C4,Ab4)	(Bb3,Gb4,Db5)	(Gb2,Gb3,A4,Eb5)
Coloring	(2)	(2)	-	(#5,b6)	(#1,9)	(#5)	(#3,#5,9,11)	(#5,d7)	-
Roman Numeral Analysis	Db: I	VII*	IV*	bIII,	II ₄ - 3	V ₇	viII ₄	III ⁷	
Structure: Phrase	B1	B1	B1	B1	B1	B1	B1	B1	B1
Structure: Measure	15	15	16	16	17	17	17	17	18
Structure: Metrical Position	0	2	0	3	0	1	2	3	0

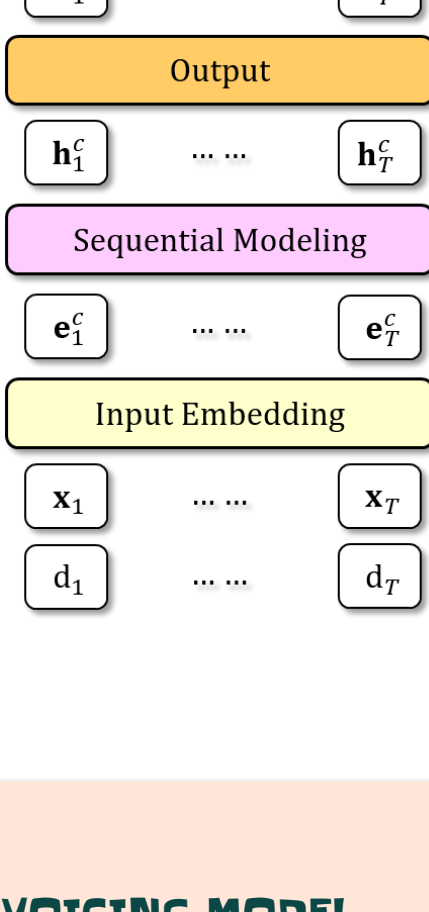
TWO-STAGE CHORD JAZZIFICATION

COLORING

Selecting a set of pitch classes (\mathbf{p}_i^c) and a bass (\mathbf{b}_i^c) for the rendering of each chord symbol

VOICING

Selecting in which octaves each pitch class is actually played on the keyboard (\mathbf{v}_i)



COLORING MODEL

Input
 $\mathbf{x}_i \in \mathbb{R}^{12}$ = chroma representation of the i th chord
 d_i = duration of \mathbf{x}_i

Output
 $\mathbf{p}_i^c \in \mathbb{R}^{12}$ = sigmoid-activated **pitch class** vector
 $\mathbf{b}_i^c \in \mathbb{R}^{12}$ = softmax-activated **bass** vector

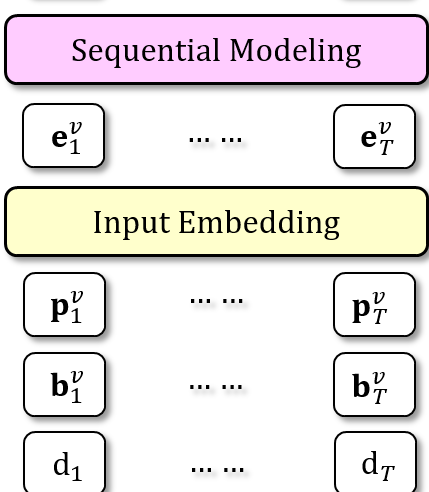
Computation
 $\mathbf{e}_i^c = \mathbf{W}^c(d_i, \mathbf{x}_i)$, (Input Embedding)
 $\mathbf{h}_i^c = f^c(\mathbf{e}_i^c | \mathbf{e}_{1:T}^c)$, (Sequential Modeling)
 $\mathbf{p}_i^c = \text{sigmoid}(\mathbf{W}^p \mathbf{h}_i^c)$, (Output)
 $\mathbf{b}_i^c = \text{softmax}(\mathbf{W}^b \mathbf{h}_i^c)$,

VOICING MODEL

Input
 $\mathbf{p}_i^v \in \mathbb{R}^{12}$ = pitch class vector
 $\mathbf{b}_i^v \in \mathbb{R}^{12}$ = bass vector
 d_i = duration of \mathbf{x}_i

Output
 $\mathbf{v}_i \in \mathbb{R}^{88}$ = **voicing** vector

Computation
 $\mathbf{e}_i^v = \mathbf{W}^v(d_i, (\mathbf{p}_i^v \oplus \mathbf{b}_i^v))$, (Input Embedding)
 $\mathbf{h}_i^v = f^v(\mathbf{e}_i^v | \mathbf{e}_{1:T}^v)$, (Sequential Modeling)
 $\mathbf{v}_i = \text{sigmoid}(\mathbf{W}^v \mathbf{h}_i^v)$, (Output)



EVALUATION ON THE PROPOSED DATASET

\mathbf{p}_i^c
PITCH CLASSES
 Multi-label classification Accuracy

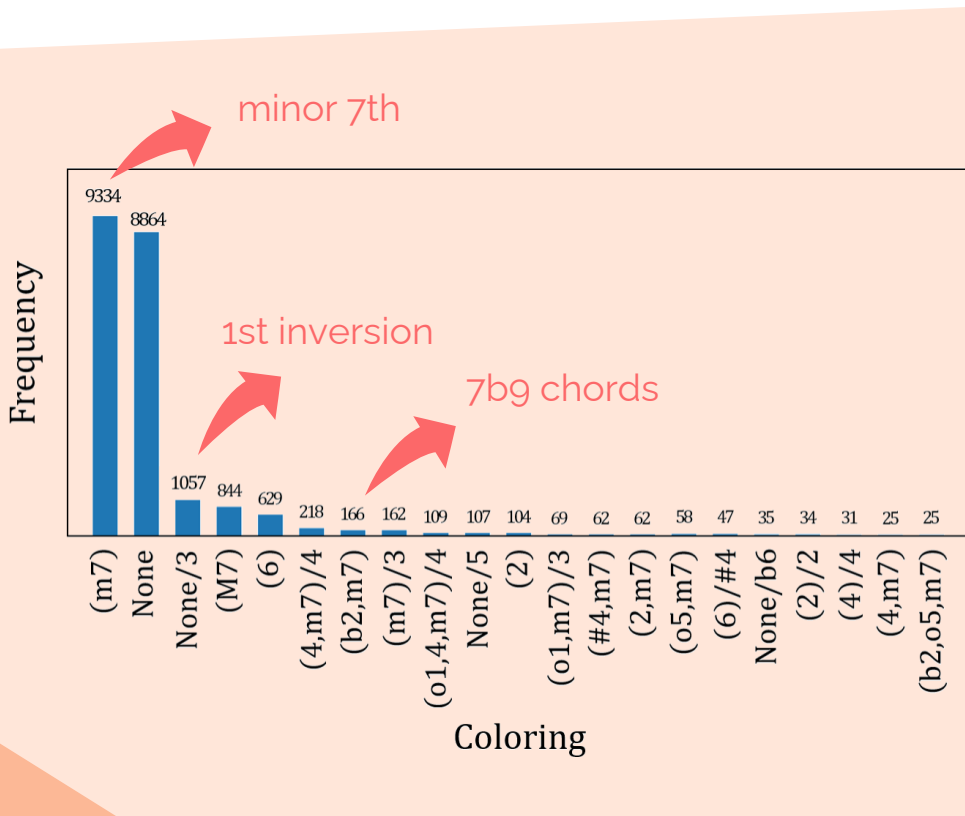
\mathbf{b}_i^c
BASS
 Multi-class Classification Accuracy

\mathbf{v}_i
VOICING
 Multi-label classification Accuracy

Result

SEQUENCE MODELLING	COLORING		VOICING
	BASS	PITCH CLASSES	
BI-DIRECTIONAL LSTM	81.87	76.52	63.64
MULTI-HEAD SELF-ATTENTION	80.78	77.02	64.86
END-TO-END	-	-	37.87

CHORD JAZZIFICATION TEST ON THE JAZZ AUDIO-ALIGNED HARMONY (JAARH) DATASET



JAZZIFICATION EXAMPLES

CONCLUSION

- ◆ By formulating the chord jazzification as two-stage classification problems, we showed that the proposed framework is capable of generating jazz voicings directly from chord symbols.
- ◆ The chord jazzification can be applied to scenarios such as music performing and composing.

THANK YOU

Do you have any questions?

Tsung-Ping Chen
 tearfulcanon@iis.sinica.edu.tw



CREDITS: This presentation template was created by [slidesgo](#), including icons by [Freemove](#), and infographics & images by [Freemove](#).