Paper ID: 3918

Difficulty-Aware Neural Band-to-Piano Score Arrangement Based on Note- and Statistic-Level Criteria

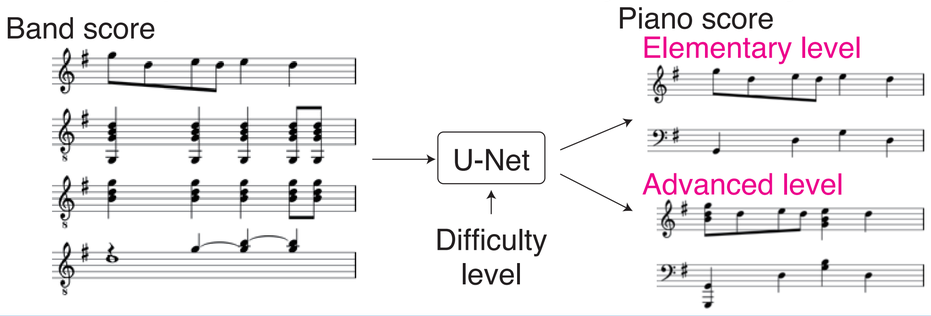
Moyu Terao[1], Yuki Hiramatsu[1], Ryoto Ishizuka[1], Yiming Wu[1], Kazuyoshi Yoshii[1,2]

[1]Graduate School of Informatics, Kyoto University, Japan
[2]PRESTO, Japan Science and Technology Agency (JST), Japan

Summary

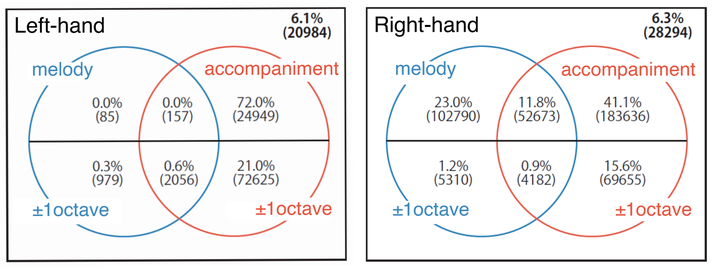
Automatic piano arrangement

- Train a U-Net conditioned by a given difficulty level in a supervised manner
- Reduce an augmented band score obtained by up- and down-shifting an original band score by one octave
- Output a piano score conditioned by a difficulty level



Problem and Approach

The “ground-truth” arrangement cannot be uniquely determined.
We train a **U-Net that estimates masks used for selecting necessary notes from an augmented band score** such that the estimated piano score is made close to the ground-truth score **at both the note- and statistic- levels** (e.g. polyphony level, polyphony width, and note density).

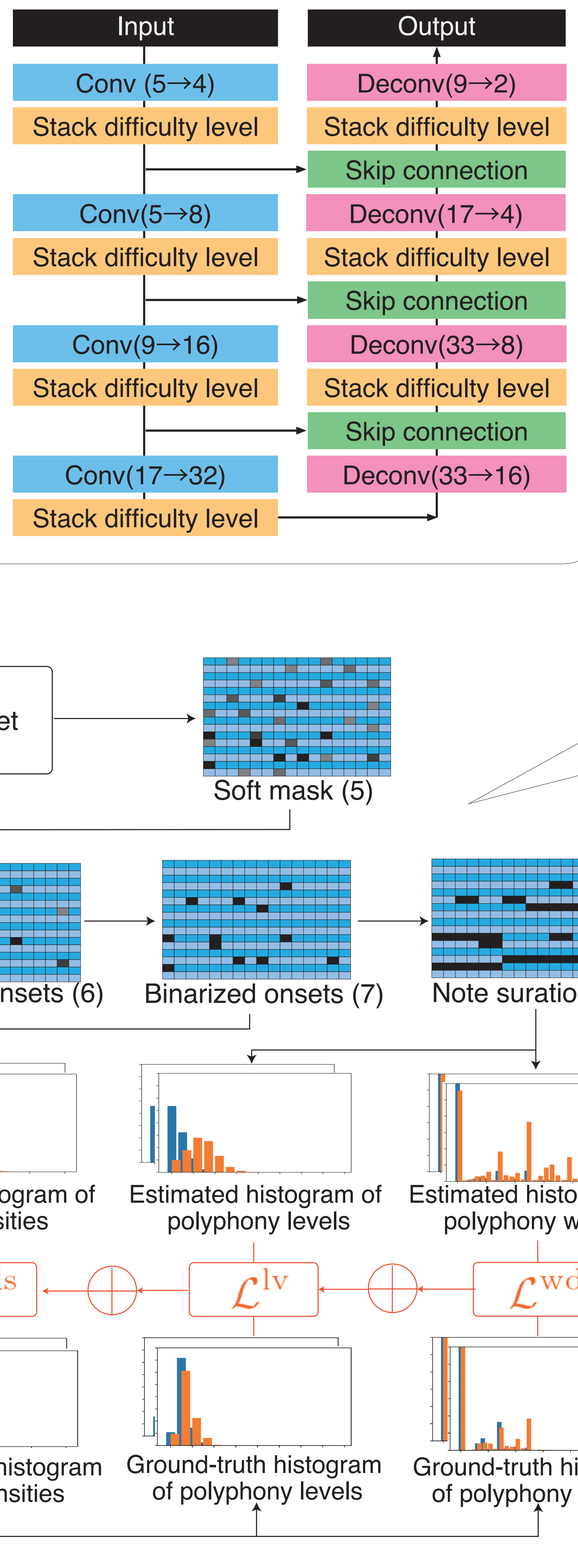


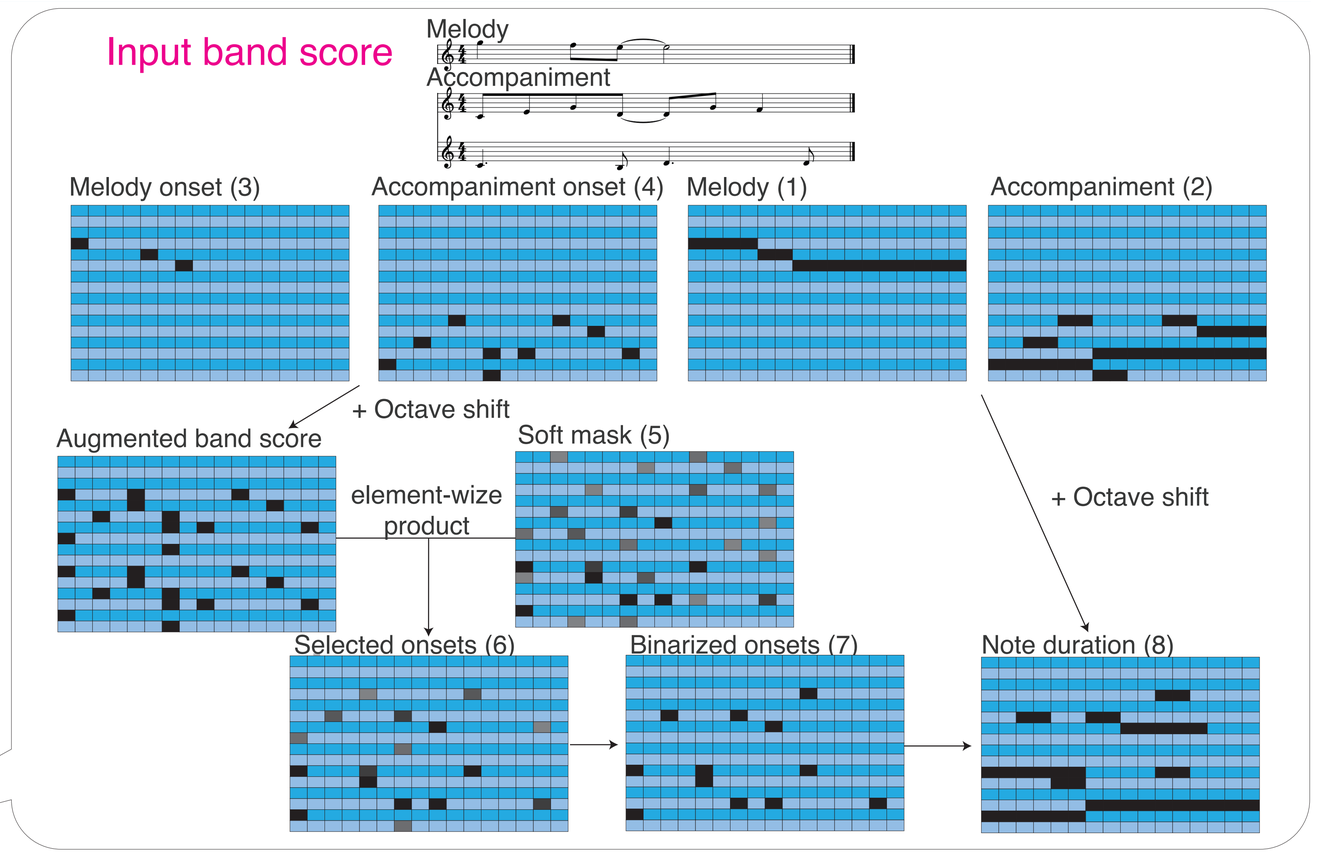
← We investigated the origins of the left- and right-hand notes of piano scores and found that a reasonable piano score can be obtained by selecting necessary notes from an augmented band score.

Proposed Method

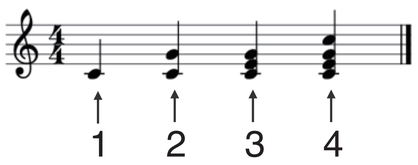
The architecture of U-Net.

We stack a difficulty level channel after convolution and deconvolution.

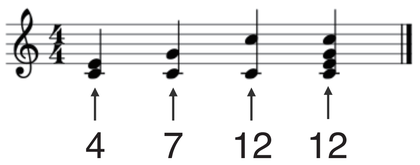




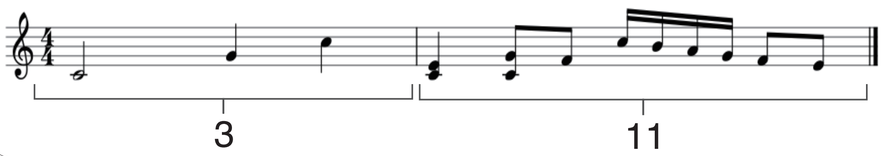
Polyphony level: Number of concurrent notes



Polyphony width: Interval between the highest and lowest pitches



Note density: Per-measure number of notes



Evaluation

Experimental Results

Dataset: 184 pairs of Japanese band and piano scores
(85: Elementary level, 99: Advanced level)

Tatum-level onset matching rates (heigher is better): $\mathcal{F}$

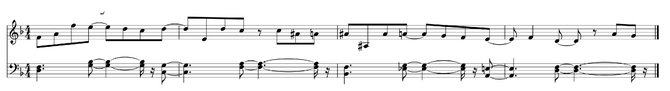
Statistic-level losses (lower is better): $\mathcal{L}^{\mathrm{lv}}$, $\mathcal{L}^{\mathrm{wd}}$, and $\mathcal{L}^{\mathrm{ds}}$

Loss function	$\mathcal{F}[\%]$ Left	$\mathcal{F}[\%]$ Right	$\mathcal{L}^{\mathrm{lv}}$ $(\times 10^4)$	$\mathcal{L}^{\mathrm{wd}}$ $(\times 10^4)$	$\mathcal{L}^{\mathrm{ds}}$ $(\times 10^4)$
$\mathcal{L}^{\mathrm{nt}}$	25.6	56.1	20	26	0.78
$\mathcal{L}^{\mathrm{nt}}+\beta^{\mathrm{lv}}\mathcal{L}^{\mathrm{lv}}$	26.6	59.3	**8**	15	0.75
$\mathcal{L}^{\mathrm{nt}}+\beta^{\mathrm{wd}}\mathcal{L}^{\mathrm{wd}}$	26.4	58.5	10	19	0.80
$\mathcal{L}^{\mathrm{nt}}+\beta^{\mathrm{ds}}\mathcal{L}^{\mathrm{ds}}$	27.2	56.4	33	42	**0.54**
$\mathcal{L}$	**27.8**	**59.7**	10	**13**	0.67

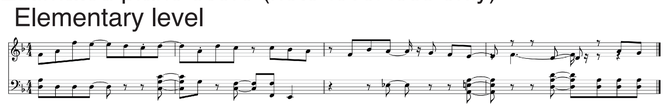
- The best matching rate was achieved when the total loss $\mathcal{L}$ was minimized
- We confirmed the effectiveness of each statistic-level loss in improving $\mathcal{F}$ and reducing $\mathcal{L}^{\mathrm{lv}}$, $\mathcal{L}^{\mathrm{wd}}$, and $\mathcal{L}^{\mathrm{ds}}$

Examples of Piano Arrangement

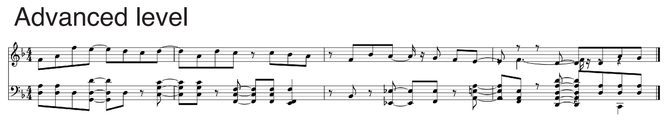
Ground-truth piano score
Elementary level



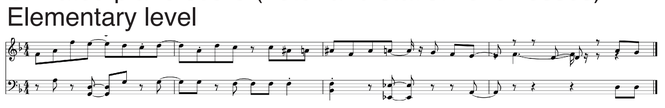
Estimated piano score (note-level loss only)
Elementary level



Advanced level



Estimated piano score (note- and statisti-level losses)
Elementary level



Advanced level