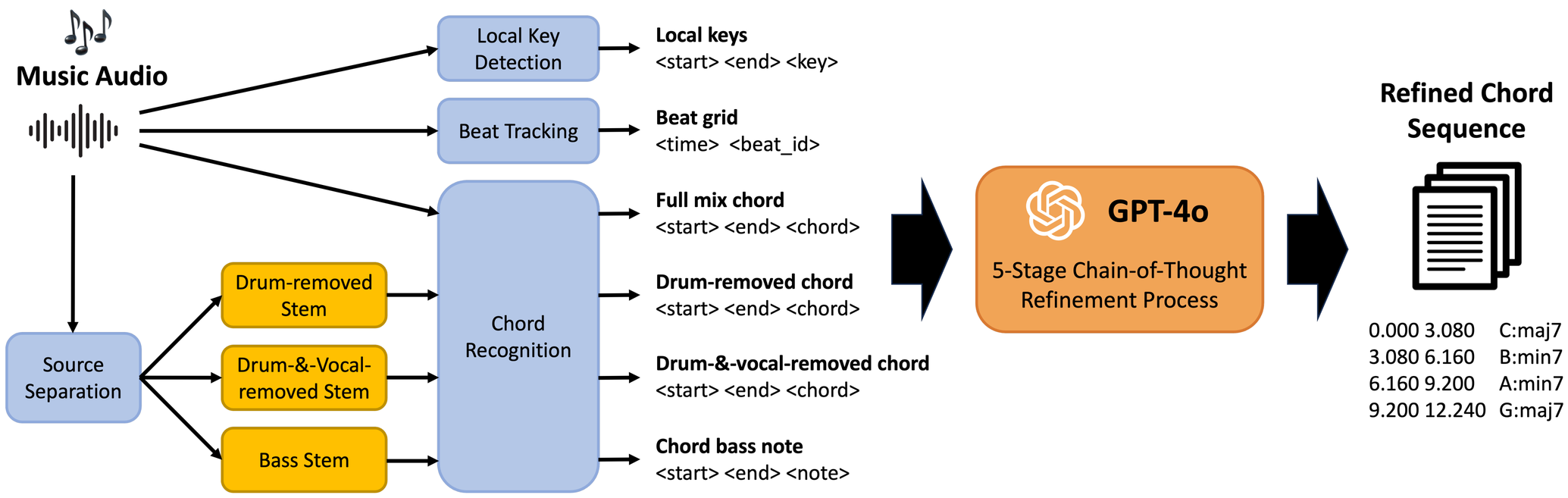
Enhancing Automatic Chord Recognition Through LLM Chain-of-Thought Reasoning

Chih-Cheng Chang[1], Bo-Yu Chen[2], Lu-Rong Chen[2], and Li Su[1]
[1]Institute of Information Science, Academia Sinica, Taiwan
[2]Rhythm Culture Corporation, Taiwan



- **Systematic Multi-Tool Integration**: Leverages diverse MIR tools—including source separation, key detection, chord recognition, and beat tracking—in a coordinated, stage-wise process.
- **Cross-Modal Transformation:** Audio-derived outputs are converted into standardized textual formats, enabling LLM to read, reason, and apply music-theoretical knowledge.
- **Reasoning-Driven Refinement:** Through a 5-stage Chain-of-Thought framework, the LLM systematically detects errors, reconciles inconsistencies, and refines chord sequences for higher accuracy.

Background

- **Limitations of Current Models**: Deep learning–based audio chord recognition systems often act as black boxes, offering good performance but limited interpretability.
- **Harmony Has Strong Theory**: Chord and harmony theory is grounded in well-established musicological principles, which current deep models struggle to explicitly reflect or explain.
- **LLMs for Reasoning**: Large Language Models excel at textual understanding and reasoning—raising the question of whether they can enhance MIR tasks while providing greater interpretability.

5-stage Chain-of-Thought Refinement

- **Music Source Separation (MSS)**

  Compare chord recognition results from different separated audio tracks; select the most reliable sequence as the base.
- **Bass Correction**:

  Refine chord roots using bass stem information, ensuring consistency with tonal context.
- **Key Correction:**

  Cross-check chords against local key estimation and reference outputs; apply conservative theory-driven corrections.
- **Anomaly Detection:**

  Identify implausible or missing chords (e.g., “N” labels) and systematically revise them with explicit reasoning.
- **Beat Alignment:**

  Align chord boundaries to detected beats for musically precise timing.

Experiment

- **MIR Tools Used**

Source separation: HT Demucs [1]
Chord recognition: Large-vocabulary ACR model [2]
Key detection: Local key estimation [3]
Beat tracking: Beat This! [4]

- **Large Language Model**

GPT-4o, used without fine-tuning

- **Dataset**

IdolSongsJp: 15 professionally composed idol-style songs [5]
UsPop2002 subset: 192 full-length Western pop songs
In-house set: 20 Western pop choruses with expert chord labels

Results

Dataset	Stage	MIREX	Root	Majmin	Thirds	Triads	Sevenths
IdolSongsJp	Baseline	79.50	80.91	80.52	77.73	74.93	65.72
	MSS	80.67	82.29	82.35	79.48	76.67	67.11
	Bass Correction	80.71	82.29	82.35	79.48	76.67	67.11
	Key Correction	80.69	82.29	82.35	79.48	76.67	67.11
	Anomaly Detection	**81.16**	**82.77**	**82.82**	**79.92**	**77.11**	**67.48**
	Beat Alignment	80.73	82.44	82.27	79.56	76.70	67.00
UsPop2002	Baseline	80.07	82.52	82.06	79.55	72.47	72.55
	MSS	80.85	83.22	82.94	80.04	73.16	73.20
	Bass Correction	80.84	83.22	82.92	80.04	73.13	73.06
	Key Correction	80.89	83.24	82.96	80.35	73.18	73.20
	Anomaly Detection	80.95	83.36	83.04	80.47	73.25	73.24
	Beat Alignment	**81.13**	**83.54**	**83.21**	**80.64**	**73.42**	**73.39**
In-house dataset	Baseline	83.29	80.25	80.79	79.48	78.89	67.17
	MSS	84.12	81.63	81.01	79.97	79.52	73.91
	Bass Correction	84.04	81.54	80.92	79.88	79.42	73.80
	Key Correction	85.29	82.29	82.28	81.13	80.67	74.15
	Anomaly Detection	85.93	83.05	82.63	81.57	80.98	74.49
	Beat Alignment	**86.06**	**83.30**	**82.81**	**81.79**	**81.17**	**74.64**

Reference

[1] S. Rouard *et al.*, “Hybrid Transformers for Music Source Separation,” ICASSP 2023
[2] J. Jiang *et al.*, “Large-Vocabulary Chord Transcription via Chord Structure Decomposition,” ISMIR 2019
[3] H. Schreiber *et al.*, “Local Key Estimation in Classical Music Recordings,” ICASSP 2020
[4] F. Foscarin *et al.*, “Beat This! Accurate Beat Tracking without DBN Postprocessing,” ISMIR 2024
[5] H. Suda *et al.*, “IdolSongsJp Corpus: A Multi-Singer Song Corpus in the Style of Japanese Idol Groups,” ISMIR 2025