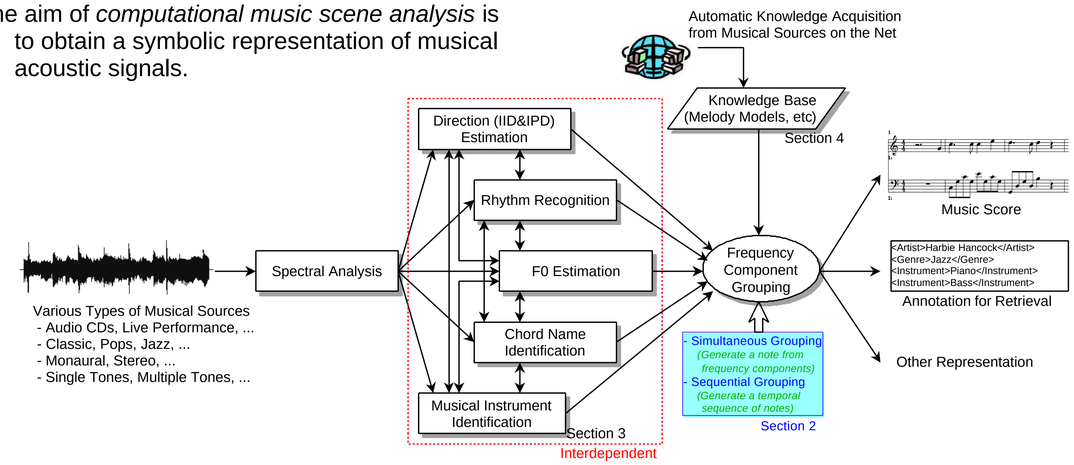
Computational Music Scene Analysis

(Okuno Laboratory, Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University)

1. What is Computational Music Scene Analysis?

The aim of *computational music scene analysis* is to obtain a symbolic representation of musical acoustic signals.



- Computational music scene analysis should support various types of musical sources.
- Computational music scene analysis should be useful for various applications.

⇒ Ontology-based Information Integration

Ontology-based Information Integration can take into consideration interdependency of processing modules.

- If visual information of musical sources is available, it may be used to disambiguate auditory information.
- If directional information of musical sources is available (musical sources is stereo), it may reduce ambiguities in frequency component grouping. ⇒ See Section 2
- Since timbres of musical instruments depend on the pitch, their identification should take into consideration the pitch dependency. ⇒ See Section 3
- A melody model defined by musical interval transitions is an important knowledge for frequency component grouping, which has a role of a language model in automatic speech recognition. ⇒ See Section 4

3. Musical Instrument Identification

～For Improving Accuracy～

Problem:

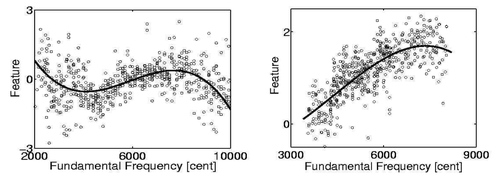
Timbres of musical instruments depend on the pitch.

⇒ Musical instrument identification should take into consideration the pitch dependency.

Solution:

(1) Approximate the pitch dependency of features as a function of F0

(2) Classify based on a discriminant function defined by the function of F0



Experimental Results:

- The use of the pitch dependency improved accuracy of musical instrument identification: 75.73% ⇒ 79.73%
- Piano's identification was the best improved: 74.21% ⇒ 83.27%

(The database of 6,247 solo tones of 19 instruments was used)

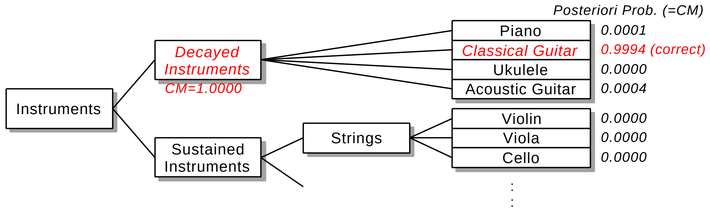
～For Flexible Identification～

Problem and Solution:

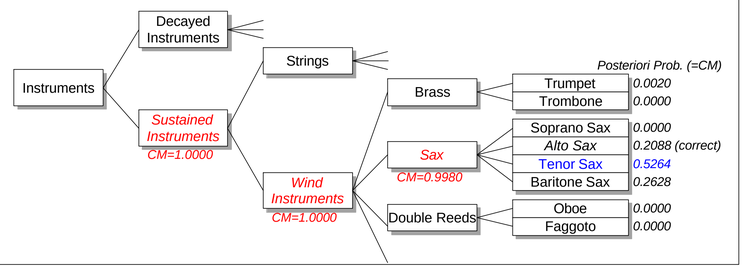
Humans recognize musical instrument sounds hierarchically.
—They can identify at higher level (e.g. strings) even if they cannot identify at lower level (e.g. violin, viola)—
⇒ Hierarchical identification with a confidence measure

Examples:

(a) A classical guitar's sound was identified correctly.
The identification result was accept. (Threshold=0.90)

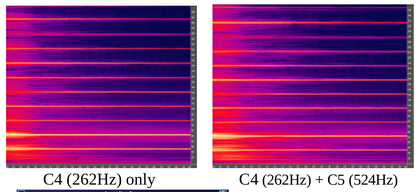


(b) An alto sax's sound was identified as "tenor sax" incorrectly.
Only the individual-level result was reject. (Threshold=0.90)



2. Frequency Component Grouping using Direction

～Simultaneous Grouping～



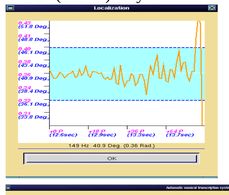
C4 (262Hz) only

C4 (262Hz) + C5 (524Hz)

Problem:

Harmonic structure is the main clue for simultaneous grouping in previous studies.

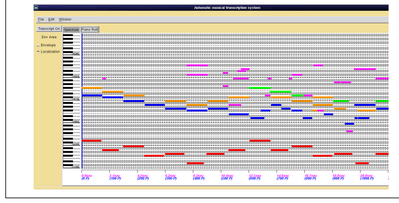
⇒ Ambiguity remains in forming notes of octave relation.



Solution:

Interaural phase difference (IPD) of overlapped frequency components has large variance.

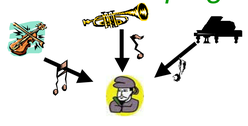
⇒ Introduce the overlap determination of frequency components based on the variance of IPD.



Experimental Results:

	Recall	Precision
Without the overlap determination	44.3%	86.7%
With the overlap determination	94.3%	75.6%

～Sequential Grouping～



Problem and Solution:

- Human listeners use directional proximity to separate multiple sounds.
- Timbre similarity have been used in previous study.

⇒Integrate directional proximity and timbre similarity by Dempster Shafer theory.

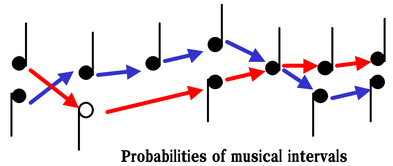
Experimental Results:

	Accuracy
Timbre similarity only	51.1%
Directional proximity only	78.3%
Integrate both the clues	89.2%

4. Melody Model (Musical Interval Transitions)

Motivation:

Human listeners use a musical context to interpret the musical sources.
⇒ A melody model, which is defined as musical interval transitions to capture the musical context, can be used in sequential grouping.

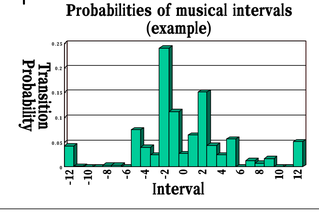


Problem:

Musical interval transitions depend on a genre, a role of each part, and an instrument.

Solution:

Construct a melody model for each genre, each role and each instrument, and then apply the most appropriate melody model to sequential grouping.



5. Conclusions and Future Works

- The framework of computational music scene analysis was presented.
  ⇒ To support various sources and various application, Ontology-based approach is important.
- As subtasks of computational music scene analysis, our studies on frequency component grouping, musical instrument identification and melody model construction were presented.
- Future works include integration of developed modules and construction of musical sound ontology.