

平成 19 年度図書館情報学海外研修助成報告書

図書館情報メディア研究科 博士後期課程 3 年

中 野 倫 靖

## 海外渡航レポート

### 1. 渡航目的

2007年9月23日から9月27日の間、オーストリアのウィーン（Vienna University of Technology）で開催された国際会議（ISMIR2007: the 8th International Conference on Music Information Retrieval）<sup>1</sup>に参加し、音楽情報検索に関する最新的话题を調査・学習するために、平成19年度図書館情報学海外研修助成を申請いたしました。

楽曲の電子化やネットワーク配信が一般化されてきた現在、膨大な楽曲（音楽情報）の集合から目的の楽曲を効率良く得るために、検索・分類・管理の新しい手法が必要とされています。「検索」は図書館情報学の重要なテーマの一つであり、様々なメディアに対して効率的かつ実用的な検索手法の実現が望まれています。すなわち、新しい音楽情報検索手法の実現は、図書館情報学における音楽メディアに対する検索の発展につながります。本会議への参加は、このような音楽情報検索の最新の研究動向を調査し、研究者と直に議論することが出来る絶好の機会といえます。また、音楽というメディアを対象とするため、論文上に書くことが出来ない実例（特に音情報）を伴った研究報告を聞くことが重要であるといえます。

申請者はこれまで、音楽的知識のない人を対象に、人が使いやすい音楽情報処理システムを音響信号処理によって実現することを目指してきました。そのために、認知心理実験によって人の音楽認知・理解を知り（Human）、音響信号処理による計算機上への実現を知り（Machine）、インタフェース構築によって人と計算機を結びつけることを知る（Interface, Interaction）、という方針で研究を行ってきました。「人」との関わりを重視したシステムを構築する、もしくは、「人」の音楽理解を計算機上に実現するという目的のためには、この3つ（H・M・I）を統一的に研究すべきだと考えています。そのような方針の下で、これまで口（くち）ドラム認識手法の実現と楽譜情報を用いない歌唱力評価手法の実現というテーマで研究を行ってきました。これらは、最終的には音楽情報検索への応用を考えており、例えば、口ドラム認識ではドラム音を真似た音声（ドンタンドドタン）を入力としたドラムパート検索への応用を、楽譜情報を用いない歌唱力評価では、歌唱力を構成する要素に着目した新しい音楽情報検索への応用を考えていました。しかし、実際は音楽情報検索のための基礎技術の実現にとどまっており、その技術を用いて具体的に検索を実現するところまでは至っていませんでした。そういった意味でも、人が音楽情報検索をどう必要としているか（Human）、それを計算機上にどのように実装するか（Machine）、どういった検索インタフェース・インタラクションが考えられるか（Interface, Interaction）を、本会議への参加を通して考察を深めたいと考えています。

---

<sup>1</sup> <http://ismir2007.ismir.net/>

## 2. ISMIR2007

本章では、今年で開催 8 回目を迎える ISMIR について、参加者・参加国、開催セッションについて述べ、また、2005 年より始まった音楽情報検索コンテスト (MIREX: Music Information Retrieval Evaluation eXchange) のタスクについて述べる。

### 2.1. 参加者・参加国

ISMIR2007 の参加者・参加国は以下の通りである。

参加者： 250 名 (うち日本 23 名)

参加国： 27 カ国

全 250 名の参加者の国別割合を図 1 に示す。

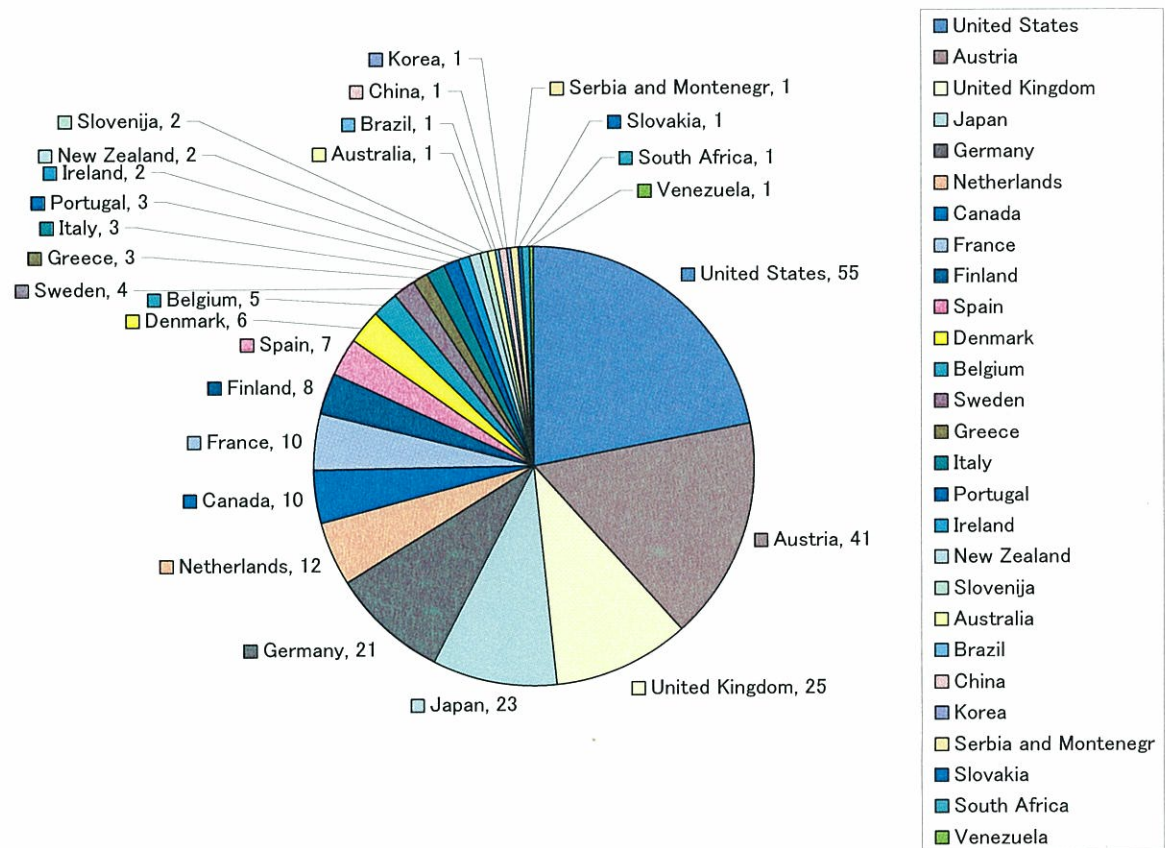


図 1 ISMIR2007 全参加者の国別割合 (全 250 名)

図 1 より、ISMIR における日本の参加率は非常に高いことが分かる<sup>2</sup>。そのような背景から、2009 年に日本 (神戸) で開催されることが、本会議で決定した (General Chairs: Masataka Goto and Ichiro Fujinaga)。

<sup>2</sup> あくまで私見であるが、日本の研究者の発表は完成度が高く、海外の研究者からの評価も高かった印象を受けた。

## 2.2. セッション数、及び論文数

どのようなセッションが開催されたかの把握は、現状のホットトピックを理解することにつながるため、本節ではセッションについて詳細を述べる。ISMIR2007での、セッション数及び論文数は以下の通りである。

セッション数及び論文数：

オーラルセッション： 15 セッション、62 件

ポスターセッション： 2 セッション、65 件（MIREX ポスターは除く）

また、セッションのそれぞれについて、セッション名と論文数を表 1 に示す。

表 1 ISMIR2007 におけるセッションの詳細

| セッション名  | 論文数 | 発表の種別 |
|---|-----|-------|
| Similarity 1  | 2   | オーラル  |
| Self-Similarity and Structure                                   | 5   | オーラル  |
| Genre Classification  | 4   | オーラル  |
| Recommendation  | 4   | オーラル  |
| Poster Session 1  | 32  | ポスター  |
| Tonality  | 4   | オーラル  |
| Instruments   | 5   | オーラル  |
| Chords  | 4   | オーラル  |
| Alignment and Segmentation                                      | 4   | オーラル  |
| Poster Session 2  | 33  | ポスター  |
| Transcription and Multipitch Estimation                         | 4   | オーラル  |
| Semantics and Ontologies  | 3   | オーラル  |
| Classification  | 5   | オーラル  |
| Content-Based Retrieval   | 4   | オーラル  |
| Similarity 2  | 4   | オーラル  |
| Performance, User Access<br>and OMR (Optical Music Recognition) | 5   | オーラル  |
| Annotation and Evaluation                                       | 5   | オーラル  |
| 計 1   | 62  | オーラル  |
| 計 2   | 65  | ポスター  |

音楽情報検索のために実現すべき項目は、次の4つがあると考えられる。(1) 音楽を理解する(解析の基礎技術、音楽構造解析)、(2) 検索のための音楽間関係を理解する(類似度の設定)、(3) 検索を実現する、(4) 検索結果を評価する。この4分類について、表1に挙げたセッションを分類して以下に示す。

- |   |                     |
|---|---------------------|
| 1) 音楽を理解する                                |                     |
| - Self-Similarity and Structures          | : 音楽構造解析            |
| - Tonality                                | : 調推定               |
| - Instruments                             | : 音源同定              |
| - Chords                                  | : コード推定             |
| - Alignment and Segmentation              | : 音楽構造解析 (楽譜との対応付け) |
| - Transcription and Multipitch Estimation | : 多重ピッチ推定           |
| - Semantics and Ontologies                | : タグ付け、音楽オントロジー     |
| - Performance, User Access and OMR        | : 演奏解析、メタデータ、楽譜認識   |
| 2) 音楽間関係を理解する                             |                     |
| - Similarity                              | : 楽曲 (メロディ、音色) 間類似度 |
| - Genre Classification                    | : 楽曲⇔音楽ジャンルの対応付け    |
| - Classification                          | : 楽曲⇔楽曲の対応付け        |
| 3) 検索を実現する                                |                     |
| - Recommendation                          | : 楽曲推薦              |
| - Content-Based Retrieval                 | : コンテンツベース検索の実現     |
| 4) 検索結果を評価する                              |                     |
| - Annotation and Evaluation               | : 知覚⇔物理量、アノテーション    |

以上の分類結果からは、以下の2つが言える。

I. 検索実現のための4タスクでそれぞれ研究がある

II. 検索実現のための基礎技術(1),(2)が盛んに行われている状況である

ただし、会議に参加した印象からは、(3),(4)も活発になってきている印象があった。例えば、Content-Based Retrieval セッションでは、市販CDのボーカルの声質に着目した検索が発表され(日本人)、検索システムも実現していた。また、興味深かったのが、Annotation and Evaluation セッションでの楽曲のタグ付け手法で、画像のタグ付けにおける ESP Game<sup>3</sup>を音楽に適用したものであった。これは、Web上で不特定多数の人にタグ付け(instruments, emotions等)させる手法であり、複数人が同じタグをつけた場合、そのタグは信頼性が高いとする考え方である。

<sup>3</sup> L. von Ahn and L. Dabbish, Labeling Images with a Computer Game, In ACM CHI, 2004. (Webサイト: <http://www.espgame.org/>)

### 2.3. MIREX におけるタスク

MIREX (Music Information Retrieval Evaluation eXchange) は、2005 年 (MIREX2005) から始まった音楽情報検索のコンテストであり、今年で 3 回目を迎える<sup>4</sup>。ISMIR においてその結果が公開され<sup>5</sup>、それぞれの技術についてポスターセッションが設けられた。テキスト情報検索分野における TREC (Text REtrieval Conference)<sup>6</sup>と同様に、共通データセットに対してその性能を評価するもので、このようなコンテストにより、検索システムの性能向上が期待される。また、MIREX におけるタスクは、セッションと同様に、音楽情報検索におけるホットトピックに結びつくため、ここで述べる。

表 2 に MIREX2005～MIREX2007 (及び、ISMIR2004) のタスクを、また表 3 に、各タスクの開催毎の変遷を示す。

表 2 MIREX におけるタスク一覧 (タスクごとに色分け)

| ISMIR2004                                      | MIREX2005                      | MIREX2006                            | MIREX2007   |
|--|--------------------------------|--------------------------------------|---|
| Genre Classification/<br>Artist Identification | Audio Artist Identification    | Audio Beat Tracking                  | Audio Artist Identification   |
| Melody Extraction                              | Audio Drum Detection           | Audio Melody Extraction              | Classical Composer Identification - Audio Artist Identification subtask |
| Tempo Induction                                | Audio Genre Classification     | Audio Music Similarity and Retrieval | Audio Genre Classification  |
| Rhythm Classification                          | Audio Melody Extraction        | Audio Cover Song                     | Audio Music Mood Classification   |
|  | Audio Onset Detection          | Audio Onset Detection                | Audio Music Similarity and Retrieval                                    |
|  | Audio Tempo Extraction         | Audio Tempo Extraction               | Audio Onset Detection   |
|  | Audio and Symbolic Key Finding | QBSH (Query-by-Singing/Humming)      | Audio Cover Song Identification   |
|  | Symbolic Genre Classification  | Score Following                      | Real-time Audio to Score Alignment (a.k.a Score Following)              |
|  | Symbolic Melodic Similarity    | Symbolic Melodic Similarity          | Query by Singing/Humming  |
|  |                                |                                      | Multiple Fundamental Frequency Estimation & Tracking                    |
|  |                                |                                      | Symbolic Melodic Similarity   |

※ 「Audio」は音響信号、「Symbolic」は MIDI などの記号形式が対象であることを示す。

<sup>4</sup> ただし、ISMIR2004 では「ISMIR2004 Audio Description Contest」という名前でコンテストがあった。

<sup>5</sup> [http://music-ir.org/mirex2007/abs/MIREX2007\\_overall\\_results.pdf](http://music-ir.org/mirex2007/abs/MIREX2007_overall_results.pdf)

<sup>6</sup> <http://trec.nist.gov/>

表 3 MIREX におけるタスク一覧（開催毎の変遷）

| タスク  | '04 | '05 | '06 | '07 |
|--|-----|-----|-----|-----|
| Audio Artist Identification                                  | ■   |     |     | ■   |
| Classical Composer Identification                            |     |     |     | ■   |
| Audio Genre Classification                                   | ■   |     |     | ■   |
| Audio Melody Extraction                                      | ■   | ■   |     |     |
| Audio Tempo Extraction                                       | ■   | ■   |     |     |
| Rhythm Classification  | ■   |     |     |     |
| Audio Drum Detection   | ■   |     |     |     |
| Audio Onset Detection  |     | ■   | ■   |     |
| Audio Beat Tracking  |     |     |     |     |
| Audio Music Similarity and Retrieval                         |     |     | ■   |     |
| Audio Cover Song   |     |     | ■   |     |
| QBSH (Query-by-Singing / Humming)                            |     |     | ■   |     |
| Score Following (Audio to Score-Alignment)                   |     |     | ■   |     |
| Audio Music Mood Classification                              |     |     |     | ■   |
| (Audio) Multiple Fundamental Frequency Estimation & Tracking |     |     |     | ■   |
| Audio and Symbolic Key Finding                               | ■   |     |     |     |
| Symbolic Genre Classification                                | ■   |     |     |     |
| Symbolic Melodic Similarity                                  |     | ■   | ■   |     |

表 2, 3 からは、MIREX2007 では、メロディー抽出やテンポ推定といった基礎技術<sup>7</sup>から、ジャンル識別やカバー曲<sup>8</sup>推定・ムード<sup>9</sup>識別などの「音楽情報検索の実現」へ主眼が置かれていると言える。また、多重音源のピッチ（基本周波数 F0: Fundamental Frequency）推定というタスクが設定され、一昔前では考えられなかった多重ピッチ推定が扱えるようになってきた現状を反映している。

また、ほとんど全てのタスクが、音響信号を対象にしていた。これは、システムを実際に運用するシーンを想定して、音響信号（市販 CD）を扱う重要性が考慮されているのではないだろうか。ただし、音響信号処理・記号処理の両方について、まだまだ発展の余地があり、今後はそれらで得た知見を融合していくことも重要であろう。

<sup>7</sup> メロディー抽出やテンポ推定は、扱う音源が多重演奏（市販 CD）であり、音楽音響信号処理としては非常に高度（かつ高水準）な技術であるが、「音楽情報検索のための」基礎技術といえるだろう。ここでは、議論の対象が音楽情報検索なので、基礎技術として記述した。

<sup>8</sup> 同一曲・異アーティストの曲。

<sup>9</sup> rollicking, cheerful, fun, sweet, amiable/good natured, 等。

### 3. 成果及び感想

実際に会議に参加することで、プレゼンテーションやポスターセッションを通して、論文を読むだけでは分からない実例（音情報）に触れることができ、論文を読むだけよりも深い理解が出来たと感じている。また、会議で発表を聞くだけでなく、Conference DinnerやBanquetへの参加において、著名な研究者の方々と話せたことで顔や名前を覚えてもらったことは、非常に有意義であったと感じている。本を好きであるだけでは一流の司書であると言えないように、一流の研究者はコミュニケーション能力に長けている必要があると考えている。ISMIR2007への参加によって、親しくなれた一流の研究者達とのコネクションをもとに、2008年はぜひとも本会議へ論文投稿してより深い議論を行えること、さらに広く名前を覚えてもらえることを目指す。そこから、2009年の日本開催では、ISMIRの中心に位置できるよう、今後も精進したいと感じた。

#### 添付写真



図2 ISMIR2007会場 (Vienna University of Technology)





図3 発表風景

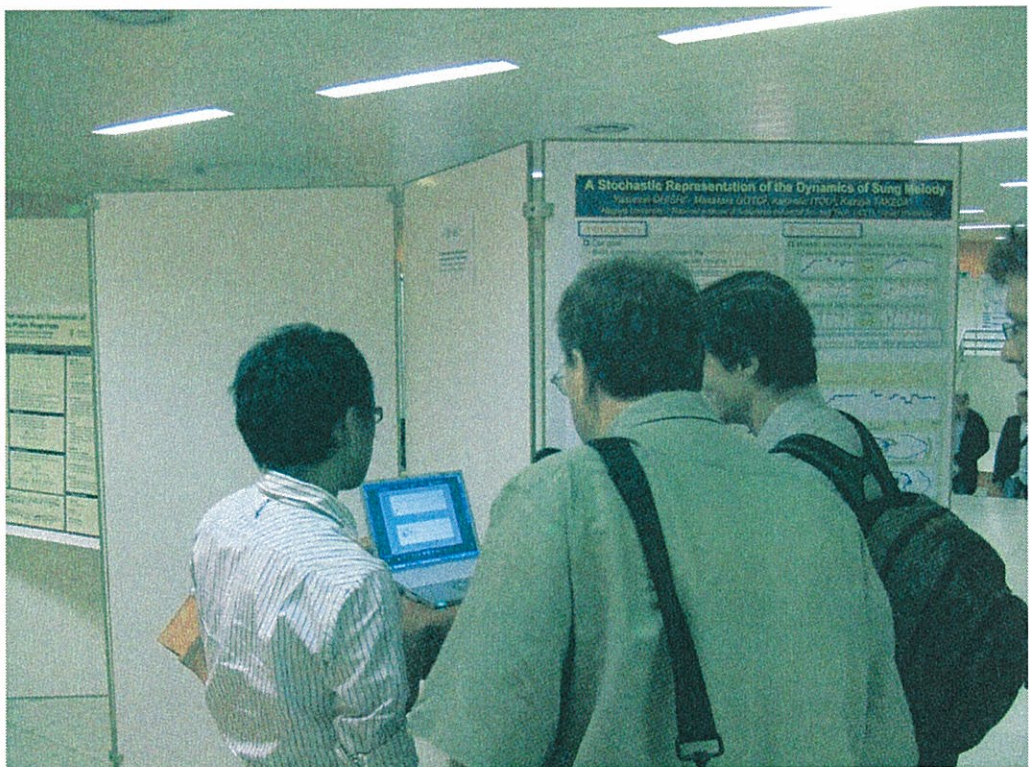


図4 ポスターセッション