

研究者海外派遣報告書（別紙）

北原 鉄朗

京都大学 大学院博士後期課程第3学年

平成18年6月5日

出張の目的

The 2006 IEEE International Conference on Acoustics, Speech, and Signal Processing（以下、ICASSP 2006）に出席し、我々の楽器音認識研究の最新成果を発表し議論するとともに、さらなる研究開発のヒントとなりうる情報を収集すること。また、ICASSP 2006と同じ時期、同じ国で行われる The 120th AES (Audio Engineering Society) Convention（以下、AES Paris 2006）に出席し、ICASSP 2006では発表の少ない音響工学に関して情報収集すること。

会議の日程と内容

ICASSP 2006は、2006年5月14～19日の6日間、フランス、トゥールーズにて行われた。14日、15日はチュートリアルのみで、本セッションは4日間であるが、オーラル6会場、ポスター8会場の平行進行となっており、1,460件の発表があった。内容は、音響信号処理全般に渡り、上述のように多数の発表があるだけでなく、当該分野で最も難しい会議と言われており、毎年採択率は50%程度に抑えられている。今年度は3,040件の投稿があつて1,460件が採択されたので、採択率は48%であつた。また、音声情報処理のみ、あるいは音楽情報処理のみに特化した会議と異なり、さまざまなバックグラウンドを持つ研究者から意見をもらうことができるという点でも重要な会議である。

AES Paris 2006は、2006年5月20～23日の4日間、フランス、パリにて行われた。AES Paris 2006は、企業によるプロオーディオ機器の展示会と通常の研究発表形式が合わさったユニークなものである。研究内容は、ICASSPに比べてよりオーディオ機器よりとなっており、スピーカーやマイクロフォンの設計、サラウンドシステムやマイクロフォンアレイの研究などが中心になっている。

ICASSP 2006の開催状況

前述の通り、ICASSPは発表件数が大変多く、特にポスターセッションは大変な盛り上がりを見せていた。多くの人が活発に議論しているため、目の前の人の発話が聞き取れないこともしばしばであつた。

音楽に関しては、前回私が参加したICASSP 2004ではポスターセッションが2つあつたにも関わらず、今回は1つ（AE-P4: Application to Music）となり、セッション数自体は減少した。これは、音楽情報検索専門の国際会議であるISMIRが定着するにつれ、投稿がそちらへ流れているものと思われる。一方、音楽と銘打っていないセッションでも音楽を題材とした研究がいくつか見られた。たとえば、“AE-L1: Audio Structure, Similarity, and Segmentation”では、Michael CaseyらやMeinard Muellerらによる音楽の類似度に関する研究発表がなされた。“MMSP-P3.5: Multimedia Database, Content Retrieval, Joint Processing and Standards”ではKeiichiro Hoashiらによる音楽情報検索に関する研究発表がなされた。“AE-L3: Audio Signal Analysis, Modeling and Coding”では、Alain de CheveigneらやGeoffroy Peeters,

Olivier Derrien による音楽音響信号処理や自動採譜に関連する研究発表がなされた。“MLSP-P4: Audio and Communication Applications”では、Christophe Rhodes らによってマルコフ連鎖モンテカルロ法を用いたオーディオセグメンテーションに関する研究発表が、Slim Essid らによって楽器音の階層的分類に関する研究発表がなされた。特に Slim Essid らとは研究のモチベーションが近く、以前我々も楽器音の階層的分類に取り組んだことから、会議前から議論を非常に楽しみにしていた。残念ながら第1著者の Slim Essid は参加していなかったが第2著者である Gael Richard とは深い議論ができた。

ポスターセッションの1つの欠点として、自分の発表と同じセッションの発表を聞けないというものがある。上述のように音楽を銘打ったセッションが1つしかなかったため、音楽に関連する多くの発表を聞けなかった。セッションの終わり際に、観客がいなくなった隙を利用していくつか聞きに行くにとどまった。Tomonori Izumitani らによる、雑音に埋もれてしまった周波数成分をマルコフ確率場を利用して修復する研究は、大変興味深かった。マルコフ確率場は画像修復に用いられており、スペクトログラムを画像とみなして画像処理技術を活用する考え方は、我々の研究にも意義深いアイデアだと感じた。



図 1: 会場前で記念写真 (左), ICASSP ポスターセッションの様子 (右)

AES Paris 2006 の開催状況

AES Paris 2006 は、前述の通り、プロオーディオ機器の展示会と通常の研究発表会とが並行に行われるユニークな会議である。しかし、展示会と研究発表会が完全に切り離されており、この特殊な状況をうまく活用しているようには見えなかった。展示会では、本当のプロオーディオに焦点を絞っており、オーディオ機器のプロではない、周辺分野の研究者にとっては楽しむのが難しいように感じた。研究発表では、スピーカーやマイクロフォン、A/D、D/A 機器の設計など、ICASSP とはかなり毛色の異なる分野が中心であった。自分とは全く問題意識の異なる研究者と知り合えたことは意義深かったといえよう。音楽情報検索に関する研究発表もいくつかあったが、こちらはあまりレベルが高いとは言えなかった。

21日の最終セッションとして、Heyser Memorial Lecture として Jim Flanagan による講演があった(題目: Spatially-Selective Sound Capture)。内容は、マイクロフォンアレイによって音の残響や雑音を除去するものが中心であった。かなり前の実験結果にも関わらず、よくできていることに感銘を受けた。

自分の発表論文の状況と反応

私は、ICASSP 2006 にて我々の楽器音認識研究の最新成果について発表した。従来の楽器音認識では、楽器を認識するに当たり、各単音(各音符に対応する処理の単位となる音)の発音時刻(いつ鳴り始め



図 2: プロオーディオ機器の展示会（左）とポスター発表会場（右）



図 3: J. Flanagan の講演（左），J. Flanagan と私（右）

たか) と基本周波数 (F_0 , どの高さの音か) を正確に推定する必要があった。しかし、さまざまな楽器音が同時に発音する状況では、これらの推定は難しく、実用化を妨げる一因となっていた。我々は、楽器音認識を「時間周波数平面における楽器存在確率の計算」と考え、この計算を、発音時刻や基本周波数を推定することなく行う方法を考案した。この楽器存在確率をグラフィック表現として可視化したものを Instrogram と呼び、楽器音認識の新たなアプローチとして ICASSP 2006 にて研究発表を行った。発表はポスター形式で行われた。

私の発表も含め、“AE-P4: Application to Music” セッションは人も多く、盛況であった。これは、音声などの非常に発表件数の多い分野と異なり、音楽関連のセッションが同じ時間帯に他のセッションでないために、人が分散しにくかったこともあろう。2時間ほぼ絶え間なく人が来て、ポスターを読んだり議論をしていた。発表内容に関しては、多くの人が面白いと言ってくれた。本研究の最も重要なポイントである「 F_0 推定をするのではなく、考えるあらゆる F_0 に対して処理を繰り返す」ことが最初伝わりにくかったので、できるだけ強調するよう努めた。本研究において PreFEst という既存技術を用いており、PreFEst の説明は完全に省いたが、知らない人もいたので、簡単な補足資料を用意するべきであった。ポスターの縮小版と論文のコピーを 20 部ほど用意し、希望者に配布した。ポスターに書ききれない細部の技術に関して議論する上で役に立った。ポスターの文字の割合をできるだけ減らし、できるだけインパクトのある図を配置し、ノートパソコンでデモを行うことで、多くの人の興味をそそるよう努めた。また、2~3 人、議

論が噛み合わない人がおり、英語力のさらなる向上が必要であることが浮彫りになった。

また、私が第2著者となっている後輩の発表も同じセッションであった。こちらも盛況であった。

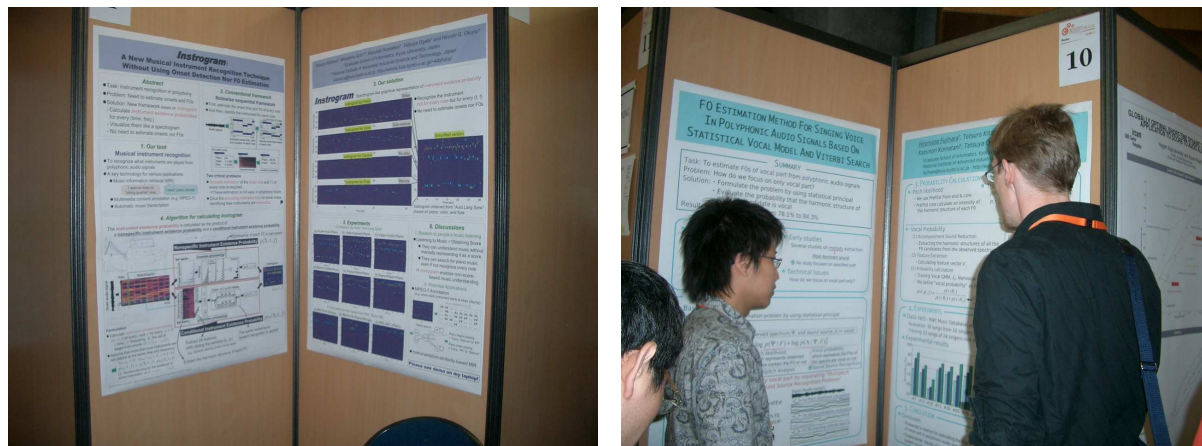


図 4: 私のポスター（左）と後輩の議論の様子（右）