

## パネル討論会「コンピュータサイエンスとしての音楽情報処理」の報告 (第 52 回全国大会シンポジウム)

片寄 晴弘 ((財) イメージ情報科学研究所), 後藤 真孝 (早稲田大学),  
堀内 靖雄 (千葉大学), 松島 俊明 (東邦大学), 村尾 忠廣 (愛知教育大学),  
志村 哲 (大阪芸術大学), 葉 孝之 (国立音楽大学), 平田 圭二 (NTT 基礎研究所)

### はじめに

1996年3月8日(金) 12:30～15:10, 電気通信大学にて開催された(社)情報処理学会第52回全国大会において, パネル討論会「コンピュータサイエンスとしての音楽情報処理」が開催された. 本資料では, そのパネル討論の内容を報告する.

本パネルの開催意図は, 文献[1]で述べた. 当日パネルでは, まず5名のパネリストのポジション発表を行った. 各パネリストの発表要旨を本資料の最後に添付する(パネリスト発表要旨-1~5(発表順)). ポジション発表に続き, フロアも交えたフリーディスカッションを行った. 次章に, そのフリーディスカッションにおける発言の要約を載せる.

### パネル討論会発言要旨

**平田(司会):**ではここで簡単に5人のパネリストのお話をまとめる. 片寄さんは自動演奏の話で, 正しい音楽, 良い音楽とは何かまだ良く分かっていない, コンテストをやって評価をしたらどうかという提案. 後藤さんのビートトラッキングの分野は, まだ個人ベースで対象や課題を広げている段階. 堀内さんの伴奏システムは, 評価は結局人間による, 今はまだ様々な曲に取り組む段階, 評価データの作成はまだ困難, 機械同士の伴奏合奏はまだ無理. 松島さんの楽譜データの表現は, 記述言語の標準化が必要だが, 記述言語で何を記述するのかという所にまだ問題が残っている. 村尾先生の音楽認知では, プログラミングすることで音楽学で言われてきた概念の検証ができるという話であった.

**志村:**楽観的には, こういった客観的評価方法から基礎的な技術が生まれて, それがいつか音楽制作, 音楽学的研究に役に立つかも知れない. 実際, 自分自身, 音情研で知り合った計算機系の研究者の方々と良い共同作業ができた. しかし悲観的には, 今日の話から工学的目標はある程度理解できたが, 音楽的な目標が理解できないし見えてこない. 音楽情報科学では, 工学的な目標と音楽的な目標がつながっているような研究テーマを設定する必要があると思う. そのためには, もっと音楽サイドの研究者がSIGMUSに加わらないと, このような目標設定はできないだろう. この研究会が私的な研究会として発足した1985年からしばらくは音楽サイドの人も多かった. しかし最近の一部の音楽サイドから会員を辞めたいという声を聞くこともあり, また多くの音楽サイドの人にとって興味の無い研究会になって来ているのではないかと

と感じる. これは音楽サイドにとって魅力的な目標が無いからであろう. 技術的な目標に関しては, 個人個人の良識と信念で決めていけばいいと思う. むしろ研究会全体としては, この会の存在意義を明確にし, 評価方法に関しては音楽的な意味, 目標も吟味した上で設定していくのが望ましいと考える. 例えば学問的な面では, 現在の音楽学や作曲の世界で何が問題となっているかということ踏まえて, また, 演奏, 伴奏についてはその今日のあるいは, 文化的な背景を考慮した上で接点となるような研究テーマを見出すべき. ただし, その目標設定の段階での判断では, 主観が前提となってくると思う.

**片寄:**自分は竹内さんと言う音楽サイドの方と共同研究をして, ツールやシステムを作っている. それらは実際に学生の表現教育にも使われている. 良いテーマを設定できれば, 上手くカップリングできると思う.

**葉:**「正しい演奏, 良い演奏」という言葉は, 音楽家から見ると言語道断である. これでは音楽サイドが逃げていってしまうだろう. 主観的な音楽を科学的に扱う難しさを認識した. しかし, ベンチマークという発想は面白い展開を生むかも知れないとも思う. 堀内さんの伴奏システムでは被験者を使ってテストしているが, 実証したり具体的な成果を出すのが必要であるというのも理解できる. 今, 音楽大学での情報処理教育について検討している. 通常, 演奏学科を主とした音楽大学で情報処理教育など必要ないと考えがちだが, 例えば, 自動伴奏システムが実用化されればオーケストラ無しに一人で協奏曲の練習などが可能になるだろうし, 記号アプリケーションを使えば声楽家は自分の声域に合わせて移調した美しいスコアを常に使う事ができるようになる. もちろん, 楽譜の自動認識システムが実用化されれば様々な目的で利用されていくであろう. しかし, 自動演奏や楽譜認識の技術には現時点ではまだまだ問題点が多いというのも事実である. 実際に数年前, MAXの「score following technique」を使って自動伴奏の実験をした. 「星に願いを」を題材に, メロディはバイオリンとサクソフォンで演奏し, コンピュータがピッチ・トラッキングをしながらMIDIピアノ(ピアノプレーヤ)で伴奏していく. ピアノプレーヤの打鍵応答時間が長かったため, シンセサイザのピアノ音に変えてみたりしたのだが, 基本的に音符情報を追っているだけでは人間のような伴奏はできないという結論に達した. 人間が伴奏している時は, 「あうんの呼吸」を感じて反応したりしているので, 音符レベルだけでなく, 音レベルや演奏時の視覚的な情報も取り入れて処理すべきである. これら伴奏システムや自動演奏

の実用化研究を促進するために、ベンチマークの設定や自動伴奏のコンクールの実施は有意義だと考える。

**片寄:**「正しい演奏, 良い演奏」という言葉で主張したかったのは、例えば 10 人が 10 人とも確実に言えるような部分を指して、客観的あるいは正しいという言葉で表現した。例えばシャッフルのノリができていのかどうかの判断などは客観的にできると考えている。

**後藤:**先ほど、菜さんの話が伴奏システムに触れた所で、演奏情報だけでは不十分で演奏者に関する視覚的な情報も必要であるということであった。これは計算機サイドの研究者も気付いている。自分は今ジャズセッションシステムを研究しているが、そこでは計算機にカメラを付けて演奏者のジェスチャーを検出したり、計算機演奏者の姿を CG で表示している。さらに将来的には、目線を合わせたりあうんの呼吸を感じたりするような高度なコミュニケーションを実現する方へ向かっていくと思う。

**平田:**堀内さんのスライドに「要素技術としての伴奏システム」という言葉があった。伴奏システムは複合システムで、さらに個々の要素技術に分解できるのではないか。その切り出した要素技術のレベルでなら、ベンチマークや工学的な評価が可能になるのではないかと思うのだが。

**堀内:**人間の行う演奏活動において、基礎的な動作の 1 つが伴奏であるから、そういう意味で使った。もちろん伴奏をさらに score following などの要素技術に分けることもできる。人間の伴奏者は独奏者の後姿だけを見ている、雰囲気を感じていろいろ反応できる。自分のシステムは、今は音の情報だけで伴奏を生成しているのでもまだ不十分だと思う。将来は、音と視覚を使って、人間のような伴奏に近づけていきたいと考えている。伴奏システムを研究開発することの音楽的な意味は、そういう所にあると思う。

**志村:**伴奏は独奏者に追従するだけのものではなく、逆に先に何かきっかけを作ることも必要である。

**堀内:**自分が実現したい合奏は正にそのようなイメージで、例えばロマン派のソナタにおける対等な独奏者と伴奏者というのが目標である。

**平賀 (フロア):**菜さんが仰っていたことの真意は、計算機システムには演奏の意図を認識したり、音楽を通して対話するという機能が根本的に欠けている、そしてそれ故に貧しい音楽になってしまっていて楽しくない、ということを描いたかったのではないかと考える。やはり、堀内さんや後藤さんのシステムでも、演奏者の意図を理解して反応するという意味では、かなり不満なのではないか。例えば、先日コンピュータと対戦した世界チャンピオンのカスパロフは対戦してもあまり楽しくなかったのではないだろうか。このギャップが象徴しているものが本質的に困難な問題なのだと思う。

**菜:**自分の真意は両方にある。後藤さんが言及してい

た計算機サイドでの研究は頼もしいと感じる一方で、本当の意味での音楽的な伴奏が可能になるかという点には疑問を持っている。

**中村 (フロア):**バイオリニストの千住真理子氏が大串先生のもとでやった卒業研究では、演奏者の演奏に込めた思いが聴取者にどれだけ伝わるかという実験が行われた。ベンチマークという観点から、このような実験をどのように考えるか。

**片寄:**演奏の表情付けのシステムにおいて、単に自動演奏を行う機能を入れるだけでは制作者のインプリメント時のアドホックな意図に縛られたものになってしまう。学習機能が重要だと思う。学習曲を決め、システムが理解できる解釈理論と特定の形容詞に基づいた演奏例を与え、その形容詞の表現法を学習するという手続きをとる。そうすれば、千住さんらの研究と同様の実験が可能だと思う。

**菜:**演奏の表情付けに関して、学習機能の研究は必須であろう。我々が以前 MAX で自動演奏をやった時はマニュアルでチューンするしかなかった。1 人の演奏家が違うタイミング違うテンポで演奏した時でも追従できるようなシステムは、ある程度までは実現できるように思う。1 つの主観に対して許容されるような自動演奏を作るシステムは可能としても、どこまで一般的に通用するか、自動化できるかは、かなり難しい問題である。

**片寄:**「正しい」というのは、提示されたスタイルと同種の表現をしている時に成立すると思う。「良い演奏」というのは、演奏に対する思い入れがある場合に関係する概念であり、かなり主観的なものだと思う。皆が同じことを感じることができるとかという点に関してはまだ良く分からない。だから、正しいということと良いということは区別して考えないといけない。

**菜:**ソロパートに対して「良い演奏」というのは無理があると思うが、伴奏に対しては、ある主観に対して工学的にフォローできるかできないかということだから、「良い伴奏」という言葉は成立すると思う。

**五十嵐 (フロア):**音楽学サイドとの共通テーマについて。MIDI プレーヤを使うと、ピアノの打鍵時刻や強弱に関するデータが正確に測定できる。それで、ブーニンや音大出身のピアニストの演奏を比較してみると、非常に似た演奏をしているのが分かる。このようなデータは音楽学サイドの方々にも興味を持ってもらえると思う。我々は数年前に声楽の伴奏システムを制作した。歌手の呼吸を計測してシステムに入力するのが良いのではないかと考えたが、センサの取り扱いが大変だったので断念した。最近アコーディオンの蛇腹操作に注目している。蛇腹は発音するより少し前に動き出すので、それを検出して伴奏システムに入力すると良いのではないかと考えている。歌曲のピアノ伴奏システムは研究として面白いと思う。自分はそのシステムを自分のコンサートに使用する予定である。また、実績作りのために、国際会

議などで伴奏システムのコンテストをやってみるのもいいだろう。村尾先生のお話を聞いていて感じたことだが、境界領域では用語の問題が大切である。例えば、拍という概念は数理科学的には確立されていない。今後は、音楽、情報の両サイドから用語に関する議論を行うべきである。

**村尾:** その拍の問題に関しては、後藤さんに質問がある。拍の持つ階層性は基本拍の認識を元に作られている。しかし速いリズムでは基本拍が認識されないということが起きる。認知心理学的には、グルーピングして 90 から 100 位のテンポになるあたりが標準的基本拍として認知されるのではないかとされている。そうすると、例えば 6/8 の曲では、どこを基本拍とするのかについてしばしば解釈が 2 通りに分かれる。ビートトラッキングの分野では基本拍という概念をどう考えているのか。

**後藤:** 音楽学で基本拍は適切に定義されていないように思えるし、ビートトラッキングの研究者の間でも基本拍の正確な定義はまだ合意されていないと思う。ビートトラッキングでは拍の持つ階層性の中のどこかのレベルが認識できれば良いという立場を取る場合もある。自分のシステムでは、人間が四分音符を感じるところという主観的な定義を与えているが、実際には心理的に抵抗を覚えないもっともらしいテンポを選ぶのに迷う場合もあるだろう。拍とは何かを探求したり、音楽心理での基本拍の概念を具体的にシステムを組んで検証するのも、ビートトラッキングの研究意義の 1 つであると考えている。

**平田:** 最後にパネリスト、コメンテータの方々から一言ずつ頂きたい。

**片寄:** このような問題の議論は継続して行いたい。結論は出ないかも知れないが、いろいろな分野の方々の意見を聞きたい。

**後藤:** AI におけるチューリングテストの音楽版、つまり、計算機の演奏なのか人間の演奏なのかをブライントテストする、ということの評価の 1 つに加えても良いと思う。平賀さんの、現状の計算機相手に演奏してもつまらないのでは、という指摘は非常に重要だと感じた。

**堀内:** 自分の伴奏システムの研究に関して、工学的にまとめたいと思う一方で、音楽的な面も追及したいと思っている。音楽、認知、計算機などの分野の人と議論の場が持てたことは良かった。音楽情報処理は、今はまだ模索しながら分野を築き上げている段階かも知れないが、こういう境界領域で総合的に進化して行くべきだと思う。システムを作って実証しながら研究を進めていきたい。

**松島:** 楽譜の認識は、他の分野に比べるとまだ客観的判断がしやすいと思う。しかし、楽譜を読むという作業には、多くの主観的な部分があるのも事実で、多くの考え方があろう。また、産学共同で楽譜の標準データベースを作る必要性も感じている。

**村尾:** 最近、演奏ルールの研究が盛んであるが、これは音楽教育に大きく貢献すると思う。音楽教育の現場で、教師は生徒の演奏に対して注意を与えるが、それらは、強く、弱く、心を込めての 3 点に集約される。演奏ルールはもちろん、そんな単純なものではない。今日、話ができなかったこととして、構成 (composite) アクセントと演奏アクセントの区別がある。例えば、構成アクセントとしての音長アクセント (durational accent) と、演奏アクセントとしての長勢アクセント (lengthening accent) は区別されなければならない。この区別は、演奏ルールの研究や自動作曲などでは特に大切な筈のだけど、あまり認識されていないように感じる。このような話を発展させて、音楽サイド、計算機サイド共通のテーマにできればと思う。

**志村:** 自分がここで話したことを否定的に受け取って欲しくない。自分の仕事にとって計算機は必須であるし、学内では情報教育の推進派であり、情報科学の重要性は認識している。しかし研究会としての SIGMUS の今後を考えると、音楽サイドと協調するのか決別するのかの選択の時期だと思う。どちらを選択しても発展する道はあると思うが、いずれにしろはつきりすべきではないか。システムのコンテストをやったとして、それが例えばゲームのようになってしまった場合、音楽に真剣に取り組んでいる音楽サイドの人間にとってはちょっと屈辱的に感じる。もしコンテストをやる場合は、多くの音楽の専門家と良く議論し、何を評価するのか、評価することの意義は何かを十分認識した上で実行して欲しい。

**葉:** ベンチマーク、標準問題という点では、楽譜認識あたりから標準データベース化すると良いと感じる。市販の楽譜認識ソフトの比較結果を知りたい。1つの分野にこだわっていても発展が望めないで、良い意味で曖昧であるべきだと思う。音楽そのものに限らず、AI や視覚情報の研究などもどんどんリンクして新しいことを始めると良いだろう。

**平田:** 最後のまとめは特に言いません。このテーマのパネルの続編は、5 月広島島の SIGMUS 研究会で開かれる予定である。そこでのパネリストは、阪大の井口先生、早大の橋本先生、NTT の小坂さん、図書館情報大の平賀先生、京都府立亀岡高校の竹内さんである。

## References

- [1] 平田, コンピュータサイエンスとしての音楽情報処理, 情報処理学会 第 52 回全国大会シンポジウム・インダストリアルセッション論文集, 1996.