

## 技術が切り拓く音楽体験の未来

フェロー 後藤 真孝  
産業技術総合研究所



技術の発展は、新たな音楽文化・表現を生み出す。技術の力で生まれたピアノやギターは多様な音楽を生んだが、楽音合成技術も既に音楽制作では不可欠となり、2007年以降は新たに歌声合成技術が注目を集めてきた。衝撃的なのは、「人間の歌声でなければ聴く価値がない」という旧来の価値観が打破され、「合成された歌声がメインボーカルの楽曲を積極的に楽しむ文化」が誕生したことである[1]。同様に今後も、技術の発展が別の既存の価値観を打破し、更に新たな文化・表現を生み出していく。いつかは「人間の作品でなければ聴く価値がない」「人間の聴衆でなければ聴いてもらう価値がない」という価値観すら変容したり[2]、[3]、人間拡張技術によって拡張された聴覚器官が楽しめるように音楽表現が大きく変化したりする可能性がある。

音楽の聴き方も技術の発展と共に大きく変わり、音楽コンテンツのデジタル化が進んで膨大に蓄積された結果、音楽配信等によってそれらを好きなときに好きなだけ楽しめる時代が到来した。しかし従来は、膨大な楽曲に受動的にアクセスできる量的な変化（楽曲の数が増える変化）が中心で、その潜在的な可能性はまだ十分には引き出されていなかった。蓄積された音楽をファイルとして配信するだけでなく、その中身に踏み込んで計算可能にし、音楽の楽しみ方がより能動的で豊かになる質的な変化（体験の質が変わる変化）こそ、デジタル化がもたらす真の価値である。そして、その変化を起こす鍵となるのが、音楽の音響信号の中身を解析する「音楽の自動理解技術」（音楽理解技術）である。

そこで、音楽のビート（時間構造）、メロディ（周波数構造）、サビ（大局的構造）等の主要な要素を推定する音楽理解技術を1990年代から時代に先駆けて研究開発してきた。更にその応用事例として、「音楽理解技術が人々の音楽体験をどのように豊かにできるか」を明らかにすべく、楽曲の構造を意識しながら盛り上がるサビだけを聴いたり、膨大な楽曲から好みの楽曲を探したりできる様々なインタフェースを、「能動的音楽鑑賞インタフェース」[4]、[5]と名付けて実現してきた。そうして「技術の使われ方も提示する」ことで、音楽理解技術が切り拓く新たな音楽体験を人々がイメージしやすくなる。しかし、インタフェースに関して論文発表をし、講演や展示、動画等で紹介することはできても、人々の日常生活で直接使ってもらうことは難しかった。

そこで「技術を直接利用可能にする」ために、一般の人々がインターネット上で気軽に利用できるWebサービス群として一般公開する研究活動を始めた。例えば、2012年にプレス発表した能動的音楽鑑賞サービス「Songle」(<https://songle.jp>)では、ニコニコ動画やピアプロ等の投稿サービス上で公開されている200万曲以上を音楽理解技術で自動解析し、音楽の中身を可視化することで人々が音楽をより深く理解して能動的に楽しめるようにしている（前述の「サビだけを聴いたり」のインタフェースも体験できる）。その上で、音楽理解技術で新たな価値を生み出して貢献するための中核となるプラットフォームを目指して、Songleの解析結果を外から利用できるWeb APIを開発・提供

することで、他の研究者・技術者・企業等が我々の技術を利用して「他の技術を創出可能にする」ことにも挑戦してきた。実際に、以下で紹介する他のサービス群や、スマホアプリ「NicoBox」等で活用されている。

楽曲の数が膨大になると、視聴者にとっては好みの楽曲の発見が難しくなり、クリエイターにとっては自分の楽曲が埋もれやすくなる問題が生じる。そこで2019年にプレス発表した音楽発掘サービス「Kiite」(<https://kiite.jp>)は、音楽推薦技術や他のユーザが薦める楽曲をサビ区間のみ効率良く試聴することで、次々と新たな楽曲に出会うことを可能にした。他にも、2013年にプレス発表した音楽視聴支援サービス「Songrium」(<https://songrium.jp>)は、膨大な楽曲の関係性を可視化して俯瞰可能にすることで新たな出会いを可能にした。

更に、音楽は聴いて楽しむだけでなく、音楽に連動させて様々な機器を制御することで、映像や照明が変化したり、ロボットが踊ったりする一体感のある音楽体験を創出することができる。そこで2017年にプレス発表した大規模音楽連動制御プラットフォーム「Songle Sync」(<https://api.songle.jp/sync>)は、多数の来場者のスマートフォンが一斉に連動してアニメーションを表示するような新たな演出や体験を可能にし、音楽ライブやイベント、花火大会等で活用されてきた。2015年にプレス発表したリリックビデオ制作支援サービス「TextAlive」(<https://textalive.jp>)は、音楽に同期した歌詞のアニメーションを手軽に制作可能にし、TextAliveを用いた作品を多くのクリエイターが発表したり、様々な音楽ライブで活用されたりしてきた。

他にも、「歌声情報処理」[6]と名付けた分野を推進し、歌声合成技術に基づく「VocaListener」[初音ミク NT]等の製品にも貢献してきた。自分の知らないあいだに、SF小説[7]に『産総研で

歌声情報処理を研究している後藤』が主人公の一人として登場していて、それを出版後に知った驚きと喜びは、研究者として最も予想しなかった出来事だった。

以上を含む「音楽情報処理」[2], [3]の研究開発を1992年から推進し、多くの方々と一緒に生み出した成果・貢献が評価され、2021年3月に電子情報通信学会から、更に同年6月に情報処理学会から、「音楽の自動理解技術とその応用に関する先駆的研究」の業績でフェロー称号を受称した。大変うれしく、励まされている。研究は決して一人の力でできるものではなく、一緒に研究開発に取り組んできた多くの仲間や、御指導・応援して下さいました多くの先生・先輩の皆様、我々の技術を活用頂く企業や利用者の皆様、学会・研究資金・所属組織等も含めた研究環境の全てに深く感謝している。これからも多くの方々と力を合わせて、ワクワクする未来を切り拓いていけるよう全力を尽くしたい。

#### 参考文献

- [1] 後藤真孝, “初音ミク, ニコニコ動画, ピアプロが切り拓いたCGM現象,” 情報処理, vol.53, no.5, pp.466–471, May 2012.
- [2] 後藤真孝, “未来を切り拓く音楽情報処理,” 情処研報 音楽情報科学, 2013-MUS-99, pp.1–9, May 2013.
- [3] M. Goto, “Frontiers of Music Information Research Based on Signal Processing,” Proc. of IEEE ICSP 2014, pp.7–14, Oct. 2014.
- [4] 後藤真孝, “音楽音響信号理解に基づく能動的音楽鑑賞インタフェース,” 情処研報 音楽情報科学, 2007-MUS-70-9, pp.59–66, May 2007.
- [5] M. Goto, “Active Music Listening Interfaces Based on Signal Processing,” Proc. of IEEE ICASSP 2007, pp.1441–1444, April 2007.
- [6] 後藤真孝, 齋藤 毅, 中野倫靖, 藤原弘将, “歌声情報処理: 歌声を対象とした音楽情報処理,” 情処研報 音楽情報科学, 2010-MUS-86, pp.1–9, July 2010.
- [7] 野尻抱介, 南極点のピアピア動画, 早川書房, 東京, 2012.