

11 音楽とロボット

水本武志 ((株) ホンダ・リサーチ・インスティテュート・ジャパン)

音楽とロボットの研究とは

音楽は聞いて楽しむだけでなく、楽器演奏や踊り、歌唱など、身体を使って表現することで楽しむこともできるメディアである。本稿では、音楽とロボットに関する研究（以降音楽ロボット研究と呼ぶ）の中でも、単独であるいは人や音楽に合わせて楽器演奏や踊りなどといった音楽表現を行うロボットの研究について紹介する。音楽ロボットが実現すれば、ロボット演奏会や、ロボットと一緒に合唱や合奏といった、新たなエンタテインメントの実現が期待できる。特に、音楽ロボットと一緒にみんなで歌ったり手拍子をしたりする参加型エンタテインメントによって生まれる非言語コミュニケーションは、言語や年齢層、文化の違いによって会話が難しい人々の交流のきっかけとなることも期待できる。

音楽ロボット研究の特徴は、身体を持つロボットを使用するという点である。身体は人の音楽の知覚と関係が深いことが知られている。たとえば、楽器演奏者の表情が聴衆の演奏楽曲の印象に影響を与えるという報告や、ピアノ演奏者の身体の動きが楽譜構造と相関を持つという報告がある（文献 1）参照。したがって、ロボットの身体表現と歌声合成や自動演奏などの音楽信号生成技術を組み合わせることで、より効果的な音楽表現が期待できる。

本稿では、音楽ロボットを音楽表現方法とインタラクションの有無から分類し、特にインタラクションを伴う音楽ロボット研究に関して、その課題と研究例を紹介する。

		インタラクションの有無	
		なし	あり
音楽表現方法	楽器演奏	独奏	合奏
	踊り	単独舞踊	足踏み・社交ダンスなど
	歌唱	独唱	合奏

表-1 音楽ロボット研究の分類

音楽ロボット研究の分類

従来の音楽ロボット研究を表-1のように音楽表現方法と外界とのインタラクションの有無で分類して紹介する。詳しい関連研究の紹介は上述の文献 1) やロボットの国際会議 IROS 2010 のワークショップ^{☆1}を参照されたい。まず音楽表現方法について、楽器演奏・踊り・歌唱のカテゴリで分類する。

• 楽器演奏

人の演奏能力の再現や新たな表現の模索を目的として広く研究されている。使用楽器も多く、たとえばフルート演奏ロボット²⁾は唇に使う素材の開発や演奏音フィードバックなどによって独奏を、ドラム演奏ロボット³⁾は人の演奏リズム構造の推定とそれに基づく即興演奏によって人との合奏を実現している。

• 踊り

力センサなどの入力から動きを実時間で推定し人と社交ダンスをするロボット⁴⁾の報告などがある。ダンスは定番のデモなので、音楽に合わせて踊るもの、人の舞踊を模倣するもの、事前準備した動作を音楽と同時に再生するものなど数多く見

☆1 <http://winnie.kuis.kyoto-u.ac.jp/RMEWS/>

られる。

•歌唱

音楽に合わせた顔動作生成によってリアルな歌唱を行うロボットの報告がある⁵⁾が、動作が必須でないため歌唱ロボットの報告は比較的少ない。さらに各カテゴリを外界とのインタラクションの有無で分類する。楽器演奏を例に挙げると、インタラクシ

ョンのない独奏では上手い演奏のための機構や制御が主な課題となり、インタラクションのある合奏では他者や音楽に合わせるための情報抽出や予測が主な課題となる。本稿では主に後者を取り上げる。

外界とのインタラクションがある場合に共通の2つの問題がある。1つめは、ロボット制御の時間的制約の強さである。ソフトウェアによる伴奏音生成やCGによるダンス動作生成と対照的に、ロボットによる表現は物理的動作を伴うので必ず遅延が発生する。したがって、ロボット制御システムには音楽信号から実時間で次のタイミングを予測し、そのタイミングちょうどに目標状態に制御できる機能が必要となる。そのためモータ命令発行から駆動までの遅延も含めて考慮しなければタイミングは合わない。2つめは、ロボットの雑音である。これは、音楽ロボット自身にマイクを装着して音楽信号を入力する状況で生じる。一般的に音楽音源よりも雑音源のモータの方がマイクに近いので、信号対雑音比は低い。音楽信号に雑音がないことを仮定する手法はそのまま使用できないため、その場合は、手法自体の改良やマイクアレイ処理による雑音抑圧などの前処理が必要となる。

リズムに合わせる音楽ロボット

リズムは音楽の基礎である。ここではリズムに合わせて動作するロボットに焦点を絞り、従来の研究

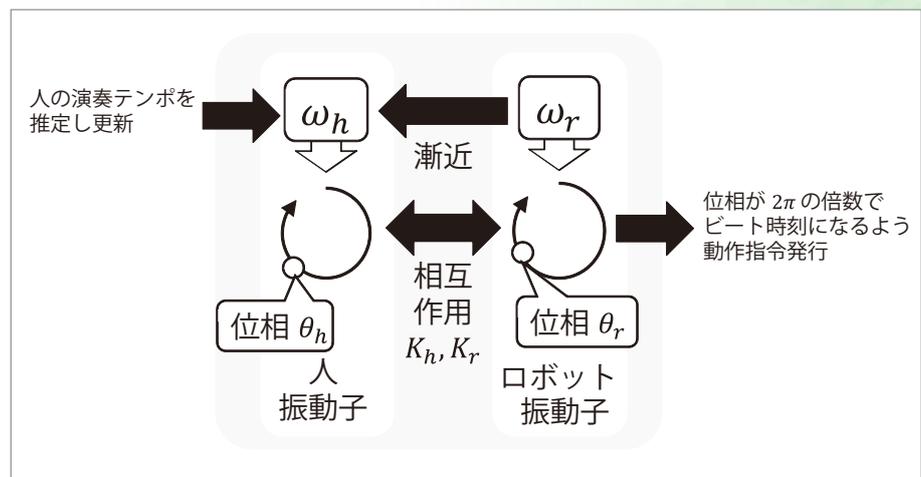


図-1 結合振動子系を用いた合奏モデル

を紹介する。音楽情報処理分野では、音楽信号からビートを検出するビートトラッキングの研究が長年行われており、以下の研究でもビートトラッキングが基礎技術として用いられている。

吉井らは音楽のリズムに合わせて足踏みをするロボットについて報告した⁶⁾。音楽信号から実時間で直前のビート時刻とビート間隔を推定し、ビート間隔を外挿することで次のビート時刻を予測する。そして、予測ビート時刻に足踏みをするよう指示する。実際の足踏み時刻と予測ビート時刻は合わないの、その誤差をフィードバックして予測を修正することで、音楽に合わせた足踏みを実現した。

足踏みロボットは再生されている音楽に対して一方的に追従する。他方、人とインタラクションを行う場合は、人のリズムもロボットに引き込まれるので、相互の引き込みを考慮しなければ互いのリズムが揺らぐため全体のリズムが安定しない。

水本らは人とロボットが互いに四拍子のリズムを合わせる状況を対象として、人と合奏するロボットについて報告した¹⁾。合奏モデル(図-1)では、各共演者(人、ロボット)のリズムを振動子で表現し、合奏全体のリズムを次式で表される結合振動子系で表現する。

$$\begin{aligned} \frac{d}{dt}\theta_h &= \omega_h + K_h \sin(\theta_r - \theta_h) \\ \frac{d}{dt}\theta_r &= \omega_r + K_r \sin(\theta_h - \theta_r) \end{aligned}$$



図-2 参加者に合わせてテルミンを演奏するロボット Nao (Aldebaran Robotics 社)

θ_h, θ_r は各共演者の振動子の位相を, ω_h, ω_r はテンポを表す. 振動子は内部クロックであり, 位相が 2π の倍数となったときに共演者はビートを発生させる (手拍子を打つ, 太鼓を叩くなどを行う) と定義する. K_h, K_r は結合強度を表し, 相手のリズムに引き込まれる強さを表す. ロボット振動子の位相 θ_r が 2π の倍数になる時刻が次のビート時刻なので, θ_r を監視することで次のビート時刻が予測できる. このように相互の引き込みを考慮したモデルによって人の揺らぐリズムに合わせた合奏を実現した.

リズムに合わせる音楽ロボットを用いて, Lim らは IROS 2011 で参加者と合奏するテルミン演奏ロボットの参加型デモを行った (図-2)^{☆2}. 観客に一方的に見せるデモはロボットの動作系列を事前にすべて作成して再生すれば実現できるが, 本デモでは観客が参加して一緒に楽しむので, より一体感が高まったようであった.

☆2 More Cowbell! A Musical Ensemble with the NAO Thereminist, Demo Session, IROS 2011.

ロボットを用いた 音楽情報処理研究の広がり

音楽ロボット研究では, 本稿で取り上げたリズムに合わせるロボット以外にも, 演奏動作や合奏中のジェスチャ認識, 楽譜を追従する合奏, 合奏中のリーダーの発見など, さまざまな研究が行われている. それには音楽情報処理分野で開発されてきたピッチ抽出, 楽譜追従, 自動伴奏などの技術が不可欠である一方, 音楽ロボット研究への適用には実時間性や耐雑音性, 予測などの改良が必要なものもある. このような改良は音楽ロボットにとどまらず音楽情報処理の広がりにも貢献する可能性がある. たとえばスマートフォンのマイクを入力に使う場合は耐雑音性が有効であろう. このように, 音楽ロボット研究を通じた音楽情報処理技術の発展によって, さらに応用範囲の広がりが期待できる.

参考文献

- 1) Mizumoto, T.: *Temporal Synchronization among Interacting Individuals in Human-Robot Ensembles and Frog Choruses*, Kyoto University: PhD Thesis (2013).
- 2) Solis, J. ほか: Development of Waseda Flutist Robot WF-4RIV: Implementation of Auditory Feedback System, *IEEE International Conference on Robotics and Automation*, pp.3654-3659 (2008).
- 3) Weinberg, G. and Driscoll, S.: Toward Robotic Musicianship, *Computer Music Journal*, 30[4], pp.28-45 (2006).
- 4) Kosuge, K.: Dance Partner Robot: an Engineering Approach to Human-robot Interaction, *ACM/IEEE International Conference on Human-Robot Interaction*, p.201 (2010).
- 5) 中野 倫, ほか: VocaWatcher: ユーザ歌唱の顔表情を真似るヒューマノイドロボットの顔動作生成システム, *情報処理学会論文誌*, Vol.55, No.3, pp.1222-1235 (2014).
- 6) Yoshii, K.: A Biped Robot that Keeps Steps in Time with Musical Beats, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.1743-1750 (2007). (2015年12月29日受付)

水本武志 (正会員) t.mizumoto@jp.honda-ri.com

2013年 京都大学大学院情報学研究所知能情報学専攻博士後期課程修了. 博士 (情報学). 現在, (株)ホンダ・リサーチ・インスティテュート・ジャパン, リサーチャ. ロボット聴覚, 生物の合唱解析, 音楽ロボットの研究に従事. IEEE, RSJ, JSAI 各会員.