

メディアテクノロジーシリーズ 2

# 音楽情報処理

後藤真孝

【編著】

北原鉄朗・深山 覚・竹川佳成・吉井和佳

【共著】

コロナ社

## メディアテクノロジーシリーズ 編集委員会

---

編集委員長	近藤 邦雄 (元東京工科大学, 工学博士)
編集幹事	伊藤 貴之 (お茶の水女子大学, 博士 (工学))
編集委員 (五十音順)	五十嵐悠紀 (お茶の水女子大学, 博士 (工学))
	稲見 昌彦 (東京大学, 博士 (工学))
	牛尼 剛聡 (九州大学, 博士 (工学))
	大淵 康成 (東京工科大学, 博士 (情報理工学))
	竹島由里子 (東京工科大学, 博士 (理学))
	鳴海 拓志 (東京大学, 博士 (工学))
	馬場 哲晃 (東京都立大学, 博士 (芸術工学))
	日浦 慎作 (兵庫県立大学, 博士 (工学))
	松村誠一郎 (東京工科大学, 博士 (学際情報学))
	三谷 純 (筑波大学, 博士 (工学))
	三宅陽一郎 (株式会社スクウェア・エニックス, 博士 (工学))
	宮下 芳明 (明治大学, 博士 (知識科学))

---

(2023年5月現在)

## 編著者・執筆者一覧

---

編著者	後藤 真孝 (6章)	
執筆者 (執筆順)	北原 鉄朗 (1, 3章)	深山 覚 (2章)
	竹川 佳成 (4章)	吉井 和佳 (5章)

---

## 刊行のことば

“Media Technology as an Extension of the Human Body and the Intelligence”

「メディアはメッセージである (The medium is the message)」というマクルーハン (Marshall McLuhan) の言葉は、多くの人々によって引用される大変有名な言葉である。情報科学や情報工学が発展し、メディア学が提唱されたことでメディアの重要性が認識されてきた。このような中で、マクルーハンのこの言葉は、つねに議論され、メディア学のあるべき姿を求めてきたといえる。

人間の知的コミュニケーションを助けることができるメディアは生きていくうえで欠かせない。このようなメディアは人と人との関係をより良くし、視野を広げ、新しい考え方に目を向けるきっかけを与えてくれる。

また、マクルーハン「メディアはマッサージである (The medium is the massage)」ともいっている。マッサージは疲れた体をもみほぐし、心もリラックスさせるが、メディアは凝り固まった頭にさまざまな情報を与え、考え方を広げる可能性があるため、マッサージという言葉はメディアの特徴を表しているともいえるだろう。

さらにマクルーハン「人間の身体を拡張するテクノロジー」としてメディアをとらえ、人間の感覚や身体的な能力を変化させ、社会との関わりについて述べている。現在、メディアは社会、生活のあらゆる場面に存在し、五感を通してさまざまな刺激を与え、多くの技術が社会生活を豊かにしている。つまり、この身体拡張に加え、人工知能技術の発展によって“知能拡張”がメディアテクノロジーの重要な役割を持つと考えられる。このために物理的な身体と情報や知識を扱う知能を融合した“人間の身体と知能を拡張するメディアテクノロジー”を提案・開発し、これらの技術を活用して社会の構造や仕組みを変革し、

どのような人にとっても住みやすく、生活しやすい社会を目指すことが望まれている。

一方、大学におけるメディア学の教育は、東京工科大学が1999年にメディア学部を設置して以来、全国の大学でメディア関連の学部や学科が設置され文理芸分野を融合した多様な教育内容が提供されている。その体系化が期待されメディア学に関する教科書としてコロナ社から「メディア学大系」が発刊された。この第一巻の『改訂メディア学入門』には、メディアの基本モデルの構成として「情報の送り手、伝達対象となる情報の内容（コンテンツ）、伝達媒体となる情報の形式（コンテナ）、伝達形式としての情報の提示手段（コンペア）、情報の受け手」と書かれている。これからわかるようにメディアの基本モデルには文理芸に関連する多様な内容が含まれている。

メディア教育が本格的に開始され20年を過ぎるいま、多くの分野でメディア学のより高度で急速な展開が見られる。文理芸の融合による総合知によって人間生活や社会を理解し、より良い社会を築くことが必要である。

そこで、このメディア分野の研究に関わる大学生、大学院生、さらには社会人の学修のため「メディアテクノロジーシリーズ」を計画した。本シリーズは“人間の身体と知能を拡張するメディアテクノロジー”を基礎として、コンテンツ、コンテナ、コンペアに関する技術を扱う。そして各分野における基本的なメディア技術、最近の研究内容の位置づけや今後の展開、この分野の研究をするために必要な手法や技術を概観できるようにまとめた。本シリーズがメディア学で扱う対象や領域を発展させ、将来の社会や生活において必要なメディアテクノロジーの活用方法を見出す手助けとなることを期待する。

本シリーズの多様で広範囲なメディア学分野をカバーするために、電子情報通信学会、情報処理学会、人工知能学会、日本ソフトウェア科学会、日本バーチャルリアリティ学会、ヒューマンインタフェース学会、日本データベース学会、映像情報メディア学会、可視化情報学会、画像電子学会、日本音響学会、芸術科学会、日本図学会、日本デジタルゲーム学会、ADADA Japan などにおいて第一線で活躍している研究者の方々に編集委員をお願いし、各巻の執筆者選

定、目次構成、執筆内容など検討を重ねてきた。

本シリーズの読者が、新たなメディア分野を開拓する技術者、クリエイター、研究者となり、新たなメディア社会の構築のために活躍されることを期待するとともにメディアテクノロジーの発展によって世界の人達との交流が進み相互理解が促進され、平和な世界へ貢献できることを願っている。

2023年5月

編集委員長 近藤邦雄

編集幹事 伊藤貴之

### 表紙・カバーデザインについて

私たちは五感というメディアを介して世界を知覚し、自己の存在を認知することができます。メディア技術の進歩によって五感が拡張され続ける中、「人」はなにをもって「人」と呼べるのか、そんな根源的な問いに対する議論が絶えません。

本書の表紙・カバーデザインでは、二値化された五感が新しい機能や価値を再構築する様子をシンプルなストライプ柄によって表現しました。それぞれのストライプは5本のゆらぎを持った線によって描かれており、手描きのような印象を残しました。

しかし、この細かなゆらぎもプログラム制御によって生成されており、十分に細かく量子化された表現によって「デジタル」と「アナログ」それぞれの存在がゆらぐ様子を表しています。乱雑に描かれたストライプをよく観察してみてください。本書を手にとった皆さんであれば、きっともう一つ面白いことに気づくでしょう。

デザインを検討するにあたって、同じコンセプトに基づき、いくつかのグラフィックパターンを生成可能なウェブアプリケーションを準備しました。下記 URL にて公開していますので、あなただけのカバーを作ってみてください。読者の数だけカバーデザインが存在するのです。世界はあなたの五感を通じて存在しているのですから。

馬場哲晃

〈Cover Generator〉ぜひお試しください

<https://tetsuakibaba.github.io/mtcg/>

(2023年5月現在)



# ま え が き

「音楽情報処理」は、コンピュータ上で音楽のあらゆる側面を扱う研究分野である。身近で魅力的な研究成果の宝庫で、とても人気が高い。スマートフォンで音楽を聴いたり、コンピュータで音楽を制作したりできるのは、音楽情報処理が発展したおかげであり、すでに音楽の鑑賞や創作の多くの場面で研究成果が利用されている。さらに、未来の音楽体験を切り拓くような新しい研究成果もつぎつぎと生まれており、研究者ではない一般の方々も音楽に関連した技術には興味を持ちやすい。音楽がデジタル化され、音楽情報処理の力をつぎつぎと新しいサービスや体験が生まれていく中で、その産業的な重要性は増している。

音楽情報処理は、学術的にも重要で多くの研究者の関心を集めてきた。それは、音楽が人の創意工夫が込められた多様で複雑な現象であり、それを情報処理するうえで、挑戦的で未解決な問題が多いからである。例えば、人間が音楽を聴いて理解できることの、まだごく一部しかコンピュータで自動的に解析できないし、人間が創作する魅力的な音楽に匹敵するような楽曲を自動的に生成することはまだできない。それでも、多くの研究者の貢献で着実に研究分野は進展し、ひと昔前から考えたら夢のような技術が数多く実現されてきた。

本書は、そうした研究分野について、「音楽情報処理」と題して日本語で初めて書かれた書籍である。音楽情報処理への学術的・産業的・社会的・文化的な関心の高さから、これまでさまざまな学会誌等で解説記事が書かれてきたが、研究分野名を書名に冠した書籍はなかった。そこで、本書では、音楽情報処理に関する基礎的な解説に加え、代表的な研究テーマに関して具体的な研究事例を交えながら紹介することとした。そうした書籍では著者も重要となるので、本書では、実際に当分野で長年研究開発をしてきて、かつ、現在も活躍し続けている現役の研究者たちが、それぞれの専門性を活かして執筆することとした。

本書は、6章で構成されている。1章では、「音楽情報処理の基礎」として、研究分野の全体像や、コンピュータ上での音楽のデータ表現を解説する（北原鉄朗担当）。2章では、コンピュータが音楽を生成する「自動作曲」について、その発展の歴史や、創造性に対する考え方と実現方法、さらには機械学習を用いたさまざまな自動作曲事例を紹介する（深山覚担当）。3章では、人間が作曲をしたり即興演奏をしたりする際のコンピュータの支援方法について、「作曲支援・即興演奏支援」の具体的な研究事例を交えて紹介する（北原鉄朗担当）。4章では、楽器に焦点を当て、その演奏を支援する「楽器演奏支援」について、楽器演奏を演奏者が学ぶ支援から、レッスンで教師が指導する支援まで含めて幅広く紹介する（竹川佳成担当）。以上の2~4章が、音楽の創作に関連したトピックだったのに対し、5~6章では、音楽の鑑賞に関連したトピックを扱う。5章では、コンピュータが音楽を自動的に解析する「自動採譜」について、その問題の定式化から出発し、ビートやコード、歌声のような主要な音楽要素の推定方法を紹介する（吉井和佳担当）。6章では、そのような音楽解析技術が、実際に人々の音楽の聴き方をどのように豊かにするのかについて、具体的な「音楽鑑賞インタフェース」の事例を交えて紹介する（後藤真孝担当）。

ぜひ、本書をパラパラとめくって、興味を持った章から読み始めてみていただきたい。本書の特長として、各章は関連しているものの独立して書かれており、その内容を理解するうえで、前の章まで読んでいないことは前提としていない。いろいろな章を読み進め、最終的に本書全体を読むことで、音楽情報処理の研究分野に対する理解を深めていくことが可能となっている。

音楽情報処理の研究開発は楽しいので、本書を読んで興味を持った方々には、ぜひ取り組んでいただきたいと願っている。本書を読んだ後に、より発展的に学びたい方々に向けたメッセージも「あとがき」に記している。最後に、本書執筆の機会を与えてくださった「メディアテクノロジーシリーズ」編集委員会ならびにコロナ社の担当者各位に、著者一同を代表して感謝の意を表する。

2023年7月

編著者 後藤真孝

# 目 次

## 第 1 章

### 音楽情報処理の基礎

<b>1.1 音楽情報処理へのいざない</b> .....	1
1.1.1 音楽情報処理とは	2
1.1.2 音楽情報処理の歴史	3
1.1.3 音楽情報処理の分類	6
1.1.4 音楽情報処理で用いられる技術やプログラミング言語	8
1.1.5 音楽情報処理の研究を知ることができる学会・国際会議	10
<b>1.2 コンピュータにおける音楽データの表現方法</b> .....	14
1.2.1 信号と記号	14
1.2.2 音響信号	15
1.2.3 スペクトログラム	17
1.2.4 MIDI	19
1.2.5 ピアノロール	21
1.2.6 MusicXML	23

## 第 2 章

### 自動作曲

<b>2.1 自動作曲とは</b> .....	24
2.1.1 創造的な行為の自動化としての自動作曲	24
2.1.2 創造性の考え方の分類	29

2.1.3 自動作曲の創造性の実現方法	30
<b>2.2 機械学習を用いた自動作曲</b> .....	<b>31</b>
2.2.1 音楽と確率・統計	32
2.2.2 自動作曲の定式化	41
2.2.3 自動作曲の解法	41
2.2.4 自動作曲の評価	45
<b>2.3 機械学習を用いた自動作曲の研究事例</b> .....	<b>49</b>
2.3.1 再帰的ニューラルネットワークを用いた研究事例	50
2.3.2 変分オートエンコーダを用いた研究事例	54
2.3.3 敵対的生成ネットワークを用いた研究事例	58
2.3.4 トランスフォーマを用いた研究事例	59

## 第 3 章

### 作曲支援・即興演奏支援

<b>3.1 作曲支援</b> .....	<b>62</b>
3.1.1 「誰」が行う「何」を支援するか	63
3.1.2 ユーザも含めて作曲システムと考える	65
3.1.3 曲線に基づく作曲支援の研究事例	69
3.1.4 ループシーケンサに基づく作曲支援の研究事例	72
<b>3.2 即興演奏支援</b> .....	<b>74</b>
3.2.1 即興演奏支援における課題	75
3.2.2 鍵盤型 UI を用いた研究事例	76
3.2.3 旋律概形を用いた研究事例	79
3.2.4 専用デバイスを用いた研究事例	81
3.2.5 音楽と無関係の行動を用いた研究事例	82
3.2.6 多様なジャムセッションを実現する	82

## 第 4 章

## 楽器演奏支援

<b>4.1 楽器演奏支援とは</b> .....	87
<b>4.2 独 習 支 援</b> .....	89
4.2.1 運指学習を考慮したピアノ学習支援	91
4.2.2 リズム学習を考慮したピアノ学習支援	95
4.2.3 視線情報を用いた打鍵支援からの離脱	97
4.2.4 曖昧情報・虚偽情報の提示によるポジショニング支援からの離脱	100
4.2.5 視覚と聴覚を用いたポジショニング学習支援とその離脱	104
4.2.6 モチベーションを考慮したピアノ学習支援システム	107
<b>4.3 レッスン支援</b> .....	109
4.3.1 気づきのアノテーション機能を持つピアノレッスン支援	110
4.3.2 悪癖発見機能を持つピアノレッスン支援	114
4.3.3 カメラスイッチング機能を持つオンラインピアノレッスン支援	116

## 第 5 章

## 自動採譜

<b>5.1 基本的な枠組み</b> .....	122
5.1.1 問題設定	124
5.1.2 統計的アプローチ	128
5.1.3 性能評価	133
<b>5.2 ビート・ダウンビート推定</b> .....	134
5.2.1 ビート・ダウンビートの定義	135
5.2.2 ビート・ダウンビート推定の手がかり	136
5.2.3 深層推論モデルに基づくアプローチ	137
5.2.4 深層生成モデルに基づくアプローチ	139

<b>5.3</b>	<b>コード・キー推定</b> .....	<b>143</b>
5.3.1	コード・キーの定義	144
5.3.2	コード・キー推定の手がかり	145
5.3.3	深層推論モデルに基づくアプローチ	146
5.3.4	深層生成モデルに基づくアプローチ	148
5.3.5	変分オートエンコーダによる統合	150
<b>5.4</b>	<b>歌声採譜</b> .....	<b>154</b>
5.4.1	ピアノロール推定と楽譜推定	154
5.4.2	歌声採譜の手がかり	156
5.4.3	深層推論モデルに基づくアプローチ	158
5.4.4	深層生成モデルに基づくアプローチ	160

## 第 6 章

### 音楽鑑賞インタフェース

<b>6.1</b>	<b>技術が切り拓く音楽鑑賞インタフェースの発展</b> .....	<b>163</b>
<b>6.2</b>	<b>楽曲の鑑賞インタフェース</b> .....	<b>164</b>
6.2.1	音楽構造解析に基づく音楽鑑賞インタフェース	165
6.2.2	自動音楽同期に基づく音楽鑑賞インタフェース	169
6.2.3	音源分離に基づく音楽鑑賞インタフェース	171
6.2.4	能動的音楽鑑賞インタフェース	175
6.2.5	音楽理解力拡張インタフェース	177
<b>6.3</b>	<b>音楽コレクションの鑑賞インタフェース</b> .....	<b>179</b>
6.3.1	音楽検索インタフェース	180
6.3.2	音楽探索インタフェース	183
6.3.3	音楽推薦インタフェース	192

あ と が き .....	197
引用・参考文献 .....	199
索 引 .....	216

- 
- 注 1) 本書の書籍詳細ページ (<https://www.coronasha.co.jp/np/isbn/9784339013726/>) から URL リンク集などの補足情報がダウンロードできます。
- 注 2) 本書で使用している会社名、製品名は一般に各社の登録商標です。  
本書では®や TM は省略しています。
- 注 3) 本書で紹介している URL は 2023 年 7 月現在のものです。



# 第 1 章

## 音楽情報処理の基礎

音楽情報処理とは、その名の通り音楽を対象とした情報処理なのだが、音楽活動（鑑賞、歌唱、演奏、作曲など）の何に着目し、誰（音楽の専門家、非専門家）をターゲットユーザとして研究開発を行うかは、さまざまである。そのため、音楽情報処理という分野にはかなり多彩な研究テーマが含まれる。本章では、音楽情報処理という研究分野を概観した後、その基礎となる音楽データのコンピュータでの表現方法についても述べる。

### 1.1 音楽情報処理へのいざない

音楽は、われわれの生活に最も根付いた芸術の一つである。街を歩けばBGMが聞こえてくるし、ひと昔前は音楽CDなどの物理メディア、いまはインターネット上の音楽配信サービスでさまざまな楽曲が流通し、老若男女問わず音楽を娯楽として楽しんでいる。中には、自ら作曲してインターネット上の動画共有サイトに投稿する人もいる。楽器演奏も、子供の習い事として定番だけでなく、趣味で演奏をする人も少なくない。

一方、コンピュータやインターネットをはじめとする情報処理技術も、われわれの生活になくてはならないものになっている。多くの人が肌身離さず使っているスマートフォンは、まさにコンピュータそのものであるし、交通機関の運行や銀行のATMなども、コンピュータで制御されている。また、インターネット上の動画共有サイトやショッピングサイト抜きには生活できないという人も少なくないだろう。

この両者、つまり音楽と情報処理技術が組み合わさったとき、何が起こるだ

ろうか。音楽を対象に情報処理技術を研究する研究分野を音楽情報処理 (music information processing)<sup>†</sup>といい、じつはすでに半世紀以上の歴史がある。本節では、音楽情報処理がどんな分野なのか、音楽情報処理によって何が可能になるのかを簡単にひも解いてみたい。

### 1.1.1 音楽情報処理とは

音楽を対象とした情報処理技術と聞いて、読者はどんなものを想像するだろうか。すぐに思いつくものは、コンピュータを活用した音楽制作だろう。テレビ番組などで、音楽制作スタジオで音楽家がインタビューを受けているシーンを見たことはないだろうか。そんな現場をよく見ると、シンセサイザーやスピーカ、ミキサなどと一緒にコンピュータが置かれていることに気づくだろう。DAW (digital audio workstation) と呼ばれるソフトウェアを使うことで、電子楽器で演奏したメロディをコンピュータに取り込んだり、音符を一つひとつコンピュータで入力し、生楽器と遜色のない音色、または生楽器では出せないコンピュータならではの音色で奏でたりすることができる。歌唱や生演奏の録音もコンピュータ上で行い、コンピュータ1台だけで楽曲を仕上げることも可能である。

一方、音楽制作をしない一般の音楽鑑賞者が直接利用する技術としては、どんなものがあるだろうか。一つは、音楽の検索や推薦であろう。街中でBGMがかかっているときに、それが誰の何という楽曲なのかを知りたいことがある。そんなときにスマートフォンで録音すると曲名とアーティスト名を調べてくれるアプリがすでに実用化されている。そのほかにも、自分の好みの楽曲を入力すると、それに似た別の楽曲を検索してくれる技術も生まれている。現在、購

---

<sup>†</sup> 「音楽情報処理」に直接対応する英語表現は music information processing であるが、ほかにもさまざまな呼び方がある。例えば、music informatics, music technology などである。computer music も用いられたことがあるが、コンピュータを活用した音楽制作 (または制作した音楽) を意味することも多い。また、音楽情報検索に特化した music information retrieval (MIR)、R の意味を変更して音楽情報処理研究全体を指すようにした music information research、あるいは音楽以外の音も含めた sound and music computing もよく用いられる。

買情報や評価情報をベースとしたソーシャルフィルタリングに基づくものが先だっって実用化されているが、楽曲の中身を解析し、その中身に基づいて検索・推薦する技術の研究も進んでいる。

このように、音楽を対象とした情報処理技術はさまざまな可能性を秘め、すでに一部は音楽制作現場やわれわれの生活に深く入り込んでいる。そんな技術の研究開発を行う研究分野、それが**音楽情報処理**である。

### 1.1.2 音楽情報処理の歴史

音楽情報処理の歴史は長く、コンピュータそのものと同じぐらいの歴史がある。1950年代にはすでに、自動で作曲された作品『イリアック組曲』が発表されている。以下、現在までの音楽情報処理のトレンドの変化を簡単に紹介する。なお、ここで述べられている歴史は網羅的ではないうえに、筆者なりのまとめに過ぎないことに留意されたい。

音楽情報処理の歴史において、初期の頃はコンピュータがまだ普及していないため、一般の音楽鑑賞者や音楽の非専門家向けというよりは作曲家が自身の表現の幅を広げるための研究や、研究者が基礎的な技術の探求を行う研究が多かった。形式的なプロセス<sup>†</sup>によって楽曲を作る試みは**アルゴリズム作曲** (algorithmic composition) と呼ばれ、さまざまなアルゴリズムによって新奇な音符列を生み出し、創作活動に用いられてきた。**楽音合成** (musical sound synthesis) も活発に研究されてきた。1980年代に世界初のフルデジタルシンセサイザー製品として発売された**FM音源**は、1970年代にスタンフォード大学の研究者によって発表されたものである。その後もさまざまな合成方式が研究され、またそれらに基づく多数の製品が発売されたことで、楽音合成技術は音楽の制作現場には欠かせない技術となっている。また、同時期(1980年代)に策定された電子楽器間の演奏情報送受信規格**MIDI** (musical instrument digital interface)

---

<sup>†</sup> ここでは、数理的な処理やアルゴリズムの実行など、作曲者の主観やアイデアを介在させないプロセスを指す。コンピュータでの実行に適しているが、人間が手動で行ってもよい。

は、研究コミュニティから生まれた技術ではないものの、演奏情報の通信プロトコルやデータフォーマットとして、音楽情報処理研究では欠かせないものとなっている。

一方、楽器音や楽曲の音楽音響信号（音の波形）の解析も活発に行われてきた。1960年代に高速フーリエ変換（fast Fourier transform, FFT）が発明されると、音声（人の声）に対して、音の高さに相当する基本周波数（fundamental frequency）を推定する研究が生まれ、楽器音に対しても同様の試みがなされるようになった。当初は単音（一つの音のみが同時に鳴っているもの）が対象であったが、1990年頃からしだいに多重音（複数の音が同時に鳴っているもの）に研究対象が広がり、自動採譜（automatic music transcription）という研究テーマが確立されていく。自動採譜の研究は奥深く、2020年代のいまでも実用化されているとは言い難いが、ディープニューラルネットワークなどの導入により、楽器を限定すれば相当高い精度が実現されつつある。

1990年代以降は、MP3（MPEG-1 Audio Layer 3）をはじめとする高品質のオーディオ圧縮・符号化技術（audio compression and coding technique）が登場した。これに加えてストレージの小型化・大容量化・低価格化が進んだことが背景となり、2000年代に携帯型の大容量音楽プレーヤーが普及した。またちょうどこの時期、インターネット上での音楽配信サービスも登場し、1曲ごとの販売から、定額制で聴き放題の音楽ストリーミングサービス（現在主流のサブスクリプションサービス）へと発展していった。

このように、音楽を扱う舞台がコンピュータ（音楽プレーヤーも小さなコンピュータである）やインターネット上に移ると、何千・何万もの楽曲をいつでもどこでも自由に聴ける分、いま聴きたい曲をどのように選び出すかが問題となる。そこで世界中で研究が活発になったのが、音楽情報検索（music information retrieval, MIR）である。音楽情報検索の研究の一環として、音楽音響信号の演奏内容（メロディ、リズムパターン、コード進行など）の自動認識や自動音楽アノテーション（automatic music annotation）、指定された楽曲または楽曲群に似た楽曲を見つける類似楽曲検索（similarity-based music retrieval）、好

みの楽曲を自動的にリコメンドする**音楽推薦** (music recommendation), 研究用の音楽データベース整備, それらの技術を活用して音楽鑑賞を高度化する**音楽鑑賞インタフェース** (music listening interfaces) など, さまざまな研究が行われるようになった。

一方, 音楽そのものの形式化を情報処理技術の力を借りて行う試みも行われた。1980年代に提唱された**GTTM** (generative theory of tonal music) や**暗意実現モデル** (implication-realization model) は, それぞれ生成文法, ゲシュタルト心理学に着想を得た音楽理論である。これらはのちに, コンピュータ上で実装する試みがなされるようになり, **計算論的音楽理論** (computational music theory) の構築という研究テーマが生まれることとなった。

2000年代には, **歌声合成** (singing voice synthesis) も大きな進化を遂げた。歌声合成技術の研究開発はその前から行われてきたが, 歌声合成技術がキャラクターを伴う形のソフトウェアとして発売されたこと, そのソフトウェアを用いて制作した音楽作品をアマチュアやプロのクリエイターがインターネット上で発表する環境が整ったことから, 爆発的なブームになった。

これらと並行して, リアルタイムに動作する音楽システムの研究も行われてきた。VR/AR 技術やウェアラブル技術を用いた演奏支援システムや**音楽表現のためのユーザインタフェース** (user interface for musical expression), コンピュータ上に構築された仮想演奏家とセッションを行う**ジャムセッションシステム** (jam session system), 人間の演奏に同期しながら伴奏を奏でてくれる**自動伴奏システム** (automatic accompaniment system) などである。これらのシステムはリアルタイムに動作することが求められるため, 低遅延処理, 安定した動作, マルチスレッドによる並行処理を行う必要がある。

上で述べたアルゴリズム作曲は, **自動作曲** (automatic music composition) と呼ばれることが多くなり, 作曲家向けの作曲支援研究のほか, 非作曲家が手軽にオリジナル楽曲を手に入れることを目指した研究, 機械学習の一応用としての研究など, 多様な研究が生まれている。特に, 2010年代の第3次人工知能 (AI) ブーム以降は, ディープニューラルネットワークの応用例として音楽の自動生

# 索引

<b>【あ】</b>	
アルゴリズム作曲	3
暗意実現モデル	5
<b>【う】</b>	
歌声合成	5
歌声採譜	154
<b>【え】</b>	
演奏表情付け	6
<b>【お】</b>	
オーディオ圧縮・符号化技術	4
音機能	79
音機能固定マッピング	79
音楽加工	174
音楽加工インタフェース	174
音楽鑑賞インタフェース	5, 164
音楽検索インタフェース	180
音楽コレクションの鑑賞 インタフェース	164
音楽情報検索	4, 143, 180
音楽情報処理	2, 3
音楽推薦	5
音楽推薦インタフェース	180, 192
音楽ストリーミング サービス	180
音楽タッチアップ	174
音楽探索インタフェース	180, 183

音楽のサイコロ遊び	26
音楽の事例検索	182
音楽表現のためのユーザ インタフェース	5
音楽理解技術	164
音楽理解力拡張インタ フェース	177
音楽理解力向上インタ フェース	179
音響信号処理	9
音響モデル	132
音高推定	126
オンセット強度信号	136
オンセットロール	126, 127

## 【か】

階層マルコフモデル	149
楽音合成	3
確率変数	124
隠れマルコフモデル	131
歌詞同期手法	169
楽曲の鑑賞インタ フェース	164
カテゴリカル分布	141

## 【き】

機械学習	9
記号表現	14
基本周波数	4
教師あり学習	130
教師なし学習	132
協調フィルタリング	192

## 【く】

クロスエントロピー	45, 147
クロマベクトル	145

## 【け】

計算音楽学	123
計算論的音楽理論	5
計算論的創造性	6
系列対系列モデル	155
ゲートタイム	21
言語モデル	132

## 【こ】

高速フーリエ変換	4
コネクションリスト	155
時間分類	19
コントロールチェーン	19

## 【さ】

再帰的ニューラル ネットワーク	50
再現率	134
最尤推定	132
サーストンの一対比較法	48
サビ	165
サビ区間検出手法	166
サンプリング周波数	15

## 【し】

時間畳み込みネットワーク	138
シグモイド関数	137

システムエクスクルーシブ  
メッセージ 20  
自動音楽アノテーション 4  
自動楽譜認識 171  
自動採譜 4, 122  
自動作曲 5  
自動伴奏システム 5  
ジャムセッション  
システム 5, 83  
償却型変分推論 153  
信号表現 14  
振幅スペクトル 18

【す】

推薦エンジン 194  
推論モデル 129  
ステップタイム 21  
スペクトル 125  
スペクトログラム  
17, 18, 124

【せ】

正解率 133  
生成モデル 131  
セミマルコフモデル 142  
ゼロ頻度問題 35  
潜在変数モデル 132  
旋律概形 71

【そ】

双方向長短期記憶 138  
ソフトマックス関数 147

【た】

対話の進化計算 67  
ダウンビート 135  
多重音高推定 126  
短時間フーリエ変換 18, 125

【ち】

チャンネルメッセージ 20  
注意機構 59, 157  
長距離依存関係 39

【て】

ディスカウンティング 37  
テイタム 127  
ディレイテッド畳み込み 138  
定Q変換 125  
適合率 134  
敵対的生成ネットワーク 58  
データオーギュメン  
テーション 36  
デルタタイム 20  
テンポ 135

【と】

統計的機械学習 123  
動的計画法 120, 134, 171  
動的ベイジアン  
ネットワーク 139  
トランスフォーマ 59, 138  
貪欲法 42

【な】

ナイキスト周波数 125  
内容に基づく音楽情報検索 180  
内容に基づくフィルタリング 193

【の】

能動的音楽鑑賞インタ  
フェース 175  
ノートオフメッセージ 19  
ノートオンメッセージ 19

【は】

バイナリクロス  
エントロピー 138  
拍節マルコフモデル 155  
パターン認識 9  
ハミング検索 180  
パワースペクトル 18  
半教師あり学習 133

【ひ】

ピアノロール 21, 122, 126  
ビタビアルゴリズム 131  
ピッチバンド 20  
ピッチロール 126  
ビート 135  
ビームサーチ 43  
ビーム幅 43  
ヒューマンコンピュータ  
インタラクション 87  
標準 MIDI ファイル 20  
標本化 15  
標本化定理 15

【ふ】

複素スペクトル 18  
フーリエ変換 9  
フレーム 18  
プログラムチェンジ 20  
プロジェクト  
マッピング 87, 88

【へ】

ベルヌイ分布 137  
変分オートエンコーダ  
54, 150

【ま】

窓関数 18

【め】

メルスペクトログラム 126

【ら】

ラプラス分布 142  
ラプラス法 36

【り】

離散フーリエ変換 17, 125  
離散変数 129  
リズム採譜 155  
量子化ビット数 15

<b>【る】</b>			<b>【れ】</b>		
類似楽曲検索	4	ループシーケンサ	64	連続変数	129
◇					
<b>【A】</b>		FFT	4	<b>【O】</b>	
Add-one スムージング	36	FM 音源	3	OMR	23, 171
Add- $\alpha$ スムージング	37	<b>【G】</b>		OSS	136
AMT	122	GAN	58	<b>【P】</b>	
Arduino	10	GTTM	5	PCA	184
AST	154	<b>【H】</b>		Processing	10
AVI	153	HMM	131	Python	9
<b>【B】</b>		<b>【I】</b>		<b>【Q】</b>	
BCE	138	IEC	67	QBH	180
BLSTM	138	<b>【J】</b>		<b>【R】</b>	
<b>【C】</b>		Julia	9	RNN	50
CE	147	<b>【M】</b>		<b>【S】</b>	
CQT	125	MATLAB	9	STFT	18, 125
CTC	155	Max	10	<b>【T】</b>	
<b>【D】</b>		MIDI	3, 19	TCN	138
DAW	2	MIDI データ	9	<b>【V】</b>	
DBN	139	MIDI メッセージ	19	VAE	54, 150
DFT	17, 125	MIR	4, 143, 180	<b>【X】</b>	
DP	120, 134, 171	MOS	47	XR(AR, MR, VR)	87
<b>【F】</b>		MP3	4		
F 値	134	MusicXML	9, 23		
FFNN	120				

—— 編著者・著者略歴 ——

後藤 真孝（ごとう まさたか）

1993年 早稲田大学理工学部電子通信学科卒業  
1995年 早稲田大学大学院理工学研究科修士課程修了（電気工学専攻）  
1998年 早稲田大学大学院理工学研究科博士後期課程修了（電子・情報通信学専攻）  
博士（工学）  
1998年 電子技術総合研究所研究員  
2001年 産業技術総合研究所研究員（改組）  
2005年 産業技術総合研究所主任研究員  
2009年 産業技術総合研究所研究グループ長  
2011年 産業技術総合研究所上席研究員  
2013年 産業技術総合研究所首席研究員  
現在に至る

北原 鉄朗（きたはら てつろう）

2002年 東京理科大学理工学部情報科学科卒業  
2004年 京都大学大学院情報学研究科修士課程修了（知能情報学専攻）  
2007年 京都大学大学院情報学研究科博士後期課程修了（知能情報学専攻）  
博士（情報学）  
2010年 日本大学講師  
2014年 日本大学准教授  
2020年 日本大学教授  
現在に至る

深山 覚（ふかやま さとる）

2008年 東京大学理学部地球惑星物理学科卒業  
2010年 東京大学大学院情報理工学系研究科修士課程修了（システム情報学専攻）  
2013年 東京大学大学院情報理工学系研究科博士課程修了（システム情報学専攻）  
博士（情報理工学）  
2013年 産業技術総合研究所研究員  
2017年 産業技術総合研究所主任研究員  
2023年 産業技術総合研究所企画主幹  
現在に至る

竹川 佳成（たけがわ よしなり）

2003年 三重大学工学部情報工学科卒業  
2005年 大阪大学大学院情報科学研究科博士前期課程修了（マルチメディア工学専攻）  
2007年 大阪大学大学院情報科学研究科博士後期課程修了（マルチメディア工学専攻）  
博士（情報科学）  
2007年 神戸大学助教  
2012年 公立はこだて未来大学助教  
2014年 公立はこだて未来大学准教授  
2022年 公立はこだて未来大学教授  
現在に至る

吉井 和佳（よしい かずよし）

2003年 京都大学工学部情報学科卒業  
2005年 京都大学大学院情報学研究科修士課程修了（知能情報学専攻）  
2008年 京都大学大学院情報学研究科博士後期課程修了（知能情報学専攻）  
博士（情報学）  
2008年 産業技術総合研究所研究員  
2013年 産業技術総合研究所主任研究員  
2014年 京都大学講師  
2018年 京都大学准教授  
現在に至る

## 音楽情報処理

Music Information Processing

© Goto, Kitahara, Fukayama, Takegawa, Yoshii 2023

2023年9月28日 初版第1刷発行

★

検印省略

編著者 後藤真孝  
著者 北原鉄朗  
深山覚  
竹川佳成  
吉井和佳  
発行者 株式会社 コロナ社  
代表者 牛来真也  
印刷所 三美印刷株式会社  
製本所 株式会社 グリーン

112-0011 東京都文京区千石 4-46-10

発行所 株式会社 コロナ社  
CORONA PUBLISHING CO., LTD.

Tokyo Japan

振替 00140-8-14844・電話(03)3941-3131(代)

ホームページ <https://www.coronasha.co.jp>

ISBN 978-4-339-01372-6 C3355 Printed in Japan

(松岡)



**JCOPY** <出版者著作権管理機構 委託出版物>

本書の無断複製は著作権法上での例外を除き禁じられています。複製される場合は、そのつど事前に、出版者著作権管理機構（電話 03-5244-5088, FAX 03-5244-5089, e-mail: info@jcopy.or.jp）の許諾を得てください。

本書のコピー、スキャン、デジタル化等の無断複製・転載は著作権法上での例外を除き禁じられています。購入者以外の第三者による本書の電子データ化及び電子書籍化は、いかなる場合も認めていません。落丁・乱丁はお取替えいたします。