

## 1章 音楽と統計的信号処理

亀岡弘和<sup>†</sup>

キーワード：音楽信号処理，多重音解析，音源分離，調・和音推定，ビート解析，楽曲構造解析

## 1. まえがき

携帯音楽プレーヤや音楽配信サービスの普及，データストレージの大容量化などに伴い，楽曲やアーティストの検索，新しいスタイルの音楽鑑賞方法，楽曲提供者の著作権保護などをサポートする技術が重要になっている．音楽音響信号から音楽的に意味のある情報を自動的に取出す音楽信号処理技術はこれらを実現する上で不可欠である．

音声と同様，音楽は人間が発し聴く音のメディアの一つである．音楽信号処理と音声信号処理の研究は関連が深い一方で三つの相違点が挙げられる．まず第一に，音声においては音韻が言語的な役割を担っているのに対し，音楽においては旋律，リズム，和声がその役割を担っている点である．例えば，音声で音韻系列が変われば異なる言語メッセージになるように，音楽で旋律，リズム，和声が変われば異なる曲になる．その意味で，音楽から音高，リズム，和音を認識するのは音声における音声認識に相当している．第二に，音声とは異なり音楽ではほとんどの場合，複数の音が混在していることが前提になっている点である．通常，音声信号処理や音声認識では対象となる音声は一つであり，それ以外の音（雑音）の影響を取り除くことが課題となるが，音楽では混在する複数の楽音が分析対象となる．後述する多重音解析と調波打楽器音分離は，多重音から各楽音の基本周波数（音高）や打楽器音成分を推定する技術である．第三に，音楽は大域的な繰り返し構造や共通構造を有している点である．例えば，ポピュラー音楽ではAメロやサビといったセクションが楽曲中に繰り返される．楽曲構造解析はこのような大域的構造を捉えるための技術である．

次節以降で，音楽信号処理の重要トピックである多重音解析・音源分離，調・和音推定，調波打楽器音分離，ビート解析，楽曲構造解析の研究事例を紹介する．また，関連する，著者らの取組みも紹介する．

## 2. 多重音解析・音源分離

ヴァイオリンなどのようにピッチのある楽音の信号は局所的に周期的である．周期信号を構成する周波数成分の中で最も低い周波数を基本周波数 ( $F_0$ ) といい，音高や知覚量であるピッチに対応する物理量である．多重音解析とは，複数の楽音が重畳した混合信号から個々の楽音の  $F_0$  を推定する問題である．音楽音響信号の  $F_0$  は，曲を特徴づける最も重要な情報の一つで，これを自動獲得できれば自動採譜，楽音分離，音楽検索などさまざまな応用に有用である．音声信号処理の分野でも  $F_0$  推定の研究は長く行われてきたが，そのほとんどは単一音が対象であった．

楽音のスペクトルは  $F_0$  が同一であっても楽器の種類によって大きく異なる．このため，単音の  $F_0$  推定であってもスペクトルの単純なピーク検出だけでは解決にならず，スペクトル全体を手がかりにする必要がある．多重音の場合，各音源の周波数成分を正確に知ることができず， $F_0$  を推定するための重要な手がかりが得られないため，単音の場合よりさらに難しい問題となる．多重音信号が各楽音に分離さえできれば単一音の  $F_0$  推定問題に帰着するため，多重音解析の問題は音源分離の問題と密接に関係している．いずれの問題においても通常は何らかの仮定が必要である．

例えば，各音源の調波性を仮定したアプローチとして，多重音信号中で最も優勢な  $F_0$  を推定するステップと対応する  $F_0$  成分と調波成分を対象の信号から減算するステップを反復的に行う手法<sup>1)</sup> や， $F_0$  をパラメータにもつパラメトリックモデルで観測信号または観測スペクトルにフィッティングする手法<sup>2)~7)</sup> などが提案されている．著書らも，音源分離， $F_0$  推定，発音時刻推定の問題を同時最適化問題として定式化し，楽音のパラメトリックなスペクトログラムモデルを用いて音源分離に相当するステップと各楽音の  $F_0$  軌跡と発音時刻を推定するステップを反復するアルゴリズムを提案している<sup>8)9)</sup>．多重音の  $F_0$  推定法の研究のより詳しい動向については他の著書<sup>10)11)</sup> を参照されたい．

また，調波性の代わりに，各音源のスペクトルが観測区間において何度も生起するという仮定をもとに楽音分離を行うこともできる．例えば，図1 (a)，図1 (c) のようなス

<sup>†</sup> NTTコミュニケーション 科学基礎研究所

"Music and Statistical Signal Processing" by Hirokazu Kameoka (NTT Communication Science Laboratories, NTT Corporation, Kanagawa)

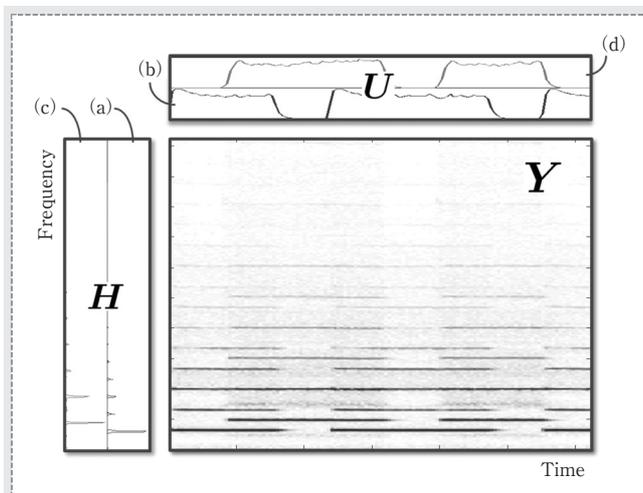


図1 NMFによるスペクトログラムの分解

ペクトルをもつ音が図1 (b), 図1 (d) のような音量軌跡で鳴っていたとする。スペクトログラムが加法的であれば、これら2種類の音の多重音のスペクトログラムは、図1 (a) と図1 (c) を横に並べた行列  $H$  と図1 (b) と図1 (d) を縦に並べた行列  $U$  の積によって表される。したがって、観測された多重音のスペクトログラムを行列と見なし、二つの行列の積に分解することにより、各音源のスペクトルおよび音量軌跡の情報を得ることができる。ただしスペクトルは非負値なので、各行列の要素が非負となるような制約が必要であることから、このアプローチは非負値行列因子分解 (Non-negative Matrix Factorization: NMF) と呼ばれる<sup>12)</sup>。また、楽音の調波構造において  $F_0$  と各倍音の間隔が対数周波数領域では  $F_0$  に依らないことを利用することで、対数周波数スペクトルのテンプレートと  $F_0$  の音量分布との畳み込みにより多重音スペクトルを表現することができる。この考え方に基づく NMF の拡張版<sup>13) 14)</sup> も提案されており、2015年のMIREX (音楽検索に関連する要素技術の国際コンテスト) において本手法が最高性能を示している<sup>15)</sup>。なお、本手法は著者らのSpecmurt法と呼ぶ多重音解析法<sup>16)</sup> に非負値制約を置いたものと見なせる。ところで、上述のように NMF ではスペクトルの加法性が前提となっているが、パワースペクトルや振幅スペクトルの算出過程は非線形であるため、スペクトルは実際には非加法的である。多重音のスペクトルを正確に表現するには、各音源のスペクトルとともに位相を考慮する必要がある。そこで、著者らは NMF と同様のコンセプトを音源信号間の加法性が成り立つ複素スペクトログラム領域および時間領域で実現するアプローチを提案している<sup>17)~20)</sup>。NMF やその拡張版について詳しく知りたい読者は文献<sup>21)</sup> を参照されたい。これらの手法を用いれば、例えば、各楽音の音高や音量をユーザが自分好みに操作可能な音楽再生システムを実現することができる<sup>22) 23)</sup>。

以上の手法は、音源分離と  $F_0$  推定を同時解決する戦略に基づいているが、音源分離は  $F_0$  以外の情報 (各楽音のスペクトルなど) も推定する必要があるため、 $F_0$  推定の問題以上に難しい問題であるといえる。したがって、目的を音高推定に限るなら、多重スペクトルに各音高が含まれているかを識別するマルチクラス分類問題と捉えたアプローチも有効である。例えば、その識別器としてサポートベクターマシンやリカレントニューラルネットワークを用いたアプローチなどが提案されている<sup>24)~26)</sup>。

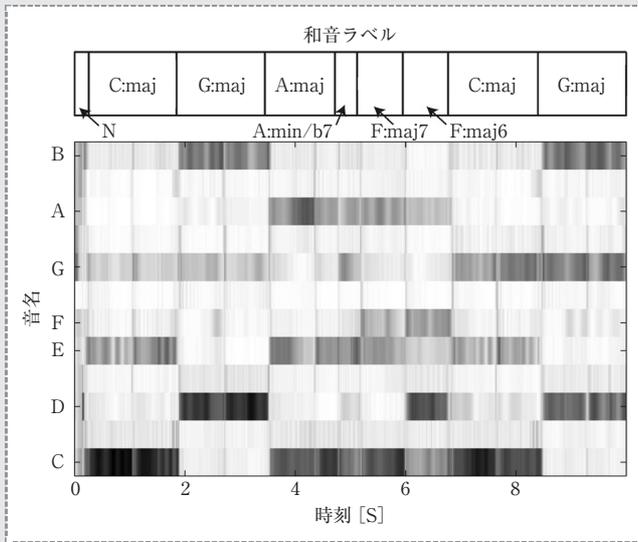
### 3. 調・和音推定

西洋音楽やポピュラー音楽などにおいて調や和音は旋律やリズムと並ぶ楽曲の重要な構成要素である。音楽音響信号から各時刻での調・和音を推定する問題をそれぞれ調推定・和音推定と呼び、楽曲の類似度や構造、ジャンルを認識する問題など幅広い音楽情報検索アプリケーションの基礎技術である。通常、調や和音が同一の区間においても各時刻では構成音の音高は多様に変化するため、各時刻周辺の観測信号のみから調や和音を一意に決定することはできない。また通常、調や和音が増えるタイミングは未知である。したがって、調・和音推定では調・和音区間推定と各区間における調・和音同定の問題を解く必要がある。もし音楽音響信号中で調や和音が同一の区間が分かれば、当該区間において出現する音高の頻度などを手がかりに調や和音を推定することができる。一方、調や和音の出現順序が既知であれば調や和音が増える時刻を推定することが可能である。このように、調・和音区間推定と各区間における調・和音同定の問題は相互依存の関係にある。

以上の性質の問題のため、隠れマルコフモデル (Hidden Markov Model; HMM) やその拡張モデルを用い、同一和音 (または調) の区間推定と各区間の和音 (または調) 推定の同時解決を目指した手法が有効である<sup>27)</sup>。和音特徴量としては、スペクトログラムを音名ごとにオクターブ間で足しあわせたクロマグラム<sup>28) 29)</sup> がしばしば用いられる。図2を見ると、同一の和音において特徴量が類似していることが確認できる。近年では、深層学習アプローチの研究が進んでおり<sup>30)~32)</sup>、MIREX2016で高い性能を発揮している<sup>33)</sup>。

### 4. 調波打楽器音分離

クラシック音楽やポピュラー音楽ではピッチのある楽音 (以後、調波音) と打楽器音が混在することが多い。前者には主に旋律や和声を表現する役割があるのに対し、後者には主にリズムを表現する役割がある。多重音解析と和音認識では音楽音響信号の中の旋律や和声、リズム解析やビート解析ではリズムに関する情報を抽出することが目的であるため、音楽音響信号をこれらの二つのタイプの音に分離する技術が有用となる場面は多い。また、調波音と打楽器音を分離できれば、それぞれの音量を変更できる音楽再生



システムを提供することもできる。これを実現する技術を調波打楽器音分離という。しかし、音源分離の問題と同様、一般に混ざり合った信号を分離する問題は容易でない。

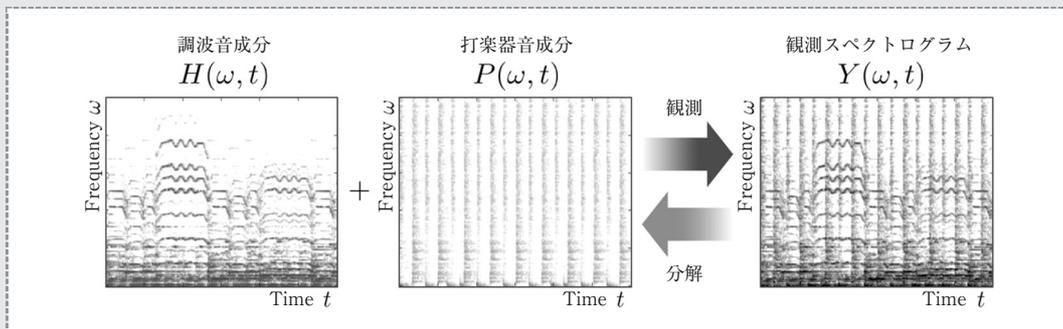
図3のとおり、調波音は周波数成分が時間方向に平行に連なる一方で、打楽器音は周波数成分が周波数方向に平行に連なる傾向にある。前者は、同一音高が一定時間持続することにより、各調波音の調波構造中のピークが時間方向に平行に連なることによる。一方後者は、広帯域におよぶスペクトルが打叩時に急峻に立ち上がりすぐに減衰するためである。著者らは、調波音と打楽器音においてスペクトログラムに現れるこれらの傾向に着目し、画像処理的なアイデアにより観測スペクトログラムを調波音と打楽器音の成分に分解する方法を提案し、Harmonic/Percussive Signal Separation (HPSS) 法と呼んでいる<sup>34)~36)</sup>。また、このアプローチを改良する検討(例えば、文献<sup>37)38)</sup>)も進められている。この他に、打楽器音のスペクトルテンプレートを用いた手法<sup>39)</sup>、独立部分空間法と呼ぶ信号分解法を用いた手法<sup>40)</sup>、前述のNMFに基づくアプローチ<sup>41)</sup>なども提案されている。

## 5. ビート解析

音楽にはほぼ等間隔に繰り返される基本的なリズムがある。これを拍(ビート)といい、音楽音響信号から各拍の時刻や拍の間隔(テンポ)を推定する問題をそれぞれビート解析、テンポ解析という。実際の演奏において、拍は必ずしも正確に等間隔に打たれるわけではなく、演奏の表情付けなどによりその間隔は揺らぐことが多い。また、すべての拍位置で音が発せられるとは限らず、拍位置以外で音が発せられることもあるため、音のありなしの情報だけではビートやテンポを推定することはできないが、拍はほぼ等間隔であること、拍位置において和音が変わりやすいこと、各音が拍位置で発せられる可能性が高いこと、などの手がかりを用いることができる。各時刻において発音された音が存在したらしいかを表すオンセット特徴量の系列から、隠れた周期的なピークを捉えるアプローチが有効である。オンセット特徴量としてはスペクトル変動量<sup>42)</sup>や深層学習により得られる特徴量<sup>43)44)</sup>、特徴量系列の周期性を捉える方法としては短時間フーリエ変換<sup>45)</sup>、マルチエージェントシステム<sup>46)</sup>、HMM<sup>47)</sup>、動的計画法<sup>48)</sup>を用いた手法などが提案されている。

## 6. 楽曲構造解析

楽曲構造解析とは、音楽音響信号をセグメントに分割し、各セグメントを何らかの категория (ポピュラー音楽のサビやAメロ、ソナタ形式の楽曲の提示部や展開部) に分類する問題である。この技術はサビの自動検出や楽曲のサムネイル(試聴用音源など)自動生成などさまざまなアプリケーションに役立つ。構造を基礎づける音楽の構成要素の関係性は「新規性」、「同質性」、「繰り返し構造」といった基準によって作られる。例えば、新規のセクションの開始時にはフィルインなど突然の変化が生じる傾向にあり、同一のセクションの間では調やテンポ、楽器編成などが一貫している傾向にある。また、ポピュラー音楽の1番と2番のサビなどのように、旋律や和音系列、リズムパターンなどが繰り返し用いられていれば、同一のセクションと見な



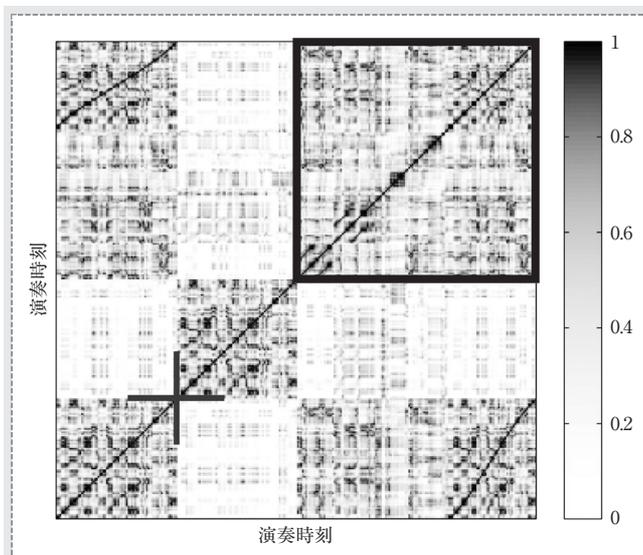


図4 自己類似度行列の例

せる。これらを手がかりに、楽曲全体に隠れた構造をいかにして見い出すが本問題の課題となる。

図4は、2時刻間の特徴量の類似度を各要素にした自己類似度行列を示している。自己類似度行列上で、近辺の類似度が高いブロック状の箇所(赤線で囲まれた部分)は同質性が高く、対角上にあるブロック同士の継ぎ目(青線同士の交点)で新規性が高い。非対角成分上で対角に走る線が繰り返し構造を表している。新規性に着目したアプローチでは、ブロック同士の継ぎ目を見つけ出す問題として定式化され、変化点検知に基づく手法<sup>49)</sup>が提案されている。同質性に着目したアプローチでは、セグメントをクラスタリングする方法<sup>50)</sup>やHMMを用いた手法も提案されている<sup>51)</sup>。繰り返し構造に着目したアプローチでは、非対角成分上の対角に走る線を動的計画法<sup>52)</sup><sup>53)</sup>や画像処理の手法<sup>54)</sup>を用いて検出する方法が提案されている。

## 7. むすび

本稿では、多重音解析、調・和音推定、調波打楽器音分離、ビート解析、楽曲構造解析など、音楽音響信号処理における重要課題と手法を紹介した。音楽信号処理研究の黎明期は、音声の分野で長く培われた方法論や技術を導入しようという事例が多く見られたが、冒頭で述べたように音楽には音声にないさまざまな固有の特徴があることから、音楽ならではの独自の信号処理技術が近年発展してきている。特に音楽信号処理では、多重音を扱うことが必須であることからNMFをはじめとした音源分離の研究が非常に進んでおり、音声信号処理(雑音・残響除去)の分野でも注目されている。一方で、最近では音声分野と足並みをそろえるかのように深層学習を各種タスクに適用する研究が盛んに進められている。音声分野と比べて研究コミュニティがまだまだ小さいことや著作権による制約もあり、学習デー

タセットを効率的に構築できる環境が整っているとは言えないが、MusicNet<sup>55)</sup><sup>56)</sup>のようなアノテーションつきのデータセットが今後増え、深層学習によるブレイクスルーが音楽信号処理分野でも起これば、まだまだ解決すべき課題の多い自動採譜、音楽検索・推薦の問題に突破口が見つかる可能性がある。

謝辞 本稿では、東京大学の客員連携講座(守谷・亀岡研究室)および同大システム情報第一研究室(嵯峨山研究室)における過去の発表資料から図の素材を一部使用した。関係各位に感謝する。また、本稿の研究の一部はJSPS科研費26730100の支援を受けた。(2017年5月16日受付)

## 〔文献〕

- 1) Klapuri, A. "Multiple fundamental frequency estimation based on harmonicity and spectral smoothness", *IEEE Trans. ASLP*, 11, 6, pp.804-816 (2003)
- 2) Goto, M. "A real-time music-scene-description system: Predominant-F0 estimation for detecting melody and bass lines in real-world audio signals", *Speech Commun. (ISCA J.)*, 43, 4, pp.311-329 (2004)
- 3) Kameoka, H., Nishimoto, T. and Sagayama, S. "Separation of harmonic structures based on tied Gaussian mixture model and information criterion for concurrent sounds", in *Proc. ICASSP*, 4, pp.297-300 (2004)
- 4) Christensen, M.G. and Jakobsson, A. "Multi-pitch estimation", in *Synthesis Lectures on Speech Audio Process*. San Rafael, CA: Morgan and Claypool (2009)
- 5) Badeau, R., Emiya, V. and David, B. "Expectationmaximization algorithm for multi-pitch estimation and separation of overlapping harmonic spectra", in *Proc. ICASSP*, pp.3073-3076 (2009)
- 6) Yeh, C. "Multiple fundamental frequency estimation of polyphonic recordings", Ph.D. dissertation, Univ. Pierre et Marie Curie (Paris 6), Paris, France (2008)
- 7) Yoshii, K. and Goto, M. "A nonparametric Bayesian multipitch analyzer based on infinite latent harmonic allocation", *IEEE Trans. ASLP*, 20, 3, pp.717-730 (2012)
- 8) Kameoka, H. "Statistical approach to multipitch analysis", Ph.D. Thesis, the University of Tokyo (2007)
- 9) Kameoka, H., Nishimoto, T. and Sagayama, S. "A multipitch analyzer based on harmonic temporal structured clustering", *IEEE Trans. ASLP*, 15, 3, pp.982-994 (2007)
- 10) de Cheveigné: "Multiple F0 Estimation", in *Computational Auditory Scene Analysis: Principles, Algorithms and Applications*, D.-L. Wang, G.J. Brown Eds., IEEE Press/Wiley (2006)
- 11) *Signal Processing Methods for Music Transcription*, A. Klapuri and M. Davy, Eds. New York: Springer (2006)
- 12) Smaragdis, P. and Brown, J. C. "Non-negative matrix factorization for music transcription", in *Proc. WASPAA*, pp.177-180 (2003)
- 13) Smaragdis, P., Raj, B. and Sashanka, M. "Sparse and shift-invariant feature extraction from nonnegative data", in *Proc. ICASSP*, pp.2069-2072 (2008)
- 14) Benetos, E., Cherla, S. and Weyde, T. "An efficient shift-invariant model for polyphonic music transcription", in *Proc. MML*, pp.1-4 (2013)
- 15) [http://www.music-ir.org/mirex/wiki/2015:MIREX2015\\_Results](http://www.music-ir.org/mirex/wiki/2015:MIREX2015_Results)
- 16) Saito, S., Kameoka, H., Nishimoto, T. and Sagayama, S. "Specmurt analysis of multi-pitch music signals with adaptive estimation of common harmonic structure", in *Proc. ISMIR*, pp.84-91 (2005)
- 17) Kameoka, H., Ono, N., Kashino, K. and Sagayama, S. "Complex NMF: A new sparse representation for acoustic signals", in *Proc. ICASSP*, pp.3437-3440 (2009)
- 18) Kameoka, H. "Multi-resolution signal decomposition with time-domain spectrogram factorization", in *Proc. ICASSP*, pp.86-90 (2015)

- 19) Kameoka, H., Kagami, H. and Yukawa, M.: "Complex NMF with the generalized Kullback-Leibler divergence", in Proc. ICASSP2017, pp.56-60 (2017)
- 20) Kagami, H., Kameoka, H. and Yukawa, M.: "A majorization-minimization algorithm with projected gradient updates for time-domain spectrogram factorization", in Proc. ICASSP, pp.561-565 (2017)
- 21) 亀岡： "非負値行列因子分解とその音響信号処理への応用", 日本統計学会誌, 44, 2, pp.383-407 (2015)
- 22) 亀岡, ルルー, 大石, 柏野: "Music Factorizer: 音楽音響信号をノート単位で編集できるインタフェース", 情報研報, 2009-MUS-81-9 (July 2009)
- 23) 亀岡: "音楽情報処理最前線! 第20回: 市販楽曲を自分好みの曲調に変えられる Music Factorizer", DTM MAGAZINE, 192, 寺島情報企画, pp.98-99 (2010)
- 24) Poliner, G.E. and Ellis, D.P.W.: "A discriminative model for polyphonic piano transcription", EURASIP Journal on Advances in Signal Processing, pp.1-9 (2007)
- 25) Böck, S. and Schedl, M.: "Polyphonic piano note transcription with recurrent neural networks", in Proc. ICASSP, pp.121-124 (2012)
- 26) Sigtia, S., Benetos, E., Boulanger-Lewandowski, N., Weyde, T., d'Avila Garcez, A.S. and Dixon, S.: "A hybrid recurrent neural network for music transcription", in Proc. ICASSP, pp.2061-2065 (2015)
- 27) McVicar, M., Santos-Rodriguez, R., Ni, Y. and Bie, T. D.: "Automatic chord estimation from audio: A review of the state of the art", IEEE Trans. ASLP, 22, 2, pp.556-575 (2014)
- 28) Wakefield, G. H.: "Mathematical representation of joint time-chroma distributions", in Proc. SPIE'99, ASPAAI, 3807, pp.637-645 (1999)
- 29) Fujishima, T.: "Real-time chord recognition of musical sound: A system using common lisp music", in Proc. ICMC, pp.464-467 (1999)
- 30) Boulanger-Lewandowski, N., Bengio, Y. and Vincent, P.: "Audio chord recognition with recurrent neural networks", in Proc. ISMIR, pp.335-340 (2013)
- 31) Korzeniowski, F. and Widmer, G.: "Feature learning for chord recognition: the deep chroma extractor", in Proc. ISMIR, pp.1-7 (2016)
- 32) Korzeniowski, F. and Widmer, G.: "A fully convolutional deep auditory model for musical chord recognition", in Proc. MLSP, pp.1-6 (2016)
- 33) [http://www.music-ir.org/mirex/wiki/2016:MIREX2016\\_Results](http://www.music-ir.org/mirex/wiki/2016:MIREX2016_Results)
- 34) 宮本, ルルー, 亀岡, 小野, 嵯峨山: "スペクトログラムの滑らかさの異方性に基づく調波音・打楽器音の分離", 音講論 (春), No.2-5-8, pp.903-904 (2008)
- 35) Ono, N., Miyamoto, K., Kameoka, H. and Sagayama, S.: "A real-time equalizer of harmonic and percussive components in music signals", in Proc. ISMIR, pp.139-144 (2008)
- 36) Tachibana, H., Ono, N., Kameoka, H. and Sagayama, S.: "Harmonic/Percussive sound separation based on anisotropic smoothness of spectrograms", IEEE Trans. ASLP, 22, 12, pp.2059-2073 (2014)
- 37) FitzGerald, D.: "Harmonic/percussive separation using median filtering", in Proc. DAFx, pp.1-4 (2010)
- 38) Driedger, J., Müller, M. and Disch, S.: "Extending harmonic-percussive separation of audio signals", in Proc. ISMIR, pp.611-616 (2014)
- 39) Yoshii, K., Goto, M. and Okuno, H. G.: "Automatic drum sound description for real-world music using template adaptation and matching methods", in Proc. ISMIR, pp.184-191 (2004)
- 40) Uhle, C., Tiddmar, C. and Sporer, T.: "Extraction of drum tracks from polyphonic music using independent subspace analysis", in Proc. ICA, pp.843-847 (2003)
- 41) Helen, M. and Virtanen, T.: "Separation of drums from polyphonic music using non-negative matrix factorization and support vector machine", in Proc. EUSIPCO, pp.1-4 (2005)
- 42) Masri, P.: "Computer modelling of sound for transformation and synthesis of musical signal", PhD thesis, University of Bristol (1996)
- 43) Eyben, F., Böck, S., Schuller, B. and Graves, A.: "Universal onset detection with bidirectional long short-term memory neural networks", in Proc. ISMIR, pp.589-594 (2010)
- 44) Böck, S. and Schedl, M.: "Enhanced beat tracking with context-aware neural networks", in Proc. DAFx, pp.1-5 (2011)
- 45) Grosche, P., Müller, M. and Kurth, F.: "Cyclic tempogram - a mid-level tempo representation for musicsignals", in Proc. ICASSP, pp.5522-5525 (2010)
- 46) Goto, M.: "An audio-based real-time beat tracking system for music with or without drum-sounds", J. New Music Res., 30, 2, pp.159-171 (2001)
- 47) Whiteley, N., Cemgil, A.T. and Godsill, S. J.: "Bayesian modelling of temporal structure in musical audio", in Proc. ISMIR, pp.29-34 (2006)
- 48) Ellis, D.P.: "Beat tracking by dynamic programming", J. New Music Res., 36, 1, pp.51-60 (2007)
- 49) Foote, J.: "Automatic audio segmentation using a measure of audio novelty", in Proc. ICME, 1, pp.452-455 (2000)
- 50) Cooper, M. and Foote, J.: "Summarizing popular music via structural similarity analysis", in Proc. WASPAA, pp.127-130 (2003)
- 51) Aucouturier, J. J., Pachet, F. and Sandler, M.: "The way it Sounds': timbre models for analysis and retrieval of music signals", IEEE Trans. MM, 7, 6, pp.1028-1035 (2005)
- 52) Goodwin, M.M. and Laroche, J.: "A dynamic programming approach to audio segmentation and speech/music discrimination", in Proc. ICASSP, 4, pp.309-312 (2004)
- 53) Jensen, K.: "Multiple scale music segmentation using rhythm, timbre and harmony", EURASIP J. Applied Signal Process., 2007, 1, pp.159-159 (2007)
- 54) Goto, M.: "A chorus section detection method for musical audio signals and its application to a music listening station", IEEE Trans. ASLP, 14, 5, pp.1783-1794 (2006)
- 55) Thickstun, J., Harchaoui, Z. and Kakade, M.: "Learning features of music from scratch", in Proc. ICLR, pp.1-14 (2017)
- 56) <https://homes.cs.washington.edu/~thickstn/musiconet.html>



かめおか ひろかず  
**亀岡 弘和** 2002年, 東京大学工学部計数工学科卒業。2004年, 同大学大学院情報理工学系研究科修士課程修了。2007年, 同大学大学院情報理工学系研究科博士課程修了。同年, 日本電信電話(株)入社。NTTコミュニケーション科学基礎研究所配属。2011年, 東京大学大学院情報理工学系研究科客員准教授。2016年, 国立情報学研究所客員准教授。音声・音楽を対象とした音響信号処理・機械学習の研究に従事。IEEE Signal Processing Society 2008 SPS Young Author Best Paper Award等受賞多数。博士(情報理工学)。