

令和 5 年 6 月 12 日現在

機関番号：14301

研究種目：基盤研究(B)（一般）

研究期間：2019～2022

課題番号：19H04137

研究課題名（和文）認識・生成過程の統合に基づく視聴覚音楽理解

研究課題名（英文）Audio-Visual Music Understanding Based on Integration of Recognition and Generative Processes

研究代表者

吉井 和佳 (Yoshii, Kazuyoshi)

京都大学・情報学研究科・准教授

研究者番号：20510001

交付決定額（研究期間全体）：（直接経費） 13,200,000円

研究成果の概要（和文）：聴覚系に関する研究成果として、歌声採譜、拍節構造推定、楽曲構造推定、コード・キー推定、ドラム採譜など、自動採譜を構成する重要なサブタスクに対して、確率的生成モデルとその統計的推論という統一的なアプローチに基づく解決法を提示することができた。いくつかのタスクで、変分自己符号化器（VAE）を構成することで、理論的には両者を同時に教師なし学習できることを示した。視覚系に関する研究成果として、画像中の人間の二次元姿勢推定において、聴覚系と同様に、確率的生成モデルとその統計的推論という統一的なアプローチに基づく解決法を提示することができた。

研究成果の学術的意義や社会的意義

人間が視聴覚を通じて音楽を理解する機構に対して、表裏一体の関係にある生成過程と推論過程を統合した計算モデルを提示することができた。このモデルは、認知科学分野で知られていたミラーニューロン仮説に着想を得ており、統計的機械学習の見地からは、変分自己符号化器（VAE）として定式化できることを示した。自動採譜のいくつかの課題や姿勢推定でこのモデルの有効性を示した。

研究成果の概要（英文）：As for auditory understanding, we have developed a unified approach based on statistical inference of probabilistic generative models to various important subtasks of automatic music transcription including singing voice transcription, music structure analysis, chord and key estimation, and drum transcription. We showed that the generative and inference models can be integrated in the VAE framework. As for visual understanding, we have developed a pose estimation method based on the same approach.

研究分野：音楽情報処理

キーワード：自動採譜 自動編曲 姿勢推定 確率的生成モデル 深層学習 償却型変分推論

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

### 1. 研究開始当初の背景

深層学習技術の発展で、時系列信号解析タスクでは目覚ましい精度向上が達成されてきた。その代表例である音声認識では、音声信号  $X$  を入力として、文字列あるいは単語列  $S$  を出力する深層ニューラルネットワーク (DNN) を教師あり学習するアプローチが主流となっていた。一方で、入力  $X$  と出力  $S$  のペアデータの存在を前提にしたアプローチの限界もまた指摘されていた。WEB 上には膨大な未アノテーションの音声信号が蓄積されているが、それらを有効活用する技術に関してはまだ研究がほとんど進んでいなかった。

音楽情報処理分野における主要タスクである自動採譜でも、音響信号  $X$  と楽譜  $S$  のペアデータを用いて DNN を教師あり学習する手法が台頭している。調波構造や時間構造を明示的に表現した音響信号スペクトログラムの生成モデルを教師なし学習する従来法より精度の点で優れているが、ピアノ曲に限定しても 80~90%程度の推定精度が限界であり、採譜結果には人間であれば見逃さないような音楽的に不自然な誤りが多く含まれていた。音楽のアノテーションはコストが高く、利用できる学習データ量は音声認識より数オーダー小さいことも問題であった。

### 2. 研究の目的

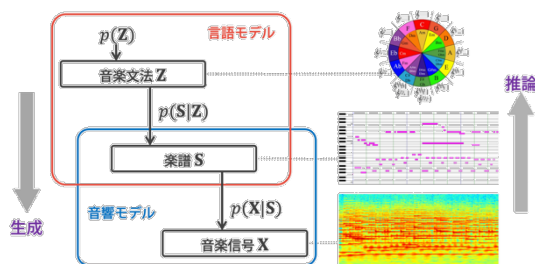
本研究では、人間が音楽を「理解」する仕組みを統計的な見地から解明することを目指した。我々は日常的に「理解が浅い・深い」という表現を用いるが、単に楽譜を書き起こしたり、楽譜を演奏したりできるのが「浅い理解」であり、抽象的な音楽・身体感覚の獲得が「深い理解」に相当するという仮説を立てた。もし、自らの体験に基づき、対象をうまく説明できるのであれば、理解が適切であると考えられる。統計的には、観測データに対する推論・生成モデルの尤度、すなわち推論・生成過程を通じた観測データの再現度を評価できる枠組みが必要であった。

### 3. 研究の方法

人間の知能の働きに着目し、表裏一体の関係にある推論タスクと生成タスクを同時に扱うための統一モデルを構築することを目指す。具体的には、音楽理解の聴覚系と視覚系に対する教師なし学習系の構築・学習を試みる。

- ・ 聴覚系：音響信号  $X$  → 楽譜  $S$  → 音楽感覚  $Z$  → 楽譜  $S$  → 音響信号  $X$
- ・ 視覚系：映像信号  $X$  → 姿勢  $S$  → 身体感覚  $Z$  → 姿勢  $S$  → 映像信号  $X$

具体的には、確率的生成モデルとその統計的推論の立場から、代表的な推論タスクである自動採譜・姿勢推定を軸として、生成タスクを統合する方法論に取り組んだ。



### 4. 研究成果

聴覚系を視覚系それぞれに分けて、主要な研究成果を以下に説明する。

#### (1) 聴覚系の計算モデル

聴覚系に関する研究成果として、歌声採譜、拍節構造推定、楽曲構造推定、コード・キー推定、ドラム採譜など、自動採譜を構成する重要なサブタスクに対して、確率的生成モデルとその統計的推論という統一的なアプローチに基づく解決法を提示することができた。特に、コード・キー推定とドラム採譜に関しては、音楽信号の生成モデル (音響モデル) を定式化したうえで、償却型変分推論法 (AVI) 推論の枠組みを用いて推論モデルと事前分布 (言語モデル) とを導入し、変分自己符号化器 (VAE) を構成することで、理論的には両者を同時に教師なし学習できることを示した (学習の安定化と性能向上のため、実際には半教師あり学習あるいは教師あり学習を行うことが有用)。これら代表的な研究成果について以下に具体的に報告する。

#### ① コード・キー推定

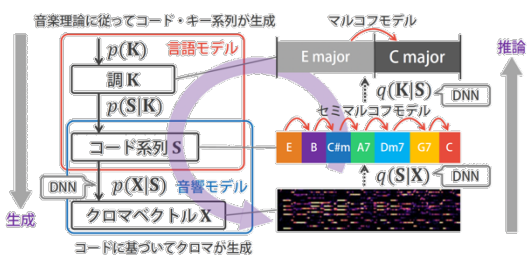
本研究では、コードやキーのアノテーションがなされていない音楽信号を有効活用できるコード・キー推定法を提案した。従来のコード・キー推定の典型的なアプローチは、アノテーション付きの音楽信号を用いて、音楽信号からコード・キーラベル系列を直接推定する深層識別モデルを教師あり学習するものであった。このような識別的アプローチでは、コード・キーラベル系列に関する音楽知識 (事前知識) はほとんど考慮されていなかった。

これに対し、本研究では、AVI の枠組みで、生成的アプローチと識別的アプローチを統合することに成功した。具体的には、自己遷移確率を高く設定したマルコフモデルあるいは事前に教師なし学習した深層生成モデルに従うと仮定したコード・キーラベル系列 (潜在変数) と、標準ガウス分布に従うと仮定したと潜在特徴量系列 (潜在変数) から、クロマベクトル (観測変数) が生成される過程を表現する階層的な深層生成モデルを定式化する。反対に、観測データとしてクロマベクトルが与えられると、潜在ラベル系列および潜在特徴量系列の事後分布は、それぞれ深

層識別モデルと深層推論モデルを用いて推定する。これら三つのモデルを内包する VAE を構成することで、HMM と同様、最尤推定の枠組みで半教師あり学習が可能になった。

実験の結果、コードラベルの生成モデル（言語モデル）とクロマベクトルの深層生成モデル（音響モデル）の両者が、識別モデルの正則化として働くことで、教師ありの条件下でコード推定精度が向上することを示した。さらに、アノテーションなしの音楽信号を用いた半教師あり学習により、精度が改善することを示した。本研究は、生成的アプローチと識別的アプローチを統合した先駆的な研究に位置づけられる。

• DNN-HSMMハイブリッドモデル+DNN推論モデル=VAE



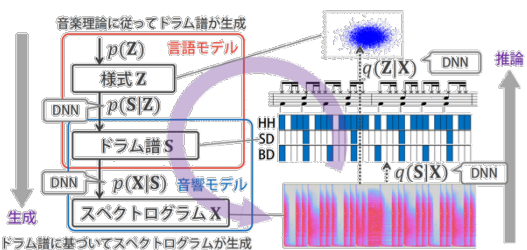
## ② ドラム採譜

本研究では、音楽信号からテイタム（16部音符）単位でドラム譜を推定する手法を提案した。従来のドラム採譜の典型的なアプローチは、アノテーション付きの音楽信号を用いて、音楽信号からドラムの発音時刻をフレーム単位で検出する深層識別モデルを教師あり学習するものであった。ただし、楽譜推定を目標とする本来の自動採譜では、各音符を楽譜上に正しく割り付けるため、時刻情報を正しく量子化する必要がある。しかし、ビート・ダウンビートからなるグリッド上で単純な量子化処理を別途行う方法には性能に限界があった。このため、ビート・ダウンビート情報を既知としつつも、量子化処理を同時学習する方法論が求められていた。

これに対し、本研究では、AVIの枠組みで、生成的アプローチと識別的アプローチを統合することに成功した。具体的には、事前に既存ドラム譜から教師なし学習した深層生成モデルに従うと仮定するテイタム単位のドラム譜断片系列（潜在変数）と、標準ガウス分布に従うと仮定したと潜在特徴量系列（潜在変数）から、スペクトログラム（観測変数）が生成される過程を表現する階層的な深層生成モデルを定式化する。反対に、観測データとしてスペクトログラムが与えられると、潜在ドラム譜断片系列および潜在特徴量系列の事後分布は、それぞれ深層識別モデルと深層推論モデルを用いて推定する。これら三つのモデルを内包するVAEを構成することで、半教師あり学習が可能になった。本研究では、多チャンネル音声強調において、音声スペクトルの深層生成モデルを事前分布として用いる研究に着想を得て、ドラム譜の深層生成モデルとして、教師なし学習で得られたVAEのデコーダを用いた点が特徴的である。

実験の結果、ドラム譜の深層生成モデル（言語モデル）とスペクトログラムの深層生成モデル（音響モデル）の両者が、識別モデルの正則化として働くことで、教師ありの条件下でドラム譜推定精度が向上することを示した。ドラム採譜において、ドラムの発音時刻検出とスペクトログラムの再構成をアドホックに組み合わせた研究は存在したが、確率的な枠組みで理論的に統合したのは本研究が初めてであった。

• DNN-DNNモデル+DNN推論モデル=VAE

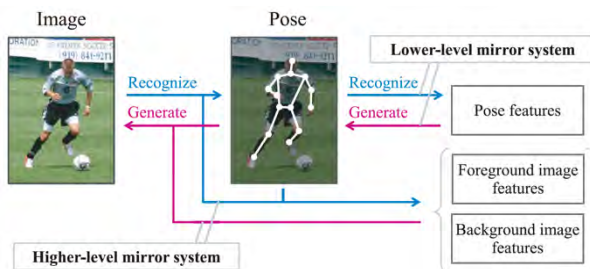


## (2) 視覚系の計算モデル

視覚系に関する研究成果として、画像中の人間の二次元姿勢推定において、聴覚系と同様に、確率的生成モデルとその統計的推論という統一的なアプローチに基づく解決法を提示することができた。本手法は、解析対象をダンス画像に限定しない汎用手法であるが、のちにダンス映像からの姿勢推定にも応用され、安定した推定が行えることを実験的に確認している。

本研究では、人物画像から姿勢推定を行うための確率的生成モデルに基づく統計的アプローチを考案した。従来の姿勢推定の典型的なアプローチは、画像から姿勢を直接推定する深層推論モデルを教師あり学習するものであった。しかし、解剖学的にはありえない姿勢がしばしば出力されるうえ、教師データの量によって性能に限界があるという問題を抱えていた。

これに対し、本研究では、他者の動作を理解するうえでは、自らの身体による動作のシミュレーションが重要な役割を果たしているというミラーニューロン仮説に着想を得た手法を考案した。具体的には、姿勢の潜在特徴から姿勢が生成される過程を表現する深層生成モデル（高次ミラーシステム）と、姿勢と画像の潜在特徴から画像が生成される過程を表現する深層生成モデル（低次ミラーシステム）を統合し、姿勢と画像の階層的深層生成モデルを定式化する。さらに、AVIの枠組みを用いて、画像から姿勢を推定する深層推論モデルを導入する。観測データとして画像が与えられると、階層VAEとしてすべてのモデルを同時学習できることを示した。



提案するMirrorNetのアーキテクチャは、姿勢に対するVAE（低次ミラーシステム）と画像に対する姿勢条件付きVAE（高次ミラーシステム）を統合した階層ベイズモデルである。生成過程

の観点では、姿勢 VAE のデコーダは姿勢の事前分布として姿勢の自然さを評価し、画像 VAE のデコーダは姿勢の尤度関数として姿勢の忠実さを評価する。推論過程の観点では、姿勢 VAE のエンコーダを姿勢の変分事後分布として使用する。

実験の結果、MirrorNet に基づく提案手法は、解剖学的に破綻のない姿勢を推定することができることを示した。さらに、アノテーションを持たない画像を併用した半教師あり学習により、姿勢推定精度が向上することを示した。姿勢推定において、これまで、画像の解析と再構成をアドホックに組み合わせた研究は存在したが、確率的な枠組みで階層的な統合に成功したのは本研究が初めてであった。

## 5. 主な発表論文等

〔雑誌論文〕 計13件（うち査読付論文 13件 / うち国際共著 0件 / うちオープンアクセス 2件）

1. 著者名 Wu Yiming, Yoshii Kazuyoshi	4. 巻 11
2. 論文標題 Joint Chord and Key Estimation Based on a Hierarchical Variational Autoencoder with Multi-task Learning	5. 発行年 2022年
3. 雑誌名 APSIPA Transactions on Signal and Information Processing	6. 最初と最後の頁 1-27
掲載論文のDOI (デジタルオブジェクト識別子) 10.1561/116.00000052	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Ryoto Ishizuka, Ryo Nishikimi, Kazuyoshi Yoshii	4. 巻 2
2. 論文標題 Global Structure-Aware Drum Transcription Based on Self-Attention Mechanisms	5. 発行年 2021年
3. 雑誌名 Signals	6. 最初と最後の頁 508-526
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/signals2030031	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Takayuki Nakatsuka, Kazuyoshi Yoshii, Yuki Koyama, Satoru Fukayama, Masataka Goto, Shigeo Morishima	4. 巻 29
2. 論文標題 MirrorNet: A Deep Reflective Approach to 2D Pose Estimation for Single-Person Images	5. 発行年 2021年
3. 雑誌名 Journal of Information Processing	6. 最初と最後の頁 406-423
掲載論文のDOI (デジタルオブジェクト識別子) 10.2197/ipsjip.29.406	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Eita Nakamura, Kazuyoshi Yoshii	4. 巻 572
2. 論文標題 Musical Rhythm Transcription Based on Bayesian Piece-Specific Score Models Capturing Repetitions	5. 発行年 2021年
3. 雑誌名 Information Sciences	6. 最初と最後の頁 482-500
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.ins.2021.04.100	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Ryo Nishikimi, Eita Nakamura, Masataka Goto, Kazuyoshi Yoshii	4. 巻 10
2. 論文標題 Audio-to-Score Singing Transcription Based on a CRNN-HSMM Hybrid Model	5. 発行年 2021年
3. 雑誌名 APSIPA Transactions on Signal and Information Processing	6. 最初と最後の頁 1-13
掲載論文のDOI (デジタルオブジェクト識別子) 10.1017/atsip.2021.4	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Kentaro Shibata, Eita Nakamura, Kazuyoshi Yoshii	4. 巻 566
2. 論文標題 Non-local Musical Statistics as Guides for Audio-to-Score Piano Transcription	5. 発行年 2021年
3. 雑誌名 Information Sciences	6. 最初と最後の頁 262-280
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.ins.2021.03.014	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Hiroaki Tsushima, Eita Nakamura, Kazuyoshi Yoshii	4. 巻 28
2. 論文標題 Bayesian Melody Harmonization Based on a Tree-Structured Generative Model of Chord Sequences and Melodies	5. 発行年 2020年
3. 雑誌名 IEEE/ACM Transactions on Audio, Speech, and Language Processing	6. 最初と最後の頁 1644 ~ 1655
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/TASLP.2020.2996088	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Ryo Nishikimi, Eita Nakamura, Masataka Goto, Katsutoshi Itoyama, Kazuyoshi Yoshii	4. 巻 28
2. 論文標題 Bayesian Singing Transcription Based on a Hierarchical Generative Model of Keys, Musical Notes, and F0 Trajectories	5. 発行年 2020年
3. 雑誌名 IEEE/ACM Transactions on Audio, Speech, and Language Processing	6. 最初と最後の頁 1678 ~ 1691
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/TASLP.2020.2996095	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Kouhei Sekiguchi, Yoshiaki Bando, Aditya Arie Nugraha, Kazuyoshi Yoshii, Tatsuya Kawahara	4. 巻 28
2. 論文標題 Fast Multichannel Nonnegative Matrix Factorization With Directivity-Aware Jointly-Diagonalizable Spatial Covariance Matrices for Blind Source Separation	5. 発行年 2020年
3. 雑誌名 IEEE/ACM Transactions on Audio, Speech, and Language Processing	6. 最初と最後の頁 2610 ~ 2625
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/TASLP.2020.3019181	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Yiming Wu, Tristan Carsault, Eita Nakamura, Kazuyoshi Yoshii	4. 巻 28
2. 論文標題 Semi-Supervised Neural Chord Estimation Based on a Variational Autoencoder With Latent Chord Labels and Features	5. 発行年 2020年
3. 雑誌名 IEEE/ACM Transactions on Audio, Speech, and Language Processing	6. 最初と最後の頁 2956 ~ 2966
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/TASLP.2020.3035001	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Aditya Arie Nugraha, Kouhei Sekiguchi, Mathieu Fontaine, Yoshiaki Bando, Kazuyoshi Yoshii	4. 巻 27
2. 論文標題 Flow-Based Independent Vector Analysis for Blind Source Separation	5. 発行年 2020年
3. 雑誌名 IEEE Signal Processing Letters	6. 最初と最後の頁 2173 ~ 2177
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/LSP.2020.3039944	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Kentaro Shibata, Eita Nakamura, Kazuyoshi Yoshii	4. 巻 566
2. 論文標題 Non-local Musical Statistics as Guides for Audio-to-Score Piano Transcription	5. 発行年 2021年
3. 雑誌名 Information Sciences	6. 最初と最後の頁 262 ~ 280
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.ins.2021.03.014	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Eita Nakamura, Yasuyuki Saito, Kazuyoshi Yoshii	4. 巻 517
2. 論文標題 Statistical Learning and Estimation of Piano Fingering	5. 発行年 2020年
3. 雑誌名 Information Sciences	6. 最初と最後の頁 68-85
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.ins.2019.12.068	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計27件 (うち招待講演 0件 / うち国際学会 27件)

1. 発表者名 Florian Thalmann, Eita Nakamura, Kazuyoshi Yoshii
2. 発表標題 Tracking the Evolution of a Band's Performances over Decades
3. 学会等名 International Society for Music Information Retrieval Conference (ISMIR) (国際学会)
4. 発表年 2022年

1. 発表者名 Tengyu Deng, Eita Nakamura, Kazuyoshi Yoshii
2. 発表標題 End-to-End Lyrics Transcription Informed by Pitch and Onset Estimation
3. 学会等名 International Society for Music Information Retrieval Conference (ISMIR) (国際学会)
4. 発表年 2022年

1. 発表者名 Keitaro Tanaka, Yoshiaki Bando, Kazuyoshi Yoshii, Shigeo Morishima
2. 発表標題 Unsupervised Disentanglement of Timbral, Pitch, and Variation Features From Musical Instrument Sounds With Random Perturbation
3. 学会等名 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) (国際学会)
4. 発表年 2022年



1. 発表者名 Moyu Terao, Yuki Hiramatsu, Ryoto Ishizuka, Yiming Wu, Kazuyoshi Yoshii
2. 発表標題 Difficulty-Aware Neural Band-to-Piano Score Arrangement Based on Note- and Statistic-Level Criteria
3. 学会等名 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (国際学会)
4. 発表年 2022年

1. 発表者名 Yuki Hiramatsu, Eita Nakamura, Kazuyoshi Yoshii
2. 発表標題 Joint Estimation of Note Values and Voices for Audio-to-Score Piano Transcription
3. 学会等名 International Society for Music Information Retrieval Conference (ISMIR) (国際学会)
4. 発表年 2021年

1. 発表者名 Takehisa Oyama, Ryoto Ishizuka, Kazuyoshi Yoshii
2. 発表標題 Phase-Aware Joint Beat and Downbeat Estimation Based on Periodicity of Metrical Structure
3. 学会等名 International Society for Music Information Retrieval Conference (ISMIR) (国際学会)
4. 発表年 2021年

1. 発表者名 Yuki Hiramatsu, Go Shibata, Ryo Nishikimi, Eita Nakamura, Kazuyoshi Yoshii
2. 発表標題 Statistical Correction of Transcribed Melody Notes Based on Probabilistic Integration of a Music Language Model and a Transcription Error Model
3. 学会等名 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (国際学会)
4. 発表年 2021年

1. 発表者名 Keitaro Tanaka, Ryo Nishikimi, Yoshiaki Bando, Kazuyoshi Yoshii, Shigeo Morishima
2. 発表標題 Pitch-Timbre Disentanglement of Musical Instrument Sounds Based on VEA-Based Metric Learning
3. 学会等名 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (国際学会)
4. 発表年 2021年

1. 発表者名 Mitsuki Hosoya, Masanori Morise, Satoshi Nakamura, Kazuyoshi Yoshii
2. 発表標題 A Real-Time Drum-Wise Volume Visualization System for Learning Volume-Balanced Drum Performance
3. 学会等名 International Conference on Entertainment Computing (ICEC) (国際学会)
4. 発表年 2021年

1. 発表者名 Andrew McLeod, James Owers, Kazuyoshi Yoshii
2. 発表標題 The MIDI Degradation Toolkit: Symbolic Music Augmentation and Correction
3. 学会等名 International Society for Music Information Retrieval Conference (ISMIR) (国際学会)
4. 発表年 2020年

1. 発表者名 Florian Thalmann, Kazuyoshi Yoshii, Wiggins Geraint, Mark B. Sandler
2. 発表標題 A Method for Analysis of Shared Structure in Large Music Collections Using Techniques from Genetic Sequencing and Graph Theory
3. 学会等名 International Society for Music Information Retrieval Conference (ISMIR) (国際学会)
4. 発表年 2020年

1. 発表者名 Keitaro Tanaka, Takayuki Nakatsuka, Ryo Nishikimi, Kazuyoshi Yoshii, Shigeo Morishima
2. 発表標題 Multi-Instrument Music Transcription Based on Deep Spherical Clustering of Spectrograms and Pitchgrams
3. 学会等名 International Society for Music Information Retrieval Conference (ISMIR) (国際学会)
4. 発表年 2020年

1. 発表者名 Go Shibata, Ryo Nishikimi, Kazuyoshi Yoshii
2. 発表標題 Music Structure Analysis Based on an LSTM-HSMM Hybrid Model
3. 学会等名 International Society for Music Information Retrieval Conference (ISMIR) (国際学会)
4. 発表年 2020年

1. 発表者名 Mathieu Fontaine, Kouhei Sekiguchi, Aditya Arie Nugraha, Kazuyoshi Yoshii
2. 発表標題 Unsupervised Robust Speech Enhancement Based on Alpha-Stable Fast Multichannel Nonnegative Matrix Factorization
3. 学会等名 Annual Conference of the International Speech Communication Association (Interspeech) (国際学会)
4. 発表年 2020年

1. 発表者名 Yoshiaki Bando, Kouhei Sekiguchi and Kazuyoshi Yoshii
2. 発表標題 Adaptive Neural Speech Enhancement with a Denoising Variational Autoencoder
3. 学会等名 Annual Conference of the International Speech Communication Association (Interspeech) (国際学会)
4. 発表年 2020年

1. 発表者名 Jeongwoo Woo, Masato Mimura, Kazuyoshi Yoshii, Tatsuya Kawahara
2. 発表標題 End-to-End Music-Mixed Speech Recognition
3. 学会等名 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) (国際学会)
4. 発表年 2020年

1. 発表者名 Yiming Wu, Eita Nakamura, Kazuyoshi Yoshii
2. 発表標題 A Variational Autoencoder for Joint Chord and Key Estimation from Audio Chromagrams
3. 学会等名 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) (国際学会)
4. 発表年 2020年

1. 発表者名 yoto Ishizuka, Ryo Nishikimi, Eita Nakamura, Kazuyoshi Yoshii
2. 発表標題 Tatum-Level Drum Transcription Based on a Convolutional Recurrent Neural Network with Language Model-Based Regularized Training
3. 学会等名 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) (国際学会)
4. 発表年 2020年

1. 発表者名 Yicheng Du, Kouhei Sekiguchi, Yoshiaki Bando, Aditya Arie Nugraha, Mathieu Fontaine, Kazuyoshi Yoshii, Tatsuya Kawahara
2. 発表標題 Semi-supervised Multichannel Speech Separation Based on a Phone- and Speaker-Aware Deep Generative Model of Speech Spectrograms
3. 学会等名 European Signal Processing Conference (EUSIPCO) (国際学会)
4. 発表年 2020年

1 . 発表者名 Kazuyoshi Yoshii, Kouhei Sekiguchi, Yoshiaki Bando, Mathieu Fontaine, Aditya Arie Nugraha
2 . 発表標題 Fast Multichannel Correlated Tensor Factorization for Blind Source Separation
3 . 学会等名 European Signal Processing Conference (EUSIPCO) (国際学会)
4 . 発表年 2020年

1 . 発表者名 Go Shibata, Ryo Nishikimi, Eita Nakamura, Kazuyoshi Yoshii
2 . 発表標題 Statistical Music Structure Analysis Based on a Homogeneity-, Repetitiveness-, and Regularity-Aware Hierarchical Hidden Semi-Markov Model
3 . 学会等名 International Society for Music Information Retrieval Conference (ISMIR) (国際学会)
4 . 発表年 2019年

1 . 発表者名 Adrien Ycart, Andrew McLeod, Emmanouil Benetos, Kazuyoshi Yoshii
2 . 発表標題 Blending Acoustic and Language Model Predictions for Automatic Music Transcription
3 . 学会等名 International Society for Music Information Retrieval Conference (ISMIR) (国際学会)
4 . 発表年 2019年

1 . 発表者名 Ryo Nishikimi, Eita Nakamura, Masataka Goto, Kazuyoshi Yoshii
2 . 発表標題 End-to-End Melody Note Transcription Based on a Beat-Synchronous Attention Mechanism
3 . 学会等名 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA) (国際学会)
4 . 発表年 2019年

1. 発表者名 Tomoyasu Nakano, Kazuyoshi Yoshii, Yiming Wu, Ryo Nishikimi, Kin Wah Edward Lin, Masataka Goto
2. 発表標題 Joint Singing Pitch Estimation and Voice Separation Based on a Neural Harmonic Structure Renderer
3. 学会等名 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA) (国際学会)
4. 発表年 2019年

1. 発表者名 Tristan Carsault, Andrew McLeod, Philippe Esling, Jerome Nika, Eita Nakamura, Kazuyoshi Yoshii
2. 発表標題 Multi-Step Chord Sequence Prediction Based on Aggregated Multi-Scale Encoder-Decoder Networks
3. 学会等名 IEEE International Workshop on Machine Learning for Signal Processing (MLSP) (国際学会)
4. 発表年 2019年

1. 発表者名 Yiming Wu, Tristan Carsault, Kazuyoshi Yoshii. Automatic Chord Estimation Based on a Frame-wise Convolutional Recurrent Neural Network with Non-Aligned Annotations
2. 発表標題 Automatic Chord Estimation Based on a Frame-wise Convolutional Recurrent Neural Network with Non-Aligned Annotations
3. 学会等名 European Signal Processing Conference (EUSIPCO) (国際学会)
4. 発表年 2019年

1. 発表者名 Takayuki Nakatsuka, Masatoshi Hamanaka, Shigeo Morishima
2. 発表標題 Audio-Guided Video Interpolation via Human Pose Features
3. 学会等名 International Conference on Computer Vision Theory and Applications (VISAPP) (国際学会)
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分担者	河原 達也  (Kawahara Tatsuya)  (00234104)	京都大学・情報学研究科・教授   (14301)	
研究 分担者	森島 繁生  (Morishima Shigeo)  (10200411)	早稲田大学・理工学術院・教授   (32689)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------