

安全安心な社会を実現し豊かな文化を創造する コンテンツ・映像処理技術研究

研究代表者 森島 繁生
(先進理工学部 応用物理学科 教授)

1. 研究課題

JST 戦略的創造研究推進事業 ACCEL「次世代メディアコンテンツ生態系技術の基盤構築と応用展開」(課題予算計 1,000,000 千円)の一環として進行している映像コンテンツ創作・鑑賞支援技術開発に関わる研究課題を遂行している。また、平成 19 年度からは、文部科学省科学研究費基盤(A)「スキルやモチベーションを向上させる現実歪曲時空間の解明」および JST 未来社会創造事業「分散型匿名化処理によるプライバシープリザード AI 基盤構築」のサポートにより研究を遂行している。これらによる研究ゴールの1つとして1枚画像からの人物キャラクターの自動生成の研究を推進している。現在は開発技術のコンテンツ・映像業界への技術展開のみならず、障害者支援など社会貢献をも視野に入れている点が特徴である。

今年度は特に、音楽情報処理に深層学習のアプローチを取り入れ、複数の楽器の演奏音楽ソースから深層クラスタリングを用いて半教師無し学習によって自動採譜する技術を世界で初めて確立した。また視覚障害者支援システムの開発に取り組み、コロナ禍において、列を発見し、ソーシャルディスタンスを保ちながら、この列に追従するシステムを、携帯電話アプリケーションとして実現した。この成果に対して早稲田大学より2名の学生が小野梓記念賞をそれぞれ受賞した。

また前者の自動採譜の研究は、ISMIR2020 および ICASSP2021 という音楽情報処理および音響信号処理のトップコンファレンスにおいて論文採択が行われ、また情報処理学会から学術奨励賞を受賞した。後者の視覚障害者支援システムは、ヒューマンコンピュータインタラクションのトップコンファレンス ACM CHI2021 で論文採択が行われ、ソフトウェア科学会主催の WISS2020 において最優秀論文賞を受賞した。この他にも、1台のカメラの動画画像から人物の3次元の動きを推定しボーンおよび関節の位置を精度よく検出する半教師無しの学習技術の研究など、多くの学術的な貢献を果たした。

2. スペクトログラムとピッチグラムの深層クラスタリングに基づく複数楽器パート採譜

本研究は任意の複数楽器で演奏された音楽音響信号に対し、各楽器パートのピアノロールを推定するための深層クラスタリングに基づく採譜手法について提案する。採譜対象の楽曲が常に特定の楽器で演奏されている場合、各ピアノロールを得るための直接的な方法は、深層ニューラルネットワーク(Depth Neural Network, DNN)を用いて各楽器ごとのピッチグラム(音高サリエンススペクトログラム)を推定する手法である。しかしながらこの手法には、事前指定外の楽器を含む楽曲を取り扱うことができないという致命的な限界がある。本研究では、楽器に依存しない音高推定器を用いてコンデンスピッチグラムを推定した後、深層球面クラスタリングによって指定した数の楽器パートに分離する。採譜精度向上のため、各楽器の音色特徴量と音高特徴量に基づくスペクトログラムとピッチグラムの同時クラスタリングを提案する。実験では、提案手法により事前指定外の未知楽器を含む楽曲に対しても、既知楽器のみで構成される楽曲とほぼ同程度かつ最先端の精度で採譜を行うことができることを確認した。

音楽音響信号中の多重音基本周波数推定 (multi-pitchestimation, MPE) は、音楽音響信号を自動的に楽譜化する自動採譜にとって基盤となる技術であり、音楽情報処理において重要な役割を担っている。従来のこの MPE の手法は、主に一種類の楽器で構成された音楽音響信号の採譜を目的としていた。この単一楽器 MPE は深層学習によって精度が大幅に改善し、近年では単一楽器 MPE を拡張し一般化を進めたものとして複数楽器 MPE が研究されている。複数楽器 MPE とは、複数楽器で構成された音楽音響信号中の各楽器に対して、各時刻の演奏音を表すピッチグラム (音高サリエンススペクトログラム) を推定するものである。複数楽器の場合、各ピッチグラムの帰属先となる楽器を推定する必要が追加される。この困難を軽減するため、先行研究では対象楽器を少数の事前指定楽器に限定してきた。この条件下では、複数楽器 MPE においてクラス分類手法を適用することで、音楽音響信号から直接各ピッチグラムを推定する方法が考えられる。

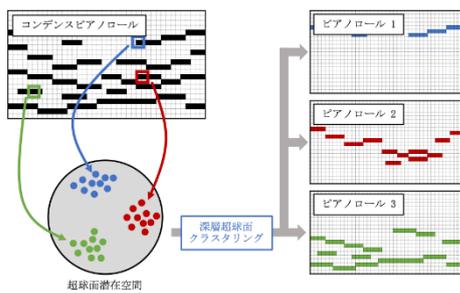


図1 パート譜採譜法の概念図

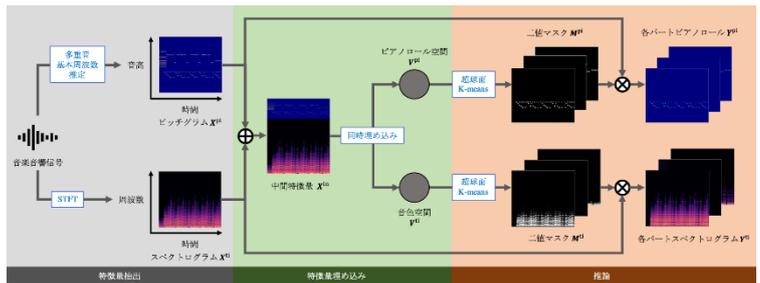


図2 提案ネットワーク全体図

これらのクラス分類に基づく手法は、特にクラシック音楽のように構成楽器がほぼ固定されている楽曲に対し、教師あり学習の枠組みにおいて成功を収めてきた。しかし、ポップスや EDM のように多くの種類の楽器が使用される現代の楽曲に対して質の高い自動採譜を達成するためには、採譜対象となる楽器に制約のないことが理想的である。音声信号処理における話者分離では、類似した研究として任意話者の音声分離がなされている。深層ニューラルネットワーク (deep neural network, DNN) で対象の任意性を扱う際には、順列に関連した技術的問題が生じる。DNN では入力を与えられると、予め定義されたラベルに対応する出力次元へ決定論的に写像される。そのため、異なる出力次元間の順列並べ替えが許容されない。この順列問題を解決するため、任意話者に対する音声分離をクラス分類ではなくクラスタリング問題として捉える、深層クラスタリングと呼ばれる手法が提案されている。この手法は類似度行列を使用することで、上記順列問題を回避すると同時に最適なクラスタリングを可能にする。本稿では、任意の複数楽器で演奏された音楽音響信号に対し、深層クラスタリングに基づいて各楽器パートのピアノロールを推定する新たな手法を提案する。楽器に依存しない音高推定器を用いて、全楽器の演奏音をサリエンスによって表すピッチグラム (コンデンスピッチグラム) を推定した後、深層球面クラスタリングによって指定した数の楽器パートに分離することで、各楽器パートのピアノロールである各パートピッチグラムを得る。クラスタリング時にピッチグラムに加えてスペクトログラムも考慮することで、楽曲に含まれる楽器の音色と音高双方の特徴に基づいた最適なクラスタリングが可能となる。また、MPE と音源分離の間には相補関係があるため、スペクトログラムとピッチグラムの同時クラスタリングを行うことで採譜精度のさらなる向上を図った。

評価実験には Slakh2100-orig を使用した。このデータセットは訓練データ 1500 曲、検証データ 375 曲、評価データ 225 曲を含む。各楽曲は複数楽器で演奏されており、データセットは MIDI データを伴う混合音および分離音で構成されている。含まれる楽器は 12 種類 (ピアノ、ベース、ギター、ドラム、弦楽器、シンセパッド、リード、金管楽器、オルガン、パイプ、シンセリード、クロマティックパーカッション) であるが、このうち採譜時に音高情報が重要でないドラムとクロマティックパーカッションは除外した。任意楽器に対す

る採譜精度を確かめるため、残る 10 楽器のうちオルガン、パイプ、シンセリードを除く 7 楽器のみを用いて学習を行った。評価時には、学習済みの 7 楽器のみを含む閉条件と全楽器を含む開条件を用意し、それぞれ評価を行った。学習データは各楽曲を 10 秒ごとに切り出した上で、正解コンデンスピッチグラムは構成楽器の MIDI データを、混合音は各楽器の演奏音を、それぞれ重ね合わせることで用意した。混合された MIDI データは二値化した後、ガウシアンフィルタを適用した。楽曲の演奏音はモノラル、サンプリングレートは 44.1kHz である。STFT の窓幅は 2048 (約 50 [ms])、シフト幅は STFT, HCQT とともに 512 (約 11 [ms]) とした。HCQT は 32.7Hz (C1) から 6 オクターブにかけて、5 倍音まで計算を行った。全体のデータ量は、訓練データ 11 時間、検証データ 3 時間、評価データ 6 時間 (開閉両条件とも) となった。

表 1 クラス分類アプローチと提案手法による実験結果

楽器	閉条件						開条件					
	[19]			提案手法			[19]			提案手法		
	P	R	F	P	R	F	P	R	F	P	R	F
ピアノ	51.28	46.50	45.87	62.02	39.61	44.07	52.51	48.04	47.37	61.87	38.90	43.64
ベース	73.75	58.79	64.04	39.72	50.78	42.24	74.27	59.66	64.67	40.59	51.88	43.23
ギター	46.64	36.72	37.69	52.91	35.45	39.46	44.59	37.12	37.25	53.45	36.50	40.32
弦楽器	55.27	56.79	52.74	66.35	48.74	52.40	53.21	56.97	52.05	65.31	48.40	52.04
シンセパッド	43.72	44.80	42.07	49.65	35.12	38.70	44.42	46.89	43.91	51.99	36.58	40.81
リード	28.53	33.90	29.27	29.87	37.37	31.53	26.92	31.72	27.53	28.87	35.46	30.04
金管楽器	35.24	25.12	24.50	37.10	30.23	29.53	37.66	25.67	25.89	36.78	30.64	30.26
オルガン	—	—	—	—	—	—	20.14	19.01	16.89	36.62	28.57	29.11
パイプ	—	—	—	—	—	—	22.62	27.13	23.02	38.37	39.49	35.22
シンセリード	—	—	—	—	—	—	20.58	17.44	17.59	29.41	25.11	24.98

and Su, L.: Polyphonic Music Transcription with Semantic Segmentation, ICASSP 2019.

表 1 に実験結果を示す。開条件下での未知楽器の採譜精度において、提案手法はクラス分類に基づく手法 [19] を上回った。また、既存手法では未知楽器の採譜精度が既知楽器の採譜精度よりも大幅に減少したのに対し、提案手法においては両者が比肩する結果となった。さらに、提案手法は閉条件か開条件かを問わず、既知楽器の採譜においても既存手法と同等の精度であった。図 4 の提案手法による採譜結果例から、提案手法が音高推定と楽器パート割り当てに成功していることが分かる。

本稿では、深層球面クラスタリングに基づく任意の複数楽器パートに対する採譜手法について述べた。提案手法ではピッチグラムとスペクトログラムの同時クラスタリングを通して、音楽音響信号中の音色と音高の特徴が同時に考慮された採譜が行われるとともに、実験により訓練データに含まない楽器を含む楽曲のパート譜採譜が可能である

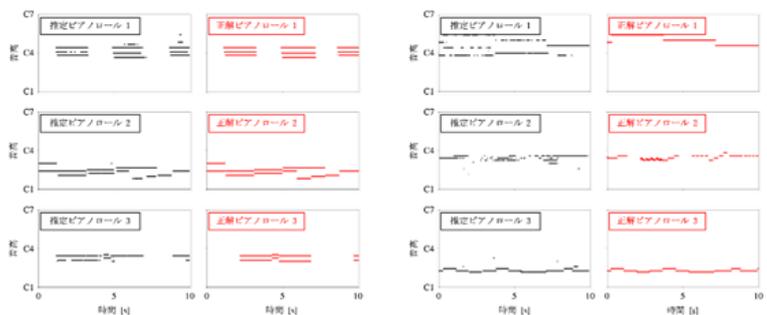


図4 パート譜採譜結果の成功例と失敗例

ことを確かめた。提案手法では各楽器パートのピアノロールに加えて分離音も得ることができるが、両者間の対応づけは手動で行う必要がある。この自動化を可能にするアルゴリズムを考案し、より統一的に音高情報と音色情報を考慮することが、最も興味深い今後の方針

として考えられる。

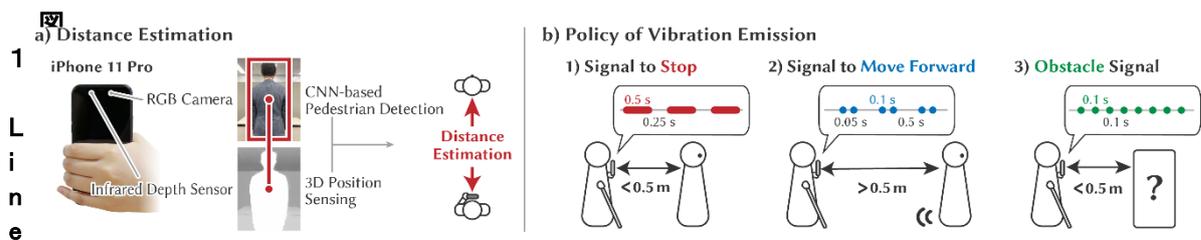
3. LineChaser: 視覚障害者が列に並ぶためのスマートフォン型支援システム

本研究は視覚障害者が列に並ぶ事を支援するシステムを提案する。列に並ぶ際は、刻々と位置が変化する列の最後尾を発見し、列の動きを検知して追従する必要があるが、どちらも視覚障害者には難しいタスクである。本研究ではスマートフォン1台で視覚障害者を列の最後尾まで案内し、列に追従できるよう支援するシステムを開発した。提案システムは列のできる場所の概形が記録された地図を使用し、地図上における周囲の人物の位置をスマートフォンのRGBDカメラを用いて検出する。検出結果に基づいて検出された人物が列に並んでいるかを判断し、音声と振動を用いてユーザに列に並ぶための指示を与える。12名の視覚障害者を対象にシステムの評価実験を121号館のロビーを利用して行ったところ、提案システムを用いることで全ての参加者が単独で列の最後尾の発見および列の追従が可能となり、普段の体験と比較して列に並ぶ際の自信度も有意に向上した。



121 号館ロビーでの視覚障害者実験の様子

空港やショッピングモールといった公共空間内では列に並ぶ機会が存在するが、視覚障害者にとって単独で列に並ぶ事は困難である。列に並ぶためには「列の最後尾の発見」および「列の追従」という2つのタスクが求められる。列の最後尾の位置は刻々と変化するため、周囲の人物が列にならんでいるか否かを判断した上で最後尾へ移動する必要がある。また列を追従する際は、いつ、どの方向へ、どれだけの距離を進むべきか判断する必要がある。本研究で視覚障害者を対象に予備調査を行ったところ、これらの列に並ぶためのタスクを視覚障害者が白杖や盲導犬といった従来の歩行支援ツールを用いて1人で行うことは極めて難しいというコメントが得られた。また予備調査では、列に追従するタスクを支援するプロトタイプシステムを実装し、評価実験を行った。プロトタイプシステムはスマートフォンに搭載されたカメラと物体検出手法を組み合わせることでユーザの前方に立つ人物との距離を計測し、距離に基づいて進むべきタイミングを振動パターンを用いてユーザに伝える(図1)。実験の結果、プロトタイプシステムを使用したユーザが追従すべき最後尾の人物以外の人物を追従してしまう結果が確認され、参加者からは前の人物との距離や方向を具体的に知りたいというコメントが得られた。

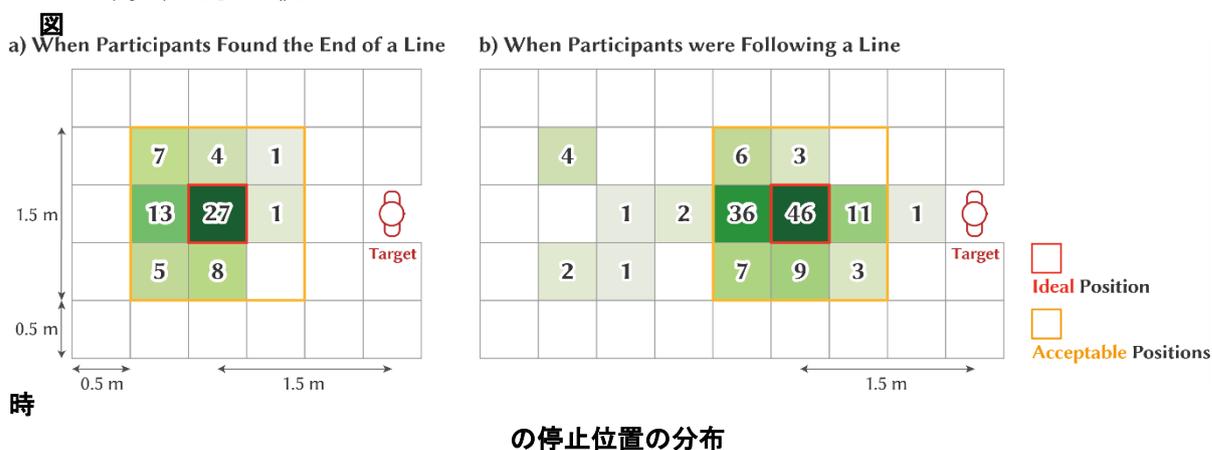


Chaser のシステム構成

予備調査の結果を基に、本研究は視覚障害者が列の最後尾を発見し、列に追従できるように支援するスマートフォン型システム **LineChaser** を提案する。提案システムは列ができる場所の概形が記録された地図を事前に用意し、用意した地図と AR マーカを用いた自己位置推定結果を元に列の入り口まで案内する。その後地図に記録された列の概形に沿って案内し、列の最後尾を目指す。その際提案手法は、RGBD カメラと物体検出手法を組み合わせることで、検出された人物が列に並んでいるか、どの人物が列の最後尾かを判断する。列の追従時はカラーヒストグラムを用いて最後尾の人物をトラッキングし、その人物の位置情報に基づいて進むタイミングと方向および距離を音声と振動を用いて伝える。本研究では 12 名の視覚障害者に対して、提案システムを用いて直線および蛇行した列に並ぶタスクを与え、提案システムの有用性を検証した。実験から

- (1) 提案システムは視覚障害者に対して列の最後尾の発見および列の追従を可能にする、
- (2) ユーザは提案システムを用いることで自信を持って列に並ぶことができる、
- (3) ユーザは音声と振動を組み合わせた案内を有用に感じる、

LineChaser の有用性を検証するために 12 名の視覚障害者（年齢：23～58（平均 43.8）歳、男性 4 名、女性 8 名）を対象に評価実験を行った。実験では直線もしくは蛇行した列の情報を記録した地図を用意し、そこに沿って 2 列を用意した。本実験では感染予防の観点から、互いにソーシャルディスタンス（1.5m）を確保するように各人物を配置した。参加者はランダムに設定された開始位置（2 種類の開始位置 S1, S2 および 3 種類の開始方向から、列の最後尾の発見および列の追従を行い、列の先頭にあるレジに見立てた机まで提案システムを用いて列に並びながら移動する。参加者が列を発見した後、列はランダムに 20, 40, 60 秒おきに 1 人ずつ進行させた。列の最後尾の位置が変わりうるという性質を実験の列でも反映させるために各試行において列に並ぶ人数を 2～4 名の 3 種類に設定し、各参加者は合計 6 回実験（直線 3 回、蛇行 3 回）を行った。参加者は右手に白杖、左手に提案システムを（ただし盲導犬ユーザは左手に盲導犬のハンドル、右手に提案システムを）持って移動した。また、人混みの環境を再現するために実験中は 60dB でスピーカーからショッピングモールの環境音を流し続けた。



評価実験の結果、図 2 に示すように全ての参加者が Acceptable Positions の領域内で列を発見することに成功した。特に Ideal Position で列を発見した割合は 40.2%（72 回中 29 回）であった。停止位置の分布を示した。全ての参加者は全試行において列の先頭にあるレジまでたどり着いた。参加者は全試行の 90.9%（144 回中 131 回）の割合で Acceptable Positions 内で停止し、34.7%（144 回中 50 回）の割合で Ideal Position 内で停止した。しかしながら参加者が Acceptable Positions の外に立ち止まる機会が 13 回観測された。Acceptable Postions の中に留まる事に失敗した 13 回は次の 4 つの理由に起因する：(1) 3 名は列が進むまでに身体の向きを修正できず、進むことができなかった、(2) 1 名はシス

テムの使い方を理解できておらず、システムが進む指示を出しているのに進まなかった、(3) 1名はシステムの指示を聞かずに前進した、(4) システムの自己位置推定の誤差が蓄積した。(1)と(2)は Accpetable Position の後方で参加者が停止した10回の原因、(3)は追従すべき人物と0.5m以下の距離しか取らず真後ろについた原因である。追従すべき人物との距離が1.7m以下となり、提案システムは振動を停止していたが、一人はそれに気づかず前進し続け、システムが発した緊急停止信号を受けて停止した。(4)は列の右側に参加者がずれた原因である。

今後は実環境で様々な条件の列(より長い列や複数本の並行した列など)を対象に実験を行うことで、ユーザの特性ごとに実環境での使用に最適な案内方法をデザインしていく。また今回の実験では、30分ほどの練習時間を確保したが、一部の参加者はシステムに慣れていなかったため列に適切に追従できない場合があった。システムを普及させていく際はシステムの提案に加え、視覚障害者の歩行訓練(O&M トレーニング)にどのように支援システムの訓練を組み込んでいくかも併せて検討する必要がある。参加者から提案システムがスマートフォン1台で完結している点を高く評価された反面、周囲にカメラを向ける動作は他人からの理解が得られない場合があるため抵抗があるというコメントも得られた。今後はスマートグラスなどのウェアラブルデバイスの使用を含め、社会受容性を損なわないシステムのデザインを検討していく予定である。

4. 共同研究者

齋藤 隼介, Facebook Reality Labs. USA

谷田 川達也, 助教, 東京大学

Angjyu Kanazawa, University of California Berkeley

Hao Li, Associate Professor, University of South California, Pinscreen Inc., CEO

浅川 智恵子, 日本科学未来館館長(2021年4月~), IBM フェロー

高木 啓伸, 日本 IBM, 日本科学未来館副館長(2021年4月~)

Hubert Shum, Associate Professor, Northumbria University

後藤 真孝, 産業技術総合研究所

齋藤 英雄, 教授, 慶応大学

中村 哲, 教授, 奈良先端大学院大学

吉井 和佳, 教授, 京都大学

5. 研究業績

5.1. 学術論文(査読付)

[1] Takashi Oya, Shohei Iwase, Ryota Natsume, Takahiro Itazuri, Shugo Yamaguchi, Shigeo Morishima, "Do We Need Sound for Sound Source Localization?", Computer Vision – ACCV 2020, Vol. 12627 pp. 119-136, https://doi.org/10.1007/978-3-030-69544-6_8, 2021.02

[2] Hideki Tsunashima, Hirokatsu Kataoka, Junji Yamato, Qiu Chen, Shigeo Morishima, "Adversarial Knowledge Distillation for a Compact Generator", Proceedings of 25th International Conference on Pattern Recognition (ICPR 2020), <https://doi.org/10.1109/ICPR48806.2021.9413150>, 2021.01.

[3] Qi Feng, Hubert P. H. Shum, Shigeo Morishima, "Resolving Hand-Object Occlusion for Mixed Reality with Joint Deep Learning and Model Optimization", Computer Animation and Virtual Worlds, <https://doi.org/10.1002/cav.1956>, Vol.31, Issue 4-5, 2020.09.

[4] Tatsuya Yatagawa, Yasushi Yamaguchi, Shigeo Morishima, "LinSSS: linear decomposition of heterogeneous subsurface scattering for real-time screen-space rendering", The Visual Computer, <https://doi.org/10.1007/s00371-020-01915-4>, 2020.07.

[5] Ryo Shimamura, Feng Qi, Yuki Koyama, Takayuki Nakatsuka, Satoru Fukayama, Masahiro Hamasaki, Masataka Goto, Shigeo Morishima, "Audio-Visual Object Removal in 360-Degree Videos", The Visual Computer, <https://doi.org/10.1007/s00371-020-01918-1>, 2020.07.

- [6] Tatsuya Koike, Shigeo Morishima, Ryoichi Ando, "Asynchronous Eulerian Liquid Simulation", Computer Graphics Forum, Vol.39, Issue 2, pp.1-8, <https://doi.org/10.1111/cgf.13907>, 2020.07.
- [7] Tatsuya Yatagawa, Hideki Todo, Yasushi Yamaguchi, Shigeo Morishima, "Data compression for measured heterogeneous subsurface scattering via scattering profile blending", The Visual Computer 36, pp.541-558, 2020., <https://doi.org/10.1007/s00371-018-01626-x>
- [8] Tomoya Yamaguchi, Tatsuya Yatagawa, Yusuke Tokuyoshi, Shigeo Morishima, "Real-time Rendering of Layered Materials with Anisotropic Normal Distributions", Computational Visual Media, 2020, <https://doi.org/10.1007/s41095-019-0154-z>
- [9] Seita Kayukawa, Hironobu Takagi, João Guerreiro, Shigeo Morishima, and Chieko Asakawa, "Smartphone-Based Assistance for Blind People to Stand in Lines", ACM CHI Conference on Human Factors in Computing Systems (CHI'20), 2020.4.
- [10] Seita Kayukawa, Tatsuya Ishihara, Hironobu Takagi, Shigeo Morishima, and Chieko Asakawa, "BlindPilot: A Robotic Local Navigation System that Leads Blind People to a Landmark Object", ACM CHI Conference on Human Factors in Computing Systems (CHI'20), 2020.4.
- [11] Shohei Iwase, Takuya Kato, Shugo Yamaguchi, Yukitaka Tsuchiya, Shigeo Morishima, "Song2Face: Synthesizing Singing Facial Animation from Audio", SIGGRAPH Asia 2020 Technical Communications, No.12 pp.1-4, 978-1-4503-8080-5, <https://doi.org/10.1145/3410700.3425435>, 2020.12
- [12] Yoshihiro Fukuhara, Takahiro Itazuri, Hirokatsu Kataoka, Shigeo Moshima, "Property Analysis of Adversarially Robust Representation", 精密工学会誌, 87 巻, 1 号, p. 83-91, <https://doi.org/10.2493/jjspe.87.83>, 2021.4.
- [13] Naoki Nozawa, Hubert P. H. Shum, Edmond S. L. Ho and Qi Feng, "3D Car Shape Reconstruction from a Contour Sketch using GAN and Lazy Learning", Visual Computer, 2020
- [14] Qi Feng, Hubert P. H. Shum, Shigeo Morishima, "Resolving Hand-Object Occlusion for Mixed Reality with Joint Deep Learning and Model Optimization", Computer Animation and Virtual Worlds, <https://doi.org/10.1002/cav.1956>, 2020
- [15] Seita Kayukawa, Tatsuya Ishihara, Hironobu Takagi, Shigeo Morishima, and Chieko Asakawa, "Guiding Blind Pedestrians in Public Spaces by Understanding Walking Behavior of Nearby Pedestrians", Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Volume 4, Issue 3, 2020.09
- [16] T. Yatagawa, Y. Yamaguchi, S. Morishima, "LinSSS: linear decomposition of heterogeneous subsurface scattering for real-time screen-space rendering", The Visual Computer, <https://doi.org/10.1007/s00371-020-01915-4>, 2020.07
- [17] Ryo Shimamura, Feng Qi, Yuki Koyama, Takayuki Nakatsuka, Satoru Fukayama, Masahiro Hamasaki, Masataka Goto, Shigeo Morishima, "Audio-Visual Object Removal in 360-Degree Videos", The Visual Computer, <https://doi.org/10.1007/s00371-020-01918-1>, 2020.7.
- [18] Tatsuya Koike, Shigeo Morishima, Ryoichi Ando, "Asynchronous Eulerian Liquid Simulation", Computer Graphics Forum, volume39 issue2 pp.1-8, <https://doi.org/10.1111/cgf.13907>, 2020.7.13
- [19] Tatsuya Yatagawa, Hideki Todo, Yasushi Yamaguchi, Shigeo Morishima, "Data compression for measured heterogeneous subsurface scattering via scattering profile blending", The Visual Computer 36, pp.541-558, 2020., <https://doi.org/10.1007/s00371-018-01626-x>
- [19] Tomoya Yamaguchi, Tatsuya Yatagawa, Yusuke Tokuyoshi, Shigeo Morishima, "Real-time Rendering of Layered Materials with Anisotropic Normal Distributions", Computational Visual Media, <https://doi.org/10.1007/s41095-019-0154-z>, 2020
- [20] Qi Feng, Hubert P. H. Shum, Ryo Shimamura, Shigeo Morishima, "Foreground-aware Dense Depth Estimation for 360 Images", Journal of WSCG, Vol.28 No.1-2 pp.79-88, 10.24132/JWSCG.2020.28.10, 2020
- [21] 栗林 雅希, 粥川 青汰, 高木 啓伸, 浅川 智恵子, 森島 繁生, "LineChaser: 視覚障碍者が列に並ぶためのスマートフォン型支援システム", WISS2020, <https://www.wiss.org/WISS2020/>, 2020.12. **最優秀論文賞**
- [22] Shohei Iwase, Takuya Kato, Shugo Yamaguchi, Yukitaka Tsuchiya, Shigeo Morishima, "Song2Face: Synthesizing Singing Facial Animation from Audio", SIGGRAPH Asia, <https://sa2020.siggraph.org/>, オンライン, 2020.12.
- [23] 石田大地, 森島繁生, 安東遼一, "GPU Smoke Simulation with Adaptive DCT Compression", Visual Computing 2020, <http://cgvi.jp/vc2020/>, 2020.12.

- [24] Shohei Iwase, Takuya Kato, Shugo Yamaguchi, Yukitaka Tsuchiya, Shigeo Morishima, "Style Controllable Facial Animation Synthesis from Singing Audio", Visual Computing 2020, <http://cgvi.jp/vc2020/>, オンライン, 2020.12.
- [25] Takashi Oya, Shohei Iwase, Ryota Natsume, Takahiro Itazuri, Shugo Yamaguchi, Shigeo Morishima, "Do We Need Sound for Sound Source Localization?", Asian Conference on Computer Vision 2020, 2020.11.
- [26] T. Yatagawa, Y. Yamaguchi, S. Morishima, "LinSSS: linear decomposition of heterogeneous subsurface scattering for real-time screen-space rendering", Computer Graphics International (CGI) 2020, <http://www.cgs-network.org/cgi20/>, 2020.10.
- [27] Ryo Shimamura, Feng Qi, Yuki Koyama, Takayuki Nakatsuka, Satoru Fukayama, Masahiro Hamasaki, Masataka Goto, Shigeo Morishima, "Audio-Visual Object Removal in 360-Degree Videos", Computer Graphics International (CGI) 2020, 2020.10.
- [28] Qi Feng, Hubert P. H. Shum, Shigeo Morishima, "Resolving Hand-Object Occlusion for Mixed Reality with Joint Deep Learning and Model Optimization", The 33rd International Conference on Computer Animation and Social Agents (CASA 2020), 10.1002/cav.1956, 2020.10.
- [29] Keitaro Tanaka, Takayuki Nakatsuka, Ryo Nishikimi, Kazuyoshi Yoshii and Shigeo Morishima, "Multi-Instrument Music Transcription Based on Deep Spherical Clustering of Spectrograms and Pitchgrams", 21st International Society for Music Information Retrieval Conference (ISMIR 2020), 2020.10.
- [30] Seita Kayukawa, Hironobu Takagi, João Guerreiro, Shigeo Morishima, and Chieko Asakawa, "Smartphone-Based Assistance for Blind People to Stand in Lines", ACM CHI Conference on Human Factors in Computing Systems (CHI'20), USA, 2020.4.
- [31] Seita Kayukawa, Tatsuya Ishihara, Hironobu Takagi, Shigeo Morishima, and Chieko Asakawa, "BlindPilot: A Robotic Local Navigation System that Leads Blind People to a Landmark Object", ACM CHI Conference on Human Factors in Computing Systems (CHI'20), <https://doi.org/10.1145/3334480.3382954>, 2020.4.

査読なし講演

- [32] 綱島 秀樹, 邱 玥, 片岡 裕雄, 森島 繁生, "Object-oriented Representation Learning の実世界データ適用に向けた最新手法の性能分析", Visual Computing 2020 - Poster, 2020.12.
- [33] 田中 啓太郎, 中塚 貴之, 錦見 亮, 吉井 和佳, 森島 繁生, "スペクトログラムとピッチグラムの深層クラスタリングに基づく複数楽器パート採譜", 情報処理学会 第 128 回音楽情報科学研究会, 2020.08.24-2020.08.25
- [34] 中村和也, 森島繁生, "個人情報保護のための写真内の指紋情報自動除去と入力画像の指紋情報が復元可能な指紋画像生成", 第 23 回 画像の認識・理解シンポジウム (MIRU2020), 2020.08.
- [35] 綱島秀樹, 大川武彦, 相澤宏旭, 片岡雄裕, 森島繁生, "Object-aware 表現学習の安定化のための KL ダイバージェンスの周期性アニーリング", 第 23 回 画像の認識・理解シンポジウム (MIRU2020), 2020.08.
- [36] 土屋志高, 夏目亮太, 片岡裕雄, 森島繁生, "音と画像を用いた演奏の音に連動した動画生成", 第 23 回 画像の認識・理解シンポジウム (MIRU2020), 2020.08.

5.2 受賞・表彰

- [1] 栗林 雅希, 粥川 青汰, 高木 啓伸, 浅川 智恵子, 森島 繁生, "LineChaser: 視覚障害者が列に並ぶためのスマートフォン型支援システム", WISS2020 最優秀論文賞, 日本ソフトウェア科学会, 2020.12.
- [2] 森島 繁生, CG JAPAN AWARD, 芸術科学会, 2020.11. <https://www.art-science.org/award/>
- [3] 森島 繁生, 岩本尚也, 加藤卓哉, 山口周悟, 中塚貴之, 羽倉賞フォーラムエイト賞, 一般財団法人 最先端表現技術利用推進協会, 2020.11. <https://soatassoc.org/hagura>
- [4] 田中 啓太郎, 中塚 貴之, 錦見 亮, 吉井 和佳, 森島 繁生, "スペクトログラムとピッチグラムの深層クラスタリングに基づく複数楽器パート採譜", 情報処理学会 第 128 回音楽情報科学研究会, ベストプレゼンテーション賞, 2020.08.

[5] 田中 啓太郎, 2020 年早稲田大学小野梓記念学術賞, 2021.3.

[6] 栗林 雅希, 2020 年早稲田大学小野梓記念学術賞, 2021.3.

5.3 学会および社会的活動

25th ACM Symposium on Virtual Reality Software and Technology (VRST2021)

Sponsorship Chair

日本顔学会理事

情報処理学会フェロー

画像電子学会フェロー

JST CREST 「人間と情報環境の共生インタラクション基盤技術の創出と展開」アドバイザー

6. 研究活動の課題と展望

コロナ禍の影響を跳ね除け、国際連携をさらに強化し、学生のモチベーションを保ちつつ、今後より高いレベルでの研究成果を蓄積していく。一方で、新しい研究領域立ち上げに尽力する。また企業との連携を強化し、開発技術の実用化を急ピッチで進めていく予定である。