

研究終了報告書

「Computational Perception Design: データ駆動手法による知覚体験設計」

研究期間: 2018年10月～2022年3月

研究者: 吉田 成朗

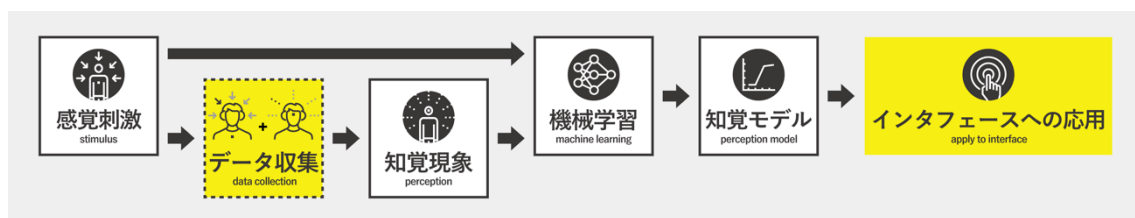
1. 研究のねらい

本研究では、機械学習に代表されるデータ駆動型の情報処理手法を用いて、人間の知覚現象と感覚刺激の対応関係を計算機で扱える形にモデル化することで、目的とする知覚現象をインタフェース上でインタラクティブに表現可能にする Computational Perception Design (データ駆動型知覚体験設計、下図) について研究する。

我々は、五感に代表される多様な感覚器官の相互作用を通して世界を感じている。人の感覚特性に考慮し、人の心や体験を中心としたインタラクションを実現するためには、人の心や行動の理解を学問の中心と据えてきた心理学の知見が活用されるべきである。

他方、心理学の知見を活用するインタラクション・インタフェースを構築する際に問題となるのが、目的とする知覚現象を再現する感覚刺激量を具体的に算出することの難しさである。心理学の知見から演繹的に導かれた感覚刺激と知覚現象の関係に対して、目的とする知覚現象を生起するのに有効な感覚刺激の範囲を同定するためには、複数の実験参加者に対して感覚刺激量を細かく設定して提示し、評価してもらう必要がある。この種のインタラクション・インタフェースを構築する上で必要な手続きとは言え、心理物理実験のように細く設定した感覚刺激量を実験参加者に対して繰り返し提示するようなやり方で、感覚刺激と知覚現象の対応関係を調べることは多大な時間と労力を必要とする。また、複数の感覚刺激が最終的な知覚現象に影響を及ぼす多感覚な知覚現象や、客観的事実と異なる知覚現象を生じる錯覚においては、感覚刺激と知覚現象の対応関係が直観的でなく、目的とする知覚現象を生起させる具体的な感覚刺激量の同定が複雑になる。

以上を踏まえ、本研究では心理学とインタラクション・インタフェース設計をつなぐ方法論として Computational Perception Design を提唱した。人間の多様な感覚に対して、目的とする知覚現象とそれを生起する。



new approach: Computational Perception Design



example_01

example_02

2. 研究成果

(1) 概要

この研究に通底する問いは、「人間の心理的・生理的特性を考慮したインタフェースをどう効果的に設計できるか？」である。人間の心理的・生理的な特性に影響を与える現象には多くの要素や多次元のパラメータが関与する。それらを効果的に組み合わせることで、人間の持つ能力や可能性を最大限に発揮可能とするインタフェースを設計可能になるはずである。そこで、こうした人間の心理的・生理的な特性を利用したインタフェースの構築を効果的にこなう統一的な方法論として、データ駆動型の知覚体験設計を提案するに至った。本研究では、この設計手法の実現性を確認するために、感覚提示装置の構築が現実的である視覚・触覚・聴覚を対象に、錯覚や多感覚などの複数のパラメータによって決定される知覚現象を搭載したインタフェースを構築した。

具体的な研究成果として、錯覚や多感覚の現象を含む触覚体験の設計に関して、バーチャルリアリティ (VR) 環境で視覚的に表示される CG 形状と同様の物体を持っているかのように知覚させる装置 (VR コントローラー) の作成を試みた。そして、把持物体の形状はその重心位置や主慣性モーメント比といった質量特性に由来するという生態心理学の知見に着目し、それら質量特性を制御することで把持物体の形状を錯覚させる装置を作成した。その際、任意の形状知覚を再現するにはどのような質量特性を作り出せばよいのかという問題に対してデータ駆動手法を適用し、装置によって制御する質量特性と、装置を把持することによって得られる形状知覚の対応関係についてデータ収集とフィッティングを行った。こうして構築した知覚モデルを装置に応用することで、様々な形状知覚を作り出すことが可能であることを示した。

また、複数の聴覚パラメータによって決定される聴覚体験の設計に関して、自身の発する声を変調してフィードバックすることで、特定の気持ちや行動を誘発する現象に注目した研究を行った。気持ちや行動の変化を目的として、コンピュータによって変調した声をリアルタイムに発声者自身の声であるかのようにフィードバックする手法を **Altered Auditory Feedback (AAF)** と呼ぶ。一方で、感情などに寄与する音声成分 (ピッチ、フォルマントなど) は多数あるため、どのような変調が目的とする気持ちや行動の変化に効果的であるか探索することは難しい。そこで、データ駆動手法を用いて声の知覚体験設計を行うとともに、設計した音声の効果的に作用するか検証を行った。クラウドソーシングに用いる音声の準備や音声変換手法の構築とともに、聴覚のパラメータ探索を行うクラウドソーシング基盤を構築した。そして、探索結果をもとに構築した、変換音声に対する自信度を推定モデルの評価実験を行った。

他にも、複数の視覚パラメータによって決定される視覚体験の設計に関して、これまでに開発した、顔写真をその本人ではない似ている顔の人のように知覚させる、そら似顔変換モデルを利用した企業との共同研究を実施した。そら似顔にメイクすることで、自分の顔であっても客観視を促し、その後に提案されるメイクに対する印象が向上することがわかった。

以上の成果やその他の成果を通して、視・聴・触覚やそれらが統合された感覚など

に対してデータ駆動手法を用いた知覚体験設計の有効性を示した。

(2) 詳細

本研究では、(A) 感覚刺激と知覚現象の対応関係を整理し、知覚データ収集のための感覚提示装置を構築した。そして、(B) 構築した装置を用いて知覚データの収集方法と、得られたデータを使って感覚刺激と知覚現象の関係をインタフェース構築に役立てられるようにモデル化する手法を検討した。さらに、(C) 実際に構築された知覚モデルを活用したインタフェースの構築と評価を行った。

研究テーマ A：感覚刺激と知覚現象の対応関係の整理と知覚データ収集のための感覚提示装置の構築

データ駆動手法を用いた知覚体験の設計が可能であるかの検証として、心理学研究を参考に知覚現象と感覚刺激の対応関係を整理した。そして、知覚現象を計算機で扱える形でモデル化するために必要なデータを集める基盤技術として、先に整理した感覚提示と知覚現象の対応関係をもとに、感覚刺激を生成する装置や仕組みを構築した。

人間の情報処理にとって優位な感覚である「視覚」に関して、提案手法の適用を試みる上で顔知覚に注目し、顔部位の大きさや位置の変形によって、知覚されるその人の印象や状態を変化させるシステムを構築した。また、視覚的な感覚提示は、特殊な環境を必要とせず個人のコンピュータでも合成・表現可能であるため、クラウドソーシングのように Web 上でデータ収集が可能なプラットフォームを用いてデータ収集を行う。

他にもクラウドソーシングを用いない場合における、データ駆動手法を用いた知覚体験設計に取り組む目的で、「触覚（形状知覚）」に注目した。把持する物体の形状知覚は、その物体の質量や慣性モーメントなどの物理特性や視覚刺激に影響を受けることがわかっている。一方で、具体的な質量の分布や慣性モーメントを求めた上で視覚の影響を加味することは難しい。そこで、多様な物理特性（質量分布や慣性モーメント）を再現する装置を作成した。これ以外にも、触覚刺激によって多様な知覚や行動を生起する装置を多数構築し、データ収集や感覚提示のための基盤となるシステム構築を行った。

さらに、「聴覚」現象として、気持ちや行動の変化を目的に、コンピュータによって変調した声をリアルタイムに発声者自身の声であるかのようにフィードバックする手法を Altered Audio Feedback (AAF) に着目した。この現象においては、感情などに寄与する音声成分（ピッチ、フォルマントなど）は多数あるため、どのような変調が気持ちや行動の変化に効果的であるか探索することは難しい。そこで、データ駆動手法を用いて声の知覚体験設計を行うとともに、設計した音声の効果的に作用するか検証を行う目的で、クラウドソーシングに用いる音声の準備や音声変換手法の構築を行うとともに、聴覚刺激のパラメータ探索を行うクラウドソーシング基盤の構築を行った。さらに、推定したパラメータを適用可能なリアルタイム加工音声フィードバック装置を構築した。また、音声以外にも楽器演奏（音色表現の習得）を支援を目的とした視覚

フィードバックを設計するシステムを構築した。

研究の当初から目的としていた、視・聴・触覚の現象や、顔知覚や声知覚のように複数のパラメータによって決定される知覚現象、視覚と形状や音色のような多感覚が統合された現象におけるデータ収集の基盤や感覚提示装置の構築について概ね網羅的に達成できた。

研究テーマ B：データ収集方法の整理と知覚現象のモデル化

感覚刺激の種類に応じて効果的にデータ収集を行う方法を整理するとともに、収集したデータをインタフェース設計に応用できる形でモデル化する方法について検討した。

視覚や聴覚に関する現象においては、クラウドソーシングを利用したブラウザベースの Web 調査を通してデータを収集した。この調査においては、ランダムな変形画像や音声の比較からどちらが目的の指標（他人っぽく見える、イキイキして見える・聞こえる、自信があるように聞こえる、など）により合致しているか答える一対比較における相対評価から、絶対評価のスコアを推定する手法を用いた（[Koyama et al. 2014]）。

また、触覚に関する現象においては、提示する感覚と知覚現象の対応関係のモデル化において、単純な回帰分析を用いた推定だけでなく、繰り返しのデータ収集を避け効率的なデータ収集を行う目的で、ベイズ推定を用いた予測モデルの構築を行った。

モデル化の手法については、上述したように単純な回帰分析からベイズ推定を用いたものまで手法の特性や利点、知覚の特性に応じて使用し、後述する研究テーマ C において知覚モデルを搭載したインタフェースの有効性を確かめることで間接的にモデルの有効性を確認した。そのため、本研究の目的であるモデル化とインタフェース応用の橋渡しに関してはある程度達成することができたと考えている。

研究テーマ C：構築された知覚モデルを活用したインタフェースの設計

構築した知覚モデルによって、目的とする知覚体験が表現可能であるか評価するため、知覚モデルを搭載したインタフェースを構築した。そして、実際にインタフェースを使用したユーザスタディを通して、生起される知覚現象が知覚モデルから推定される知覚体験と同様であるか検証を行った。ここでは代表的な成果についていくつか紹介する。

把持物体の形状はその重心位置や主慣性モーメント比といった質量特性に由来するという生態心理学の知見に着目し、それら質量特性を制御することで把持物体の形状を錯覚させる装置“Transcalibur”（図 1）を作成した。その際、任意の形状知覚を再現するにはどのような質量特性を作り出せばよいのかという問題に対してデータ駆動手法を適用し、装置によって制御する質量特性と、装置を把持することによって得られる形状知覚の対応関係についてデータ収集とフィッティングを行った。こうして構築した知覚モデルを装置に応用することで、様々な形状知覚を作り出すことが可能であることを示した。この成果は CHI2019 において発表し、上位 5%の研究成果に与え

られる Honorable Mention Award を受賞した (5.(1)1.)。



図1 質量特性の制御によってCGと一致した形状知覚を作り出す装置“Transcalibur”

聴覚現象における知覚体験設計として、上述した AAF が与える影響や効果に関して検証を進めた。本研究においては、音声変換パラメータが特定されていない「自信声」に対する変換パラメータの推定を行った。クラウドソーシングを用いた音声変換パラメータ推定を実施し、自信声に変換するパラメータを特定するとともに、自身や他人の声としてリアルタイムにフィードバックする装置を構築した。成果の一部は CHI2019 において併催されたワークショップにて発表している (5.(3)3.)。

また、視覚と聴覚が統合された場合におけるインタフェース設計について検討を行った。楽器演奏技術習得支援の研究として、楽器の音色に対してじっくり感じる図形を視覚的にフィードバックする手法について取り組んだ (図3)。こうしたフィードバックによって、耳で音を聞くよりも、演奏している音色の状態を直観的に把握して修正可能になるのではないかと考えた。そして、音色と対応すると直観できる視覚フィードバックの対応関係を計算し、そのフィードバックによって実際に楽器演奏能力が向上することを示した。この成果は、VRST2021 にて発表した (5.(3)1.)。

産学連携の成果として、資生堂グローバルイノベーションセンターと共同で、自身の顔を他人のように知覚させる顔変形技術 (他人のそら似顔) を活用した研究を実施した。日本人女性に多い悩みとして、自分の今のメイクから変えることができない、どういったメイクをしたら良いのかわからないということがある。一方で、新たなメイクを提案されても、それが自身に似合っているかどうか判断できず納得感を得ることが難しい。そこで、自分の顔に対して客観視を促すことができれば、提案されたメイクの魅力に気づきやすくなるのではないかと考えた。そして、そら似顔にメイクすることで自分の顔への客観視に繋がり、その後に見る提案メイクに対する印象が向上するということを実験を通して確かめた。

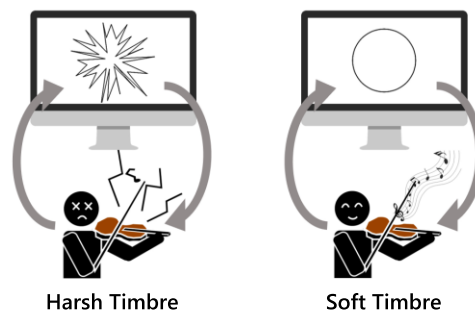


図3 音色に対してじっくりくる図形をリアルタイムに表示することで音色の状態を視覚的に把握
 視覚・触覚・聴覚やそれらの組み合わせについて知覚モデルを活用したインタフェース設計および知覚体験設計を行い、その有効性に関して確認することができ(図4)、研究目的は概ね達成できたと考える。

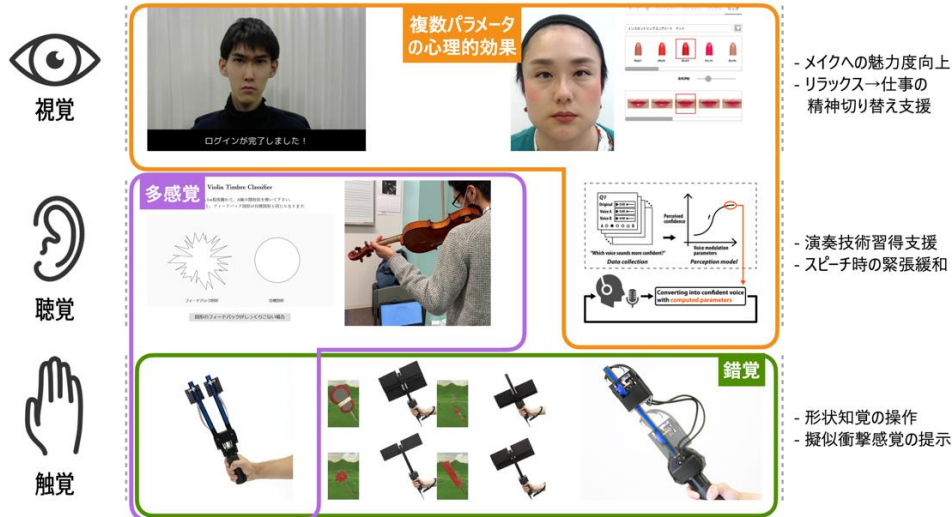


図4 視・聴・触覚やそれらの組み合わせにおいてデータ駆動手法による知覚体験設計の有効性を検証

3. 今後の展開

研究を通して「データ駆動型の知覚体験設計」というビジョンを示し、実際に動作するインタフェースを構築したことによって、それらの有効性を体験を通して評価・認識できるようにしてきたことは本研究課題の成果であり貢献であると考えます。一方で、データ駆動型の知覚体験設計という点で個別の研究が採用しているアプローチは一貫しているが、触覚・視覚・聴覚などに試みた個別の設計手法を包括的に扱う理論の構築には至らなかった。今後の展開として、包括的な理論の構築に向け、視覚・聴覚・触覚以外の感覚を対象とした知覚体験設計や、身体性(身体イメージ・身体スキーマ)を対象とした知覚体験設計をこれからも実施・収集していき、デザインパターンの構築や、設計の限界や制約を探索していきたい。

本研究の成果が将来的な社会実装につながるために必要な展開に関して、これまでの取り組みや類似の取り組みを総説論文としてまとめることや、研究職以外へのアプローチ(主にデザイナーに向けて)としてブログにまとめることを予定している。設計に関する背景や、狙い、知見を公開することで、申請者以外へこうした方法論を普及することや、方法論を応用したインタフェース・インタラクション設計を促すことを目的としている。こうした知見の整理は1-2年の短期的な展開である。また、比較的長期的な展開(5-10年)に関しては、個々の開発したインタフェースに関する説明をWeb上で公開したり、展示などの体験できる機会を提供する機会を通して、本研究のアプローチや個々のインタフェースで実現できることに対する理解者を増やしていく。

4. 自己評価

当初予定していた視覚・触覚・聴覚に関して知覚モデルを活用したインタフェース設計を行うことができ、研究目的は概ね達成できたと考えている。研究を進めるにあたって、独立して研究を進めることができる安定した研究費のおかげで、研究実施に必要な装置の構築や計測を高いレベルで実施することができた。しかし、新型コロナウイルス感染拡大の影響により、研究期間の後半には人を対象とした実験を行うことが難しかった期間があり、それが構築したインタフェースの評価に影響した。対面を伴わない実験への計画変更や、対策を徹底しての実験実施、ワクチンの普及により感染拡大がある程度抑えられてからの実験実施などの対策を通して研究を進めることになり、研究に多少の遅れがみられた。

この先、情報世界と物理世界の統合に伴う情報の氾濫に人間はさらされていくと考えている。そのような状況においては、心理的・生理的な裏付けのもと、人間の認知的・身体的負荷を軽減、もしくは、人の持つ能力や可能性を十分に発揮可能にするインタラクションやインタフェースがますます重要となると確信している。そうした世界や社会の実現に先立って、本研究を提案できたことや、提案を示す取り組みに邁進できたことを嬉しく思う。

一方で、DNN や CNN、GAN などを用いた最新の機械学習手法や、ベイズ推定やベイズ最適化を用いた確率的モデルを活用した知覚体験設計手法について深く検討することができなかった。これは、3年半というさきかけの期間において当初予定していたインタフェース設計にほぼ時間を費やしてしまい、他の手法や方法論を取り入れる余裕がなかったことや、自身の数理の知識やアルゴリズムの理解が及ばなかったことが原因である。こうした手法を深く知り活用できるようになることで、人間の知覚をより精確にモデル化したり、人間の知覚体験をより効果的に表現できることに繋がり、直観的で人間の認知負荷を必要とせず、人間の能力を存分に発揮することを支援するインタフェース・インタラクションを実現することが可能になると考える。

5. 主な研究成果リスト

(1) 代表的な論文(原著論文)発表

研究期間累積件数: 7件

1. Jotaro Shigeyama, Takeru Hashimoto, Shigeo Yoshida, Takuji Narumi, Tomohiro Tanikawa, and Michitaka Hirose. 2019. Transcalibur: A Weight Shifting Virtual Reality Controller for 2D Shape Rendering based on Computational Perception Model. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19). Association for Computing Machinery, New York, NY, USA, Paper 11, 1-11.

形状知覚錯覚を利用して、把持物体の形状知覚を自在に編集可能なインタフェースを設計する方法を示した。物体の質量特性(慣性モーメント、質量中心)と知覚する形状の関係性が不明であるため、データ駆動型アプローチを使用。実験参加者には、提案するインタフェースによって表現した質量特性とそれを把持した際に感じた知覚形状を回答してもらう。この対応関係を学習した知覚モデルと提案インタフェース(Transcalibur)によって任意の形状知覚錯覚を引き起こすことを可能にした。

2. S. Shimizu, T. Hashimoto, S. Yoshida, R. Matsumura, T. Narumi and H. Kuzuoka, "Unident:

Providing Impact Sensations on Handheld Objects via High-Speed Change of the Rotational Inertia,” 2021 IEEE Virtual Reality and 3D User Interfaces (VR), 2021, pp. 11–20, doi: 10.1109/VR50410.2021.00021.

重り位置を高速に移動させて慣性モーメントを変化させることで、それを擬似的な衝撃力として知覚させる装置(Unident)を提案。重りの停止位置と、感じた衝撃強度の対応関係を収集し学習することで、提案装置を用いて任意の衝撃感覚を表現可能にした。実験を通して、変化させ慣性モーメント量に応じて、様々な大きさの衝撃感を与えることができることを示した。また、提案装置によって実現可能なアプリケーションを示した。

3. 池田華子、吉田成朗、新井智大、鳴海拓志:加工自己顔へのメイクアップ実施体験—提案されたメイクアップスタイルへの積極的受容促進方法の提案、日本バーチャルリアリティ学会論文誌、Vol.26 No.1、p42–51、2021年4月

本研究では、自分に似合うメイクとして推奨されたメイクを肯定的に受け止めることができる手法を提案した。実験参加者は、自分の顔に対する客観視を促す目的で、自分の顔を他人の顔のように知覚させることが可能な顔画像(他人のそら似画像)を用いてメイクアップシミュレータでメイクを行った。他人のそら似顔でメイクを繰り返すと、推奨されるメイク方法に対する印象が向上した。しかし、元の自分の顔でメイクを繰り返した参加者には、印象の改善が見られなかった。この結果は、修正された自己顔へのメイクが、メイクの変化に対する自発的な肯定的受容を改善する可能性を支持するものである。

(2)特許出願

なし

(3)その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

[主要な学会発表(査読あり)]

1. Kota Arai, Mone Konno, Yutaro Hirao, Shigeo Yoshida, and Takuji Narumi. 2021. Effect of Visual Feedback on Understanding Timbre with Shapes Based on Crossmodal Correspondences. In 27th ACM Symposium on Virtual Reality Software and Technology (VRST '21), December 8–10, 2021, Osaka, Japan. ACM, New York, NY, USA, 3 pages. (to appear) <https://doi.org/10.1145/3489849.3489912>
2. Yuhu Liu, Takeru Hashimoto, Shigeo Yoshida, Takuji Narumi, Tomohiro Tanikawa, and Michitaka Hirose. 2019. ShapeSense: a 2D shape rendering VR device with moving surfaces that controls mass properties and air resistance. In ACM SIGGRAPH 2019 Emerging Technologies (SIGGRAPH '19). Association for Computing Machinery, New York, NY, USA, Article 23, 1–2. <https://doi.org/10.1145/3305367.3327991>
3. Kana Naruse, Shigeo Yoshida, Shinnosuke Takamichi, Takuji Narumi, Tomohiro Tanikawa, Michitaka Hirose: Estimating Confidence in Voices using Crowdsourcing for Alleviating Tension with Altered Auditory Feedback. Asian CHI Symposium: Emerging HCI Research Collection in ACM Conference on Human Factors in Computing Systems (CHI) 2019, May 4–9, 2019. <https://doi.org/10.1145/3309700.3338418>

[受賞]

1. Augmented Humans Conference 2022 Best Paper Award, Mar 2022.
2. ACM CHI 2019 Honorable Mention Award, May. 2019.