

## 研究課題別事後評価結果

1. 研究課題名： 多様なデータへのキャプションを自動で生成する技術の創出

2. 個人研究者名

牛久 祥孝（オムロンサイニックス株式会社 リサーチアドミニストレイ  
ティブディビジョン プリンシパルインベスティゲーター）

3. 事後評価結果 担当アドバイザー：土井アドバイザー

（以下、2020年度課題事後評価時のコメント）

本研究は、画像のような静的なデータから動画のような動的なデータまで、多様な形態のデータを説明するキャプションを生成する問題に対し、(i)個人の属性や好みへの対応、(ii)詳細な表現への対応、(iii)教師キャプションを持たないデータへの対応と、加速フェーズで追加した(iv)非言語的に埋め込まれた文法構造の獲得の4点において着実に研究成果を創出し、詳細なキャプションを生成する一連の技術を確立した。その学術成果は国際会議LREC 2020、ICCV 2019等で発表されている。さらに、他のACT-I研究者達と連携してVisual & Languageに関連した国際会議ワークショップ企画を主催したり、企業・大学等との共同研究にも幅広く取り組んだりした点も高く評価できる。今後も、画像・動画のキャプション生成研究を社会実装までつなげられる研究者として大きなインパクトを生み出していくことが期待される。

（2021年9月追記） 担当アドバイザー：土井アドバイザー

短い延長期間中であつたが、作業映像理解中の材料の視覚的な変化と動画からのキャプション生成を改良して、動画キャプション生成技術としての作業手順生成手法を確立し、さらに、異なるドメイン間で作業者の手領域の切り出しを頑健に行う教師なしドメイン適応技術を開発した。ACM MM2021 等に採択され、対象を調理から生化学実験に広げるなど、ACT-I後の更なる発展が期待される。

以上

## 研究課題別事後評価結果

1. 研究課題名： 視覚に基づく言い換えのセマンティック類型

2. 個人研究者名

チョ シンキ (京都大学 学院情報学研究科 特定准教授)

3. 事後評価結果 担当アドバイザー：中小路アドバイザー

(以下、2020年度課題事後評価時のコメント)

本研究は、画像上で視覚的に表現された概念に対応する言い換え（パラフレーズ）であるVGP (Visually Grounded Paraphrase)という概念を核として、その基礎技術の開発と応用展開を目指したものである。基礎技術として、視覚的言い換への同定技術、言い換への類型化とアノテーション技術、多言語間の視覚的言い換え技術の開発に成功した。その成果は、AAAI 2020といった国際会議をはじめ、Neurocomputing、IEEE Accessなどの海外論文誌で発表され、国際的なスケールで大きな成果をあげたことが高く評価できる。VGPに関わるこれらの個々の基礎技術の開発に加えて、基礎技術、技術課題、技術展開、応用、という四つのレイヤーからなるVGP研究のフレームワークを構築したことで、自然言語処理とコンピュータビジョンを融合するVGPという新たな研究領域を自ら立ち上げることに成功したとも位置付けられる。今後は、VGP分野のさらなる深化と共に、より広い自動翻訳研究におけるVGPの応用展開が期待される。

(2021年9月追記)

担当アドバイザー：中小路アドバイザー

本課題は、新型コロナウイルスの影響を受け6ヶ月間期間を延長した。延長期間においては、当初の研究期間中に着手した絵画データの分類手法を発展させたことに加え、研究成果であるVGPの基礎技術をベースとして、ビデオデータを活用した機械翻訳の高度化や、視覚表現と言語表現の変換手法などを開発した。また、マルチメディア関連の著名な国際会議を含む英語論文4本として研究成果を発表し、海外の研究者らとの協力関係をするなど、積極的に研究成果の伝播に努めた。延長により、自動翻訳研究におけるVGPの応用展開への頑健な足掛かりとなる成果が得られたと考えられる。

以上

## 研究課題別事後評価結果

1. 研究課題名： 衛星リモートセンシングから得られる時空間ビッグデータの機械学習による地震被害の判別

2. 個人研究者名

宮本 崇 (山梨大学 大学院総合研究部 准教授)

3. 事後評価結果 担当アドバイザー：内田アドバイザー

(以下、2020年度課題事後評価時のコメント)

本研究は、衛星画像を用いて建造物の震災被害状況推定を行うため、機械学習モデルの開発を目指す、情報学と土木工学の融合分野の研究である。正解付きデータ数も少なく、また解像度も非常に低いという困難な状況で、深層畳み込みニューラルネットワークを適用する点に本研究者の独創性があると言える。具体的には、過去の衛星画像系列の利用や、住宅の築年代と構造種別（木造・モルタル等）など、使える情報を積極的に取り込むことで、この難題に取り組んだ。その結果、90%精度で倒壊の有無を検出できることを実証した点はインパクトがあり、土木工学における貢献は大きい。また本研究者は学術的成果を出すことに留まらず、土木工学分野における機械学習活用に関する講演や執筆で活躍したことも高く評価できる。実際、土木学会誌という当該分野最大の学術誌において、AI×土木の特集号を組むなど、情報学と土木工学の融合を図るべく活躍しており、今後も分野全体をけん引していくことが期待される。

(2021年9月追記) 担当アドバイザー：内田アドバイザー

本課題は、新型コロナウイルスの影響を受け、6ヶ月間期間を延長し、複数のことを行った。第一に、これまで行ってきた被害推定に関する論文投稿を行い、採択に至っている (Journal of Selected Topics in Applied Earth Observations and Remote Sensing, JCR impact factor 3.827)。第二に、自身の専門分野における機械学習によるアプローチと物理モデルに基づくアプローチの融合の重要性に気づき、土木学会の「AI・データサイエンス論文集」にポジションペーパーを投稿し、採択に至っている。第三に、4月よりドイツ人工知能研究所 (DFKI) に滞在しており (コロナ禍ではあったがドイツ入国は問題なく認められた)、同研究所の研究者とのコラボレーションの輪を広げていると聞いている。延長により、これまでの総括的な仕事、今後採るべき方策に関する見通し、そして自身の研究活動を国際展開する経験、を獲得・実施することができた。

以上

## 研究課題別事後評価結果

1. 研究課題名： 水泳プール中の水の流れを3次元計測する技術の開発

2. 個人研究者名

山下 聖悟

(株式会社エクサウィザーズ 技術統括部 AI Vision グループ 事業部長補佐)

3. 事後評価結果 担当アドバイザー：原アドバイザー

(以下、2020年度課題事後評価時のコメント)

本研究では、プールのように水中に人が存在する環境における流体計測技術の実現を目指して、人体に悪影響を与えないトレーサー粒子と光源を用いた流体計測手法を開発した。加速フェーズでは特に、水泳プール中の水の流れを3次元計測する技術をより深める研究を行い、既に関発したトレーサー粒子(人工いくら)が沈む問題を解決するために、ナノセルロースをトレーサー粒子として用い、その流動性複屈折という特性(流れによって複屈折が生じる性質)を利用した光学系を提案した。この光学系により、計測が必要な部分のみが可視化されるため、流体計測を可能とする範囲が飛躍的に向上した。これらの一連の成果は、複数の国際会議・国内会議で公表されて高く評価されている。さらに、スポーツや水流計測の専門家など異分野との連携やアウトリーチ活動にも積極的に取り組んでいる。今後、情報学と光学やバイオメカニクス等の融合分野の発展に貢献する研究者として活躍していくことが期待される。

(2021年9月追記) 担当アドバイザー：原アドバイザー

本課題は、新型コロナウイルスの影響を受け、6ヶ月間期間を延長し、3次元的な流れを定量的に計測するための光学系や、特定の種類の流れのみを撮影可能にする基礎技術の開発を行った。

その結果、平面光源ではなく、シート状に広がる白色LED光源を用いて側面から光を差し込んだ場合の特性を観察することで、シート状の光が照射された限られた奥行き範囲のみに「流のある部分のみ明るくなる」特性を発見し、3次元的な流れを復元できる可能性が得られた。さらに、流れのある部分のみが明るくなる特性は、明るさの変化だけではなく、流れの方向に応じて特定の発色があることが分かり、光学フィルタをカメラに装着することで色情報から流れを方向ごとに取り出せる可能性が得られた。

以上のように、延長により、今後のイノベーションに向けた展開をより一層後押しする成果が得られた。

以上

## 研究課題別事後評価結果

1. 研究課題名：適材適所システム：広葉樹林業で発生する多様な小径木の家具および建築への応用

2. 個人研究者名

吉田 博則（東京大学 大学院情報理工学系研究科 特任研究員）

3. 事後評価結果 担当アドバイザー：川原アドバイザー

(以下、2020年度課題事後評価時のコメント)

本研究は、小径木を家具や建築部材として活用するために、各々の材の特徴をデジタル化したデータを活用し、適材適所を誰もが実現できるワークフローの確立に取り組んだ。プロダクトスケールでのデザイン支援ツール「Tsugite」を開発し、組み立て可能性、加工可能性、強度などを考慮し、接合部の性能についてリアルタイムに解析結果を表示することで木工継手・仕口の設計と製作を支援することを可能にした。その研究成果は国際会議 ACM UIST 2020に採択されて高く評価されている。さらに、建築スケールでは三次元形状を考慮した枝を組み合わせるための、交差部のジョイントを生成し加工するシステムを開発した。それを利用して製作されたインスタレーション「Swirled Branches」は日本科学未来館に展示されている。本研究者の研究成果は、HCIおよび情報学の研究成果を持続可能な社会の実現につなげることを目指したものであり、今後も分野を牽引する研究者として活躍していくことが期待される。

(2021年9月追記) 担当アドバイザー：川原アドバイザー

本課題は、新型コロナウイルスの影響を受け、6ヶ月間期間を延長し、形態最適化の取り入れとロボットアームを用いたアセンブリタスクの支援技術の開発に取り組んだ。

形態最適化 (Topology Optimization) をデータセットの生成に取り入れた結果、形状だけでなく構造的にも理にかなった形態の生成が可能となった。また、アセンブリ支援技術においては、ロボットアームを介してロボットと共同でアセンブリ作業を遂行する新規手法の開拓を行なった。また延長により、海外の現代美術ギャラリーでの展示や国内外のアーティスト・デザイナーとの共同作品制作にもつなげることができ、魅力的な成果展開につながる成果が得られた。

以上

# 研究報告書

## 「多様なデータへのキャプションを自動で生成する技術の創出」

研究期間： 2019年4月～2021年3月

(新型コロナウイルス感染症の影響を受け 2021年9月まで延長)

研究者： 牛久 祥孝

### 1. 研究のねらい

画像や動画のキャプション生成は、データ内の事物とそれぞれの関係性を理解して自然言語で表現する、メディア理解の究極の形態の一つである。

この研究では、画像や動画と言った多様な形態のデータそれぞれを説明するキャプションを自動で生成する技術の確立をめざす。特に本研究では、キャプション生成に本質的に必要であり未解決である、(i)個人の属性や好みへの対応、(ii)詳細な表現への対応、(iii)教師キャプションを持たないデータへの対応と、加速フェーズではさらに、(iv)非言語的に埋め込まれた文法構造の獲得を実現し、多様なデータへのキャプション生成を可能にする大きなブレークスルーをもたらす。

まず、画像を説明するキャプションは、(i)何を説明するか、どう表現するか、ということがユーザの属性や好みに応じて変化するべきである(図1)。

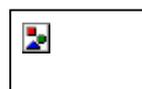


図1 個人の属性や好みに応じたキャプションの生成。

また、これらのキャプションは(ii)データ中の小さな領域をも詳細に表現できる必要がある。現在の画像キャプション生成は、「猫がこちらを見ている」など、画像中に大きく映った事物のみを、抽象的な語彙で表現する傾向がある。本研究では、注目されそうな部分に応じて図1のように「凶悪な目つき」という小さな領域と「猫」という大きな領域の対象を同時に扱える技術を確立する。

さらに、(iii)教師キャプションを持たないデータへの対応が必要である。現在のキャプション生成技術は、教師キャプションが付与された画像を数万枚の規模で作成したデータセットを用いる。しかし、実世界のマルチメディアデータには動画や音声と言った多様なモダリティが存在し、さらに画像でも撮影位置や時刻といったメタデータが付随する場合も多い。しかしながら、大規模データ収集の高いコストにより、上記のような様々なモダリティのマルチメディアデータについて、他の要求機能を満たすようなキャプション生成を実現するために都度大規模なデータセットを収集するのは困難である。

最後に加速フェーズでは、レシピや組み立て書のような図とテキストが混在したマニュアル、料理などの作業動画といった手順が埋め込まれたデータを理解する技術も確立する。そのためには、工程を抽出してその前後関係を学習する必要がある。この様な(iv)非言語的に埋め込ま



[れた文法構造の獲得](#)は、より一般的なコンテンツのストーリー理解にも必須の機能である。

## 2. 研究成果

### (1) 概要

本研究が画像キャプション生成において満たしたい要求機能は、[\(i\)個人の属性や好みへの対応](#)、[\(ii\)詳細な表現への対応](#)、[\(iii\)教師キャプションを持たないデータへの対応](#)、[\(iv\)非言語的に埋め込まれた文法構造の獲得](#)の4点であった。その成果として、各要求機能への取組とその統合的な評価を並列で進めてきた。

具体的には以下の通りである。[\(i\)個人の属性や好みへの対応](#)の解として**研究テーマ A「転移学習の研究」**に取り組み、既存研究より現実的かつ難しい問題設定におけるドメイン適応についての技術を確立し、特許を出願中である。[\(ii\)詳細な表現への対応](#)の解としては**研究テーマ B「ユーザーフレンドリーな参照表現生成」**を遂行し、詳細な差異に注目できるようなキャプション生成基盤技術としてトップ国際会議での発表に至っている。[\(iii\)教師キャプションを持たないデータへの対応](#)については**研究テーマ C「教師なしキャプション生成の研究」**として疑似教師データを活用しながらの教師なし学習によるキャプション生成を実現し、トップ国際会議での発表を予定している。[\(iv\)非言語的に埋め込まれた文法構造の獲得](#)については、**研究テーマ D「作業手順の構造的理解」**に取り組んだ。調理の様子をおさめた動画に埋め込まれた作業の手順をグラフ構造として理解し、そのレシピを生成する技術を開発し、国際論文誌での発表に至っている。

またこれらを統合して評価するために**研究テーマ E「料理動画・レシピデータセットの構築と統合評価」**として新規データセットの収集と、上記技術を統合したキャプション生成技術の評価に取り組んだ。

新型コロナウイルス感染症の影響を受け 6 ヶ月間研究期間を延長し、作業映像を中心としたデータセットの収集とそれらの映像理解としてのキャプション生成を更に積極的に展開した。結果として、**研究テーマ D と E**において後述の通り、複数の成果が生まれるに至っている。

### (2) 詳細

**研究テーマ A「転移学習の研究」**は、個人に依らない基本的なキャプション生成モデルを個人ごとの属性に応じた画像や自然言語に転移させることで、教師データが少ない状態 [\(iii\)](#)での個人適応 [\(i\)](#)を実現するための研究である。そもそも従来の転移学習は、その評価実験に明示的に/暗黙的に仮定されている条件が複数あり、実用性に疑問があった。本研究では、キャプション生成への適用を見据えて現実に即した評価条件設定に改めた深層転移学習手法を開発した。本成果は特許出願中かつ、コンピュータビジョン分野でのトップ国際会議である CVPR 2021 に採択された。

**研究テーマ B「ユーザーフレンドリーな参照表現生成」**は、「データ中の小さな領域をも詳細に表現できる」という [\(ii\)](#)の要求機能だけを切り出したものに対応する。参照表現生成では、画像内にある複数の物体の中から、どれに言及しているのかが一意に定まるような

キャプションを生成する。例えば図1に対して「物が写った写真の一部」というキャプションを生成しても、間違っていないもののどの領域についての記述かが全く分からない。そこで「猫の首輪について」と言うと、画像中央下部にある金色の物体に定まる。単に一意に定まることだけを目指す「魚」だけでも良いことになってしまうが、それだと画像のあらゆる領域を探さないといけないような参照表現生成になってしまうため、人に対して生成するキャプションとしては利便性が少ない。そこで本テーマでは、「猫の首輪」のように画像内で視覚的に顕著性の高い領域を紐づけながら、冗長性の少ないキャプションを生成する技術をアテンション機構とランキング学習を組み合わせた深層学習モデルで実現した。この成果は代表者が東京大学に在籍している際にデンソー、デンソーアイティラボラトリとの共同研究として進めたものであり、コンピュータビジョン最高峰の国際会議である ICCV 2019 でデータセットの公開と併せて発表済みである。

**研究テーマ C「教師なしキャプション生成の研究」**は、転移学習が難しい場合でも(iii)を解決するための手段として進めている。教師データとなる画像/動画とキャプションの対が無い状態でも、そうした画像/動画といったデータやキャプションのような文のデータといった、ペアになっていないデータであれば容易に収集可能である。そこで本テーマでは、そうした画像と文が独立に存在するデータについて、一部の画像から尤もそうな画像・キャプションの疑似ペアを少しずつ増やして教師あり学習する手法を開発した。この成果は NAIST および理研 AIP と共同で取り組み、自然言語処理分野においてメジャーである国際会議 EACL 2021 に採択され、発表予定である。

**研究テーマ D「作業手順の構造的理解」**は、動画のような動的なデータの中での非言語的に埋め込まれた文法構(iv)を獲得するためのテーマである。長い動画のようなデータの中では、時々刻々を言語で記述するキャプションの上位に、そのストーリーのような遷移が存在する。例えば調理動画であれば、各調理ステップに対応する記述と、それらの系列を同時に理解する必要がある。そこで本テーマでは、調理動画の理解を各調理ステップのキー画像からなる画像列の理解と捉え、途中経過の食材や初出の材料がどのステップで組み合わされているのかを示す木構造を定義し、そうした木構造を用いたレシピ生成技術を開発した。画像列と材料のリストから木構造を生成し、その木構造からレシピを生成する。また生成されたレシピからも木構造を推定し、前述の木構造と一致するかどうかを評価させてレシピと動画列の一貫性を担保している。この成果は京都大学との共同研究として進め、速報誌である IEEE Access に掲載されている。

コロナ延長期間中においては、上記技術が画像列からのキャプション生成であったものを動画からのキャプション生成とすること、そして作業映像理解中に材料の視覚的な様子  
が変化する点をモデリングすること、この2点をさらに改良した動画キャプション生成技術  
としての作業手順生成手法を確立し、マルチメディア分野においてトップ国際会議である  
ACM MM 2021 に採択されている。

またコロナ延長期間中においては更に、作業中の人間を切り出す映像理解において、  
その照明条件や環境の差異などの違いに頑健な作業理解の技術に取り組んだ。具体的  
には作業者の手領域を切り出すことが作業者の行動認識や作業対象の材料・器具の

認識に有効であるが、異なるドメイン間でこうした作業者の手領域の切り出しが頑健に行えるような教師なしドメイン適応技術を開発し、速報誌であるIEEE Accessに掲載されるに至っている。

**研究テーマE「料理動画・レシピデータセットの構築と統合評価」**は、テーマDのような調理動画を対象として、その多様性への対応を正しく評価できるようなベンチマークを作成し、実際にそれぞれの研究テーマから得られた技術を統合したキャプション生成技術を適用して評価を行うことを目的としている。料理レシピのデータベースとそれらに基づいたサービスで定評のあるクックパッドと、京都大学や東京大学の研究者との合同でデータセットの構築にあたった。

本研究テーマでは2段階でデータセット収集とその公開を進めている。まず、クックパッドが既に所持している画像列データとレシピのデータについて、その作業手順構造をグラフとして付与したデータセットを構築、こうした自然言語を含めたリソースについて公開して共有する国際会議としてはもっとも著名なLREC 2020で発表済みである。

またコロナ延長期間中においてはこうした作業の対象を調理構造から生化学実験へと広げ、一人称視点での実験作業映像と作業言語記述を整備したデータセットを構築した。コンピュータビジョン分野でのトップ会議であるICCV 2021のワークショップに採択され、ACT-I後の発展としてより一般的な作業理解としての動画キャプション生成への足掛かりを築いた。

### 3. 今後の展開

この2年間によって、画像のような静的なデータから動画のような動的なデータまで、個人に応じて詳細なキャプションを生成する一連の技術が確立された。これらは画像や動画と言った非言語情報を自然言語で検索・閲覧したり、更に近年盛んに研究されているデータ生成ネットワークに応用したりと言った形で、自然言語によってマルチモーダルなデータを活用する基盤技術となる。例えば、本研究で最終的に実現した手順映像の言語理解は、作業記録の自動生成や非熟練者への言語サポートに活用できる。

最近では自己注意機構を活用したTransformerと呼ばれるネットワーク構造が自然言語処理のみならずマルチモーダルなデータの理解についても有効性を認められつつある。本研究でもこのネットワーク構造の活用について試行中であるが、Transformerを用いる場合はより大量な学習データが必要となる。そのため、より少ないデータでもそのバイアスを活用してキャプション生成を可能にする畳み込みニューラルネットワークと再帰ニューラルネットワークの組合せや、そうしたバイアスをデータセット間で吸収する転移学習など、本研究で取り組んできた諸技術については引き続ききわめて有効な技術であり続けると考えている。

さらに、本研究は本質的に、画像/動画といった画素数列データと自然言語列というグラフのような構造を持つデータの変換に取り組むものであった。今後もこうしたデータ構造を超えるモダリティ変換の研究を別の系に適用し、効率的なデータ活用基盤の研究を続行する予定である。

#### 4. 自己評価

研究目的の達成状況および研究費執行状況については、要求仕様(i)~(iv)それぞれを満たす技術の開発に成功し、それぞれ特許出願や国際論文誌/会議での発表に至っている。

研究実施体制については、各研究テーマの遂行にあたって研究代表者と相乗効果の見込める研究者と産学問わず連携して進めたことにより、複数の研究成果を並行して創出できた。

研究成果の波及効果については、企業との共同研究によって各企業が持つデータに対する社会実装が期待される。また、各研究テーマは世界的にも広く注目されているような論文誌/国際会議で発表しており、積極的に国内外での招待講演を通じてその更なる周知に努めている。更に、研究テーマ A や E で述べたようなデータセットの公開を通じて広く世界の研究者に本テーマでの更なる研究を促しており、今後より本研究領域が盛んに研究されるようになることを期待している。

研究課題の独創性・挑戦性としては、画像/動画キャプション生成という研究領域を真に社会実装するために必要となる機能を網羅的に押さえた研究課題であると自負している。現在もこうした画像/動画と自然言語を含めたメディア理解は Vision and Language というテーマでコンピュータビジョン、自然言語処理、機械学習といった諸分野で取り組まれているが、今も同一のデータセットに対する精度競争を注意機構や Transformer などの新たな関連技術によって更新する取り組みが多い。本研究では一貫して、基本的に既存の研究課題をより現実的な設定に改善しながらその解となる技術を提案し続けており、新規かつ有用な課題に挑戦し続けている。

#### 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: 9件

1. Mikihiro Tanaka, Takayuki Itamochi, Kenichi Narioka, Ikuro Sato, Yoshitaka Ushiku, and Tatsuya Harada. Generating Easy-to-Understand Referring Expressions for Target Identifications. The IEEE/CVF International Conference on Computer Vision. 2019, pp. 5794-5803.

「2.研究成果」の研究テーマ B で述べたユーザーフレンドリーな参照表現生成技術。十分詳細なキャプションとはキャプション生成対象の画像やその領域を一意に特定できるキャプションであるとして、そのようなキャプション生成課題である参照表現生成に取り組んだ。特に画像や映像の参照表現生成では、対象をすぐに見つけられる参照表現と時間のかかる参照表現があり得る。本研究ではユーザがすぐに対象を発見できるような参照表現生成に取り組んだ。

2. Taichi Nishimura, Atsushi Hashimoto, Yoshitaka Ushiku, Hiroataka Kameko, Yoko Yamakata, and Shinsuke Mori. Structure-Aware Procedural Text Generation from an Image Sequence. IEEE Access. Vol. 9, pp. 2125-2141, 2021.

「2.研究成果」の研究テーマ D で述べた作業天順の構造理解のためのキャプション生成技

術。調理作業の映像から、どの工程でどの材料が統合されて最終的な作業結果の料理に至ったかという材料統合木構造を内部的に推定し、キャプションを生成させる技術としている。更にこのキャプションから材料統合木構造を再度推定させ、前述の材料統合木構造との差異を最小化させることで、キャプション記述の過不足を最小化させる手法となっている。

3. Qing Yu, Atsushi Hashimoto, and Yoshitaka Ushiku. Divergence Optimization for Noisy Universal Domain Adaptation. The IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021, pp. 2515-2524. ※コロナ延長時の成果

「2.研究成果」の研究テーマ A で述べた転移学習手法。キャプション生成への適用を見据えて現実に即した評価条件設定に改めた深層転移学習手法を開発したもので、これまで広く取り組まれている教師なしドメイン適応という研究では仮定されないラベルノイズやドメイン間のクラス集合のずれを取り入れたノイズ有ユニバーサルドメイン適応という課題を提言し、これにダブルヘッド型のクラス分類器と情報量最適化による手法で取り組んだものである。

4. Takehiko Ohkawa, Takuma Yagi, Atsushi Hashimoto, Yoshitaka Ushiku, and Yoichi Sato. Foreground-Aware Stylization and Consensus Pseudo-Labeling for Domain Adaptation of First-Person Hand Segmentation. IEEE Access. Vol. 9, pp. 94644-94655, 2021. ※コロナ延長時の成果

「2.研究成果」の研究テーマ D で述べた作業手順の構造的理解のための映像認識技術。作業映像において作業者の手を逐次認識しておくことは、作業者の行動や行動の対象である器具・材料等の理解のためにも必須のモジュールである。一方でドメインが異なる作業映像では、その照明条件や環境などの差異から、共通して作業者の手を認識するのは難しい。そこで転移学習手法を取り込み、複数環境で頑健に手を認識可能な技術を確立した。

(2)特許出願

研究期間累積件数:3件(特許公開前のもも含む)

1	発明者	郁青、牛久 祥孝、橋本 敦史
	発明の名称	推定システム、推定装置および推定方法
	出願人	オムロン
	出願日	2019/11/14
	出願番号	2019-206384
	概要	「5.主な研究成果リスト(1)」の 3.に対応する特許で、ラベルノイズやドメイン間のクラス集合のずれを取り入れたノイズ有ユニバーサルドメイン適応という課題についてダブルヘッド型のクラス分類器と情報量最適化による手法で解決を図るもの。
2	発明者	橋本 敦史、牛久 祥孝、森 信介、西村 太一
	発明の名称	構造化データ表現の獲得方法
	出願人	オムロン
	出願日	2020/03/06
	出願番号	2020-038785
	概要	「5.主な研究成果リスト(1)」の 2.に対応する特許で、作業行程中の映像から材料の統合過程を木構造で表現してから、その木構造を映像中の作業工程の言語記述に変換するもの。

(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)

1. 【受賞】牛久祥孝. 第8回ヤマト科学賞. 2021/4/21.
2. 【基調講演】牛久祥孝. 深層学習によるビジョン&ランゲージの世界. ビジョン技術の実用ワークショップ(ViEW), 2019/12/5.
3. 【招待講演】Yoshitaka Ushiku. Deep Learning for Natural Language Processing and Computer Vision. Tutorial on Asian Conference on Machine Learning, Nagoya, Japan, 2019/11/17.
4. 【招待講演】牛久祥孝. 画像・映像理解と自然言語への架け橋. 情報処理学会連続セミナー第3回, 2019/9/26.
5. 【招待講演】牛久祥孝. 多様なデータへのキャプションを自動で生成する技術の創出. 情報科学技術フォーラム(FIT), 2019/9/3.
6. 【招待講演】牛久祥孝. ビジョン&ランゲージ~「意図」をどのようにモデリングするか? 画像センシングシンポジウム(SSII), 2019/6/12.

# 研究報告書

## 「視覚に基づく言い換えのセマンティック類型」

研究期間: 2019年4月~2021年3月  
 (新型コロナウイルス感染症の影響を受け2021年9月まで延長)  
 研究者番号: 50226  
 研究者: チョ シンキ

### 1. 研究のねらい

Visually grounded paraphrases (VGPs) are different phrasal expressions describing the same visual concept in an image. Our research until now treats VGP identification as a binary classification task, which ignores various phenomena behind VGPs. E.g., Figure 1 (a) “field hockey” and “lacrosse” are linguistic paraphrases; Figure 1 (g) “competitors” and “a group of bicyclist” describe the same visual concept from different aspects; however, these two pairs of VGPs have been treated equally, which is obvious undesirable. In the acceleration phase, we aim to **create typology of VGPs to elucidate the phenomena behind VGPs**. Studying typology for linguistic paraphrases is not new in natural language processing. The paraphrase typology focuses on lexicon, syntax and discourse phenomena in paraphrases. VGPs differ from paraphrases that they focus on the semantic aspects of images to describe the same visual concept. Therefore, we believe that the paraphrase typology is unsuitable for VGPs, and **propose to create VGP typology based on semantics**.

In natural language inference, seven basic semantic relations between two phrases have been defined, i.e., *equivalence*, *forward entailment*, *reverse entailment*, *negation*, *alternation*, *cover* and *independence*. As VGPs describe the same concrete visual concept, *negation* and *cover* relations should not exist in VGPs.

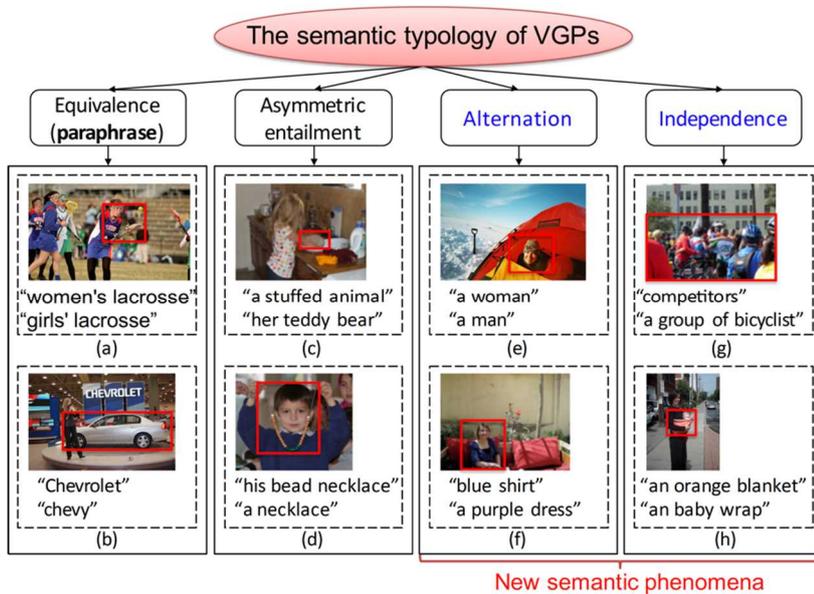


Figure 1: The semantic typology of VGPs.

Detailed analyses on the Flickr30k entities dataset indicate that the other five semantic relations cover VGPs. Figure 1 shows our semantic VGP typology (note that forward and reverse entailment are combined, but these can be further classified according to applications). Figure 1 (a) and (b) are linguistic paraphrases; Figure 1 (c) and (d) are entailment VGPs where one phrase

contains more fine object attribute description of the same concept, i.e., “teddy bear” and “bead.” Knowledge such as “teddy bear *is a* stuffed animal” can be mined from this type. Figure 1 (e) and (f) use alternate phrases to describe the same visual concept, which may come from the difference in human recognition. The phrases in Figure 1 (g) and (h) are linguistically independent but become VGPs upon grounding. We believe that **alternation and independence VGPs are new semantic phenomena that cannot be explained only with language**, and visual inference is required to automatically identify them. More interestingly, when we look at a combination of two noun phrases, **the semantic relation can change from the sub-noun phrases**. E.g., “A woman speaking in front of a chevy” and “A woman promoting the Chevrolet” of Figure 1 (b), although both the sub-noun phrases are *equivalence*, the relation of the entire phrase pair is *independence*.

In the acceleration phase, **we plan to annotate the VGP typology on the Flickr30k entity dataset and design novel language and vision models to automatically classify VGPs**. The creation of the semantic VGP typology will not only elucidate the phenomena behind VGPs but also **open up novel ways of utilizing VGPs for various language and vision tasks, which require semantic understanding**. E.g., it can significantly boost the performance of textual entailment via visual grounding, which is a fundamental but very challenging natural language understanding problem. Understanding the semantic relation via visual grounding is also crucial towards machine reading, which is the main challenge in the Todai robot project. Therefore, we believe that this work will **significantly deepen and promote the research in both language and vision understanding**.

## 2. 研究成果

### (1) 概要

In summary, we mainly studied the following VGP topics in the acceleration phase.

1. **VGP identification improvement.** We improved the VGP identification model proposed in the ACT-I research period via a gated network and visual grounding. Our model can adaptively use visual and language features based on different VGPs.
2. **VGP typology annotation and classification.** Following the VGP typology explained in the previous section, we annotated a large-scale VGP typology dataset based on the original VGP dataset and designed a model for VGP classification.
3. **Cross-lingual visual grounding.** As the core technology of VGP, visual grounding has been studied for English only. We created a French visual grounding dataset and created a cross-lingual visual grounding model for French visual grounding via transferring knowledge from English.

Besides the core VGP research, we also studied its applications of visual question answering and human prediction. We published our work at international journals of Neurocomputing [1], IEEE Access [2], and international conferences of AAAI, WACV etc. [3–7]

During the extended 6 months research period due to COVID-19, we further studied

artwork classification, video-guided machine translation, and visually grounded supervision for vision-and-language transformers. We published our work at international conferences of ICMR and ACL-IJCNLP-SWR [8–10].

## (2) 詳細

We detail the main VGP topics in the acceleration phase that we studied.

### 1. VGP identification improvement

Previous studies have developed models to identify VGPs from language and visual features. In these existing methods, language and visual features are simply fused. However, our detailed analysis indicates that VGPs with different lexical similarities require different weights on language and visual features to maximize identification performance. This motivates us to propose a gated neural network model (Figure 2) to adaptively control the weights. In addition, because VGP identification is closely related to phrase localization, we also propose a way to explicitly incorporate phrase-object correspondences. From our evaluation in detail, we confirmed our model outperforms the state-of-the-art model.

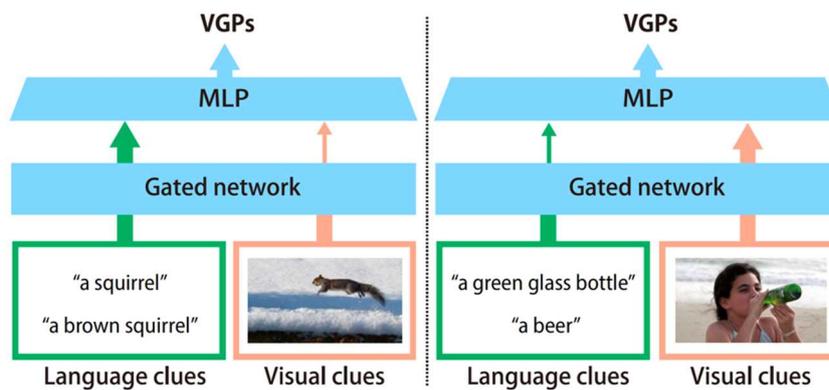


Figure 2: Our gated network for VGP identification.

### 2. VGP typology annotation and classification

Previous studies treat VGP identification as a binary classification task, which ignores various phenomena behind VGPs such as linguistic paraphrases and VGPs from different aspects. In this work, we propose semantic typology for VGPs, aiming to elucidate the VGP phenomena. We construct a large VGP dataset that annotates the class to which each VGP pair belongs according to our typology. In addition, we present a classification model that fuses language and visual features for VGP classification on our dataset (Figure 3). Experiments indicate that joint language and vision representation learning is important for VGP classification.

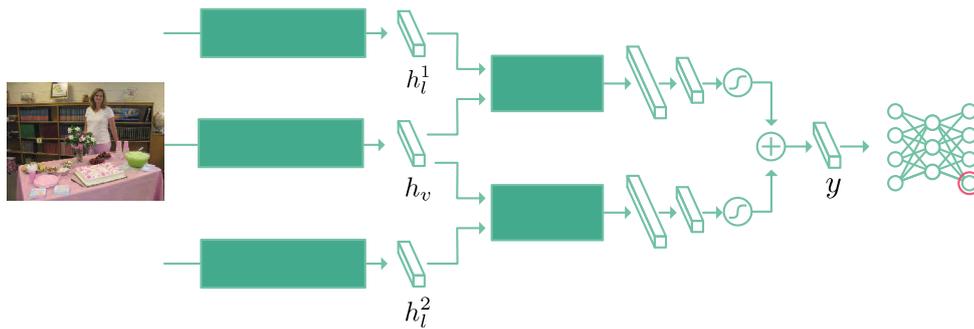


Figure 3: Our VGP classification model.

### 3. Cross-lingual visual grounding

Visual grounding is a vision and language understanding task aiming at locating a region in an image according to a specific query phrase. However, most previous studies only address this task for the English language. In this work, we present the first work on cross-lingual visual grounding to expand the task to different languages to study an effective yet efficient way for visual grounding on other languages. We construct a visual grounding dataset for French via crowdsourcing. Our dataset consists of 14k, 3k, and 3k query phrases with their corresponding image regions for 5k, 1k, and 1k training, validation and test images, respectively. In addition, we propose a cross-lingual visual grounding approach that transfers the knowledge from a learnt English model to a French model (Figure 4). Despite that the size of our French dataset is 1/6 of the English dataset, experiments indicate that our model achieves an accuracy of 65.17%, which is comparable to the accuracy 69.04% of the English model.

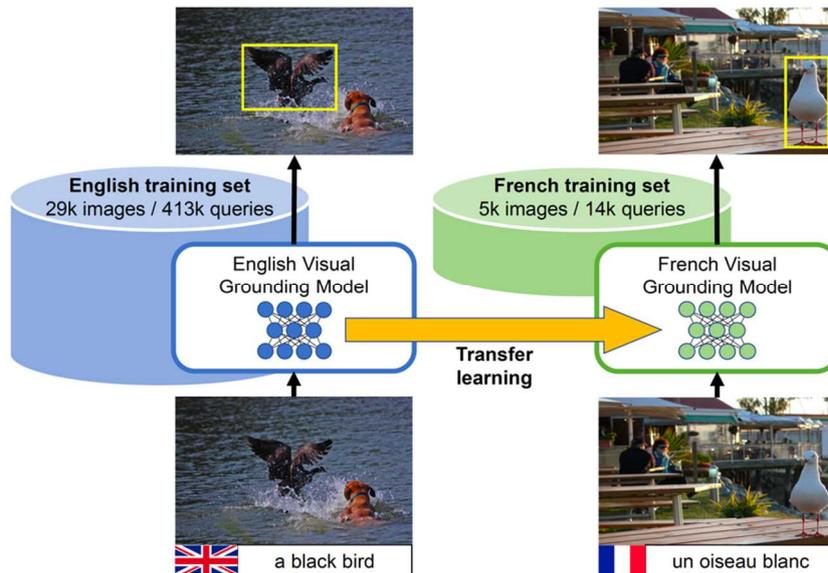


Figure 4: Cross-lingual visual grounding.

### 3. 今後の展開

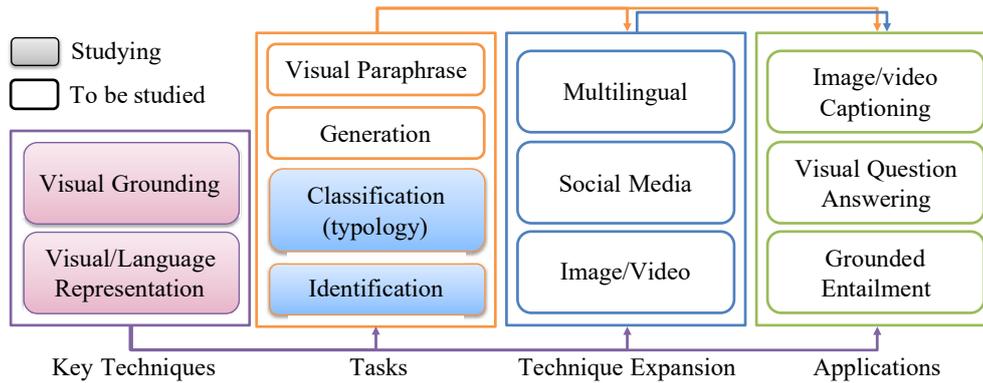


Figure 5: The VGP research field that we aim to establish (arrows denote the technical hierarchy between each module category).

Although we originally created the concept of VGP and its typology in our JST ACT-I project, this is just a **start point**. There is a large amount of problems to be studied, by addressing which we can: **1) establish a research field of VGP; 2) establish vision and language processing techniques based on VGP understanding; 3) it has the potential to elucidate how humans learn and process paraphrases in vision and language understanding.** Figure 5 summarizes these problems and shows the VGP research field that we aim to establish.

### 4. 自己評価

We have successfully achieved our research goal, which was creating typology for VGP. We further improved VGP identification and proposed cross-lingual visual grounding.

#### ・研究の進め方(研究実施体制及び研究費執行状況)

Regarding to the research implementation system, because of the multimodality of the project, we closely collaborated with researchers in the computer vision (CV) field. We also employed RA students for creating VGP related datasets.

Regarding to the research funding execution status, we reasonably used it for purchasing GPU servers, dataset annotation, RA employment, and travel expenses. However, due to COVID-19, we have to apply for an extension of half a year for our project.

#### ・研究成果の科学技術及び学術・産業・社会・文化への波及効果

VGP identification and classification that we have started in our ACT-I research period is a great start point for VGP research. VGP research is crucial for both language and vision understanding. We believe that our pioneering work on VGPs in the ACT-I research period has the potential to make VGPs a new language and vision research field in both NLP and CV fields. We are eager to promote the VGP research field with other researchers further.

#### ・研究課題の独創性・挑戦性

We originally proposed the concept of VGPs, which is an achievement inspired by our study in

both NLP and CV. Due to the pioneer and multimodality characteristics, this work is very challenging. We overcame the problems and systematized VGP research via close collaboration between NLP and CV researchers.

## 5. 主な研究成果リスト

### (1) 論文(原著論文)発表

[1]. Wenjian Dong, Mayu Otani, Noa Garcia, Yuta Nakashima, Chenhui Chu. Cross-Lingual Visual Grounding. IEEE Access, (2021).

[2]. Mayu Otani, Chenhui Chu, Yuta Nakashima. Visually Grounded Paraphrase Identification via Gating and Phrase Localization. Neurocomputing, Volume 404, Pages 165–172, (2020).

### (2) 特許出願

研究期間累積件数:0 件

### (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)

[3] Yuta Kayatani, Zekun Yang, Mayu Otani, Noa Garcia, Chenhui Chu, Yuta Nakashima, Haruo Takemura. The Laughing Machine: Predicting Humor in Video. In Proceedings of the IEEE 2021 Winter Conference on Applications of Computer Vision (WACV 2021), (2021).

[4] Zekun Yang, Noa Garcia, Chenhui Chu, Mayu Otani, Yuta Nakashima, Haruo Takemura. BERT Representations for Video Question Answering. In Proceedings of the IEEE 2020 Winter Conference on Applications of Computer Vision (WACV 2020), (2020).

[5] Noa Garcia, Mayu Otani, Chenhui Chu, Yuta Nakashima. KnowIT VQA: Answering Knowledge-Based Questions about Videos. In Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI 2020), (2020).

[6] Mayu Otani, Chenhui Chu, Yuta Nakashima. Adaptive Gating Mechanism for Identifying Visually Grounded Paraphrases. ICCV 2019 MDALC Workshop, (2019).

[7] Chenhui Chu. Visually Grounded Paraphrase Identification. Third International Workshop on Symbolic-Neural Learning, (2019).

[8] Cheikh Brahim El Vaigh, Noa Garcia, Benjamin Renoust, Chenhui Chu, Yuta Nakashima, Hajime Nagahara. GCNBoost: Artwork Classification by Label Propagation through a Knowledge Graph. In Proceedings of the ACM International Conference on Multimedia Retrieval 2021 (ICMR 2021), (2021). (コロナ延長時の成果)

[9] Weiqi Gu, Haiyue Song, Chenhui Chu and Sadao Kurohashi. Video-guided Machine Translation with Spatial Hierarchical Attention Network. In Proceedings of the ACL-IJCNLP 2021 Student Research Workshop, (2021). (コロナ延長時の成果)

[10] Jules Samaran, Noa Garcia, Mayu Otani, Chenhui Chu and Yuta Nakashima. Attending Self-Attention: A Case Study of Visually Grounded Supervision in Vision-and-Language Transformers. In Proceedings of the ACL-IJCNLP 2021 Student Research Workshop, (2021).

## 研究報告書

### 「衛星リモートセンシングから得られる時空間ビッグデータの機械学習による地震被害の判別」

研究期間：2019年4月～2021年3月

(新型コロナウイルス感染症の影響を受け2021年9月まで延長)

研究者番号：50228

研究者：宮本 崇

#### 1. 研究のねらい

大地震が生じた際の人命救助などの初動対応を円滑に行う上で、建物の倒壊状況などの被害情報の収集は重要となる。被害情報は現状では人的に収集されているが、近年は人工衛星画像の分析が被害情報の把握手段として注目を受けている。衛星画像は高速に広範囲を撮影する一方で、その解像度の限界から住宅1棟単位の被害判別が困難とされており、重要な研究課題とされてきた。そこで本研究では、近年に発展の著しい深層学習をはじめとした機械学習手法を用いることによって、衛星画像からの住宅1棟単位での地震被害判別を高精度することを目的とする。

被害判別の高精度化を達成するにあたり、本研究では以下の2つのアイデアを導入する。

1 点目は、地震後に撮影された画像だけでなく、平常時から撮影・ストックされている衛星画像を入力データとして用いることにより、時系列での画像の変化を捉えて建物の被害判別に利用する点である。2 点目は、衛星画像に加えて建物築年代などの被害有無と関連の高い情報を補助的に用いることにより、被害判別の確度を高めようとする点である。

上記のような考えを反映させた被害判別のシステムを実現することにより、今後もハードウェア開発の継続が期待される人工衛星によって周期的に撮影・蓄積される、大規模な撮影画像群を災害時に効果的に利用するデータ処理プログラムを構築する。衛星から得られる災害情報は、都市内の監視カメラや UAV・SNS などから得られる情報などと併せて多角的に分析されることにより、災害直後の混乱期において客観的に高精度な被害情報を得る手段として活用されることが期待される。本研究提案では、そのような次世代の防災体制構築のための基幹的な要素技術の一つを開発することを狙いとする。

#### 2. 研究成果

##### (1) 概要

多時期に渡る衛星画像の系列データを元に、地震によって異常の生じた箇所を検出する手法として、時系列異常検知の手法と3次元畳み込みニューラルネットワークの比較を行い、後者の有効性を確認した。また、衛星画像と構造物情報という異なるモードの情報を統合して被害有無を判断する手法として、マルチモーダルニューラルネットワークモデルを設計し、衛星画像単体による被害有無の判断よりも精度が大きく向上する結果を得た。

上記の2つの構造を統合した深層学習モデルによって、住宅の倒壊・非倒壊の判別精度を約90%まで向上させる結果が得られた。判別に失敗した10%のデータの分析からは、住宅の傾きや1階部分のみの破壊など、直上からの撮影では把握の困難な破壊形態が半数

以上を占めていることが分かった。このことから、本研究で開発したモデルは上空からの衛星画像から判別可能な住宅の倒壊を十分に把握する性能を有していると結論付けた。

開発した技術について都道府県庁や内閣府へとヒアリングを行い、初動対応や激甚災害の指定といった場面における開発技術の適用性を確認するとともに、ヒアリング先の協力の下で今後は実証実験を行うこととなった。

また、本研究を通して得られた、防災上の問題に対する機械学習手法の応用に関する知見を体系的に整理し、国内外の事例に関するレビューを行いながら論文として公表を行った。同論文では、本研究のような被災状況の把握や災害事象の予測といった課題において、データの不均衡性、機械学習モデルの説明性、データのない領域での外挿的予測性能の向上、の3点の克服が重要であることを指摘し、これら3点に対応する情報学的手法の整理を行うと共に、今後の見通しについて論述した。

新型コロナウイルス感染症の影響を受け6ヶ月間研究期間を延長し、衛星データを用いた被害推定結果に関する上述の分析結果を取りまとめ国際誌へと投稿し採択された。また、防災上の問題に対する機械学習手法の応用において前年度に指摘したデータの不均衡性・機械学習モデルの説明性・外挿的予測性能の向上、の3つの課題に対処する方法として、物理的モデリング手法とデータ科学的モデリング手法の統合の重要性を指摘し、現在の研究動向を調査して整理したうえでレビュー論文として取りまとめを行った。

## (2) 詳細

### 研究テーマA「衛星画像の系列データからの地震被害判別手法の検討」

平常時から撮影・保存されている衛星画像群と地震後に緊急撮影された衛星画像の比較から、地震によって異常のあった住宅を同定・検出する手法の検討を行った。大きく、

①平常時の撮影画像を正常データとして学習し、地震後の画像が正常データの範囲を外れたかどうかを判断する異常検知モデル

②地震前・地震後の画像ペアから画像特徴の時系列的変化を抽出し、被害有無を判別する時系列畳み込みニューラルネットワークモデル

の2種の機械学習モデルを比較検討した結果、後者の方が優位であることが分かった(図1)。

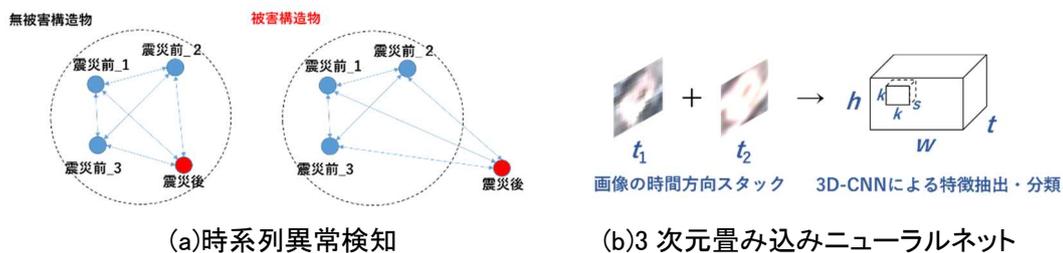


図1 衛星画像の系列データからの地震被害判別モデル

### 研究テーマB「衛星画像を補完するデータを活用したマルチモーダル機械学習による被害判別モデルの検討」

地震直後から被害判別のために活用可能なデータとして、住宅の築年代と構造種別の 2 種のデータに着目し、これらのデータと衛星画像から統合的に地震被害の有無を判別するマルチモーダルニューラルネットワークを設計した(図 2)。

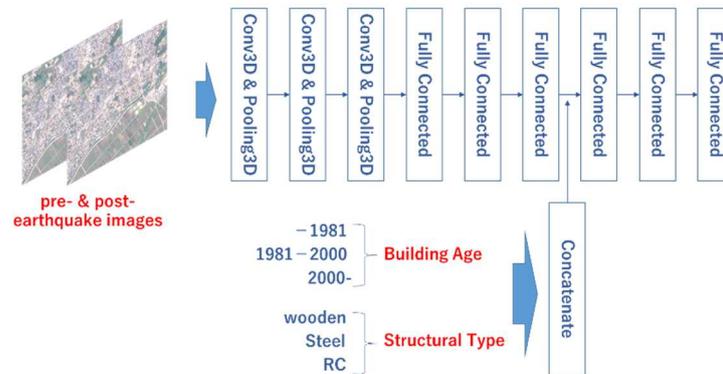


図 2 マルチモーダルニューラルネットワークによる被害判別モデル

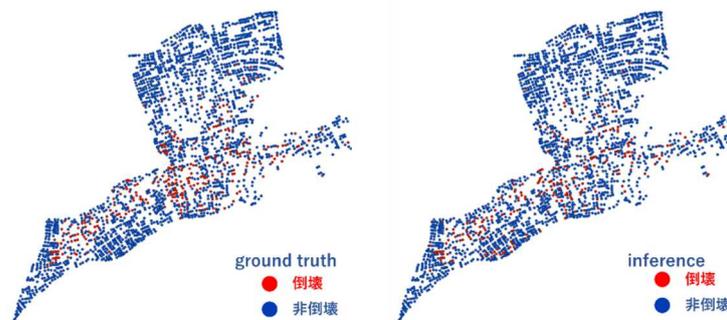


図 3 実際の被害分布(左)と提案手法による被害推定結果(右):  
2016 年熊本地震における被災地域を分析

検討手法によって、約 90%の精度で倒壊住宅を検出可能であることが確認された(図 3, および研究成果リスト:論文 1)。また、判別を誤った 10%について詳細に分析を行った結果、その半数以上は住宅の傾きや 1 階部分のみの倒壊といった、直上撮影からは原理的に判別の難しい倒壊形態であることが分かった。そのため、本研究で開発したモデルは、上空からの衛星画像から判別可能な住宅の倒壊を十分に把握する性能を有していると考えられる。

【誤判別したデータの内訳】

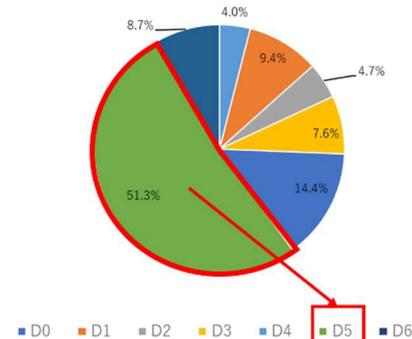


図 3 誤判別したデータの内訳:半数以上は D5(部分的な倒壊や傾き)に属している

### 研究テーマ C「防災上の課題に対する機械学習手法応用のレビューと体系化」

本研究を通して得られた、防災上の問題に対する機械学習手法の応用に関する知見を体系的に整理した。本研究のような被災状況の把握や災害事象の予測といった課題におい

て、データの不均衡性、機械学習モデルの説明性、データのない領域での外挿的予測性能の向上、の 3 点の克服が重要であることを指摘し、これら 3 点に対応する情報学的手法の整理を行うと共に、今後の見通しについて論述した(研究成果リスト:論文 2,3)。またここで挙げた 3 つの課題に対処する方法として、物理的モデリング手法とデータ科学的モデリング手法の統合の重要性を指摘し、現在の研究動向を調査して整理した上でレビュー論文として取りまとめを行った。

### 3. 今後の展開

本研究を通して開発した地震被害把握の技術は、我が国における衛星ビッグデータや構造物データといった情報資産を最大限に活用しながら、地震直後の災害対応の円滑化を行うものであり、1 応歳の実務上大きな意義を有するものである。また学術的な観点からは、先端的なセンシング手法や機械学習技術を防災学上の課題に適用している点において、日本の目指す超スマート社会の将来像に沿った防災の在り方の実現に資するものと考えられる。

今後は、大きく以下の 2 点の方向性において、研究の展開が考えられる。1 点目は、衛星画像は直上撮影という手法を採るために、住宅内部の破壊やわずかな傾きといった破壊形態を原理的に検知することが難しい。そのため、他のセンシング手法との相補的な運用による、より詳細な被害把握手法の確立が望まれる。2 点目は、国際社会に目を向けた際に、日本のように各構造物の詳細なデータといった情報資産が豊富な国は多くない。したがって、全球的にデータ取得可能な衛星画像のみから、地震による被災地域の被害判別を行う精度を向上させることは、国際的な防災力の向上に必要と考えられる。

### 4. 自己評価

衛星データからの住宅 1 棟単位での被害判別という当初に掲げた研究目的について、具体的な判別モデルの提案、衛星データから判別可能な被害形態の限界に関する検証と考察、という 2 点を達成することができた他、今後の方針として、上記の今後の展開に述べたように他のセンシング手法との相補的な運用などの発展的な研究課題が明らかになった。また、本研究課題を通して得られた知見を元に、防災学と情報学の接点に関するレビューを行い、論文やパネルディスカッション、学会誌への寄稿といった様々な形で広く本取り組みを周知することができた。こうしたことから、当初の目的については十分に達成することができたと考えている。

本研究は、ACT-I 研究プログラムにおける領域会議やサイトビジット等の行事をマイルストーンとして計画的に行うことができた。最終年度にはコロナ禍の影響を受け、特に一部物品費や旅費の執行、および研究の実施に困難が生じたが、研究期間を延長することによって、上記の今後の展開に述べた一部の発展的研究課題について取り組む機会とする計画である。

地震をはじめとする様々な自然社会に度々見舞われる我が国において、先端的なセンシング手法やデータ分析技術を活用することにより、災害対応をより円滑にすることは学術・実務の両面から大きく重要視されている。本研究は、そうした潮流の下での先駆的な研究成果と位置付けられる。

これまでに、衛星データから地震被害を分析しようとする研究は多くあった中で、震災前後の衛星画像を時空間データとして捉え 3 次元畳み込み構造を適用する着想、および多数のデー

タから被害判別を行うマルチモーダル学習を適用する着想はいずれも本研究が独自に提案・検証したものであり、研究の独自性は高いものである。

## 5. 主な研究成果リスト

### (1) 論文(原著論文)発表

研究期間累積件数: 7件

1. Takashi Miyamoto and Yudai Yamamoto: Using multimodal learning model for earthquake damage detection based on optical satellite imagery and structural attributes, Proceedings of 2020 IEEE International Geoscience and Remote Sensing Symposium, 2020, pp.6623-6626

地震前後の画像を詳細に比較することで地震による建物形状の変化を検知する 3D convolution を導入した AI 技術を開発し、従来の画像処理技術よりも高い精度で建物の倒壊を検知できることを検証した。また、日本全国で社会のデジタルツインとして整備が進んでいるデジタル都市データから、地震時の建物被害と関連の深い建物の築年代と建築材料の 2 種の情報を抽出し、衛星画像の画像処理結果と統合した分析を AI に行わせることで、更に倒壊した建物の検知精度を向上させた。

2. 宮本崇, 浅川匡, 久保久彦, 野村泰稔, 宮森保紀: 防災応用の観点からの機械学習の研究動向, AI・データサイエンスシンポジウム論文集, 2020, Vol.1, No.J1, pp.242-251.

機械学習モデルにおける研究動向からこの 2 点へ個別に対処するための方法論, およびそれら 2 点の課題に同時に対処する数理モデルとデータ駆動モデルを統合したアプローチについて, それぞれの考え方や具体的な手法, 応用事例を調査した。

3. 宮本崇: パターン認識と法則発見のデータサイエンス, AI・データサイエンスシンポジウム論文集, 2020, Vol.1, No.J1, pp.270-277.

データ科学的手法の工学的応用において, データの取得が本質的に難しい場合などに求められる計算モデルの外挿能力や, 意思決定を伴う問題などにおいて重要視される計算モデルの説明性といった問題へ対処する方針の 1 つとして, 現象の背後にある法則をデータ駆動的に発見し計算モデルを構築する法則発見の考え方について, その概要と現在の研究動向, 今後の発展性を論じた。

4. Takashi Miyamoto, Yudai Yamamoto: Using 3D Convolution and Multimodal Architecture for Earthquake Damage Detection Based on Satellite Imagery and Digital Urban Data, Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, in press ※  
コロナ延長時の成果

3D convolution とマルチモーダル構造を導入した AI 技術によって, 衛星画像と都市データから 92%の精度での被災建造物の検知に成功した。また, 誤検知した 8%のデータについて詳細な分析を行い, その過半数は直上撮影からの把握が困難な破壊形態の建造物であることを確認し, 地震被害分析における衛星画像の適用範囲について明らかにした。

### (2) 特許出願

ありません。

(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)

1. 受賞: 2020 AI・データサイエンス論文賞
2. 受賞: 2020 応用力学シンポジウム講演賞
3. 学会発表: Takashi Miyamoto and Yudai Yamamoto: Using multimodal learning model for earthquake damage detection based on optical satellite imagery and structural attributes, 2020 IEEE International Geoscience and Remote Sensing Symposium, Online, September 2020.
4. 学会発表: Takashi Miyamoto: Earthquake Building Damage Detection Using 3D Spatio-Temporal Convolutional Neural Network and Multi-Temporal Satellite Imagery, AAI 2019 Artificial Intelligence for Natural Disasters Workshop, Hawaii(USA), January 2019.
5. プレスリリース: 「大地震時の被害状況を即時に把握する AI 技術を開発」, 2021/8/23

## 研究報告書

### 「水泳プール中の水の流れを3次元計測する技術の開発」

研究期間：2019年4月～2021年3月

(新型コロナウイルス感染症の影響を受け2021年9月まで延長)

研究者番号：50229

研究者：山下 聖悟

#### 1. 研究のねらい

水中での推進効率や速度を向上させるためには、水流を理解することが大事であることが知られている。例えば、水中で生活する魚や昆虫は泳ぐ際に後方に複数の渦状の水流を作り、その力により水中を素早く移動している。人の泳動作による水流を広範囲で計測することができるになれば、水中での人の推進を妨げる抵抗の発生原因や位置を新たに明らかにできる。また、3次元的水流が計測できれば、人の推進メカニズムの解明やより効率のよい泳動作の発見ができる可能性がある。しかし、既存の流体計測手法には、人が存在する環境での流体計測においていくつかの問題があった。流体計測には、流体に微小な粒子(トレーサー粒子)を散布し、レーザー照射による粒子の散乱光によって粒子を周囲より明るくし、カメラによる水の流れを計測可能にする粒子イメージ計測法(PIV)が用いられることが多い。しかし、広範囲の粒子に十分な光を照射するためには強力なレーザー光源が必要になる。もし人全体をカバーするような範囲での流体計測が必要である場合、人の肌や目に重大な悪影響を与える可能性がある強力なレーザー光源が必要とされる。本研究では、人体に悪影響を与えないトレーサー粒子と光源を用いた流体計測手法を開発する。提案流体計測手法には、計測範囲の拡大が従来の手法に比べて容易であり、トレーサー粒子が均一に分散する・沈殿しないという特徴も持つ。本研究では、トレーサー粒子として偏光面を回転させる能力(複屈折あるいは旋光性)のある粒子(セルロースなど)を用いる。トレーサー粒子を可視化するためには、偏光した光を照射する室内灯程度の光量のバックライトと、カメラのレンズに設置された偏光板を用いる。2枚の偏光板(右・左円偏光板)を重ねた場合、後部からの光が透過しなくなる遮光が起こる。2つの偏光板に挟まれた位置に偏光面を回転する能力を持つ粒子がある場合、偏光板の外から見ると粒子のみが明るくなる。粒子のみが周囲に比べて大幅に明るく見える場合、カメラを用いた位置計測が可能となる。

## 2. 研究成果

### (1) 概要

ACT-I 加速フェーズでは、ACT-I の研究課題である“水泳プール中の水の流れ 3 次元計測する技術の開発”をより深める研究を行った。具体的には、水泳、流体計測を専門とする研究者との共同研究の推進、実用性の向上、他研究領域への応用を目指した。

申請者は、水泳動作と水中での抵抗・抗力や水泳動作による水流について高い専門性を持つ鹿屋体育大学(現筑波大学)の研究者らとの共同研究を開始した。このコラボレーションは本研究課題を飛躍的に加速させる。本研究では、水流計測を水泳のために用いている専門家からの意見を継続的に得ながら、より学術的な貢献が大きくなる方向へ研究を的確に進めた。

プール中の水の動きを計測するためには、トレーサー粒子が万遍なく長期間水中に漂っている状態が好ましい。しかし、ACT-I 研究期間で開発したトレーサー粒子(人工いくら)はゆっくりと底に向かって沈むため、計測の際には水をかき混ぜる等の方法により水中に漂わせる必要があった。また、粒子径が大きいため、大まかな水の流れしか計測できないという制限もあった。加速フェーズでは、実用性に関わるこれらの問題を解決した。申請者はこれまでの基礎実験から、セルロースが既存のトレーサーの問題(沈殿速度の早さ、粒子径の大きさ)を解決する糸口になることを明らかにした。提案した光学系は複屈折を持つ部分だけを周囲に比べて明るく可視化することができる。また、ナノセルロースのトレーサー粒子は、流れによって複屈折が生じる流動性複屈折という光学特性を持つ。つまり、水泳動作などにより水が流れた場合、その部分のみが周囲に比べて明るくなる。計測が必要な部分のみが可視化されるため、既存の流体計測手法の問題点とされていた粒子同士による遮蔽が少なくなり、より広範囲で流体計測が可能になると考えられる。

新型コロナウイルス感染症の影響を受け 6 ヶ月間研究期間を延長した。延長された期間内には、3次元的な流れを定量的に計測するための光学系や、特定の種類の流れのみを撮影可能にする基礎技術の開発を行った。ACT-I 加速フェーズの期間では、様々な角度から撮影することにより3次元的な流れの概要を把握できるようにしたが、定量的な計測は実現できていなかった。ACT-I 加速フェーズにて提案した手法では、偏光を照射する平面光源と偏光板を取り付けたカメラの間に水槽(流れを観測する環境)を置くような光学系が用いられている。延長期間の実験では、平面光源ではなく、シート状に広がる白色 LED 光源を用いて側面から光を差し込んだ場合の特性を観察した。結果として、シート状の光が照射された限られた奥行き範囲のみに”流れている部分だけが明るくなる”特性が得られた。この特性を用いると、3次元的な流れを復元することも可能であると考えられる。加えて、流れの方向に応じて特定の発色や偏光角の変化が観測されることがわかった。カメラに光学フィルターを追加することによって特定の流れのみを可視化する仕組みが実現できる可能性がある。

### (2) 詳細

研究テーマ A 「水泳、流体計測を専門とする研究者との共同研究の推進」

申請者は、大型のガラス水槽(幅 2m、奥行き 3m、高さ 1.5m)を用いて人が泳ぐ際に起こす水の流れを観測する実験を共同研究者と共に行った。実験の際には、0.02wt%のセルロース

ナノクリスタル懸濁液 5 トンを水槽中に準備した。また、水泳経験者に水槽中で様々な泳法を用いて泳ぐように依頼し、その水流を観測した(図 1)。例を挙げると、スカーリングと呼ばれる水中でその場で留まる際に用いられる手の動き、クロールの手の動作などによって発生する手の周りの水流を観測した。また、平泳ぎに伴う足の動きやバタ足、ドルフィンキックの際の足周辺の水の流れについても観測した(図 2)。結果として、過去の研究により予測されたような水の流れが観測された。例えば、スカーリングは手を左右に振るような動作を行いその場に留まるための浮力を作り出すためのものである。過去の研究から、手で作り出した水の流れ(渦)を横方向に投げ、手のひらでそれを受けることで下方向に安定した水流を作り出していることがわかっている。本実験においても、手のひらが渦を作り出していること、その渦が横方向に移動すること、渦を手のひらで受けることによって結果的に下方向に流れが発生していることが観測できた[5]。ドルフィンキックにおいても、揃えた足が上から下に移動する際に足の裏の上部に強い渦が発生していることと、その渦が足の動きについていくように下方向に移動して、結果として体に対して後方下方向に大きな水の流れが発生していることが観測された。しかし、大型水槽を用いた本実験では、スペースの制限のため、連続した泳動作を行うことができない。水泳に関する学術的な貢献を提供するためには、実際のプールや回流水槽のような連続的な泳動作が可能な環境において計測を実現する必要がある。本研究では、計測範囲を更に拡大できる光学系の開発や、遠隔地への計測環境の実装を考慮した携帯性の実現、更に水の容量の多い環境にて十分な濃度のトレーサー粒子を散布する方法の確立などを目指した。



図 1 人が泳ぐ際に起こす水の流れを可視化し観測する実験

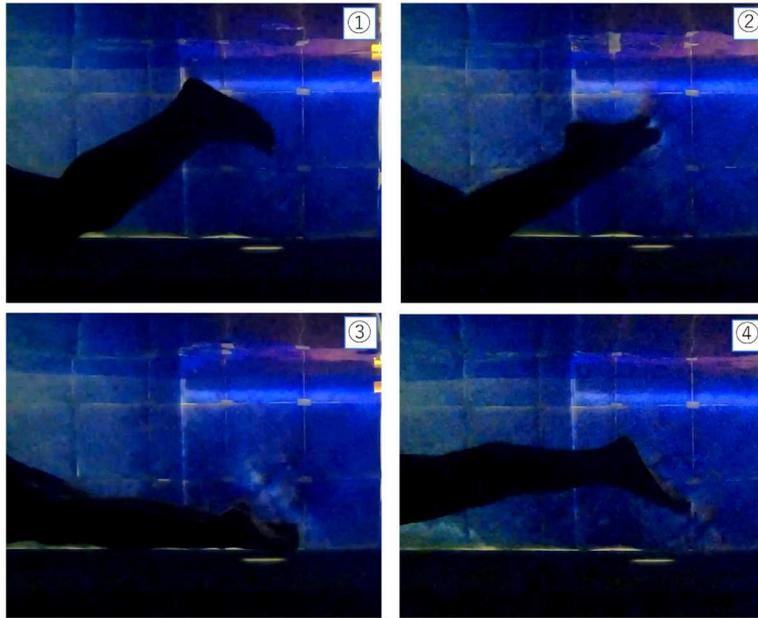


図 2 ドルフィンキックによる水の流れを可視化した例

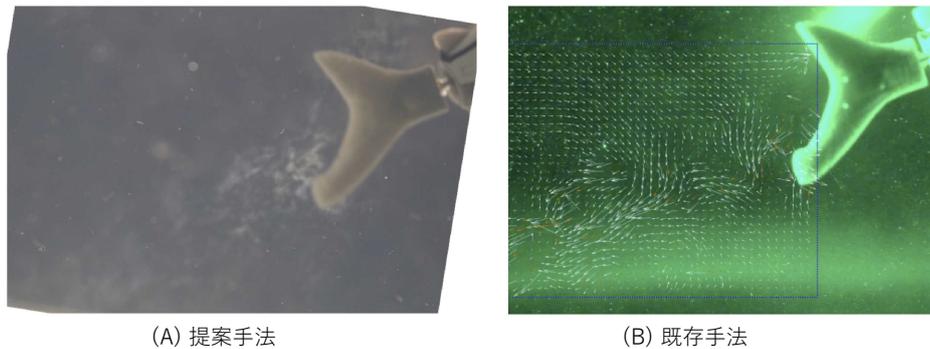
#### 研究テーマ B 「トレーサー粒子の均一分散・長時間浮遊の実現と計測可能範囲の向上」

ACT-I の期間中に、セルロース微粒子や人工いくらなどによるトレーサー粒子を提案した。しかし、これらの粒子には沈殿速度が速く粒子径が大きいため、微少な水の動きに追従することが難しいという問題があった。本研究では、ナノサイズのセルロース粒子が、既存のトレーサー粒子の問題を解決する糸口になることを基礎実験から明らかにした。セルロースは植物繊維の主成分であり健康食品としても販売されている。セルロースナノファイバー(SNF)は、ナノサイズのセルロース繊維である。SNFは最大で数ヶ月の間、沈殿または浮上せず水中を浮遊し続ける特性を持つ。また、水全体に万遍なく分散した状態を保つことも知られている。加えて、SNFは粒子サイズが微小であり光を散乱させないため水に散布した場合も水をほとんど白濁させないSNFを含む水は、流れなどにより力(せん断力)を受けた部分だけが強い複屈折を持つ特徴を持つ。また、提案する光学系は複屈折を持つ部分だけを周囲に比べて明るく可視化することができる[3]。これらを組み合わせることで、水泳動作などにより水が流れた場合、その部分のみが周囲に比べて明るく観測される。計測が必要な部分のみが可視化されるため、既存の流体計測手法の問題点とされていた粒子同士による遮蔽が少なくなり、より広範囲で流体計測が可能になると考えられる。セルロースの加工や製造は、これらに高い専門性を持つ大阪大学能木研究室の春日氏らと共同で行った。セルロース及び材料合成に精通している研究協力者の存在は、本研究をさらに加速した。

#### 研究テーマ C 「精度や制限の明確化」

セルロースナノファイバーをトレーサー粒子に用いる場合、粘度が剪断力(渦や乱流の際に

発生)に応じて変化するという特性を考慮に入れて流体計測を実現する必要がある。本研究では、従来のトレーサー粒子を用いて流体計測を行なった場合との比較実験を行い精度や制限の明確化を試みた(図 3)。具体的には、レーザーを用いてトレーサー粒子を可視化しカメラで流れを観測する既存手法を用いて水の流れを計測した。また、回流水槽中に一定の形のオブジェクトを置いた際の水の流れをシミュレーションにより求め、提案手法との比較検証のための足掛かりとした。



(A) 提案手法

(B) 既存手法

図 3 既存の光学系との比較実験の例

### 3. 今後の展開

2020 年度には、実際のプールを用いた検証実験を実施予定であった。新型コロナウイルスによる緊急事態宣言などの影響により 2020 年度の実験が延期されたため、2021 年度は引き続き、上記した検証実験の実施を目標に研究活動を行う。また、2019 年度、2020 年度に引き続き、計測範囲を更に拡大できる光学系の開発や、遠隔地への計測環境の実装を考慮した携帯性の実現、更に水の容量の多い環境にて十分な濃度のトレーサー粒子を散布する方法の確立などを目指す。2020 年度に引き続き、セルロースナノファイバーをトレーサー粒子に用いる場合と、従来のトレーサー粒子を用いて流体計測を行なった場合との比較実験を行う。比較実験により、精度や制限を明確化する。例えば、セルロースナノファイバーをトレーサー粒子に用いる場合、粘度が剪断力(渦や乱流の際に発生)に応じて変化するという特性を考慮に入れて流体計測を実現する必要がある。本研究の目標は、水の流れを 3 次元的に計測することである。そのため、ステレオカメラを用いて粒子を追跡する従来手法と、提案する粒子と可視化手法に特化した、3 次元的に流れを計測するため新たな光学系と計測手法の研究開発を行う。例えば、機械学習を用いて粒子の形状やブレを学習して、既存手法よりも正確に粒子を追跡する手法や、粒子のボケ方などの見た目の情報から、その 3 次元位置を推測する手法など考えられる。加えて、提案技術を学校やスイミングスクール等の一般社会においても使用可能な製品にすることを旨とする。また、提案技術の他研究領域への応用も促進する。例えば、本技術の光学系は透明な身体を持つゾウリムシやミジンコなどを可視化することもできるため、水中微生物の行動観察に役に立つことが明らかになっている。流体だけではなく気体への応用も可能であることが予測される。

#### 4. 自己評価

##### ・研究目的の達成状況

トレーサー粒子を可視化して水の流れを計測可能にする光学系の開発や、トレーサー粒子の検証についてはおおよそ計画通りに進行している。しかし、粒子を3次元的に計測するための機構についてはまだ多くの課題が残されている。例えば、現段階では、トレーサー粒子の数を従来手法に比べて大幅に制限した場合にのみ3次元計測が可能である。水の流れの3次元的な計測や、様々な角度からの水の流れの可視化はバイオメカニクス領域に大きな学術的な貢献をもたらすため、引き続き誠意的に研究開発を進める。また、人が連続的に泳動作を行うことができる大型の水槽またはプールにおける計測実験は、水泳のメカニズムに関する理論の証明などに必須である。新型コロナウイルスの感染拡大の影響により年度内の大規模な実験はほとんど断念されてしまったが、来年度は、実際の水泳環境を用いた人の起こす水の流れの可視化に関する実験を遂行できるように準備を進めている。

##### ・研究の進め方(研究実施体制及び研究費執行状況)

2019年度は、当初予定された通りに研究が実施された。検証実験のための費用が予定よりも多くかかることが予測されたため、増額措置に申請した。2020年度は、新型コロナウイルスの感染拡大の影響で、ほとんどの実験が来年度以降に延期された。実際の水泳環境を用いた大規模な実験を来年度に行うために予算の執行は制限された。既存の流体計測手法との比較実験のために想定よりも多くの費用が必要になることがわかったため、増額措置に申請した。

##### ・研究成果の科学技術及び学術・産業・社会・文化への波及効果(今後の見込みも重視してください。)

流体計測における提案手法の有効性がある程度証明されたが、既存の手法との定量的な比較や制限の明確化が不足している。そのため、本研究成果に関心がある企業や研究者らとともに、提案手法をより実用的にするための研究開発を続ける。例を挙げると、セルロースナノファイバーをトレーサー粒子に用いる場合、粘度が剪断力(渦や乱流の際に発生)に応じて変化するという特性を考慮に入れて流体計測を実現する必要がある。また、従来のトレーサー粒子を用いて流体計測を行なった場合との比較実験を行い精度や制限を明確化する必要がある。3次元的に流れを計測するための新たな仕組みも必要とされる。

##### ・研究課題の独創性・挑戦性

本研究課題の進行には光学、情報処理、バイオメカニクスなどの総合的な知識と技術が必要である。申請者は、水泳動作と水中での抵抗・抗力や水泳動作による水流について高い専門性を持つ研究者との共同研究を推進している。加えて、材料の開発や検証に専門的な知識を持つ研究者らも協力しながら研究開発を進めている。これらのコラボレーションは、情報工学や光学の知見を他の領域に応用し、様々な領域で学術的な貢献を生み出すことに寄与している。

#### 5. 主な研究成果リスト

##### (1)論文(原著論文)発表

1. Shogo Yamashita, Shunichi Suwa, Takashi Miyaki, Jun Rekimoto, "Feasibility Study on Water Flow Visualization Using Cellulose Particles and Pervasive

Display”, The 8th ACM International Symposium on Pervasive Displays(PerDis 2019), June 12-14, 2019.
2. Shogo Yamashita, Shunichi Suwa, Takashi Miyaki, Jun Rekimoto, “Water Flow Measurement for Swimmers using Artificial Food-grade Roe as Tracer Particles”, The 5th International ACM In Cooperation HCI and UX Conference(CHIuXiD 2019), April 8-9, 2019
3. Shogo Yamashita, Takaaki Kasuga, Shunichi Suwa, Takashi Miyaki, Masaya Nogi and Jun Rekimoto, “Fluid-Measurement Technology using Flow Birefringence of Nanocellulose”, SIGGRAPH '19: ACM SIGGRAPH 2019 Posters, July 13, 2019
4. Shogo Yamashita, Shunichi Suwa, Takashi Miyaki, Jun Rekimoto, “A Human-Friendly Fluid Measurement Technology using Artificial Fish Eggs as a Tracer”, ACM CHI 2019 ASIAN CHI SYMPOSIUM: EMERGING HCI RESEARCH COLLECTION, May 2, 2019
5. Shogo Yamashita, Shunichi Suwa, Takashi Miyaki, Jun Rekimoto, “Water Flow Visualization for Swimming using Nanocellulose”, ACM CHI 2020 ASIAN CHI SYMPOSIUM: EMERGING HCI RESEARCH COLLECTION, April 25, 2019

(2)特許出願

研究期間累積件数:1件

1.

発明者: 山下 聖悟

発明の名称 : 3次元流体計測法

出願人: 東京大学

出願日: 2019/07/12

出願番号: 2019-13324 (2018-12298)

(2)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)

Best Poster Award at CHI 2019 Asia CHI symposium (The 5th Asian CHI Symposium Emerging HCI Research Collection)

## 研究報告書

### 「適材適所システム: 広葉樹林業で発生する多様な小径木の家具および建築への応用」

研究期間: 2019年4月～2021年3月

(新型コロナウイルス感染症の影響を受け2021年9月まで延長)

研究者: 吉田 博則

#### 1. 研究のねらい

戦後の針葉樹林化に伴い落葉広葉樹林の減少また、安価な輸入材により人工林の放置が問題となっている。広葉樹林を維持管理するには択伐による小径木や枝の処理が必要であるが、これらはチップ化され燃やすしかなかった。本研究の狙いはこれらの使いづらい多様な形状の木材を、より高付加価値を持つ家具や建築に活用できるようにすることである。落葉広葉樹は堅く密度も高いので成木だと高価だが、択伐から出る材なので安価に入手可能である。これらの材は使い方を工夫すれば、高付加価値を持つ家具や建築構造に十分応用可能である。

従来の伝統工法では経験ある職人によって材の特徴を見極める適材適所が行われていたが、人手不足や国内林業の衰退もあり今後ますますこのような材を使う機会は減っていくと考えられる。本研究では各々の材の特徴をデジタル化したデータを活用し、適材適所を誰もが実現できるワークフローを社会実装することを目指す。このようなワークフローの実現に向けて、1. 使いやすいデザインインターフェースを実現し、これらの材の需要を高める 2. これらのデザインと実際の林業から発生する枝をリンクし、生産工程に落とし込む包括的なワークフローの2点に着目し開発を進めていく。

適材適所が実現されると小径木にも需要が生まれる。需要があると落葉広葉樹林に恒常的に手が入り、本来の森林の姿である適地適木にもつながる。最終目標として持続可能な循環型社会の実現に貢献したい。

#### 2. 研究成果

##### (1) 概要

加速フェーズ期間ではプロダクトスケールと建築スケールの二つのスケールに意識して研究を遂行した。プロダクトスケールでは、以前のAct-I期間では決め打ちだった接合部のデザインを多様化するようなデザイン支援ツール「Tsugite」を開発した。家具スケールでは人間の距離が近く、より部材の接合部にディテールデザインが求められデザインの差別化要素となる。建築スケールでは前回のAct-I期間の後半および加速フェーズ前半で開発した三次元枝材の適材適所アルゴリズム、および3次元加工支援システムを用いて科学技術未来館の常設展示インスタレーション「Swirled Branches」を製作・提供した。また「Tsugiki」プロジェクトでは、これまでの直接部材にジョイント加工するのではなく、接合部の型を三次元出力し、レジンでキャストす

る方向で進めている。

(以下コロナ延期後の進捗)

対外発表として論文執筆だけでなく、コロナの影響で延期になったが LA の現代美術ギャラリー「Baert」にて 2022 年 2 月の展示に向けて、アーティスト Daniel Silva と準備を進めている。また今年 2 月には Tsugiki の対外発表として Wood Design Award で展示し、Tsugite および Tsugiki を用いたテーブルの設計をよび製作をデザイナーの甲斐氏と進めている。

## (2) 詳細

### テーマ 1. プロダクトスケールでのデザイン支援ツールの開発および検証

釘を使わない接合部である木工継手・仕口を新たに設計して製作することは、熟練者でないユーザにとって困難で時間がかかる作業である。「Tsugite」システムは、切削加工機による加工と、計算機を用いたインタラクティブな形状モデリングを組み合わせることによって、カスタムデザインの木工継手・仕口の設計と製作を支援するものである。本システムには、マニュアル編集モードと、ギャラリーモードの2つのモードがある。マニュアルモードでは、リアルタイムに更新される接合部の性能に関する解析結果やシステムからの提案を見ながら、ユーザが手作業で接合部の形状を編集する。性能としては、組み立て可能性、加工可能性、強度などを考慮している。ギャラリーモードでは、あらかじめ計算済みの接合部形状が多数画面に提示され、ユーザはその中から好みのものを選択する。接合部の設計が完成したら、角を丸めるなどの処理をして、3軸CNCフライス盤で製造する。本研究ではユーザスタディによる評価の結果や、実際に設計し製作した接合部や家具の例を示した。



図1 左: 開発したシステムを利用して設計・製作した接合部の例。右: 開発したシステムを利用して設計・製作した椅子の例。釘や接着材を使わずに、組み立て、分解、再組立てが可能である。

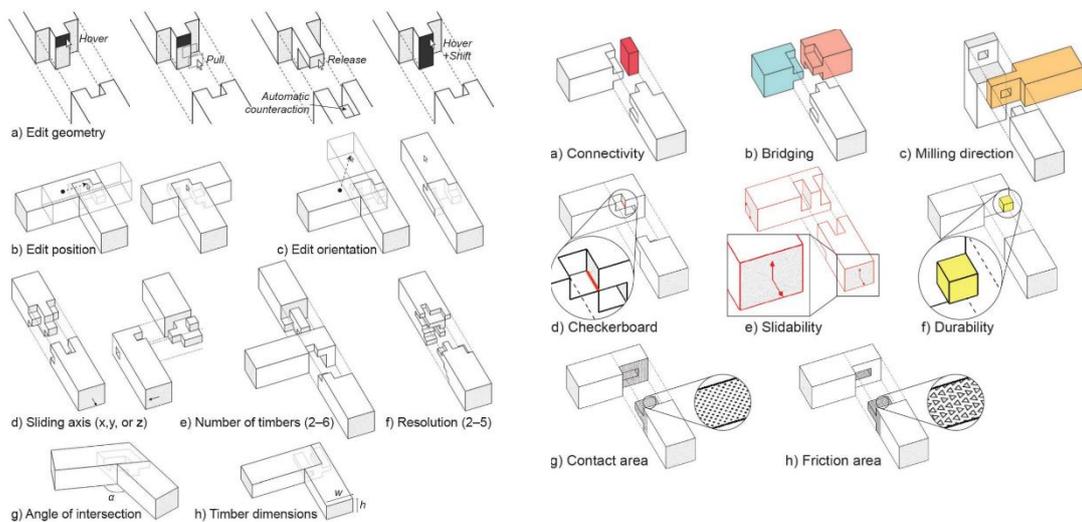


図2 左:ユーザーによる操作の例、右: システムからのフィードバックの例

## テーマ2. 建築スケールでの適材適所システムの開発および検証

三次元形状の枝を三次元的に配置し、自動で交差部のジョイントを生成し加工するシステムを開発した。テーマ1と違いジョイント部の基本的な設計は固定でパラメータを変化させるのみである。三次元的にジョイントを加工するため、従来の2.5次元CNC切削機では加工が困難である。そこでユーザーが加工ベッドに枝を3次元的に配置する際に正確に三次元形状の枝を配置をサポートする音声ガイダンスシステムを開発し、実際に使用することで科学技術未来館のインスタレーション「Swirled Branches」を製作した。

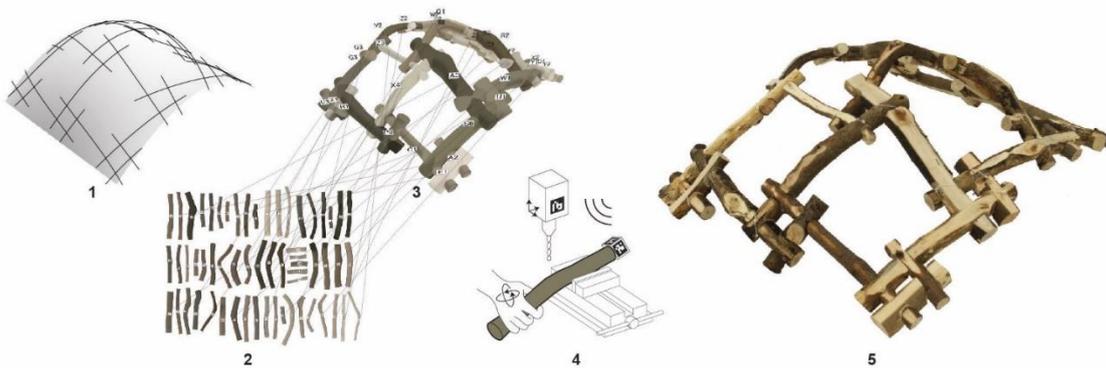


図3 設計および製作のワークフロー。自由曲面を入力に、あらかじめ3次元スキャンしておいた枝群を自動で適切に配置し、加工はユーザーに配置を指示することで2.5次元加工機でも3次元加工が可能となる。



図 4 開発したシステムを利用した例。科学技術未来館に展示している「Swirled Branches」のデザイン(左)と実際に組上がった構造(右)。

### 3. 今後の展開

「Tsugite」システムは Github でも公開されており、すでに多数の問い合わせがある。これらの中には CAD のベンダー等もあり、既存のシステムに組み込むことも可能である。

#### ※コロナ延長後の記載

Wei 氏と共に Github のレポジトリを整理しており、より再利用できる形で公開準備中である。また適材適所システムにおいては、データ駆動型手法を取り入れることで、より多様な形態に対応できつつある。

### 4. 自己評価

#### ・研究目的の達成状況

当初提案した研究テーマ1は達成され、さらに研究テーマ2も現在開発中の Tsugiki は型を使用したシームレスな接合を目指しており、このような研究は他に類を見ない。

#### ・研究の進め方(研究実施体制及び研究費執行状況)

国内外での発表を通じて研究体制は整いつつある。研究費は主に加工機械や材料費および場所代として執行され、当初計画していた通り適切に処理されたと考えられる。

#### ・研究成果の科学技術及び学術・産業・社会・文化への波及効果

海外ではリサイクルよりもアップサイクルという考え方が広まりつつある。リサイクルが材料を粉々にするのにに対し、アップサイクルは材料をできるだけそのままの形で別の用途に用いる。ACM で Upcycle を検索してもヒットする論文数は限られるが、本研究の考え方は今後 HCI および情報学が持続可能な社会の実現に貢献する際に、重要かつ先駆的な研究例として挙

げられると期待できる。また文化的にも本研究のテーマである、適材適所はそもそも日本の伝統構法から発想を得ており、日本の伝統文化の海外への発信という意味でも期待できる。参加型設計製作と合わせて、SNS のような情報と人をつなげる手法に実際の素材や経験を介して情報、人、モノをつなげるようなサービスに発展することが期待できる。

## 5. 主な研究成果リスト

### (1) 代表的な論文(原著論文)発表

- |  |
|--|
| 1. <b>Human-in-the-loop fabrication of 3D surfaces with natural tree branches.</b> <i>Maria Larsson, Hironori Yoshida, and Takeo Igarashi.</i> SCF '19: Proceedings of the ACM Symposium on Computational Fabrication (June 2019).                       |
| 2. <b>Tsugite: Interactive Design and Fabrication of Wood Joints.</b> <i>Maria Larsson, Hironori Yoshida, Nobuyuki Umetani, and Takeo Igarashi.</i> UIST '20: Proceedings of the ACM Symposium on User Interface Software and Technology (October 2020). |
|  |

### (2) 特許出願

研究期間累積件数: 1 件

1.

発 明 者: Larsson Maria, Hironori Yoshida, Nobuyuki Umetani, Takeo Igarashi

発明の名称: Information Processing System, Information Processing Method And Program

出 願 人: 東京大学

出 願 日: 2020/09/30

出 願 番 号: 63,085,401

### (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)

Wood change award: <https://awrd.com/award/woodchangeaward>

東京大学プレスリリース <https://www.u-tokyo.ac.jp/content/400148405.pdf>