

令和 2 年 6 月 29 日現在

機関番号：33908

研究種目：基盤研究(C) (一般)

研究期間：2017～2019

課題番号：17K06471

研究課題名(和文) センサ情報からの熟練技能の自動抽出とロボットへの転移手法に関する研究

研究課題名(英文) Research on a method to transfer high-level skills extracted from sensing information to robots

研究代表者

橋本 学 (Hashimoto, Manabu)

中京大学・工学部・教授

研究者番号：70510832

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：本研究では、熟練作業者が持っている「技」を自動抽出し、これをロボットシステムに転移するための技術に関する検討をおこなった。研究成果は次の2つである。(1)作業対象物としての3次元物体と手の動作の両方を扱える作業センシング・記述・分析手法を提案した。具体的には、視線情報と作業時間に着目した分析法、視線と手の動きの連動性に着目した分析法、および視線と手の時系列情報に着目した分析法の3つを提案した。(2)多様な物体を認識しロボット把持パラメータを決定するための技術として、曲率ベースの局所特徴量、プリミティブ近似を利用したモデルレス物体認識手法および概形モデル利用物体認識手法を提案した。

研究成果の学術的意義や社会的意義

本研究の成果は、熟練作業者の「技」という、従来は暗黙知とされてきた情報を、視線や手の動きという物理計測可能な情報の分析によってデータ化し、熟練者の技が具体的にどのように数値的に表現できるかを示した。この点に最も大きな学術的意義がある。また、年々少子化が進み、労働力人口の減少を余儀なくされている日本の社会において、人に代わって高度な製品を製造するロボットの実現に向けた有用な情報を提供できた点に、社会的な大きな意義があると考えている。

研究成果の概要(英文)：In this research, we have studied a method that can extract "skill" of an expert person automatically and a method that can transfer the extracted "skill" to the robot system. There are two major research results as follows. (1)We proposed a technique of sensing, description and analysis that treat both 3D objects and human hand motions. Concretely, a method based on information about view-direction and working-time, a method based on co-occurrence between view-direction and hand-motion, and a method based on spacial temporal information about view-direction and hand-motion. (2)We also proposed a curvature-based local features, a model-less object recognition method by utilizing approximating objects to prepared primitive shapes such as hexahedron, cylinder and sphere, and an object recognition method by using rough shape models.

研究分野：画像情報処理

キーワード：作業動作分析 熟練作業者 視線計測 動作計測 深層学習 LSTM 3次元物体認識 スキル分析

## 様式 C-19、F-19-1、Z-19（共通）

### 1. 研究開始当初の背景

本研究では、生産・流通分野におけるヒト（作業員）の動作とモノ（製造物）の動きを捉える画像認識技術、およびロボットへの「技」の転移を取り扱う。想定する適用先は、工場でのセル生産における製品組み立てや、配送センターにおける商品取り出しである。近年、特にこれらの分野において熟練作業員の不足が深刻な問題となっていることから、「匠の技」の有形資産としての記録や、ロボットによる熟練者級のものづくりの実現が強く期待されている。

### 2. 研究の目的

本研究では、生産工程において熟練作業員が持っている「技」を自動抽出し、これをロボットシステムに転移するための技術を確立する。最終的にはこの技をロボット動作として実現することを目指しつつ、ヒトの手と、把持されているモノの移動軌跡をセンシングして記録し、さらに初心者と熟練者のデータを比較することによって、無駄の無い動きに代表される「熟練の技」を自動的に抽出する方法を提案し、効果を実証する。

### 3. 研究の方法

本研究においてブレークスルーすべきメイン課題は、作業対象物たる 3 次元物体と手の動作を捉える作業センシング・記述・分析手法を提案することである。さらにサブ課題として、作業対象になり得る多様な物体を認識できる 3 次元物体認識技術も必要になる。以上のことから、本研究では、作業動作認識と分析を 3 年間にわたって継続的に実施する中において、実験に必要な要素技術として、物体認識技術の検討についても 1~2 年目に並行実施した。

### 4. 研究成果

当初の計画通り、本研究の成果として、作業員が取り扱う組み立て対象物を認識するための物体認識に関する成果、および本研究の主眼たる作業員動作のセンシング、記述、および分析法に関する成果を得ることができた。以下、(1)にて物体認識に関する成果を、また(2)にて(3)の前提となる実験用の各種データベースの構築について報告し、最後に(3)にて作業動作分析手法の提案と分析結果に関する成果をまとめる。

#### (1) 3次元物体認識技術に関する成果

物体認識に関して、モデル有りモデル無し（モデルレス）の2つの場合について検討した。前者として、曲率ベースの局所特徴量の検討については、対象物に顕著な凹凸が存在しない場合でも適用できる新手法として曲率ベースの特徴量を提案した。具体的には、マルチケール曲率ベースの特徴量を提案し、従来法の組み合わせ手法より高性能であることを確認した。また後者については、本研究の最終ゴールが熟練の技能をロボットに転移させ、ロボットでのハンドリングを想定していることに関連している。すなわちモデルを持たない一般物体にも対応できるロボット把持位置自動決定手法が必要となることから、対象物を直方体等のプリミティブ形状で近似して対象物の把持パラメータを決定する手法を提案するとともに、実験により有用性を実証した。さらに、対象物の形状が準備したプリミティブ形状（直方体、円柱、球形）と異なる際には利用しにくいという問題への対応として、物体の概形のみが既知で寸法が未知という状況に対応しうる仮説検証型の物体認識手法を提案して効果を実証した。

#### (2) 作業動作データベースの構築

本研究で使用することを目的に、作業動作分析用の3種類のデータベースを作成した。それぞれ、キューブ型のおもちゃの組立作業における作業員の視線位置情報が記録されているもの、飛行機型のおもちゃの組立作業における作業員の視線と手の位置情報が記録されているもの、そして、2つ目と同じ対象物に対して視線の計測方法を変更し、視線位置情報が改良されたデータである。視線の計測には Tobii Pro Glasses2（以下、Tobii と記載する。）を、また手の計測には光学式モーションキャプチャを用いた。1つ目と2つ目のデータベースについては、視線位置を Tobii から得られる一人称視点画像中の 2 次元座標として格納した。これに対して3つ目のデータベースは、Tobii とモーションキャプチャを同期させるソフトウェアを用いて、3次元視線位置をモーションキャプチャと同一の座標系にて格納した。3つ目のデータベースのデータフォーマットの例を図1に示す。データベースには、視線と両手の3次元座標が時系列で格納されており、各時刻に対応する作業工程を人手でアノテーションした工程ラベルも格納されている。そのため、各工程における視線と手の動きの分析が可能である。

#### (3) 作業動作分析手法の提案と分析結果

本節では、3種類の作業動作分析手法とその分析結果について記す。1つ目は、作業員の視線と作業時間に着目した分析方法、2つ目は、作業員の視線と手の連動性に着目した分析方法、3つ目は、視線と手の時系列情報に着目した分析方法に関する研究成果である。

	A	B	C	D	E	F	G	H	I	J	
1	右手 x	右手 y	右手 z	左手 x	左手 y	左手 z	視線 x	視線 y	視線 z	工程ラベル	
2	0.744972	0.110486	0.121911	0.671583	0.032241	0.486616	0.621627	0.518812	0.726196	1	工程ラベル: 1: 工程1の動作 2: 工程2の動作 3: 工程3の動作 4: 工程4の動作 5: 工程5の動作 6: 工程6の動作
3	0.745056	0.11854	0.121274	0.671611	0.032333	0.486711	0.632183	0.526669	0.699583	1	
4	0.744084	0.127211	0.120411	0.671648	0.032364	0.486825	0.643605	0.539188	0.662488	1	
5	0.741034	0.135813	0.119598	0.671681	0.032324	0.486901	0.634806	0.533525	0.684257	1	
6	0.735303	0.144263	0.118903	0.671693	0.032204	0.486879	0.628378	0.526701	0.703259	1	
7	0.725914	0.152365	0.118399	0.671693	0.032029	0.486799	0.626405	0.523548	0.710582	1	
8	0.714618	0.159848	0.118871	0.671784	0.031637	0.486602	0.622896	0.522028	0.715683	1	
9	0.701442	0.16639	0.11978	0.671785	0.0315	0.486521	0.620246	0.521886	0.716725	1	
10	0.684734	0.171821	0.120296	0.67177	0.031417	0.486455	0.618954	0.521235	0.717547	1	
11	0.667328	0.174885	0.121004	0.67188	0.031366	0.486312	0.618702	0.523108	0.713513	1	
12	0.648554	0.178092	0.121211	0.671872	0.031279	0.486307	0.615104	0.520242	0.721082	1	
13	0.628639	0.1805	0.121139	0.67174	0.031138	0.486428	0.61443	0.520601	0.720324	1	
14	0.608236	0.182007	0.12055	0.671696	0.030722	0.486191	0.609346	0.516917	0.732276	1	
15	0.588053	0.182225	0.119989	0.671544	0.030561	0.485974	0.604348	0.512482	0.745063	1	

原点からの距離（メートル）で表現

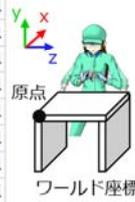


図1. データフォーマットの例

① 視線情報と作業時間に着目した作業動作分析

本手法では、各スキルレベルの作業動作の特徴を分析するために、作業者の視線移動と各作業動作に要する作業時間に着目した。視線情報から視線の移動方向と移動量を表す視線移動パターン数と視線停留の出現率を算出する。また、作業動作の履歴から組立作業を構成する各作業動作の作業時間を算出して、これらの情報を統合的に分析することによって、各スキルレベルの作業動作の特徴を明確にした。

具体的に、一人称視点画像中の x 座標と y 座標の二つの時系列データをもとに視線移動パターン数と停留率を算出するために、図 2(a) のように各フレーム t と前フレーム t-1 から移動量を算出した。次に大小の視線移動と停留を表現するために、負のしきい値を含めた 4 種類のしきい値を設け、移動量を同図 (b) のように 5 段階で量子化した。最後に量子化された x, y 座標の移動量を統合して視線移動ベクトルとしてコード化した。ここで、視線移動コードは全部で 25 種類あり、これらを用いて視線移動パターン数と停留率を表現している。

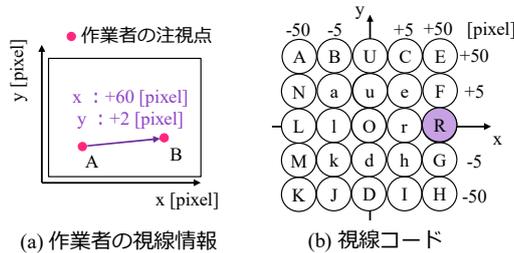


図2. 視線情報のコード化

提案手法による実験結果を分析した結果、以下のことがわかった。

1. 初級者が中級者のスキルを習得すると、作業時間が大幅に減少する。
2. 中級者が熟練者のスキルを習得すると、視線の動きが大幅に変化する。
3. 初級者が熟練者のスキルを獲得する過程で、習熟の要因が異なる。

② 視線と手の動きの連動性に着目した作業動作分析

次に、各スキルレベルの作業動作の特徴を分析するために、作業者の視線と手の動きの連動性に着目した手法を提案した。作業者の視線と手の位置情報を量子化し、視線、利き手、非利き手のコード列を生成する。この手法では、視線と利き手、視線と非利き手のコードを統合して扱うことによって、視線と手の連動性を表現できる。

視線と手の位置情報のコード化については、図 3(a) のようにそれぞれの位置を 2 次元で表現し、同図 (b) のように 18 個の領域に量子化したうえで、各座標が位置する領域のコードをエンコードする。作業動作分析のための視線と手の連動性を表現する視線・動作統合特徴量の抽出としては、同フレームにおける視線と利き手、視線と非利き手のコードのペアを全フレームに渡って算出することとした。図 4 のように視線と利き手および視線と非利き手のコードのペアの頻度を表した 2 種類の共起ヒストグラムを生成し、これらを視線・動作統合特徴量と呼ぶ。本研究では、この特徴量を用いて作業動作の特徴を分析した。

提案手法による各スキルレベルにおける作業動作を分析したところ、図 5 に示す結果を得ることができた（作業者は右利きであった）。初級者と中級者の作業動作の特徴としては、両者ともに 5 種類以上の利き手コードが生じたことから、両者ともに部品を把持する動作において

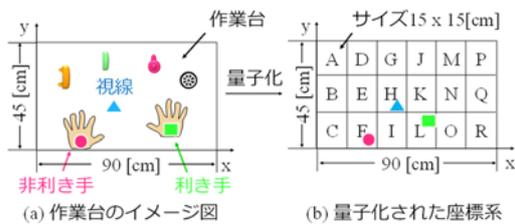


図3. 視線と手の位置情報のコード化の流れ

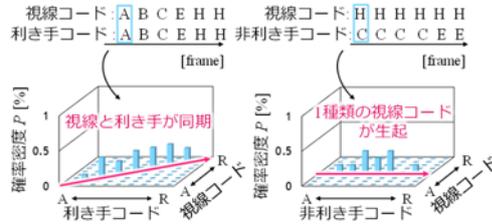


図4. 視線・動作統合特徴量の抽出方法

は利き手を多くの領域に動かす傾向にある。また、非利き手コードについては、初級者は、2種類のみコードが生じたことから部品を把持する動作において、非利き手を動かさず利き手だけを利用して作業する傾向にあるが、中級者は、5種類のコードが生じたことから、5つの領域に非利き手を動かしており、両手を効率的に利用して作業できるようになることが判明した。また、視線コードについては、両者ともに7種類以上のコードが生じたことから、部品を把持する際に両者の視線が部品の位置まで動き、部品の位置を目視で確認してから部品を把持する傾向にあると言える。次に、中級者と熟練者の作業動作の特徴としては、利き手および非利き手コードについては、両者ともに5種類以上のコードが生じたことから、両手を効率的に利用して作業する傾向にある。また、視線コードについては、中級者は、さまざまな領域に視線が移動したことにに対して、熟練者は視線コードが4種類のみ生じたことから、熟練者のスキルを習得すると視線からの情報を利用せずに作業台の中心に視線を停留させて作業する傾向にあることを確認した。

以上を含む一連の分析結果から、以下のことが判明した。

1. 初級者と中級者間では、非利き手の動きが特徴的に異なる傾向にある。
2. 中級者と熟練者間では、視線の動きが特徴的に異なる傾向にある。

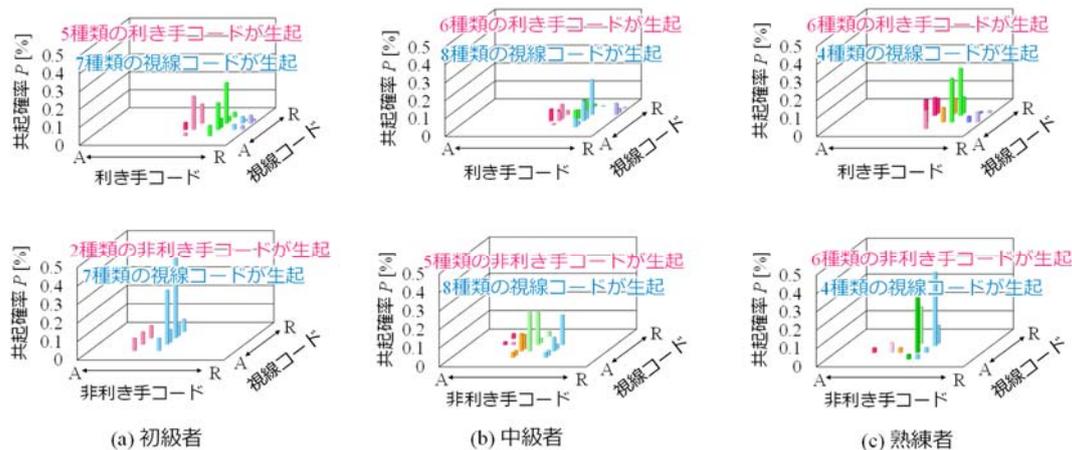


図5. 部品を把持する動作における視線・動作統合特徴量

### ③ 視線と手の時系列情報に着目した作業動作分析

さらに、各スキルレベルの作業動作分析のために利用した情報として、作業者の視線と手の時系列情報に着目した。ここでは「初級者か中級者」の2クラス、および「中級者か熟練者」の2クラスのスキルレベルを識別する機械学習ネットワークを生成する。これがもしスキルレベルを識別できた場合には、そのネットワークはスキルレベル間の作業動作の特徴を捉えていると仮定し、ネットワークのパラメータ等を解析することとした。利用した機械学習ネットワークの構造を図6に示す。視線と両手の3次元の位置情報を時間ごとに2つのストリームのLSTM (Multi-Stream LSTM) に入力する。3×N [frame] の膨大な時系列の入力データから表現能力の高い特徴を抽出するために、中間層の次元を増量させながら3層のLSTMを重ね、LSTMから抽出された時系列の特徴量から全結合層を用いて2クラスのスキルレベルを識別した。

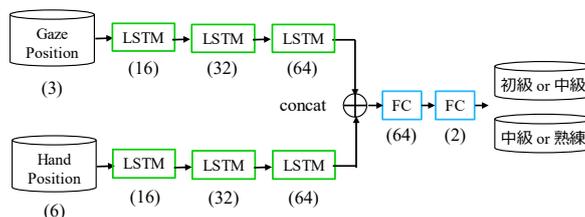


図6. Multi-Stream LSTMのネットワーク構造

ネットワークの解析方法にはアテンション機構を用いた。時系列情報を伝搬して識別をおこなう LSTM において、どの時刻の入力データが識別に有効であったかを 0 から 1 の実数値で表現する。例えば、N フレームの入力データが LSTM に入力された場合、N 個のアテンションが算出され、アテンションが 1 に近ければその時刻の入力データが識別に有効であったと判断できる。

この分析方法を用いて評価実験をおこなった。本実験の対象動作は視線・両手位置時系列 DB (連動) 中の部品を把持し、その部品を組み付けてねじを締めるという一連の作業動作である。その結果、スキルレベルの識別性能は、初級・中級識別では F 値が 0.75、中級・熟練識別では F 値が 0.72 であった。次に、アテンション機構を用いたネットワーク解析について記述する。図 7 に、初級者と中級者の識別テストにおいて、初級者のデータを入力とした時のアテンションを示す。視線と手のアテンションの共通点として、最終フレーム付近のアテンションが高くなっている。これは、LSTM の性質として時系列情報を考慮するものの、より近い時刻の情報の寄与率が高くなることから、最終フレーム付近のアテンションが高くなったと考えられる。次に相違点として、手の初期フレームのアテンションが高くなっていることが確認できるが、初級者と中級者の動作を確認したところ、初級者は利き手だけを利用して作業しているのに対して、中級者は両手を利用して作業する傾向にあることが確認されたことから、これが原因となっており、手のアテンションが高くなったと考えられる。図 8 に、中級者と熟練者の識別テストにおいて、中級者のデータを入力とした時のアテンションを示す。視線と手のアテンションの共通点として、最終フレーム付近のアテンションが高くなる傾向にあった。この理由は、前述と同様であると考えられる。視線と手のアテンションの相違点としては、視線の初期フレームのアテンションが高くなっている。中級者と熟練者が部品を把持しに行く動作を確認したところ、中級者は部品の位置まで視線を動かしているのに対して、熟練者は作業台の中心位置で視線を停留させることが確認された。また両者の手の動きを確認したところ、部品の位置に近い手を使用していた。このことから、部品を把持しに行く動作における視線の動きが特徴的に異なるため、視線の初期フレームのアテンションが高くなったと考えられる。

以上の一連の実験および分析結果から以下のことが判明した。

1. スキルレベル識別性能としては、初級・中級識別では F 値が 0.75、中級・熟練識別では F 値が 0.72 であった。
2. アテンションの分析より、部品を把持しに行く動作において、初級者と中級者では、手の動きが異なる傾向にある。
3. アテンションの分析より、部品を把持しに行く動作において、中級者と熟練者では、視線の動きが異なる傾向にある。
4. 各スキルレベルの作業動作特徴を時系列情報に基づいて分析できる。

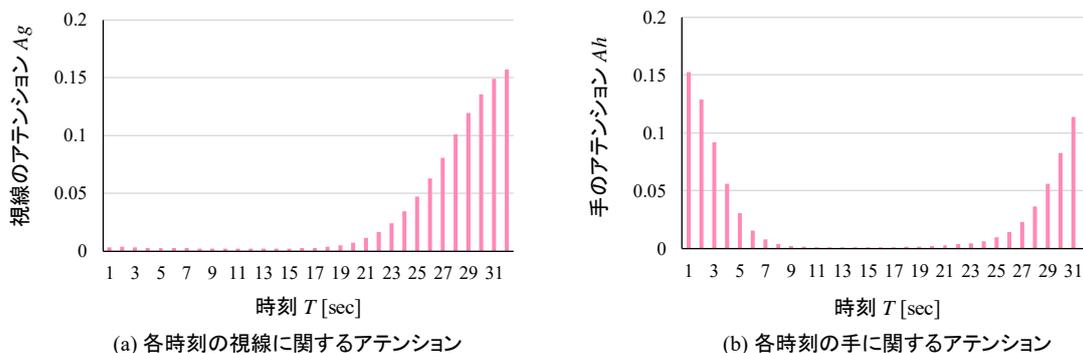


図7. 初級・中級識別時のアテンション

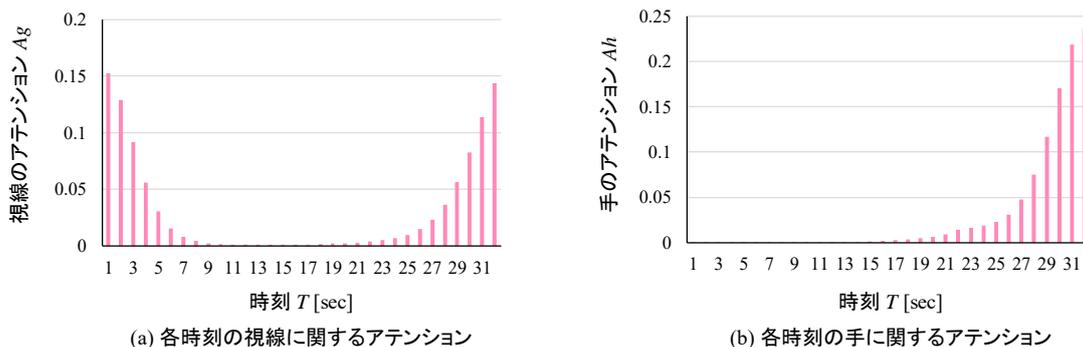


図8. 中級・熟練識別時のアテンション

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計16件（うち招待講演 0件 / うち国際学会 3件）

1. 発表者名 川瀬陽平, 三好遼, 橋本学
2. 発表標題 作業者の習熟要因分析のためのMulti-Stream LSTMを用いた熟練度合い推定
3. 学会等名 動的画像処理実利用化ワークショップ (DIA2020)
4. 発表年 2020年

1. 発表者名 Yohei Kawase, Manabu Hashimoto
2. 発表標題 Analysis of Skill Improvement Process Based on Movement of Gaze and Hand in Assembly Task
3. 学会等名 The 18th International Conference on Computer Analysis of Images and Patterns (CAIP2019), vol 11679. Springer, pp.15-26 (国際学会)
4. 発表年 2019年

1. 発表者名 Yohei Kawase, Koichi Taguchi, Manabu Hashimoto
2. 発表標題 Integrated Analysis of Position of Gaze/Hand for Skill-up Process Analysis of Assembly Tasks
3. 学会等名 International Workshop on Advanced Image Technology (国際学会)
4. 発表年 2019年

1. 発表者名 鈴木貴大, 橋本学, 大西剛史, 藤吉弘巨
2. 発表標題 部分的なサイズ変動を許容できる多面体の位置姿勢認識手法の提案
3. 学会等名 動的画像処理実利用化ワークショップ (DIA2019)
4. 発表年 2019年

1. 発表者名 永野史也, 鳥居拓耶, 橋本学, 大西剛史, 藤吉弘巨
2. 発表標題 面情報を手がかりとした仮説検証による複数サンドイッチの同時認識
3. 学会等名 動的画像処理実利用化ワークショップ (DIA2019)
4. 発表年 2019年

1. 発表者名 Takuya Torii, Manabu Hashimoto
2. 発表標題 Reliable Primitive Approximation for Estimation of Robot Grasping Parameters Using 3D-Deep Neural Network
3. 学会等名 IS&T International Symposium on Electronic Imaging 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 川瀬陽平, 田口皓一, 橋本学
2. 発表標題 組立作業における視線と身体動作の連動性分析のための視線・動作に関する統合特徴量の提案
3. 学会等名 電子情報通信学会バイオメトリクス研究会 (BioX2018)
4. 発表年 2018年

1. 発表者名 川瀬陽平, 田口皓一, 橋本学
2. 発表標題 組立作業のスキル獲得プロセス分析のための視線・動作統合特徴量の提案
3. 学会等名 第24回画像センシングシンポジウム (SSI12018)
4. 発表年 2018年

1. 発表者名 川瀬陽平, 田口皓一, 橋本学
2. 発表標題 組立作業の習熟プロセス分析のための視線と身体動作の統合的分析手法
3. 学会等名 精密工学会IAIPサマーセミナー2018 (SS2018)
4. 発表年 2018年

1. 発表者名 川瀬陽平, 田口皓一, 橋本学
2. 発表標題 組立作業の習熟プロセス分析のための視線 / 動作のコード化特徴量
3. 学会等名 動的画像処理実利用化ワークショップ (DIA2018)
4. 発表年 2018年

1. 発表者名 川瀬陽平, 田口皓一, 橋本学
2. 発表標題 組立作業の習熟プロセス分析のためのコード化視線特徴量
3. 学会等名 電気学会知覚情報研究会
4. 発表年 2018年

1. 発表者名 鳥居拓耶, 藤田祐真, 橋本学
2. 発表標題 3次元物体認識のための特徴点検出・LRF生成・特徴量記述の最適組合せの自動決定法
3. 学会等名 動的画像処理実利用化ワークショップ (DIA2018)
4. 発表年 2018年

1. 発表者名 鳥居拓耶, 秋月秀一, 橋本学
2. 発表標題 DNNの尤度統合によるロボットの把持位置推定のための対象物のプリミティブ近似
3. 学会等名 第20回画像の認識・理解シンポジウム (MIRU2017)
4. 発表年 2017年

1. 発表者名 鳥居拓耶, 橋本学
2. 発表標題 3D-DNNを用いたロボットの把持位置推定のための未知物体のプリミティブ近似
3. 学会等名 精密工学会IAIPサマーセミナー2017
4. 発表年 2017年

1. 発表者名 鳥居拓耶, 橋本学
2. 発表標題 3D-DNNによるロボットピッキングのための高信頼なプリミティブ物体認識
3. 学会等名 平成29年度電気・電子・情報関係学会東海支部連合大会
4. 発表年 2017年

1. 発表者名 鳥居拓耶, 橋本学
2. 発表標題 ロボットピッキングにおけるモデルレス把持パラメータ推定のための高信頼プリミティブ物体近似
3. 学会等名 第18回計測自動制御学会システムインテグレーション部門講演会 (SI2017)
4. 発表年 2017年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

<p>【受賞】 BioX研究会奨励賞，電子情報通信学会 バイオメトリクス研究専門委員会，2018/11/21 .</p> <p>【受賞】 Best Paper Award, IWAIT2019, 2019/1/9 .</p> <p>【研究成果公開ホームページ】 組立作業における作業者の視線と手の動きに着目した習熟プロセスの分析 <a href="http://isl.sist.chukyo-u.ac.jp/projects/#H4-1">http://isl.sist.chukyo-u.ac.jp/projects/#H4-1</a></p>
--

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	秋月 秀一  (AKIZUKI SHUICHI)  (40796182)	慶應義塾大学・理工学部(矢上)・助教    (32612)	本研究分担者の所属は，2019年度は，中京大学工学部に変更されています。(2017～2018年度は慶應義塾大学)