

令和 2 年 7 月 2 日現在

機関番号：34315

研究種目：基盤研究(C)（一般）

研究期間：2017～2019

課題番号：17K00372

研究課題名（和文）センサ配置の変化など、環境要因の変化にロバストな、プロセスにおける行動推定・評価

研究課題名（英文）A Robust Behavior Estimation and Evaluation Method against Sensor Position Variation and Environmental Changes

研究代表者

李 周浩（Lee, Jooho）

立命館大学・情報理工学部・教授

研究者番号：80366434

交付決定額（研究期間全体）：（直接経費） 3,400,000円

研究成果の概要（和文）：本研究では一連の作業をプロセスと定義し、一定のプロセスを持つ作業に対してカメラを用いてある人物を撮影し、その人物が行っている行動がプロセスのどの部分（フェーズ）であるかをロバストに推定するものである。最初に模範となるプロセスを記録したときのカメラの位置と異なる位置にカメラを設置しても推定精度を落とさずにプロセス推定が行えるように世界トップ水準の高精度な行動分類器を実現した。これにより異なる環境で記録した別人物の同行動を比較することが可能となる。

研究成果の学術的意義や社会的意義

本研究の成果によって、調理法、組み立て作業など、決まった一連の作業を行う先生役の人の動作をカメラで収めることによって、異なる空間で異なる人がその動作をまねて行った場合どの動作が作業全体におけるどの部分であるかを判別できると同時に2者間の動作の違いが比較できるようになる。これまでの手法はカメラを用いた場合できるだけ同じ位置にカメラを配置しないと再現性が低かったが本研究はその問題を概ね解決した。また本研究成果を活用することで、遠隔教育、ビデオ教育などがもっと容易にできるようになる。

研究成果の概要（英文）：In this research, a series of work is defined as a process. A camera is used to record a work having a certain process, and it is estimated robustly that what part of the process is the action that the person is performing. We have realized a world-class high-precision action classifier so that process estimation can be performed without lowering the estimation accuracy even if the camera is installed at a position different from the position of the camera when the model process was first recorded. This makes it possible to compare the same actions of different persons recorded in different environments.

研究分野：空間知能化、ロボティクス、機械学習

キーワード：フェーズ推定 行動クラスタリング ロバスト推定

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

人の行動を正しく推定する研究は様々な応用が期待される分野である。例えば、ロボットへのやさしい教示手段、無形伝統文化スキルのアーカイブ、手術のフェーズ自動推定など多様な応用先がある。特に手術のフェーズの自動推定は、医療従事者の不足、不十分な医療訓練時間、熟練医の技能伝授困難など様々な問題の改善に繋がると期待されており、日本と欧州がこの分野の研究をリードしているとも言える。Stauder ら [IPCAI, 2014] は、各手術器具に固定したセンサから取得したデータに基づいて手術工程を分析する方法を提案し、約 65% 正解率が得られている。また、その他の多くの研究では、医師の体または手術道具にデバイスを取り付けて手術のフェーズ推定を行っている。(用語の定義に関しては Fig.1 をご参考)

研究代表者は、古くからセンサと処理機能を分散配置して空間内の事象を把握して適切な支援を行う新しい空間概念として知能化空間(Intelligent Space)を提案してきており、蓄積された技術を基に 2013 年から本提案の基礎となる腹腔鏡下手術のフェーズ自動推定研究を行ってきた。開発したフェーズ自動推定手法を Fig.2 をもって簡単に紹介する。(A)空間内に設置された 3 台の画像センサから取得した連続画像データを(B)前処理を施してノイズを除去した後、(C)Optical Flow を求めて一定時間の 3 次元空間上の動きを 2 次元平面上に射影し、Bag-of-Words を作る。(D)Latent Dirichlet Allocation モデルを用いて次元削減を行い、(E)その結果に対して正規化を行う。(F)最後に、Hidden Markov Model を用いて基準となるモデルデータ(学習データ)を学習し、新たな観測データが入力されると自動的にフェーズ推定が行われる。本手法を用いた実験の結果、胆嚢摘出手術の全 12 フェーズ(CO2 Inflation, Trocar Insertion, Dissection など)に対して推定正解率、約 84% を記録した。この手法の主な特徴は、医師や道具にマーカやデバイスを取り付けなため手術に一切の邪魔をしない点である。また、画像データから手、顔、または、道具を認識してそれらの動きをトラッキングするのではなく、対象を明かさず動きだけを捉えて処理を行っているため人物や衣装などが変化しても比較的安定して高精度の結果が出せる。

研究代表者は、以上の背景と成果を踏まえて、手術のような特定のプロセスだけではない、より一般的なプロセスにおけるフェーズ推定の実現と模範モデルを基準として現作業者のプロセスを定量的に評価できる手法の実現を目指した。さらに、以上のような目標を達成するために、センサの配置や様々な環境要因が変化しても影響を受けにくい行動推定の仕組み解明に挑んだ。

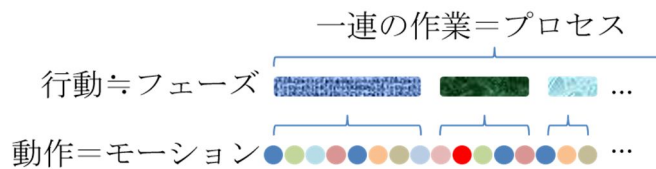


Fig.1 用語の整理

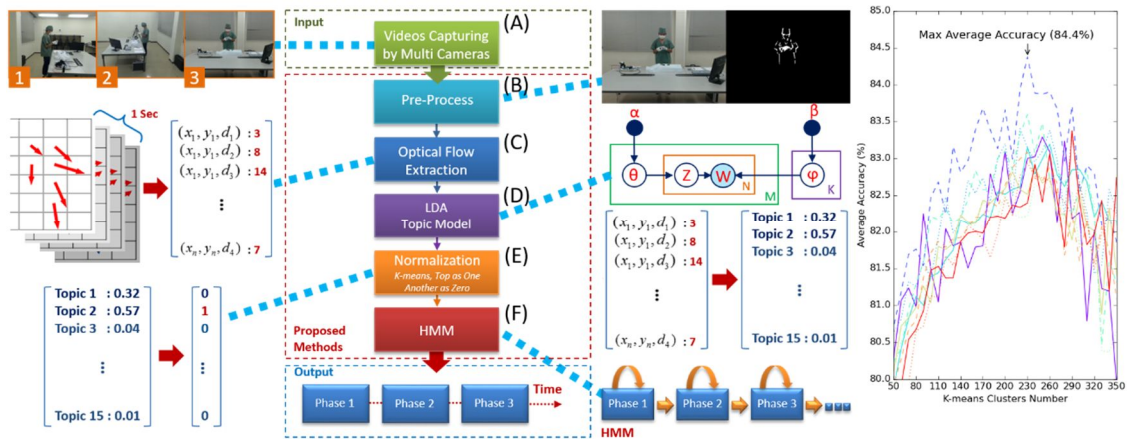


Fig.2 これまで申請者が開発してきた手術フェーズ推定手法の概要図(左)と推定正解率(右)

2. 研究の目的

人の連続動作で構成される一連の作業(プロセス)においてフェーズ(順序のある行動)をロバストに推定・評価できるようにするための要因と条件の解明を行う。研究代表者はこれまで知能化空間技術と画像センサによる手術のフェーズ推定技術を研究してきた。本研究では、これまでの研究成果を発展させ、基準モデルの作成時と観測対象の推定時のセンサ配置、照明条件、ノイズなど環境要因が変化してもシステムの精度に大きな影響が出ない、ロバストな行動推定・評価手法の実現を目指す。行動推定手法の一連の流れをモジュール別に分けて、各々を研究分析し様々な環境変化に影響を受けにくい行動推定・評価の仕組みを解明することを本研究の目的とする。

3. 研究の方法

本研究は、画像センサを用いて手術のフェーズ推定を行うアルゴリズムを基に、センサの配置の変化などの環境要因の変化にロバストな行動推定・評価アルゴリズムにさせるための必要要因を解明すると同時に特定環境ではなく一般環境で使用可能にすることを目標としている。本研究では、一連の作業を行う際に現在がどのフェーズに当てはまるかを自動推定する手法を拡張した。

標準の計測空間を製作して、常にその施設を用いてプロセスにおけるフェーズ推定と定量評価を行うのであれば問題の難易度は低くなるが、非現実的である。従って、計測装置を実環境に持って行って推定・評価作業を行うのが現実的であるが、各々の空間の形状は様々であるため同じ位置と角度にセンサを配置して対象を捉えるのは困難である。そのため、センサの配置が異なっても精度が落ちない行動推定および定量評価手法が必要である。また、特集装置を使用したり、複雑な装置を使用したりすると実用性がなくなるため安価なセンサで簡単に使える手法にしなければいけない。このような要求条件を満たすために行う研究内容の詳細を以下にまとめる。

(1) センサの配置およびその他の環境要因に影響を受けにくい動作データの変換および処理

本提案のベースとなるフェーズ推定手法は、研究代表者が2013年に提案してから、持続的に研究を行い性能と完成度を高めた手法に発展させた。しかし、手術のようなプロセスに限らず、料理、リハビリテーション、製品の組み立てなど手順をもつ作業に幅広く応用するためにはモデルデータと異なる環境で計測されたデータであっても推定精度が高くなければならない。この問題を解決するため Fig.3 のように、同じプロセスに対してセンサの位置を変えたり、センサの数を変えたりするなど条件を変えながら研究用のデータセットを作成し、環境要因の変化が結果に及ぼす影響を調査し、その影響が最小限になるようにデータ変換と処理内容を変えて比較検討を行った。過去の研究では、データ変換として入力画像内の動きを画像センサのイメージフレーム上の2次元の動きとして射影しているが、どこまでの環境変化であれば許容された精度範囲に留められるかを定量的に調べると同時に、大雑把な画像センサ間の相対姿勢と時間同期を用いた、環境変化の影響を受けにくい方法を生み出した。

(2) プロセスのフェーズ推定の一般化

本研究では、様々なプロセスに応用可能なフェーズ推定を目指した。従って、医療、介護、製造などの様々な分野のフェーズ推定の要求精度を調べると同時に、環境要因の変動範囲について調べた。さらに、様々な空間に適用可能にするために与えられた空間における限定的最適センサ配置を試みた。

(3) プロセスの階層的な定量評価

あるフェーズを構成する動作は必ずしも一通りではない。たとえば、はさみで切ってテープを貼る動作をモデルフェーズでは3回行ったとして、推定対象のフェーズでは、その動作を4回繰り返しても同一フェーズである。現在の手法では、はさみで切ってテープを貼るという動作を認識しているわけではなく、空間上の動きとして対象を捉えているためこのようなものにも対応している。さらに、モデルデータに対する評価を行う際にはこのような違いに対して、どこに違いがあるかを探して、その違いの程度を数値化する。また、フェーズ間の違いだけでなく、プロセス全体としての評価も行い、部分と全体に対する階層的な定量評価の実現を試みた。

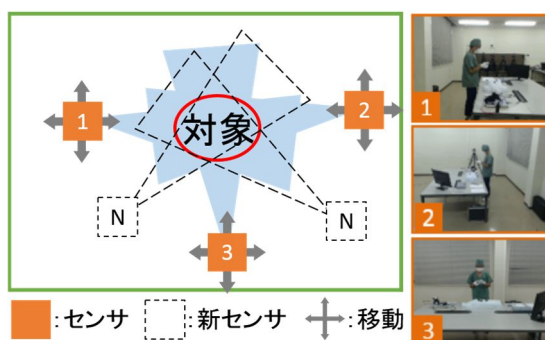


Fig.3 センサの移動と新しいセンサの追加

4. 研究成果

隠れマルコフモデル (HMM)、主成分分析 (PCA)、潜在ディリクレ配分 (LDA) を組み合わせ、指定されたワークフローのフェーズ (タスク) を自動的にセグメント化する2つの方法を提案した。2.研究の目的で述べた通り、手術室での外科ワークフローなど、ワークフローの特定の時点ごとに適切なフェーズラベルを出力することが目標である。提案方法では、一連のフェーズで構成されるすべてのワークフローの特定時点で適切なフェーズラベルを推定できる。この作業の基本的な仕組みは、PCA 圧縮ヒストグラムまたは人間、機器、材料などの一般的な作業コンテキストのオプティカルフローからの特徴をまとめる LDA トピックモデルによって取得された観測値に基づいて HMM を構築することである。このような作業コンテキストの認識は、ワ

ワークフローをキャプチャするために複数の簡易的に同期されたカメラを使用することによって達成される。さらに、メソッドの堅牢性は、各ワークフローの平均長が 12.8 分の最大 12 フェーズを含むデータセットを制作して実験を行い検証した。これらの結果に基づいて、The 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI 2017)にて成果報告を行った。さらに、信頼性を高めるために、部分的なオクルージョン、雑然とした背景、動的な環境、照明の影響、対象人物の外観の影響などによる感度低下を防ぐ方法について IEEE / SICE International Symposium on System Integration (SII 2019) で発表を行った。

特にカメラの視点の変動の問題を適応的に処理できる 2 つのアプローチも提案した。フェーズ推定の実験を行う前に、アクション認識で評価した。最初のアプローチでは、ORB (oriented FAST and rotated BRIEF)のバイナリ特徴とオプティカルフロー (OF) モーション特徴の組み合わせである、結合された ORB-OF 特徴を提案した。これは、ORB バイナリ記述子ベクトルをクラスに編成および正規化するバイナリベクトルクラスタリングアルゴリズムを考案し実現した (Fig.4)。提案手法は、データセットを使用して実験を行いその性能が検証できた。

2 番目のアプローチでは、単一視点から一連の異なる視点を生成し、複数のスケールでのアフィン変換に基づく空間サンプリングを行うアルゴリズムを提案した (Fig.5)。高密度オプティカルフローのヒストグラムが、生成された特定視点の固定サイズパッチにわたってローカルフィーチャ記述子として抽出される。特徴ベクトル空間を使用して、特徴的な情報を持つ次元のみを保持し、不要な次元を破棄する次元選択手法も提案して精度向上を図った。CG によるデータセットと現実データセットの両方を使用して実験を行い、提案手法の有効性を検証した。提案手法は、幅広い視点変化において精度を維持できることが確認できた。さらに、本提案手法は、対象のスケール、対象の位置、動作速度、部分的なオクルージョン、およびバックグラウンドの変動に対する影響が少ない。評価実験で、提案手法が広い範囲の視点変化でも精度を維持できることが確認でき、不変範囲は、シミュレーションデータセットでは約 65 度、現実的なケースでは 45 度であった。i3DPost というパブリックデータセットを使用して、提案手法とこれまで公開された他研究グループ手法との比較を行った。6 つのクラスのサブセットで、提案手法の精度 99.74% であり、これまで最も高精度を出した Angelini ら [IEEE SigPort, 2018]の手法とよりも高い結果であった。この成果を Applied Intelligence、Springer、1468-1486、2020 に掲載した。

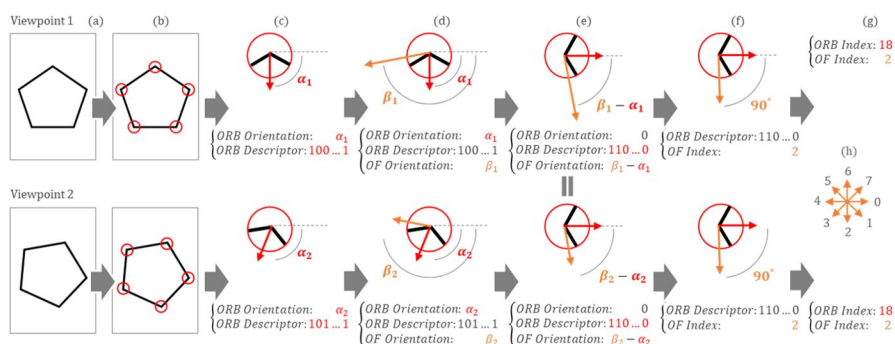


Fig.4 - Overview of the proposed action recognition method based on ORB-OF feature.

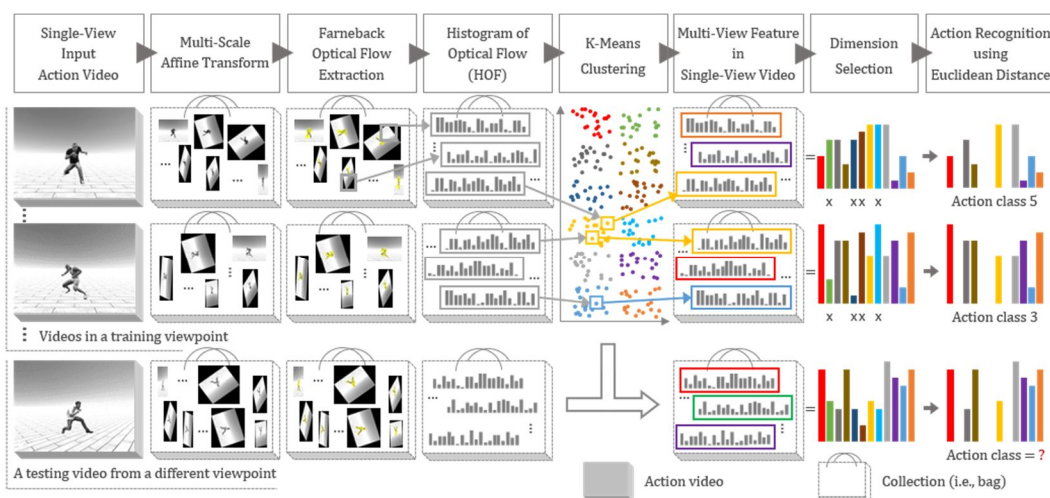


Fig.5 - Overview of the proposed action recognition method for both training and testing phases (different camera viewpoint for each phase).

5. 主な発表論文等

〔雑誌論文〕 計8件（うち査読付論文 8件/うち国際共著 1件/うちオープンアクセス 5件）

1. 著者名 Dinh Tuan Tran, Hirotake Yamazoe, Joo-Ho Lee	4. 巻 50
2. 論文標題 Multi-Scale Affined-HOF and Dimension Selection for View-Unconstrained Action Recognition	5. 発行年 2020年
3. 雑誌名 Applied Intelligence	6. 最初と最後の頁 1468-1486
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s10489-019-01572-8	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 Sakurai Ryuhei, Yamane Sasuke, Lee Joo-Ho	4. 巻 15
2. 論文標題 Restoring Aspect Ratio Distortion of Natural Images With Convolutional Neural Network	5. 発行年 2019年
3. 雑誌名 IEEE Transactions on Industrial Informatics	6. 最初と最後の頁 563 ~ 571
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/TII.2018.2803041	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 Ryuhei Sakurai, Taiki Shimba, Hirotake Yamazoe, Joo-Ho Lee	4. 巻 13
2. 論文標題 Synthesis of Expressive Talking Heads from Speech with Recurrent Neural Network	5. 発行年 2018年
3. 雑誌名 Journal of Korea Robotics Society	6. 最初と最後の頁 16-25
掲載論文のDOI（デジタルオブジェクト識別子） 10.7746/jkros.2018.13.1.016	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 Tadashi MATSUO, Nobutaka SHIMADA	4. 巻 100-D
2. 論文標題 Construction of Latent Descriptor Space and Inference Model of Hand-Object Interactions	5. 発行年 2017年
3. 雑誌名 IEICE Trans. on Information and System	6. 最初と最後の頁 1350-1359
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transinf.2016EDP7410	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 -

〔学会発表〕 計20件（うち招待講演 2件 / うち国際学会 14件）

1. 発表者名 里岡 樹、山添 大丈、李 周浩
2. 発表標題 再構成可能な知能化空間における突起とレールを用いた Mobile Module の動作改善
3. 学会等名 第20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(SI2019)
4. 発表年 2019年

1. 発表者名 Makoto Sei, Akira Utsumi, Hirotake Yamazoe, Joo-Ho Lee
2. 発表標題 Network Structure for Personalized Face Pose Estimation Using Incrementally Updated Face-Shape Parameters
3. 学会等名 The 5th Asian Conference on Pattern Recognition (ACPR 2019) (国際学会)
4. 発表年 2019年

1. 発表者名 Sasuke Yamane, Hirotake Yamazoe, Joo-Ho Lee
2. 発表標題 Human motion generation based on GAN toward unsupervised 3D human pose estimation
3. 学会等名 The 5th Asian Conference on Pattern Recognition (ACPR 2019) (国際学会)
4. 発表年 2019年

1. 発表者名 里岡 樹、山添 大丈、李 周浩
2. 発表標題 再構成可能な知能化空間における突起を用いたMobile Moduleの開発
3. 学会等名 第37回 日本ロボット学会学術講演会(RSJ2019)
4. 発表年 2019年

1. 発表者名 山根 佐介, 山添 大丈, 李 周浩
2. 発表標題 GANを用いた姿勢既知の教師データを必要としない3次元人物姿勢推定
3. 学会等名 第22回 画像の認識・理解シンポジウム (MIRU2019)
4. 発表年 2019年

1. 発表者名 Makoto Sanada, Tadashi Matsuo, Nobutaka Shimada, Yoshiaki Shirai
2. 発表標題 Recalling Candidates of Grasping Method from an Object Image using Neural Network
3. 学会等名 The 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2019) (国際学会)
4. 発表年 2019年

1. 発表者名 福井 尚卿, 松尾 直志, 島田 伸敬
2. 発表標題 力入力に対する剛体群の運動応答予測と静力学的構造安定性の推定
3. 学会等名 ロボティクス・メカトロニクス講演会2019(ROBOMECH2019)
4. 発表年 2019年

1. 発表者名 施 真琴, 内海 章, 山添 大丈, 萩田 紀博, 李 周浩
2. 発表標題 CNN を用いた眼球中心・虹彩中心推定に基づく視線検出手法の検討
3. 学会等名 日本機械学会ロボティクス・メカトロニクス講演会2018 (ROBOMECH2018)
4. 発表年 2018年

1. 発表者名 Tatsuki Satooka, Hirotake Yamazoe, Joo-Ho Lee
2. 発表標題 Barb based Fast Movement of Mobile Module for Deploying Devices in Reconfigurable Intelligent Space
3. 学会等名 2018 15th International Conference on Ubiquitous Robots (UR) (国際学会)
4. 発表年 2018年

1. 発表者名 Makoto Sei, Akira Utsumi, Hirotake Yamazoe, Norihiro Hagita, Joo-Ho Lee
2. 発表標題 Investigation of gaze detection method based on eyeball center and iris center estimation using CNN
3. 学会等名 The 14th joint Workshop on Machine Perception and Robotics (MPR2018) (国際学会)
4. 発表年 2018年

1. 発表者名 Joo-Ho Lee, Tatsuki Satooka, Hirotake Yamazoe
2. 発表標題 Toward Efficient Environment: Reconfigurable Intelligent Space
3. 学会等名 AMSM2018 (招待講演) (国際学会)
4. 発表年 2018年

1. 発表者名 DINH TUAN TRAN, Hirotake Yamazoe, Joo-Ho Lee
2. 発表標題 View-Invariant Human Activity Recognition Using Topic Model on Combined ORB-OF Feature
3. 学会等名 2019 IEEE/SICE International Symposium on System Integration (国際学会)
4. 発表年 2018年

1. 発表者名 Ryuhei Sakurai, Sasuke Yamane, Joo-Ho Lee
2. 発表標題 Correcting Aspect Ratio Distortion of Natural Images by Convolutional Neural Network
3. 学会等名 The 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI 2017) (国際学会)
4. 発表年 2017年

1. 発表者名 Dinh Tuan Tran, Ryuhei Sakurai, Hirotake Yamazoe, Joo-Ho Lee
2. 発表標題 PCA-based Surgical Phases Estimation with a Multi-Camera System
3. 学会等名 The 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI 2017) (国際学会)
4. 発表年 2017年

1. 発表者名 Dinh Tuan Tran, Hirotake Yamazoe, Joo-Ho Lee
2. 発表標題 Camera Pose-Independent Action Recognition in Operating Room
3. 学会等名 The International Conference on Information and Communication Technology and Digital Convergence Business (ICIDB-2018) (国際学会)
4. 発表年 2018年

1. 発表者名 Joo-Ho Lee
2. 発表標題 Coexistable robotics : relief, safety and convenience
3. 学会等名 The 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI 2017) (招待講演) (国際学会)
4. 発表年 2017年

1. 発表者名 Masaki Yano, Tadashi Matsuo, Nobutaka Shimada
2. 発表標題 Description and recall of the object using processes with scene change using LSTM
3. 学会等名 The 2nd Workshop on Machine Learning Methods for High-Level Cognitive Capabilities in Robotics 2017 (国際学会)
4. 発表年 2017年

1. 発表者名 Tadashi Matsuo, Hiroya Fukuhara, Nobutaka Shimada
2. 発表標題 Transform Invariant Auto-encoder
3. 学会等名 The 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2017) (国際学会)
4. 発表年 2017年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

<p>環境変化にロバストな画像データを用いたフェーズ推定及び評価 http://www.aislalab.org/index.php/phase-estimation/</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	島田 伸敬 (Shimada Nobutaka) (10294034)	立命館大学・情報理工学部・教授 (34315)	