

| | | | |
|--|--|----|------|
| 京都大学 | 博士 (工学) | 氏名 | 陳 龍飛 |
| 論文題目 | Analysis and Modeling of Machine Operation Tasks using Egocentric Vision (エゴセントリックビジョンを用いた機械操作タスクの分析とモデリング) | | |
| <p>(論文内容の要旨)</p> <p>本論文は、機器操作のマニュアルを自動生成することや利用者ガイダンスを行うための教示データを得ることを目的として、機器を操作する様子を操作者が身につけたカメラによって記録し、得られたデータを自動的に解析することによって、機器操作の手順を自動的に獲得する手法、および、初心者から熟練者までを操作者の対象とし、習熟過程も分析することによって、網羅的かつ質の良い操作モデルを得る研究をまとめたものである。本論文は以下の5章からなっている。</p> <p>第一章は序論であり、一人称視点映像 (egocentric vision) の現在の活用状況や将来の可能性について、また、AR 技術などを用いたユーザ支援の現状や課題について基礎的な議論を進め、機器操作の支援が重要なトピックとなっていることを述べている。このような背景を踏まえ、以後の章で述べる、接触を基にしたに操作タスクの表現、多様な操作者の操作記録からのモデル獲得、習熟過程の分析に基づく操作モデル獲得手法などの考え方を紹介している。</p> <p>第二章では、RGB-D カメラによって得られる一人称視点映像から機器操作のタスクモデルを教師なし学習によって獲得するための手法について述べている。まず、操作者の手が機器と接触した部分を映像から自動的に抽出し、接触時間や手指の形より何らかの操作を行っているとは推定されるものをホットスポット (hotspot) とする。位置、接触時刻、手先の形状から推定される操作の種類をその属性として与えることにより、機器操作の基本的なインタラクション列を得る。このようなインタラクション列は操作者によって部分的に異なるが、十分に習熟した者の操作であれば、重要な操作ステップが漏れなく含まれていることを期待できる。本研究では実際に、意味的に同一な hotspot の統合、機器による手のオクルージョンへの対処、操作目的以外の接触に対する処理などを加えることによって、隠れマルコフモデル (HMM) を用いた教師なし機械学習が良好に働き、それによって確率的時系列モデルとして良い操作モデルが得られることを示している。実験としては、ミシン操作の比較的複雑なタスク 3 つ (そのうち 2 つは 10 ステップ以上の操作を含む) を設定し、各 20 系列程度与えられた操作記録を統合することによって、機器付属のマニュアルに示された手順内容を十分に含んだ操作モデルを得ることができることを示した。</p> <p>第三章では、初心者を含めた多様なユーザに有用な操作モデルを得るために様々な習熟レベルの操作者の操作記録を統合する手法について述べている。これは、熟練者のやり方が初心者にとって必ずしも適していないこと、例えば、熟練者は初心者にとっては必要な操作や確認を省いていたり、初心者には難しすぎる操作を行っていることなどが問題となるためである。そのため、初心者を含めた多様な操作者の操作記録を統合してタスクモデルを構成することが望まれる。しかし、初心者には操作誤りや不必要な動作も多く、二章で述べた方法を直接用いてタスクモデルを自動獲得することは難しい。そのため、熟練者の操作から得られるタスクモデルをベースラインとし、</p> | | | |

| | | | |
|--|---------|----|------|
| 京都大学 | 博士 (工学) | 氏名 | 陳 龍飛 |
| <p>そこに初心者などから得られる操作系列を統合していくブートストラップ的手法を提案している．そのために，新しい操作ステップ，繰り返し，代用となる操作ステップ，省略，順序変更などの差異がある場合の操作系列を統合するアルゴリズムを新しく提案している．これにより，重要な操作を含むが多様性に欠ける熟練者のモデルに対して，初心者や習熟途中の者に適した方法などを付加していくことにより，よりガイドンスに適した操作モデルが得られる．実験では，IH ヒータやミシンの操作タスクに対し，それぞれ 2 人の熟練者と 7 人の非熟練者の操作を記録し，提案手法を用いることによって良好な操作モデルが得られることを示している．具体的には，初心者のみがよく使う易しい方法，確認のために必要な操作，また，初心者がおかしやすい誤りとその修正方法などを含んだ操作モデルが得られることを示した．</p> <p>第四章では，操作記録から操作者のスキルレベル（習熟の度合い）を推定する方法について述べている．これは，種々のレベルの操作者を無作為に集めたり，クラウドソーシングなどによって多くの人から操作記録を集める場合に，操作者の熟練度をあらかじめ評価することが難しいためである．さらに，個々の操作者の特性として，ある部分の操作には熟練していても，他の部分に対しては初心者レベルであるような場合も想定される．そのため，多くの操作記録に対し，単純に第三章で述べた手法を適用することが難しい．この問題に対し，操作者のスキルレベルを自動的に推定することができれば，熟練者の操作記録を選び出し，前章のベースラインを構成することができる．これを可能にするために，この研究では初心者が熟練していく過程を詳細に記録し，その振る舞いの変化を分析することにより初心者と熟練者を識別するための特徴を洗い出している．具体的には，個々の作業ステップをさらにサブステップに分割し，それぞれのサブステップにおける手，視線，頭部それぞれの動き，それにかかる時間などが習熟につれて変化する傾向を調査した．その結果，様々な傾向，例えば，頭部の動きが最も習熟度を良く反映するパラメータであること等を見出している．これらの特徴を用いて，スキルレベルが高い操作者の操作記録を選び出し，ベースラインとして用いることによって，無作為にベースラインを選んだ場合に比べて良い操作モデルが得られること，また，それが人手で構成した理想的な操作モデルと良く一致していることなどを実証した．</p> <p>第五章は結論であり，本論文で得られた成果について要約している．さらに，第二章～第四章で述べてきた手法の課題や，さらに発展させるためのアプローチなどについて提案しながら，この分野の研究や応用についての展望について述べている．</p> | | | |