

# 複数センサによる卓上作業認識と作業支援環境の構築

## Recognizing Tabletop Works and Building Support Environment of Manual Assembly Working Using Multiple Sensors

高塚洋平<sup>†</sup> 香川雄平<sup>†</sup> 石井愛弓<sup>‡</sup> 飯島正<sup>‡</sup>  
Yohei Takatsuka<sup>†</sup> Yuuhei Kagawa<sup>†</sup> Ayumi Ishii<sup>‡</sup> Tadashi Iijima<sup>‡</sup>

<sup>†</sup> 慶應義塾大学大学院 理工学研究科

<sup>‡</sup> 慶應義塾大学 理工学部

<sup>†</sup> Graduate School of Science and Technology, Keio Univ.

<sup>‡</sup> Department of Science and Technology, Keio Univ.

### 要旨

近年のセンサ技術やネットワークの普及、情報処理能力の向上により、それらを活用した新しいサービスが提案されている。また、それは計算機の中に閉じられたものだけでなく、新しいタイプの入力デバイスやAR(Augmented Reality; 拡張現実感)サービスのよう、実世界に拡張されたものもある。

本論文では、カメラをはじめとするセンサを用いて作業者の行っている作業を認識し、作業者に対して支援を行う手法を提案する。作業者が行っている作業を認識することで、事前に定義された完成図に基づいて、次に行うべき作業内容を提示することが可能になる。また、習熟した作業者を対象として作業パターンを抽出することで、明示されていない作業パターンの構築を行い、習熟していない作業者に対して同様な作業支援を行う。

本論文では、そのような作業支援機能を実現する手法およびその環境を提案し、現在構築中の評価用プロトタイプに関して報告する。

### 1. はじめに

日常生活において卓上(tabletop)での作業は多く行われ、読み書き、調理、食事など多岐にわたる。近年では、直観的な入力を実現する卓上インタフェースの研究や、卓上インタフェースとAR技術を組み合わせた新しい利用法も提案されている。このような傾向から、卓上は、人間の作業空間として一般的な環境にあり、情報技術との親和性の高い環境であると考えられる。

本研究では、卓上で行われる組立作業に焦点を当て、その作業を認識し、作業者に対して支援する環境を提案する。卓上環境に可視光カメラをはじめとする様々なセンサ機器を配置し、センサ間の情報を統合することで作業者の行う行動を認識する。一般に、作業者の特性や習熟が作業の効率や成否に大きく影響する。本研究では、認識された作業者の行動を元に、事前にモデル化された一連の作業に合わせて次に行うべき行動を提示することで作業者に対する支援を行う。また、習熟した作業者や作業を十分に理解している作業者は、指示を必要とせずに作業を行うことが可能であると考えられる。そのような作業者の一連の作業パターンを系列パターンマイニングによって抽出し、事前に定義することなく作業モデルを構築する。これにより、習熟した作業者の経験や知識を、未熟な作業者に対するフィードバックとして活用することが可能となる。

本研究と同様、作業を認識し支援するシステムの提案として、調理作業を対象としてレシピに基づいて支援を行うもの[1]がある。これは、調理作業に伴う食材などの物体の近くに、画像処理によって解析可能なマーカーを取り付けることで物体の変化や位置をデータとして扱っており、手の動き等によって作業者の行動を認識し、モデル化されたレシピと照らし合わせる。そして、拡張現実感として作業者に対して提示支援を行っている。モデル化された作業群を用いることや視覚的な支援を行う点などで本研究と一致するが、本研究では、より作業空間内の制約を減らし、より自然な作業を行えるような環境を目指す。

以下では、第2節において卓上作業の認識手法およびその環境について提案し、続く第3節で、そのシステムのより具体的な設計を述べる。また第4節では提案システムの評価のための実験について述べる。

### 2. 作業認識のためのセンシング技術

各種カメラ、床圧力センサといったセンサ機器からの情報を把握し、作業行動を認識する手法について提案する。次に行うべき作業を推定し、視覚的に作業の支援を行う。

## 2.1. 卓上作業認識環境

本研究で提案する環境には、あらゆるセンサが作業台を中心として設置されていることを想定している。図1に作業台と各種センサの配置を示す。本研究での作業はレゴブロックを用いた組み立て作業を題材として扱う。組み立て作業では、赤、青、緑、灰色のブロックを使用し、灰色ブロックを組み立ての土台となるブロックとする。

今回使用する可視光カメラ(Logicool社, HD Pro Webcam C910)は、500万画素の画像フレームを取得可能なもので、今回扱うレゴブロック等を含む画像に対して画像処理を行うのに十分な解像度を持つものである。また、熱画像カメラ(Artray社, ARTCAM-320-THERMO)は、あらゆる物体から放射されている赤外線を計測することで、温度の検出を行う。床圧力センサ(Xiroku社, LL Sensor)は、1cm<sup>2</sup>あたり1素子のセンサが埋め込まれており、細かな圧力データを扱うことができる。

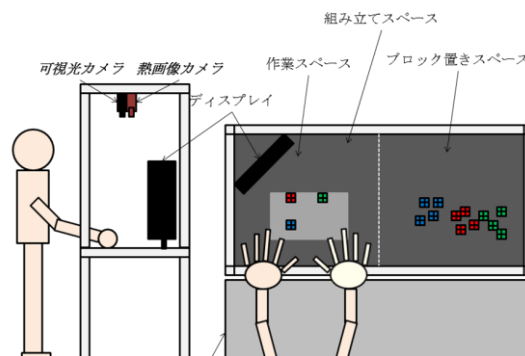


図1 作業環境の概要

## 2.2. 画像処理技術によるパターン認識

### 2.2.1. 色情報に基づく物体の検出

作業の認識を行う上で、どの物体を扱った作業なのかを検出する必要がある。本研究では、物体認識のためには、その物体の色と大きさが必要な情報と仮定している。物体認識を行う上で、事前に作業中に取り扱う物体のテンプレートとなる画像を用意する。このテンプレート画像から各物体に関する色ヒストグラムを生成する。このヒストグラムから入力画像中の各ピクセルがその物体のピクセル分布とどの程度適合しているかを計算することで、入力画像中にその色の物体の領域を見つけ出すことができる。

今回の作業素材となるレゴブロックには、表面に多数の突起があるため、多数のノイズを含む検出結果になってしまうため、まずテンプレート画像や入力画像に対して平滑化処理を施し、ノイズを低減を図る。また、平滑化処理に加えて、検出画像に対してオープニング演算及びクローズング演算を行うことで、より正確な物体の検出を実現する。

### 2.2.2. 手領域をトリガとした作業結果の認識

一連の作業を行う上で、一般に、作業前と作業後で作業対象物体に加工によって何らかの変化が起こる。本研究では作業の前後での変化に注目し、先述の手法によって検出された色領域における作業前後の差分を計算する(作業の残後の画像の排他的論理和を求める)ことで、その作業によって起こった変化を検出する。組立作業では、各作業の際に手の領域を含む画像フレームが作業前後に含まれるため、手領域を含む画像フレームをトリガとして作業結果の差分処理を行う(図2)。

このように手領域を含む画像フレームを利用するためには、正確に手領域の検出を行う必要がある。先述の色ヒストグラムに基づいた検出手法でも手領域の検出を行うことはできるが、(肌)色情報のみの検出では、他の肌色物体と誤検出してしまいう可能性がある。そこで、本研究では熱画像カメラを用いて手領域の検出を行う。作業者の体温は素材であるレゴブロックを含めた作業対象物体よりも温度が高いという前提の下、素材と作業者の手を区別する。

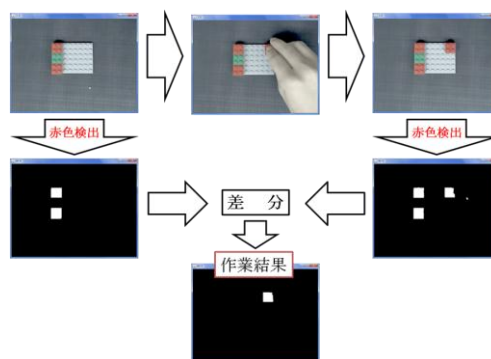


図2 作業結果の検出

### 2.2.3. 矩形領域の検出

図2に示されるように、各色で検出した物体の領域および作業結果の画像は正しく矩形にならず、やや歪んでいたり欠損のある矩形となる。作業結果等を正確に取得するためには、そのような画像から正確な矩形領域を推定する必要がある。そこで色検出画像に対して矩形領域を探索するためにMean Shift法を用いる。Mean Shift法はデータ中の最頻となる地点を探索する手法で、探索範囲に矩形領域を指定することで、色検出画像中で最も矩形らしい部分を矩形として出力することができる。本研究では、探索範囲を可変

の矩形領域として扱うことが可能な CAMSHIFT 法[2]を用いて、様々な大きさのレゴブロックに対応する。

### 2.2.4. 指先検出

作業中の指先の位置情報は、作業認識を行う上で重要な情報源となる。本研究では、画像情報から指先の位置を検出するために以下に示す3つの手法を実装し、評価中である。

まずは、熱画像カメラから取得された手領域に対して凸包となる領域を検出し、領域中の頂点となる特徴点を指先とする手法である(図2)。

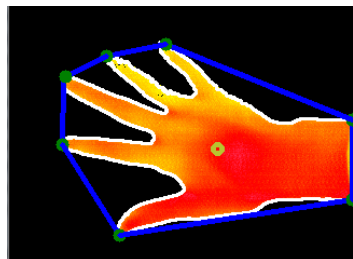


図3 凸包領域による指先検出



図4 輪郭ピクセルに基づく指先検出



図5 距離変換画像中の特徴ピクセル

第二の手法は、Shahzad ら[3]による手領域の輪郭ピクセルのベクトルを用いた手法である。肌色領域のみの二値化画像に対して輪郭を検出し、 $k$ 番目の輪郭ピクセルと $k+n$ 番目および $k-n$ 番目の輪郭ピクセルからなる2つのベクトルを計算する。この2ベクトルからなる $k$ 番目の輪郭ピクセル回りの角度に注目し、この角度が閾値以下の輪郭ピクセルを抽出する。更にそれら2ベクトルから成る $2 \times 2$ の行列の行列式の正負によって指先か指の谷間かを判定する(図3)。第三の手法は、Le ら[4]による距離変換処理を用いた特徴ピクセルに基づく手法である。肌色領域のみの二値化画像に対して距離変換処理を行い、各画像中のピクセルに対して局所最大かつ閾値を満たすピクセルのみを抽出する(図4)。この抽出されたピクセルと手の中心との距離を取ること

### 2.2.5. 床圧力センサを用いた作業位置の推定

一般的に卓上での立ち作業の場合、作業者がその場にとどまって作業をし続けることは稀である。そこで本研究では、作業台周りに設置された床圧力センサによって、作業者の作業している位置を推定する。この推定により、複数の作業スペースを行き来する作業の認識に活用することができる。図4に示すように床圧力センサを用いて、足の向きや立ち位置を推定する。

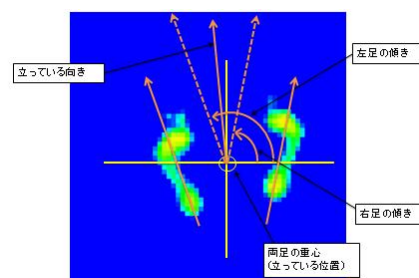


図3 足の向きと位置の推定方法

## 2.3. 照合による作業の認識と系列パターンマイニング

### 2.3.1. ワークモデル

多くの作業は、ある目的を達成するための順番を伴った一連の小さな作業系列に分割できる。

したがって、事前に順序立てた作業群を定義することで、作業者が次に行うべき作業を特定することが可能となる。本研究では、そのような作業群を各ワークが先行作業を持つ有向非巡回グラフ(DAG)によるワークモデルとして構築する。表1にワークモデルの例を示す。

表1 ワークモデルの例

作業名	先行作業	色	大きさ	位置
A	なし	赤	2x2	(0,0)
B	なし	赤	2x2	(6,0)
C	なし	緑	2x2	(6,2)
D	なし	赤	2x2	(6,4)
E	なし	緑	2x2	(0,2)
F	なし	赤	2x2	(0,4)
G	B,C	青	2x2	(6,1)
H	F,E,F	青	6x2	(1,3)

### 2.3.2. 作業パターンの抽出

系列パターンマイニング(Sequential Pattern Mining)手法を用い、作業パターンを抽出する。系列パターンマイニングとは、複数のアイテムの順序列の集合をデータセットとして、頻出する系列を取り出す手法である。系列パターンマイニングでは、最小サポートと呼ばれる値を閾値として、それを満たす全ての系列を発見することが問題となる。

ここでのサポートとは、その系列がデータセット中に出現する確率を意味しており、最小サポートを満たすシーケンスは頻出系列として抽出される。本研究では、2001年に Pei ら[5]が提案した PrefixSpan アルゴリズム

ムを利用する。PrefixSpan では、データセット中の各系列に対して投射と呼ばれる処理を行い、1 アイテムずつ頻出系列を伸ばしていくことで探索を行う。投射とは、接頭のアイテムより後ろに続く系列を取り出す処理であり、PrefixSpan では、各アイテムを順番に接頭のアイテムとし、データセットに対して再帰的に投射を行うことで、全ての頻出系列の抽出を実現している。

本研究では、各作業をアイテムとしてみなし、習熟した作業者による一連の作業を系列として記録していくことでデータセットを生成する。十分な大きさのデータセットが生成された後に、PrefixSpan を適用することで頻出する作業パターンを抽出し、その作業パターンを元に、先述のワークモデルを構築する。この処理によって、事前にワークモデルを定義することなく、習熟した作業者の経験や知識に基づいたワークモデルを構築することが可能となる。

### 2.3.3. 情報提示による作業支援

ワークモデルによって推定された作業者が次に行うべき行動を、視覚的に作業者に提示することで、作業支援を行う。図 1 の作業環境中の作業台の上にあるディスプレイに、上部に設置されたカメラから取得された作業スペースの画像をリアルタイムに映し、その画像上にマーキングすることで作業支援とする。

## 3. 作業支援システムの構築と評価実験

本提案における作業認識手法を実証するために、先述の手法によって物体の色検出を行い、手領域画像をトリガとした差分によって作業結果を取得するモジュールを構築した。また、XML 形式のファイルを解析しワークモデルを構築するメカニズムと、上からの卓上の画像に重ね合わせる形でマークを表示することによる作業支援モジュールも構築した。これらのモジュールの統合によって作業支援システムを構成する。なお、今回のシステムでは、画像処理に関連する手法を数多く使用しているが、それらはオープンソースの画像処理ライブラリである OpenCV[6]を用いている。

提案した手法によって作業者の作業が認識され、正しく支援を行っているかどうかを評価するために、評価実験を行った。今回は、図 6 に示す二通りの完成図を目指す組立作業を行い評価したところ、

Type1 に関しては、正しく作業の認識を行うことができ、視覚的に作業指示を提示できた。しかしながら、Type2 に関しては、作業認

識において誤りが多く、正しく支援を行うことができなかった。これは、色情報を元に差分を行っていることによるもので、同じ色のブロックを重ねるような作業の場合に生じるものである。

## 4. おわりに

卓上での組立作業を認識する手法について提案し、事前に定義したモデルと認識された作業を照合することで作業に対して視覚的な支援を行う環境について述べた。今後の課題としては、評価実験によって明らかになった問題を解決するために距離画像などを用いてブロックの高さを取得することや、より具体的でわかりやすい作業指示手法を提案することなどが挙げられる。

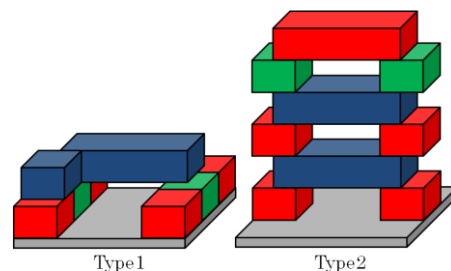


図 4 実験用の完成図

## 参考文献

- [1] Younes Fadil, Satoru Mega, Arata Horie and Kuniaki Uehara, "Mixed Reality Cooking Support System Using Content-Free Recipe Selection", Proceedings of Multimedia the Eighth IEEE International Symposium, pp.845-850(2006).
- [2] Gary R. Bradski, "Computer Vision Face Tracking For Use in a Perceptual User Interface",
- [3] Shahzad Malik and Joe Laszlo, "Visual Touchpad: A Two-handed Gestural Input Device", Proceedings of Multimodal interfaces the 6th international Conference, pp.289-296(2004).
- [4] Le Dung and Makoto Mizukawa, "Fast Fingertips Positioning Based on Distance-based Feature Pixels", Proceedings of Third International Conference on Communications and Electronics(ICCE), pp.184-189(2010).
- [5] Jian Pei, Jiawei Han, Mortazavi-Asl B., Pinto H., Qiming Chen, Dayal U. and Mei-Chun Hsu, "PrefixSpan: Mining Sequential Patterns Efficiently by Prefix-Projected Pattern Growth", Proceedings of Data Engineering 2001 17<sup>th</sup> International Conference, pp.215-224(2001)
- [6] Open Computer Vision Library (<http://sourceforge.net/projects/opencvlibrary/>)