

人物大装置の組み立て作業改善活動支援のための 俯瞰距離画像からの人物姿勢分類

高橋 典宏*・山澤 一誠*・生雲 公啓**
野田 賢*・横矢 直和*

Posture Classification from Bird's-eye View of Range Data for Supporting Improvement Activities in
Assembly Work of Human-sized Product

Norihiro TAKAHASHI*, Kazumasa YAMAZAWA*, Masahiro IKUMO**,
Masaru NODA* and Naokazu YOKOYA*

Improvement activity called IE (Industrial Engineering), which eliminates inefficiency, is always performed in factories by specialists of IE and plant foremen. IE aims to promote the efficiency of overall production process, and such an activity enables us to upgrade products and boost productivity. Conventional IE is performed in order of Plan, Do, Check, and Action. Check takes long time, and depends on worker's skill. On the issues, many methods that aim to support the improvement activities are proposed. This paper proposes a method of human posture classification for inefficiency detection in an assembly work of a human-sized product. In such work, many posture changes of a worker happen. Recognizing human posture is important for inefficiency detection. Factory environments usually restrict positions of sensors such as a ceiling for avoiding the interference with workers, and require simple instruments. The proposed method uses a range image sensor that is installed on a ceiling and captures bird's-eye range images. The method computes features from the range images, and classifies human postures into predefined categories by using the Support Vector Machine (SVM). In the simulation of assembly work, appropriate parameters of SVM learning are searched for and the accuracy of posture classification is evaluated. Furthermore, experiments are carried out with a real factory environment and the validity of the proposed method for improvement activity support is evaluated by an IE specialist's hearing.

Key Words: range image sensor, bird's-eye view, posture classification, SVM, improvement activity support

1. はじめに

製造工場では、製品品質や生産性の向上を目的とした作業のムダを排除するための改善活動が常に行なわれている。このような作業の改善活動を IE (Industrial Engineering)¹⁾といい、IE の専門家 (以下、IEr) や現場の監督などの改善活動従事者により行なわれる。

IE は Fig.1 に示すように、まず現場を直接またはビデオカメラなどにより撮影することで観測し、観測データを分析することで改善を必要とする作業や物の配置などを抽出する。分析結果から改善案の案出を行ない、改善案を現場へ反映させて改善の前後における効果について検討するといったサイ

クルで行なわれている。IE で行なわれる分析手法の代表的なものとして、時間研究法 (Time Study)¹⁾が挙げられる。時間研究法とは、 unnecessary 作業の発見、動作や手待ち時間の削減を行ない、作業の効率化による稼働率や生産性の向上を目的としている。時間研究法では、改善を行なうための観測回数は一般的に 10 回程度必要であり、観測・分析で必要とする時間は膨大なものとなっている。近年では、製品ライフサイクルの短縮化などに伴い、より効率的な作業改善が望まれており、改善活動を支援する様々な手法が開発されている^{2)~6)}。これらの改善活動支援システムを用いることで、従来の IE と比較して改善活動を効率良く正確かつ容易に行なうことが可能となる。

しかし、製品の製造開始から全工程が完了するまでの時間 (以下、リードタイム) が長い作業においては、前述したような改善活動を支援する手法を適用しても効率的な改善活動を行なうことは困難であり、このような作業の改善に取り組んでいない製造工場も少なくない。また、リードタイムの長い作業には複雑な多数の工程があり、同じ工程を同じ作業者が行

* 奈良先端科学技術大学院大学 情報科学研究科

** オムロン株式会社 技術本部 コアテクノロジーセンター

* Graduate School of Information Science, Nara Institute of Science and Technology, Nara

** Core Technology Center, OMRON Corporation, Kyoto

(Received July 11, 2008)

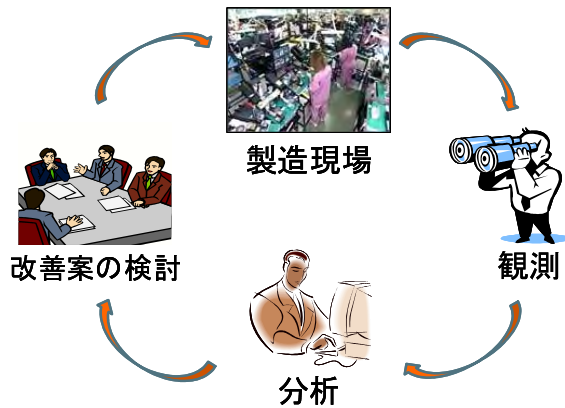


Fig. 1 IEにおける活動サイクル

なっても毎回動作が異なるため動作に繰り返し性がないという特徴があり、作業改善がさらに困難なものとなっている。リードタイムの長い作業における分析手法として、ワークサンプリング法 (Work Sampling)¹⁾が挙げられる。ワークサンプリング法とは、作業状態を効率よく観測するために対象作業を任意の時間間隔で観測する分析手法であり、リードタイムの長い作業を効率的に改善することが可能となる。しかし、ワークサンプリング法では時間を空けて観測するため、多くの観測漏れが発生するという問題がある。このため、リードタイムの長い作業を対象とした改善活動では、長時間の作業を漏れなく観測し、かつ改善ポイントを効率的に発見することが望まれる。

リードタイムの長い作業としては、現金自動預け払い機 (ATM) や券売機、医療機器、工場で用いられる基板検査装置などの人物大の装置を組み立てる作業などがある。人物大の装置の組み立て作業では、作業者の移動や立ち上がり動作の繰り返し、無理な姿勢での作業などのムダが発生しており、作業者の移動や姿勢を認識することでこれらのムダの発見を効率的に行なうことができる。人物の移動検出によるムダの発見は従来から研究されているため⁵⁾、本研究では、作業者の姿勢を推定することで立ち上がり動作の繰り返しや無理な姿勢での作業などのムダを発見するための支援に取り組む。本研究で対象とする作業は、従来のIE支援手法を用いても改善活動を行なうことが困難であったセル生産方式による人物大の装置の組み立て作業とし、改善活動を効率的に行なうための支援を目的とする。ここでセル生産方式とは、一人あるいは数人のチームで多工程を担当しながら、部品の取り付けから組み立て、検査までの全工程を行なう生産方式であり、部品や工具を配置したセルと呼ばれる作業場内において作業を行なう。本研究では、姿勢を推定することができるだけの情報を得るために、奥行き情報を取得することができる距離画像センサを天井に取り付け、上方より作業者を見下ろすように作業の様子を観測する。上方より撮影した距離画像 (以下、俯瞰距離画像) から特徴量を抽出し、パターン識別手法のひとつであるサポートベクターマシン (以下、SVM)^{19), 20)}を

用い、改善活動従事者が直感的に作業者の姿勢を把握できるように予め定義した教示姿勢のいずれかに分類する。この提案手法を模擬実験および実際の製造工場において実験し、量的評価およびIErへのヒアリングにより、改善活動支援に対する有用性を示す。

2. 製造工場における作業改善活動と本研究の位置付け

2.1 作業改善活動の支援に関する従来手法

近年、より効率的な作業改善が望まれているが、従来のIEでは観測や分析に多くの手間と時間がかかっていた。そこで、IEにおける観測・分析作業を支援する様々な手法の研究・開発が盛んに行なわれている²⁾⁻⁶⁾。

従来のIEでは、ストップウォッチなどにより計測された動作時間や作業内容を紙に手書きで記入するなどしていたため、分析に多くの手間がかかっていた。このような問題に対し、改善活動における分析作業を支援するシステムも開発されている^{3), 4)}。これらのシステムは、作業の様子を記録した映像を入力とし、その映像をパソコン上で再生しながら作業分析をマウス操作で容易かつ正確に行なうことができる。作業分析は、組み立て・検査・梱包といった作業を要素作業ごとに区切り、各要素作業の所要時間などを記録する。得られたデータは標準作業と比較し、時間のかかった作業や異常作業などを分析する。実際の作業では、1秒程度の要素作業が連続する様な作業も少なくない。しかし、これらのシステムを用いることで、改善活動の初心者でも改善が必要な作業を見落とすことなく、効率的な作業分析を行なうことができる。これらのシステムの用途として、作業分析、作業熟練者と初心者の映像比較による作業者の訓練、改善活動従事者のIE技術習得などが挙げられる。

板倉らは、滞在作業者の手や頭の移動軌跡からムダが発生したサイクルを自動で検出する手法⁶⁾を提案している。ここで、滞在作業者とは一定の場所にとどまって組み立て作業や梱包作業を行なう作業者のことであり、作業の手順は作業手順書などにより定められている。また、滞在作業は予め定められた作業を繰り返し行なうという特徴があり、この繰り返し作業の一巡をサイクルとしている。滞在作業では、作業の動作に含まれるムダの発見が望まれる。この手法では、作業者の両手と頭にカラーマーカを装着し、ビデオカメラにより撮影する。ムダの含まれるサイクルの検出は、作業の繰り返しによりカラーマーカを追跡して得られるマーカ軌跡の再現性を利用して行なう。この手法を用いることで作業中のムダを自動的に発見することができ、効率的な改善活動を行なうことができる。

2.2 本研究の位置付けと方針

本研究で対象とする作業は、セル生産方式による人物大の装置の組み立て作業である。人物大の装置の組み立て作業においてIEを行なうことが困難である理由として、リードタイムが非常に長い作業であるため、観測・分析に時間がかかりすぎるものがあげられる。この問題は、2.1節で述べたIE

を支援する手法を用いても解決には至らない。

人物大の装置の組み立て作業における IE では、作業者の動作に着目したムダの改善が行なわれている。このような作業における作業者の動作は、移動、全身の上下動などの大きな動作、肩から先の腕や手で行なう細かな動作に分類できる。このうち、移動と大きな動作に含まれるムダは、細かな動作に含まれるムダと比較してムダを改善することによる生産性の向上などの効果が大きく、優先的な改善が望まれる。作業者の移動に含まれるムダの発見はすでに取り組みられているため⁵⁾、本研究では大きな動作に着目し、IE の支援を行なう。人物大の装置の組み立て作業における大きな動作は、作業者の姿勢変化としてとらえることができ、大きな動作に含まれるムダを発見する上で作業者の姿勢を認識することは重要である。作業者の姿勢を把握することで、各工程における特定の姿勢の発生頻度、姿勢変化の様子と発生時刻、同一姿勢の継続時間など、ムダの発見に役立つ情報を改善活動従事者に提供することができ、リードタイムの長い作業において効率的に改善を行なうことができる。

人物大の装置の組み立て作業においては、作業者の動作は全身動作が多く移動範囲も広いいため、姿勢推定に用いるセンサの設置位置は作業に支障を来さない場所に制限される。製造工場において姿勢推定に用いるセンサは、作業者に物理的・心理的な拘束を発生しないものが望まれる。さらに、セル内には人物大の装置や部品棚などの遮蔽物も多く存在するため、センサを天井に取り付けるなどそれらを考慮して設置する必要がある。上方から作業者を観測する場合、単眼カメラでは姿勢推定に必要な情報を取得することが困難である。そこで本研究では、俯瞰距離画像を用いた作業者の姿勢分類を行なうことで、従来の IE 支援手法を用いても改善活動を行なうことが困難であったセル生産方式による人物大の装置の組み立て作業における改善活動支援を行なう。

3. 俯瞰距離画像からの SVM による人物の姿勢分類

本研究では、セル生産方式による人物大の装置の組み立て作業を想定し、作業中の作業者の姿勢を推定することによる作業改善活動の支援を目的としている。本章では、人物大の装置の組み立て作業における作業者の姿勢を推定する手法について述べる。

3.1 人物の姿勢推定に関する従来研究

人物の姿勢推定手法は接触型と非接触型に大別できる。接触型の人物姿勢推定手法は、機械式⁷⁾や光学式⁸⁾、磁気式⁹⁾などがある。接触型の人物姿勢推定手法は高い計測精度が得られるが、センサやマーカを装着するため身体に物理的・心理的な拘束を伴う。一方、センサやマーカを身体に装着せずに人物姿勢の計測を行なう非接触型の人物姿勢推定手法がある。非接触型の人物姿勢推定手法は、一般に画像情報を用いて行なわれ、一台のカメラを用いる手法^{10)~12)}やステレオカメラなどにより得られる距離画像を用いる手法^{17), 18)}、異なる位置に置かれた複数台のカメラを用いる手法^{13)~16)}に分類

できる。しかし、これらの手法は姿勢の特徴をとらえるためにカメラを対象人物の正面に置く必要があり、センサの設置位置が制限される製造工場に適用することは困難である。そこで、本研究では天井など上方に設置した距離画像センサにより作業者を見下ろすように作業の様子を撮影し、得られた俯瞰距離画像から人物の姿勢を分類する。

3.2 姿勢分類の対象とする教示姿勢の定義

作業時の身体動作の最適化に関する考え方として、「動作経済の原則」¹⁾がある。動作経済の原則とは、作業動作における人的資源の有効使用に関する事柄であり、人間のエネルギーをムダなく有用な作業へ転化し、有効な仕事量を増すためにムダのない作業動作を実現しようとする法則である。動作経済の原則では、作業を効率的に行なうために不要な動作を排除し、動作は身体の末端部位(手指)で行なうことが望ましいとされている。そのため、作業中に発生する姿勢の変化は、大きな重心移動を伴うため不要な動作と考えられる。

作業姿勢の評価法として、専用の計測器を必要とせず比較的簡易に作業姿勢を評価することができるため、世界中で利用されている OWAS (Ovako Working Posture Analysing System) 法^{21), 22)}がある。OWAS 法は、作業姿勢を背部、上肢、下肢、把持重量の4項目でとらえ、姿勢を4桁の数字で表した姿勢コードとして記録し、このコードを用いて作業姿勢の負担度を評価する。ここで、頭部と骨盤を結ぶ線(体幹)を背部、股関節-膝-足首を下肢とする。背部は鉛直方向に対して平行な状態を0度とし、腰を曲げるに従って角度が大きくなるものとする。また、下肢は鉛直方向と平行な状態を180度とし、膝を曲げるに従って角度が小さくなるものとする。

本研究では、人物大の装置の組み立て作業の様子を撮影した映像を観察し、OWAS 法を参考に後述する実験対象の作業中に見られる作業者の姿勢を以下に示す五つの姿勢に要約し、これらを教示姿勢として定義した。ここで、問題の簡単化のため把持重量については考慮しないものとする。

- 立位 (Fig.2(a))
背部のなす角度が20度以下、下肢のなす角度が180度から150度までの状態
- 踏み出し (Fig.2(b))
「立位」から片足を任意の方向に肩幅以上踏み出し、片足に荷重がかかった状態
- 前屈 (Fig.2(c))
背部のなす角度が20度以上の状態、もしくは下肢のなす角度が150度以下の状態
- しゃがみ込み (Fig.2(d))
「立位」における膝の高さより骨盤が低い位置にあり、足のみで姿勢が保てる位置に重心がある状態
- 四つん這い (Fig.2(e))
「立位」におけるひざの高さより骨盤が低い位置にあり、足のみでは姿勢が保てない位置に重心がある状態

後述する実験対象の作業中には Fig.3 に示す教示姿勢間での姿勢遷移が見られた。図中の実線は発生する頻度が高い姿

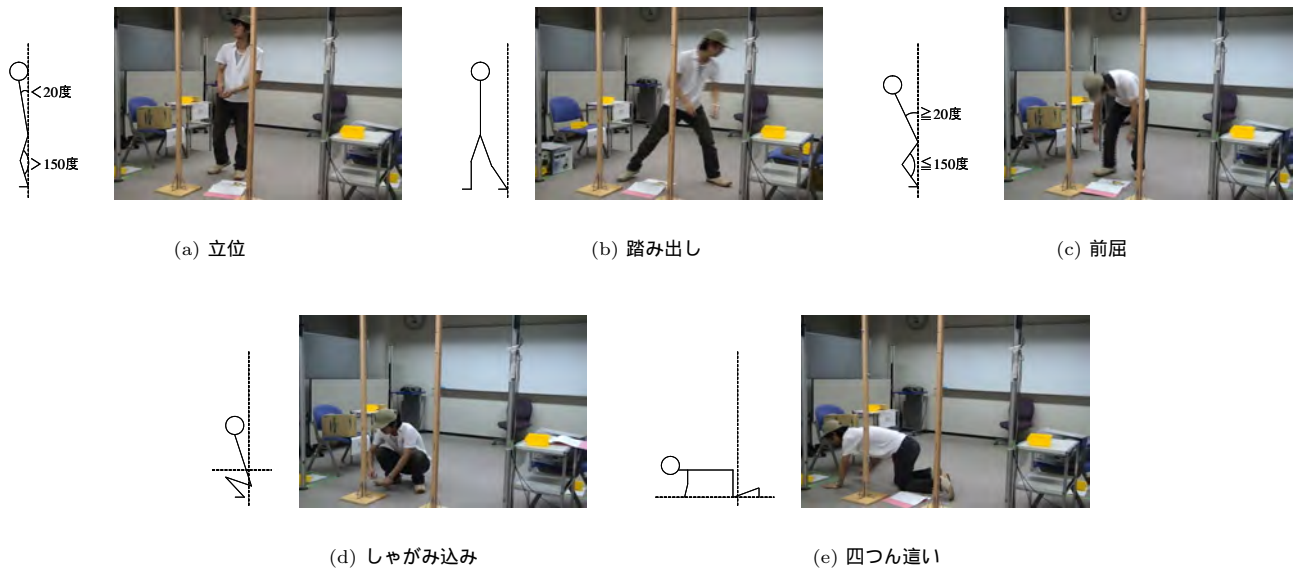


Fig. 2 教示姿勢の概念図と横から見た様子の例

勢遷移を示し、点線は発生する頻度が極めて低い特異な姿勢遷移を示す。

3.3 SVMを用いた人物の姿勢分類手法の概要と処理手順

本節では、俯瞰距離画像からのSVMを用いた人物の姿勢分類手法の概要について述べる。Fig.4に人物の姿勢分類手法の処理の流れを示す。製造工場では現場の環境制限により作業の様子を観測するセンサの設置位置は天井などに制限される。本研究では、上方に設置した距離画像センサにより作業員を見下ろすように作業の様子を撮影し、Fig.5(a)に示すような俯瞰距離画像を取得する。Fig.5(a)と同じ姿勢を横から見た様子を Fig.5(b) に示す。学習モデル生成の処理 (Fig.4(a)) では、俯瞰距離画像に対して背景差分を施すことで人物領域を抽出する。実際の製造工場において、セル内では製造物の形状や工具箱の配置などが時間的に変化する。そこで本研究では、人物領域の抽出手法として動的背景更新による物体検出手法である森田らの手法を採用した²³⁾。本研究では、輝度値の代わりに俯瞰距離画像の距離データを用い、動的背景更新による人物領域の抽出を行なう。

つぎに、抽出された人物領域から 3.4 節で述べる特徴量の抽出を行なう。また、俯瞰距離画像に対して 3.2 節で定義した教示姿勢のうちの一つを手動で与え、これを教示データとする。抽出された特徴量および教示データを用いてパターン識別器の学習を行ない、学習モデルを生成する。

パターン識別手法には、ニューラルネットワークや k 最近傍法、決定木、SVM などがある。このうち、SVM^{19), 20)} は他のパターン識別手法に対して以下のような特徴を持つ。

- マージン最大化という基準により最適解が一意に定まる
- パラメータの次元数が大きい少数の教示データからでも過学習を起こしにくい
- 最適化すべきパラメータの数が少ない

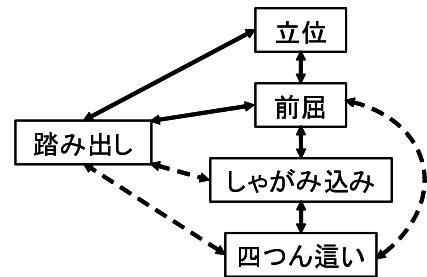


Fig. 3 教示姿勢間の遷移図

IEを行なう際、学習に用いる教示データを多数用意することは改善活動従事者の手間を増やし、改善活動の効率化という目的に反するため、本研究では姿勢分類に用いるパターン識別手法として、少数の教示データからでも過学習を起こさず最適解を求めることができるSVMを用いる。SVMは、線形しきい素子を用いた2クラスのパターン識別手法であり、汎化性能に優れており様々な分野で用いられている。(1)式に示す最小化問題を解くことで最適解 w^* , b^* を導出する。

$$\begin{aligned}
 \underset{w, b, \xi}{\text{Minimize}} \quad & G(w, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \\
 \text{s.t.} \quad & \forall_i, y_i \cdot (w^T x_i + b) - (1 - \xi_i) \geq 0 \\
 & \forall_i, \xi_i \geq 0
 \end{aligned} \tag{1}$$

ここで、 $x_i (i = 1, \dots, n)$ は n 個の学習データ、 w は識別平面の法線ベクトル、 $\xi_i (i = 1, \dots, n)$ は正変数とする。また、右辺第一項はマージンを大きく取るためのものであり、右辺第二項はマージンからはみ出した学習パターンに対するペナルティ項である。係数 C は右辺第一項と第二項のバランスを調整する定数であり、小さくするとマージンが大きくなる。最適な C は実験により決められる。SVMは、カーネル関数を導

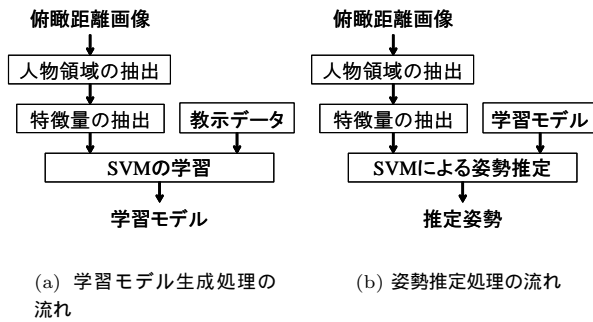


Fig. 4 人物の姿勢分類処理の流れ

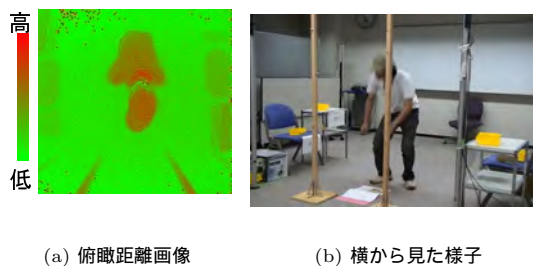


Fig. 5 同一姿勢の撮影方向による見え方の違いの例

入ることにより非線形識別手法に拡張することができ、3.4節で述べるような特徴量に適用できる。また、複数のSVMを組み合わせることにより複数の姿勢に分類するような複数クラス識別問題への対応が容易に行なえる。姿勢推定の処理(Fig.4(b))では学習モデル生成の処理と同様に、俯瞰距離画像に対して背景差分により人物領域を抽出し、人物領域から特徴量を抽出する。SVMの学習により得た学習モデルを用い、教示姿勢のいずれかの姿勢が推定結果として出力される。

3.4 俯瞰距離画像から抽出する特徴量の定義

本節では、作業者の姿勢を推定するために、SVMの学習に用いる特徴量を定義する。SVMは入力される特徴量のうちどの特徴量が重要かをマージン最大化により計算し学習を行なう。そのため様々な特徴量を多く利用したほうが結果が向上するが、特徴量の次元数が大きくなると膨大な学習時間がかかる。そこで本研究では背景差分により得られた人物領域から抽出する特徴量として、人物領域の大きさ S 、重心の三次元座標 G 、人物領域における x, y, z 各軸方向の最小点・最大点の三次元座標 $E_{x_{min}}, E_{y_{min}}, E_{z_{min}}, E_{x_{max}}, E_{y_{max}}, E_{z_{max}}$ 、局所領域の平均距離を定義する。ここで、原点は距離画像センサの光学中心であり、 x 成分と y 成分は床と水平な面上で直交した各方向を表し、 z 成分は距離画像センサの光軸(ほぼ鉛直)方向を表す。距離画像センサは天井に下向きで取り付けられていることから、天井から床に向かう方向を z の正方向とする。以下では、これらの特徴量の詳細について述べる。

人物領域の大きさ

人物領域の大きさ S は、俯瞰距離画像に写っている人物領域の実空間での面積を表し、以下の式で与える。

$$S = \frac{\sum_{i \in N} z_i^2}{n_{max} z_{max}^2} \quad (2)$$

ここで、 i は人物領域 N に含まれる画素を表し、 z_i は $z_{min} \leq z_i \leq z_{max}$ の範囲で与えられる。ただし、最も値の小さい z_{min} は距離画像センサの取り付けられている高さを、最も値の大きい z_{max} は床の高さを表している。また、 n_{max} は距離画像センサの画素数を表している。

重心

重心 $G = (G_x, G_y, G_z)$ は、人物領域における重心の三次元座標を表す。ただし、天井に取り付けられた距離画像センサから見ることで見られる人物の部位は、頭や肩、背中など比較的床から高い位置にある部位が多いため、人物領域の重心の三次元座標 G は、実際の人物の重心よりも高い位置となる。三次元の各座標値 G_x, G_y, G_z それぞれを特徴量とし、以下の式で与える。

$$G_x = \frac{\sum_{i \in N} z_i^2 (x_i - x_{min})}{(x_{max} - x_{min}) \sum_{i \in N} z_i^2}$$

$$G_y = \frac{\sum_{i \in N} z_i^2 (y_i - y_{min})}{(y_{max} - y_{min}) \sum_{i \in N} z_i^2} \quad (3)$$

$$G_z = \frac{\sum_{i \in N} z_i^2 (z_i - z_{min})}{(z_{max} - z_{min}) \sum_{i \in N} z_i^2}$$

ここで、 x_i, y_i はそれぞれ $x_{min} \leq x_i \leq x_{max}, y_{min} \leq y_i \leq y_{max}$ の範囲で与えられ、 $[x_{min}, x_{max}]$ は x_i の値域(距離画像センサに写る左右の端点座標)を示し、 $[y_{min}, y_{max}]$ は y_i の値域(距離画像センサに写る上下の端点座標)を示す。

端点

端点は、人物領域における x, y, z 各軸方向の最小点・最大点の三次元座標 $E_{x_{min}}, E_{y_{min}}, E_{z_{min}}, E_{x_{max}}, E_{y_{max}}, E_{z_{max}}$ であり、各点での三次元座標値を特徴量として用いる。以下では、例として x 軸方向における最小点 $E_{x_{min}}$ および最大点 $E_{x_{max}}$ における三次元座標値を求める式について示す。 x 軸方向の最小点の三次元座標 $E_{x_{min}} = (E_{x_{min_x}}, E_{x_{min_y}}, E_{x_{min_z}})$ は以下の式で与える。

$i_{x_{min}} = \underset{i \in N}{\operatorname{argmin}}(x_i)$ となる $i_{x_{min}}$ について、

$$E_{x_{min_x}} = \frac{x_{i_{x_{min}}} - x_{min}}{x_{max} - x_{min}}$$

$$E_{x_{min_y}} = \frac{y_{i_{x_{min}}} - y_{min}}{y_{max} - y_{min}} \quad (4)$$

$$E_{x_{min_z}} = \frac{z_{i_{x_{min}}} - z_{min}}{z_{max} - z_{min}}$$

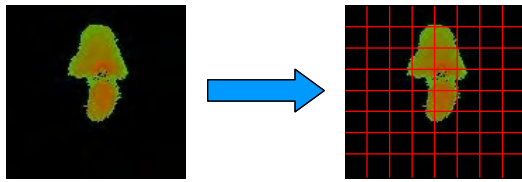


Fig. 6 背景差分画像の局所領域化

x 軸方向の最大点の三次元座標

$\mathbf{E}_{x_{max}} = (E_{x_{max}x}, E_{x_{max}y}, E_{x_{max}z})$ は以下の式で与える。
 $i_{x_{max}} = \underset{i \in N}{\operatorname{argmax}}(x_i)$ となる $i_{x_{max}}$ について、

$$\begin{aligned} E_{x_{max}x} &= \frac{x_{i_{x_{max}}} - x_{min}}{x_{max} - x_{min}} \\ E_{x_{max}y} &= \frac{y_{i_{x_{max}}} - y_{min}}{y_{max} - y_{min}} \\ E_{x_{max}z} &= \frac{z_{i_{x_{max}}} - z_{min}}{z_{max} - z_{min}} \end{aligned} \quad (5)$$

同様に、 y および z 軸における最小点・最大点の三次元座標 $\mathbf{E}_{y_{min}}, \mathbf{E}_{y_{max}}, \mathbf{E}_{z_{min}}, \mathbf{E}_{z_{max}}$ を求める。

局所領域の平均距離

対象作業の変更などによる分類対象となる姿勢の変更に対応するために、一般的な特徴量として局所領域の平均距離を導入する。局所領域の平均距離は、Fig.6 に示すように背景差分画像を一定サイズの局所領域画像に分割し、各局所領域画像において距離の平均値を算出し、これを特徴量とする。本研究では、 128×123 画素の距離画像を 32×32 の局所領域画像群に分割した。

4. 実験と考察

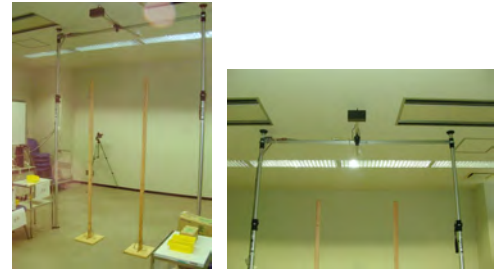
本研究では、まず提案手法の姿勢分類精度を確認するために、人物大の装置の組み立て作業を模擬した予備実験を行ない、つぎに実際の製造工場での組み立て作業に対して提案手法を適用し、改善活動における提案手法の有用性について IER による評価を行なう。

4.1 人物大の装置の組み立て作業を模擬した実験

本実験では、まず SVM の学習で用いるパラメータの最適値を実験的に求める。つぎに、得られたパラメータを用いて複数の被験者の実験データから姿勢分類を行ない、その精度について考察を行なう。

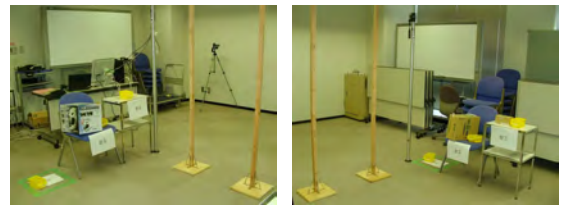
4.1.1 模擬実験の環境と条件

人物大の装置の組み立て作業を模擬した実験環境の様子を Fig.7 に示す。模擬実験で行なう作業は、ボルトとナットの取り付け作業および箱の移動作業とし、全工程が 8 ~ 10 分程度で終わる作業内容とした。ボルトとナットの取り付け作業は、異なる高さに複数の穴を開けた二本の木の棒を 70cm 離して鉛直に立て (Fig.7(a))、指定されたサイズのボルトとナットを設置された部品棚から取り、指定された穴に取り付ける作業である。また、箱の移動作業は、大型・中型・小型の三種類の箱を指定された場所まで移動する作業である。被験者は、これらの作業内容を記した作業工程表を見ながら作業を行な



(a) 模擬実験環境全体の様子

(b) 距離画像センサの設置の様子



(c) 部品棚 (左側) の様子

(d) 部品棚 (右側) の様子

Fig. 7 模擬実験環境

Fig. 8 距離画像センサ²⁴⁾

う。作業の様子を観測する距離画像センサは、床から 266cm の高さに人物を見下ろすように下向きで設置した (Fig.7(b))。ボルトやナット、箱の置かれている棚は、Fig.7(c)、7(d) に示すように三段階の高さで左右に設置した。本実験では、センサの解像度や画角、測距範囲、応答速度などから、Fig.8 に示す松下電工製の EKL3101K²⁴⁾ を距離画像センサとして用いた。

被験者は、第一著者 (被験者 A: 身長 172cm, 作業に慣れている) および他の男性九名 (被験者 B ~ J: 身長 168cm ~ 178cm, 作業に慣れていない) である。このうち、被験者 A ~ E の五名については二回測定を行ない、一回分のデータを学習用、一回分のデータを評価用として用いた。また、被験者 F ~ J の五名については一回測定を行ない、評価に用いるデータとした。

SVM で用いるカーネル関数は、汎化性能が高く調整するパラメータが少ないことから (6) 式で示す RBF カーネルを用いた。

$$K(\mathbf{x}, \mathbf{y}) = \exp(-\gamma \|\mathbf{x} - \mathbf{y}\|^2) \quad (6)$$

また、複数クラスの識別法として、高い識別性能が得られる

ことで知られている one-against-one 法²⁵⁾を用いた。

本実験では、俯瞰距離画像および横からビデオカメラによって撮影したカラー映像を目視により確認し、撮影された俯瞰距離画像の全フレームに対して手動で教示姿勢のいずれかを正解姿勢として与えた。ここで、人物の位置はあらかじめ分かっているものとし、作業者の全身が映っていないフレームは学習・分類の対象外とした。作業中に被験者が取った姿勢には、「前屈」と「踏み出し」を同時に行なっている姿勢(以下、混合姿勢)が見られた。この姿勢においては、「前屈」よりも「踏み出し」の方が「立位」からの重心移動が大きく改善効果の大きいムダと判断し、本実験では「踏み出し」を正解として与えた。SVMの学習は、被験者A~E五名分の学習用データを用いて行なった。学習用データにおいて、姿勢変化の途中で見られるどちらの姿勢か判断が困難な姿勢は学習に用いないこととした。姿勢分類評価には、被験者A~Eの学習用データとは異なるシーケンスの評価用データおよび被験者F~Jのデータを用いた。

4.1.2 模擬実験の結果と考察

本実験ではまず、SVMの学習で用いるパラメータの最適値について検討する。RBFカーネルを用いたSVMの調整可能なパラメータは、はずれ値に対するペナルティ乗数 C およびモデルの複雑さを表すパラメータ γ の二つである。パラメータの最適値は、被験者Aの評価用データを用いて再現率(以下、正解率)を算出することで求めた。ここで再現率を用いるのは、実際の作業において四つん這いなどのムリな姿勢の出現率は非常に低く、このような出現率の低い姿勢の検出を重要視したためである。SVMの学習は被験者A~Eの学習用データを用いるが、各教示姿勢の作業における出現頻度はTable.1に示すように大きく異なるため、本実験では学習に用いる各教示姿勢の教示データのフレーム数を統一して学習を行なった。こうすることで、出現頻度の低い教示姿勢も出現頻度の高い姿勢の影響を受けることなく推定することができ、このことはIEを行なう観点からも望ましい。各被験者から教示姿勢毎に20フレーム(約1.3秒)をランダムで選択し、各教示姿勢につき100フレームの教示データをSVMの学習に用いる。被験者によってシーケンス全体で教示データが20フレーム得られない姿勢があるため、その場合は他の被験者からデータを補う。ここで、実際に改善活動で提案手法を用いることを考えた場合、3.3節でも述べたように学習で用いるデータを多くするには手間がかかり、効率的な改善活動の妨げになるため、本実験では教示データとして用いるフレームの数を少なく設定した。本実験では、SVMの学習で用いるパラメータの最適値を求めるため、作業の様子を撮影したシーケンス全体において、各教示姿勢毎の正解率の平均を(7)式を用いて算出した姿勢平均正解率を用いる。

$$\text{姿勢平均正解率} = \frac{\text{各教示姿勢の正解率の総和}}{\text{教示姿勢数}} \quad (7)$$

Fig.9にパラメータ C および γ の変化に対する姿勢平均正解率の変化の様子を示す。以降の実験では、姿勢平均正解

Table 1 各被験者の評価用データにおける各姿勢のフレーム数

	踏み出し	立位	前屈	しゃがみ込み	四つん這い
フレーム数	642	13065	9385	31661	1889

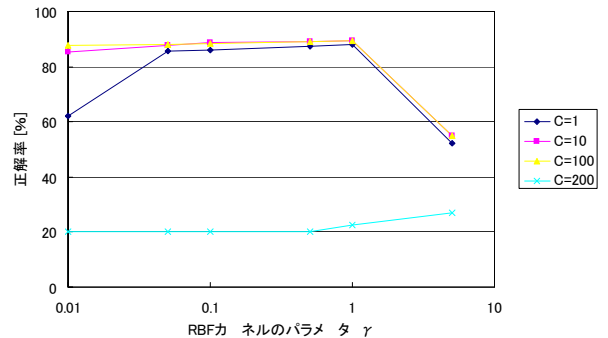


Fig. 9 パラメータ変化に対する姿勢平均正解率の推移

Table 2 全被験者を対象とした各教示姿勢の正解率

	推定姿勢 [frame]	推定姿勢 [frame]					正解率 [%]
		踏み出し	立位	前屈	しゃがみ込み	四つん這い	
正解姿勢 (frame)	踏み出し (642)	561	30	43	0	8	87.38
	立位 (13065)	1721	9658	1632	36	18	73.92
	前屈 (9385)	925	980	6693	731	56	71.32
	しゃがみ込み (31661)	327	346	1754	28309	925	89.41
	四つん這い (1889)	72	3	175	262	1377	72.90

率が最も高くなる $C = 10$ 、 $\gamma = 1.0$ をSVMの学習に用いるパラメータ値として採用する。

つぎに、被験者十名の評価用のデータを用いて提案手法の姿勢分類精度の評価を行なう。各被験者の全工程における姿勢の遷移を色別で表したものをFig.10に示し、Fig.10における被験者Aの工程開始2分から3分の部分を拡大し、グラフ化したものをFig.11に示す。ここで、Fig.10, 11における横軸は時間を表す。また、全被験者を対象とした各教示姿勢における推定された姿勢のフレーム数と正解率をTable.2に示す。Table.2において赤で示された部分は正しく推定されたフレーム数を示し、青で示された部分はFig.3で示した姿勢遷移において対象となる教示姿勢と隣り合う姿勢を示している。

Fig.10では、全工程での姿勢遷移の様子や各教示姿勢の出現頻度、被験者毎の作業に要した時間などを把握することができ、実際の改善活動で改善活動従事者に対して提供する最も基本的な情報とする。Fig.10からでは姿勢の変化の様子を詳しく知ることは困難なため、Fig.11のように拡大して見ることでこれを可能とする。

Table.2において、「踏み出し」「しゃがみ込み」では87%以上の正解率が得られており、「立位」「前屈」「四つん這い」では72%前後の正解率となっている。ここで、各教示姿勢において学習者-非学習者間でt検定を行なったところ、すべて

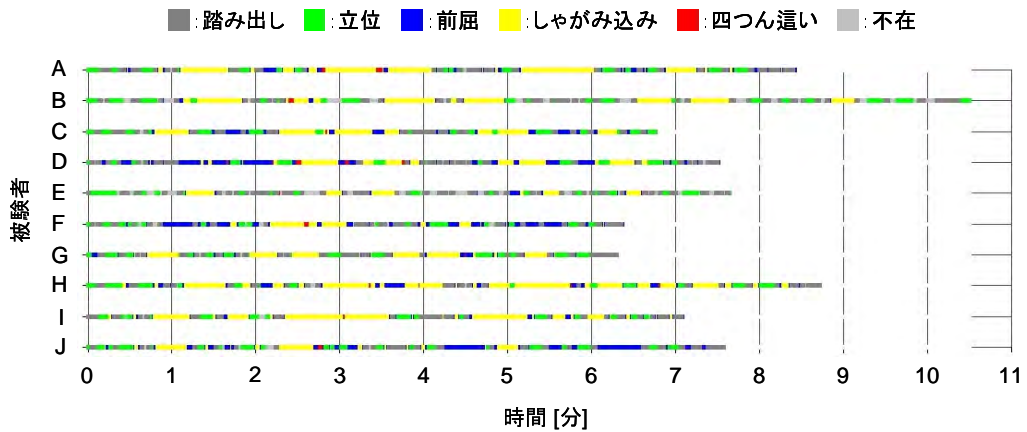


Fig. 10 各被験者の全工程における姿勢遷移

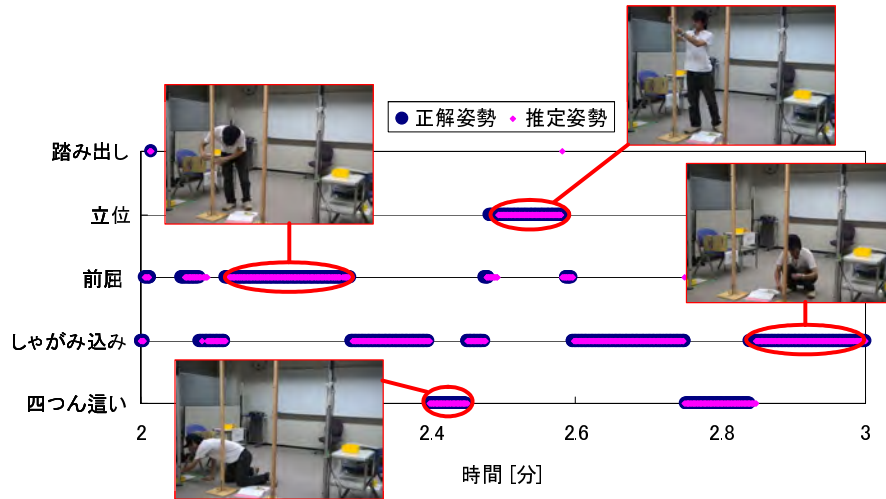
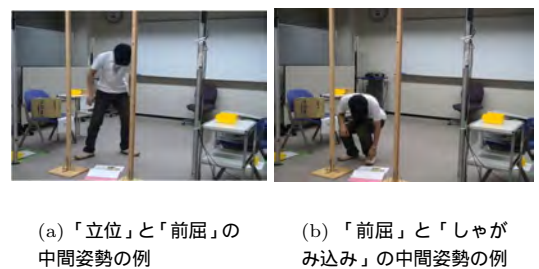


Fig. 11 姿勢遷移の詳細な様子 (被験者 A)

の教示姿勢において学習者-非学習者間に有意差は認められなかった。誤推定されたフレームの多くは青で示された部分に含まれ、正解姿勢と近い姿勢で誤推定が起きていることが示されている。これは、Fig.12 に示すようなふたつの姿勢の中間的な姿勢 (以下、中間姿勢) における誤推定が多く含まれている。また、先に述べた混合姿勢では「踏み出し」を正解として与えているため、これによる誤推定も含まれていると考えられる。これらの誤推定は人が直接見ても判断が難しく、改善活動を行なう上で IER に対してさほど影響を与えるものではないと考えられる。また、本研究で対象とした人物大装置の組み立て作業のようなリードタイムの長い作業では、改善活動そのものがあまり行なわれていないのが現状であり、Table.2 における赤と青の部分を含めた各姿勢の正解率は 80 %を超え、十分な効果があると考えている。さらに、Fig.13 の頭髪部分が赤くなっているように、頭髪部分の距離の誤計測による誤推定が考えられる。本実験で用いた距離画像センサは、4.1.1 でも述べたように近赤外光を対象シーンに照射し、その反射が戻ってくるまでの時間から距離を計測するため、頭髪部分では乱反射などにより距離を正し



(a) 「立位」と「前屈」の中間姿勢の例 (b) 「前屈」と「しやがみ込み」の中間姿勢の例

Fig. 12 誤推定の起こりやすい姿勢

く計測することができない。実際の製造工場では作業中の着帽を義務付けている工場が多く、大抵の姿勢における計測は問題ないが、「前屈」や「四つん這い」など後ろ髪が写る場合は、その部分での誤計測が生じることが考えられるため、このようなノイズへの対策が必要である。

4.2 製造工場における組み立て作業での実験

本実験では、実際の製造工場において組み立て作業を行なっている様子を距離画像センサで撮影し、提案手法により姿勢

