

# 距離画像センサを用いた俯瞰画像による人物の姿勢推定 Posture Estimation from Bird's-eye View of Range Image Sensor

高橋 典宏<sup>†</sup>                      山澤 一誠<sup>†</sup>                      生雲 公啓<sup>‡</sup>  
Norihiro TAKAHASHI    Kazumasa YAMAZAWA    Masahiro IKUMO

野田 賢<sup>†</sup>                      横矢 直和<sup>†</sup>  
Masaru NODA    Naokazu YOKOYA

## 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の改善が常に行われており、作業の効率を分析するうえで作業者の姿勢を知ることは重要である。作業の改善活動をIE(Industrial Engineering)といい、IErと呼ばれる専門家などにより行われている。従来のIEは、IErなどが作業の様子を直接または撮影された映像を観察しながら改善方法の案出を行っており、IErの熟練度に依存している。また、人物大の製品の組立作業など完成までに数日を要する場合、作業のすべてをIErが観察するのは困難である。本研究では、セル生産方式による人物大の装置の組立作業を想定し、IErによる改善活動の支援を目的としている。作業の様子を撮影した映像から作業のムダの改善に役立つ情報を自動で抽出しIErに提示することで、短時間で効果的な改善活動が行えると考えている。改善に役立つ情報であるかを判断するうえで、作業者のとる姿勢の意味を推測することは重要である。人の姿勢とその意味の関連を利用した研究として、公共スペースにおいて人が置かれた状況により取り得る行動はある程度限定されるという考えの下、人のいる場所や姿勢などから人が欲しているサービスを推測しロボットを用いてサービスを提供する研究[1]が行われている。本稿では、改善活動を支援するシステムの一環として、作業のムダと思われる人物の姿勢を距離画像センサを用いて推定する手法を提案する。

従来、製造ラインにおける作業のムダを検出する手法として、作業者の頭部および両腕に取り付けたカラーマーカの軌跡を利用した手法[2]が提案されている。この手法では、繰り返し作業が行われていることを前提とし、作業の繰り返しによるマーカの軌跡の再現性を利用してムダな作業を自動で検出している。しかし、人物大の装置の組立作業など全身動作を伴う場合や、作業に用いる工具や取り付ける部品などの配置が一定でない場合などは繰り返し作業でない場合が多く、このようなシーンへの適用は困難である。また、マーカの装着など作業者に對する拘束が発生している。

人物大の装置の組立作業における作業者の動作には全身動作が多く移動範囲も広いため、監視装置の設置位置は作業に支障を来さない天井などに制限される。そこで本研究では、天井に距離画像センサを取り付けることで上方より作業の様子を撮影する。センサから得られる距離情報から特徴量を抽出し、評価関数を用いて教示姿勢と比較することで作業者の姿勢を推定する。

## 2. 人物の姿勢推定手法

本研究では、撮影環境の制限により人物の上方に距離画像センサを設置する。上方から撮影した距離画像(以下、俯瞰距離画像)の例を図1(a)に示す。距離画像センサより得られる距離情報から2.1節に示す姿勢を反映した特徴量を抽出し、評価関数を用いて教示姿勢に対する入力された俯瞰距離画像中の人物姿勢の評価値を算出する。人物の姿勢は評価値が最小となる教示姿勢として推定する。本研究で対象とする姿勢は、図2に示す「立位」

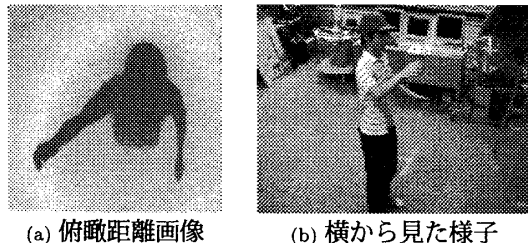


図1: 同姿勢の撮影方向による見え方の違いの例

「しゃがみ込み」「上半身前傾」「四つん這い」「踏み出し」の5つであり、これらを教示姿勢とした。動作経済の原則より、作業を効率的に行うためには身体の動作を末端部位(手指)で行うことが理想とされる。このことより、教示姿勢間の姿勢変化は重心移動を伴うため作業のムダと考えられる。

### 2.1 評価関数の定義

前述した5つの姿勢を推定するための特徴量として、人物領域の“大きさ”、“重心の三次元座標”、“最高点と重心の三次元座標間の差”を俯瞰距離画像から抽出する。まず、人物領域の抽出を行うため入力された俯瞰距離画像に対し背景差分を施し、ノイズ除去のため面積が閾値以下の小領域を除去する。これにより得られる人物領域の画素数を大きさ $S$ として、また、人物領域の重心の三次元座標 $\mathbf{G} = (G_x, G_y, G_z)$ を算出する。人物領域中の最高点は、ノイズが発生することを考慮して実際の最高点から一定の画素を除いた後の人物領域における最高点を $\mathbf{H} = (H_x, H_y, H_z)$ とし、最高点と重心の三次元座標において各成分の差を最高点と重心の差 $\mathbf{H}' = (H'_x - G_x, H'_y - G_y, H'_z - G_z)$ として取得する。また、 $\mathbf{G}$ ,  $\mathbf{H}$ ,  $\mathbf{H}'$ における $x$ ,  $y$ が水平方向、 $z$ が垂直方向を表す。これらの特徴量を用い、教示データとして与えた各姿勢に対する入力された姿勢の評価値を以下の評価関数を用いて算出する。

$$F(S_i, \mathbf{G}_i, \mathbf{H}_i, S_t, \mathbf{G}_t, \mathbf{H}_t) = w_1 S_f + w_2 \cdot \mathbf{G}_f + w_3 \cdot \mathbf{H}'_f \quad (1)$$

ただし、

$$\begin{aligned} S_f &= |S_i - S_t| \\ \mathbf{G}_f &= (|G_{ix} - G_{tx}|, |G_{iy} - G_{ty}|, |G_{iz} - G_{tz}|) \\ \mathbf{H}'_f &= (|H'_{ix} - H'_{tx}|, |H'_{iy} - H'_{ty}|, |H'_{iz} - H'_{tz}|) \end{aligned}$$

ここで、 $S_i$ ,  $\mathbf{G}_i$ ,  $\mathbf{H}_i$ ,  $\mathbf{H}'_i$ は入力された俯瞰距離画像から得られる特徴量、 $S_t$ ,  $\mathbf{G}_t$ ,  $\mathbf{H}_t$ ,  $\mathbf{H}'_t$ は教示データからの俯瞰距離画像から得られる特徴量を示す。また、 $w_1$ ,  $w_2$ ,  $w_3$ は各項への重みである。

### 2.2 人物の姿勢推定

本研究では、2.1節で定義した評価関数により算出される評価値が最小となる教示姿勢を入力された俯瞰距離画像中の人物姿勢と推定する。ただし、作業の効率を分析するうえでの作業者の姿勢推定において、姿勢変化の途中における姿勢は変化前と変化後どちらの姿勢としても捉えることができる。そこで、全教示姿勢のうち二番目

<sup>†</sup>奈良先端科学技術大学院大学 情報科学研究科  
Nara Institute of Science and Technology (NAIST)  
<sup>‡</sup>オムロン(株)技術本部 センシング&コントロール研究所  
OMRON Corporation Sensing & Control Technology Lab.



図 2: 上方から見た教示姿勢

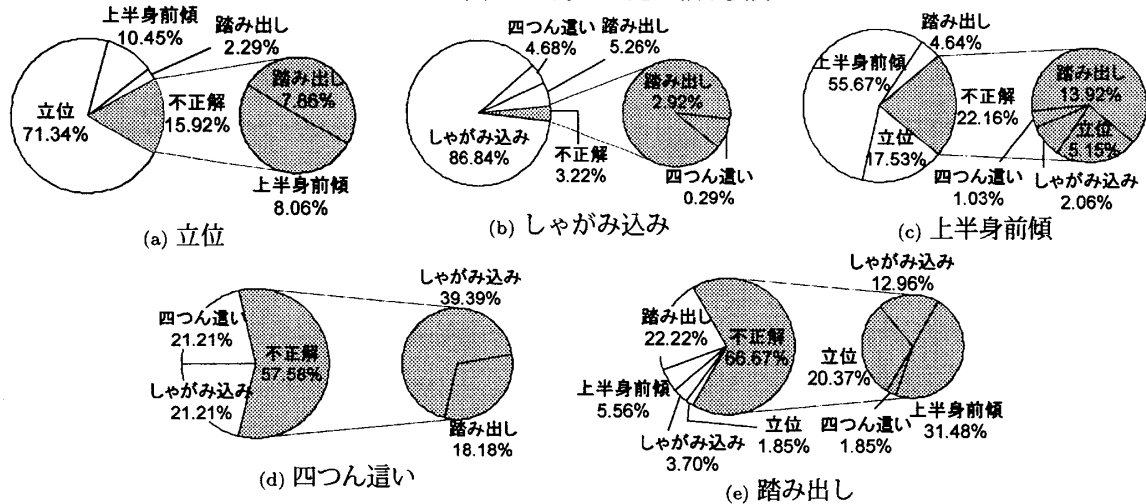


図 4: 正解データに対する人物姿勢の推定結果

に評価値が小さくなる姿勢において、その評価値が最小の評価値から一定の範囲内に含まれる場合にのみ二番目に評価値が小さくなる姿勢についても入力画像中の人物姿勢として推定する。つまり、入力画像中の人物姿勢として推定される姿勢の数は1つもしくは2つとなる。

### 3. 実験

人物大の製品の組立作業を模擬した実験環境で、提案手法の有用性を検証した。実験環境の様子を図3に示す。実験は松下電工製の距離画像センサ EKL3101K[3]を床から253cmの高さに人物を見下ろすように設置した。被験者1名が5つの教示姿勢を含む様々な動作を5分間行い、その様子を上方から距離画像センサにより撮影した。教示姿勢ごとに撮影したデータ中から教示データとして1フレームずつ取得した(図2参照)。ここで、2.1節で示した最高点  $H$  は人物領域において5画素目に高い点とした。また、式1の重みを経験的に  $w_1 = 0.0062$ ,  $w_2 = (0.025, 0.025, 0.20)$ ,  $w_3 = (0.25, 0.25, 0.25)$  とした。入力の全フレームに対し、教示姿勢のうちの1つを手動で与え正解データとし、推定の正解率を求めた。ただし、人物が移動によりフレームアウトする際や立ち位置などにより頭部が写っていないフレームは対象としない。2.2節で述べたように、推定される姿勢は最大で2つとなるので、この場合、推定された姿勢のうちどちらかが正解データと同じであれば正しく推定したものとす。

姿勢推定の結果を図4に示す。図は、手動で与えた正解データに対し評価値が最小となった各姿勢の割合を示す。円グラフ中において白で示された推定姿勢は、正解とされる姿勢が評価値の最も小さい2つの姿勢に含まれているために正解となり、灰色で示された推定姿勢は、正解とされる姿勢が評価値の最も小さい2つの姿勢に含まれておらず不正解であることを示す。白で示された正しく姿勢推定が行われた確立は「しゃがみ込み(図4(b))」において90%以上、「立位(図4(a))」、「上半身前傾(図4(c))」では80%程度となった。しかし、「四つん這い(図4(d))」、「踏み出し(図4(e))」ではいずれも40%程度の正解率であった。これは、入力である俯瞰距離画像における人物の見え方が人物の立ち位置により大きく異なる

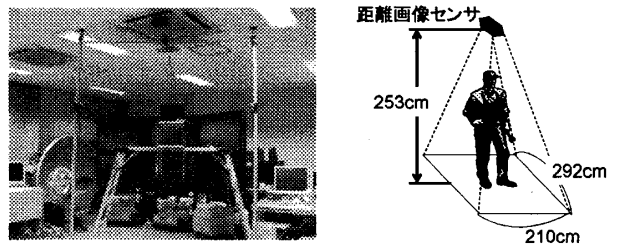


図 3: 実験環境

ためと考えられる。

### 4. まとめ

本稿では、セル生産方式による人物大の装置の組立作業を想定し、IErによる改善活動の支援を目的としたシステムの一環として、作業のムダと思われる人物の姿勢を上方に設置した距離画像センサを用いて推定する手法を提案した。本研究では、「立位」「しゃがみ込み」「上半身前傾」「四つん這い」「踏み出し」を教示姿勢として与え、人物領域の“大きさ”、“重心の三次元座標”、“最高点と重心の三次元座標間の差”を特徴量として評価関数を定義し、入力データとして教示姿勢を含む様々な姿勢を5分間撮影した俯瞰距離画像を与え、人物の姿勢推定を行った。その結果、「しゃがみ込み」では90%以上、「立位」「上半身前傾」では80%程度の確率で正解と同じ姿勢が推定されたものの、「四つん這い」「踏み出し」では高い正解率を得ることができなかった。今後の課題として、重みの最適化を行うなど人物の見え方に依存しない評価関数について検証することなどが挙げられる。

### 参考文献

- [1] 見持圭一, 古結義浩, 宅原雅人: “ネットワークロボットによる人物行動と状況の認識—アンコンシャスロボットとビジュアルロボットとの連携—”, 信学技報, NR-TG-2-09, pp.9-14, 2006-11.
- [2] 板倉豊和, 服部可奈子, 折原良平: “工場作業者の作業軌跡データからのムダ作業発見”, 信学技報, PRMU2006-116, pp.71-76, 2006-10.
- [3] 松下電工株式会社: “距離画像センサ あかりと電気設備 Ebox”, <http://biz.national.jp/Ebox/kyorigazou/index.html>.