

3次元人物姿勢情報を用いた睡眠時行動認識への取り組み

Action Recognition of the Sleeping Situation using 3D Human Pose Information

指田 岳彦* 工藤 雄太** 青木 義満**
Takehiko SASHIDA Yuta KUDO Yoshimitsu AOKI

要旨

近年の高齢化の進行に伴い医療介護施設は増え続けている。そうした施設では入居者の就寝時にスタッフが定期的に見回るなどして事故の防止に努めている。しかし人員不足から見回りの作業負担は非常に大きく、また常時見守っている訳ではないため、事故が起きた時に即発見出来ないこともある。また、基本的に事故は起きてから発見されるため、何故そうした事故が発生したのかを知ることは非常に難しい。その様な状況から、睡眠時のベッド上の事故を認識する見守りシステムが求められている。こうした需要を受け、既に幾つかのベッド見守りシステムが市場に出ているが、それらが認識できるのは主に“ベッド上にいる／いない”の様な状態であり、“座ろうとしている”という様な状態間を繋ぐ“行動”は推定できていない。

行動を推定するためには人の位置や姿勢を捉えることが重要である。しかし、一般に良く知られる人の姿勢推定は監視カメラを前提としているため、ほぼ立位のみを対象としており、重力の制約がなく姿勢の自由度が高いベッド上の見守りには適さない。また、睡眠時には布団や毛布による遮蔽が発生するため、既存の姿勢推定では対応が難しい。

本研究ではこれらの問題を鑑みて、3次元データを用いて姿勢推定を行い、それにより得られる特徴量を用いることでベッド及びその周辺での睡眠時行動を認識することを目指している。

距離画像情報から人の姿勢推定を得ると共に、姿勢情報から特徴量を抽出、Random Forestで学習して行動の認識を行った結果、正答率は姿勢推定で65.9%、行動認識で94.6%となり、行動の境界付近での判断の揺れを考慮し、3フレームの認識ズレを許容した場合、行動認識の正答率は97.8%となった。

これらの技術を用いることで高齢者の怪我の予防や介助者の見守り負担の低減に貢献することが出来ると考えている。

Abstract

As Japanese aging population grows, the number of nursing homes is increasing accordingly. Care workers, however, feel high burdens in nighttime safety confirmation operations, and have troubles in finding accidents immediately because they cannot care the elderly all the time. Also, care workers cannot know the cause of accidents because they find accidents after they happen. Thus, automatic accident monitoring systems are needed to automatically notify accidents around beds to care workers. Although there are some accident monitoring systems are on the market, they can only detect states of whether the elderly is on the bed but actions which involve state transition, such as sitting down. It is important to recognize human positions and poses for detecting actions. As most of human detection algorithms, however, assume that a person is standing, they are not adequate for bed monitoring systems which require no restrictions on human poses.

To solve this problem, we use 3D image data for human pose detection, and then detect human pose by features extracted from the pose information. After these steps, we classify and detect actions using the“Random Forests”algorithm. We achieved a success rate of 65.9% in human pose detection and a success rate of 94.6% in action recognition.

We believe that our work can prevent injury of elderly people and reduce the burdens of scare workers.

*開発統括本部 システム技術開発センター データサイエンス技術室
**慶応義塾大学

1 概要

介護・見守り分野への適用を目的として、“ベッドに座る”、“寝返り”といった、人が寝起きする際の行動の認識に取り組んだ。距離画像情報から人の姿勢推定を得ると共に、姿勢情報から特徴量を抽出、Random Forestで学習して行動の認識を行った。その結果、正答率は姿勢推定で65.9%、行動認識で94.6%となり、行動の境界付近での判断の揺れを考慮し、3フレームの認識ズレを許容した場合、行動認識の正答率は97.8%となった。

2 背景・課題・従来技術

近年、高齢化の進行に伴い、医療介護施設は増え続けている。そうした施設では入居者の就寝時にスタッフが定期的に見回るなどして事故の防止に努めている。しかし、人員不足から見回りの作業負担は非常に大きく、また常時見守っている訳ではないため、事故が起きた時に即発見出来ないこともある。また、基本的に事故は起きてから発見されるため、何故そうした事故が発生したのかを知ることが非常に難しい。そのような状況から、睡眠時のベッド上の事故を認識する見守りシステムが求められている。

こうした需要を受け、既に幾つかのベッド見守りシステムが市場に出ている[1][2]。これらのシステムを用いることで事故や事故に綱が得る行動を検知し通知することは出来るが、認識できるのは主に“ベッド上にいる／いない”の様な状態であり、“座ろうとしている”という様な状態間を繋ぐ“行動”は推定できていない。

より多くの行動を推定するためには、人の位置や姿勢を捉えることが重要である。しかし、一般に良く知られる人の姿勢推定は監視カメラを前提としているため、ほぼ立位のみを対象としており、重力の制約がなく姿勢の自由度が高いベッドの見守りには適さない。また、睡眠時には布団や毛布による遮蔽が発生するため、既存の姿勢推定では対応が難しい。

本研究では、これらの問題を鑑みて、3次元データを用いて姿勢推定を行い、それにより得られる特徴量を用いることでベッド及びその周辺での睡眠時行動を認識することを目指している。

3 提案の流れ

本提案では、Fig. 1 に示した流れで人の睡眠時行動認識を考える。

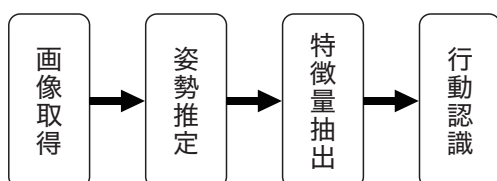


Fig. 1 Outline of our system.

センサーを通じて距離画像を入手。距離画像情報を用いて人の姿勢情報を取得した後、特徴量化し、行動毎の特徴を機械学習によって学習し、推定を行う。

以下で具体的に述べる。

3.1 画像取得

今回、姿勢推定や特徴量として3次元情報を用いる。そのため、画像としては距離画像を用いる。距離画像をとりこむセンサーとしては、Depth Sense社のTOFカメラDS311を用いた。大まかなスペックをTable 1に示す。

Table 1 Spec of Time Of Flight camera DS311.

方式	Time Of Flight (TOF)
距離画像画素数	160*120 (QQVGA)
画角	57.3*42.0
Frame Rate	25 fps (今回の評価では実測16fps程度であった)
測定範囲	1.5m-4.5m
距離精度	< 3cm

見守りの際にカメラを部屋の上部に設置することを想定しセンサー設置高さを250cmに設定、更に寝ているときに人の全身が入る様にベッド中心から230cmの位置にカメラを固定し、睡眠時行動の撮影を行った。Fig. 2にカメラ配置のイメージ、Fig. 3に実際に得られる距離画像の例を示す。

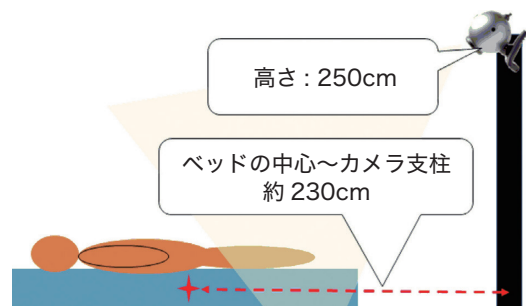


Fig. 2 Outline of camera installation.

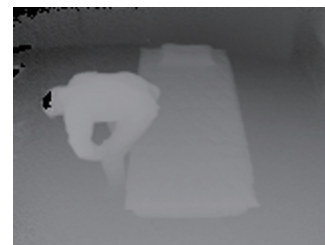


Fig. 3 Depth image.

3.2 姿勢推定

既に述べたとおり、ベッド上での姿勢推定においては既存の立位を前提とした推定技術はそのままでは適用し難い。そのため今回は、工藤らによるベッド上での姿勢認識手法を用いる[3]。同手法では、Oriented Normal Vectors (ONV) 特徴量およびHistograms of ONV (HONV)

特徴量を用いて、人物領域とその姿勢を推定する。同手法ではまずSVMによる頭部検出器を用いて頭部を検出した後、頭部を基点に胴体領域を探索、形状モデルを基に胴体を部位毎に領域分割し、姿勢を求めている。また、同手法ではONV特徴を用いるため、布団などによる遮蔽に対してもある程度の耐性が期待できる。工藤らは同手法により、布団による遮蔽の有無に依らずベッドにおける人物姿勢を70%超の精度で可能としている。

本論文では、同手法を用いて人物の姿勢を推定し、それによって得られる頭部、体幹部、大腿部、下腿部といった人体部位の位置関係を特徴量として用いる。

3.3 特徴量抽出

後述の認識対象行動を区別する特徴量を想定しリストアップを行い、大別すると以下の様なものに分けられた。

- ① 各人体部位の絶対位置
- ② 人体部位やベッド間の相対位置
- ③ 各人体部位の法線情報
- ④ 人体部位間が成す角度
- ⑤ ①～④の時間変化

これらの特徴化し、最終的に125個の特徴量を抽出した。

3.4 行動認識

3.3で抽出した特徴量を学習し、行動の推定を行う識別器を構成する。

認識すべき行動としては、本提案は人が寝起きする際の行動を認識することを目的としているため、寝起きする際に通常起こりうる状態・行動を想定し、Table 2で示す行動を認識対象行動としてリストアップした。

Table 2 List of recognition target action.

無し(無人)	立っている	座る
座っている	寝る	寝ている
寝返り	起き上がる	立ち上がる

これらの内、“立っている”、“座っている”、“寝ている”は、それぞれその状態を保っている事を表している。

一方、“座る”、“寝る”、“起き上がる”、“立ち上がる”はそれぞれ対象の状態への遷移を表している。

一般に、ベッドでの事故は何もしていないのに起こるといことは少なく、例えば起き上がろうとするなど、対象が何らかの行動を起こした際に発生することが多い。そのため、これら遷移行動を認識することは事故の要因を認識することにも繋がり、非常に重要であると言える。

各行動の特徴量を学習する手法として、簡便でありながら高い性能を有するとして、Breimanらに提案されたRandom Forestを用いた[4]。後述する学習用データに特徴量抽出を適用し、結果をRandom Forestに入力し、識別器を構成した。

4 評価

4.1 評価対象

一般の人物行動認識では、既に多様なデータセットが提案されている。しかし、今回対象としている様な睡眠時に特化したデータセットは公開されていないため、自ら作成する必要がある。

今回はDepthカメラによる撮影環境を構築し、入床～離床までの流れを撮影し学習評価画像を作成した。実際の睡眠時に行われる行動を考慮し、以下のような流れで入床～離床までの動作を演じてもらい、撮影を行った。

- ① スタート (人は写っていない)
- ② 画面内に進入
- ③ ベッドの縁に座る
- ④ 身体の向きを変え、寝る
- ⑤ 寝返りを打つ
- ⑥ 上体を起こし、座った姿勢になる
- ⑦ 体の向きを変えベッドの縁に座る
- ⑧ 立ち上がる
- ⑨ 画面外に移動する

Fig. 4 にイメージを示す。

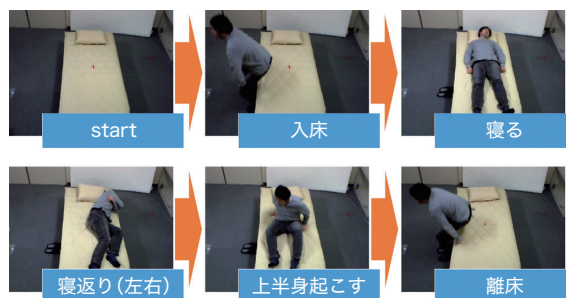


Fig. 4 Outline of test data (Color Image).

得られた画像について、一枚ずつ目視で正解姿勢として人体の各部位を囲う矩形情報と行動のラベルを付与し、学習・評価用データとして用いることとした。最終的に各行動の画像枚数はTable 3の様になった。

Table 3 Amount of test data for individual action.

	学習用	評価用	合計
無人	74	131	205
立っている	170	51	221
座る	76	26	102
座っている	422	130	552
寝る	85	21	106
寝ている	866	792	1658
寝返り	208	223	431
起き上がる	65	14	79
立ち上がる	53	12	65
全体	2019	1400	3419

これらの学習・評価用データおよび同データから抽出される特徴量を用い、姿勢推定・行動認識の評価を行った。

5 評価方法

姿勢推定と行動認識それぞれの評価方法について述べる。評価に用いる画像は、姿勢推定においては、工藤らの発表に準じる形で、推定した各部位の重心座標が正解座標に対し±6画素以内に収まっている場合を正解とし、正答率を求めた。工藤らの手法は頭部推定後に他の部位の位置推定を行うため、頭部以外の部位に関しては頭部推定正解時の正答率という形で求めることとした。

次に行動認識においては認識対象行動毎に正解行動と認識結果の一致率を求めることとした。

6 結果と考察

姿勢推定の結果を Table 4, Table 5 に示す。

Table 4 Accuracy of head detection.

部位	正答率
頭部	65.9%

Table 5 Accuracy of body detection when head detection is succeed.

部位	正答率
体幹部	95.6%
大腿部	63.3%
下腿部	33.2%

結果を見ると、下腿部の正答率が他の部位に比べて劣る。工藤らの手法では人体の各部位をそれぞれ一つの四角形で表現している。しかし人の脚部は2本に分かれており、動きや姿勢の自由度が非常に大きく、一つの四角形で表現することは難しい。用意しているモデルに上手く当てはまらないことが精度低下の原因の一つと考えられる。また、頭部推定後は上半身から順番に各部位の推定を行っていくため、体幹部や大腿部推定の誤差が伝播し下腿部推定に影響を与えている可能性も考えられる。

次に行動認識の結果を Table 6 に示す。Table では横軸に正解行動、縦軸に認識結果をとり、セルには各正解行動と認識結果の組合せについて該当件数を示している。また、最下部には各行動について認識の正答率を示している。

結果を見ると、正解と完全一致の場合で正答率94.6%となっており、ある程度分離できていることが分かる。

個々の行動についてどの様に間違えているのかを見ると、“座っている”と“寝る”、“寝ている”と“寝返り”の様に、ある行動とその前後の状態・行動の間で判定を間違えている場合が多いことが伺える。すなわち、行動と行動の境界を正しく認識できていないと考えられる。

しかし行動と行動の間の境界は非常にあいまいである。例えば、寝ている状態と寝返り状態を明確に線引きすることは人間にとっても難しい。

Table 6 Result of action recognition.

		正解行動								
		無人	立っている	座る	座っている	寝る	寝ている	寝返り	起き上がる	立ち上がる
認識結果	無人	128	0	0	0	0	0	0	0	0
	立っている	0	51	3	0	0	0	0	0	4
	座る	0	0	23	0	0	0	0	0	0
	座っている	0	0	0	130	6	0	0	0	0
	寝る	0	0	0	0	15	0	0	0	0
	寝ている	0	0	0	0	0	761	28	0	0
	寝返り	0	0	0	0	0	30	195	0	0
	起き上がる	0	0	0	0	0	1	0	14	0
	立ち上がる	0	0	0	0	0	0	0	0	8
個別正答率		97.7%	100.0%	88.5%	100.0%	71.4%	96.1%	87.4%	100.0%	66.7%
全体正答率		94.6%								

そうした状況を鑑みて、正解行動との一致率を算出する際、±3フレームのズレを許容し、Nフレーム目の行動がXであったときにN±3フレーム内に行動をXと認識した場合は正解とした場合の値も算出した。なお、3フレームは今回の検討において0.2秒程度に相当する。結果を Table 7 に各フレーム完全一致の場合と合わせて示す。

Table 7 Result of action recognition when we do not allow recognition lag and result of action recognition when we allow 3 frame recognition lag.

行動		完全一致		3frame内一致	
		サンプル数	正答数	正答率	正答数
無人	131	128	97.7%	131	100.0%
立っている	51	51	100.0%	51	100.0%
座る	26	23	88.5%	26	100.0%
座っている	130	130	100.0%	130	100.0%
寝る	21	15	71.4%	18	85.7%
寝ている	792	761	96.1%	777	98.1%
寝返り	223	195	87.4%	211	94.6%
立ち上がる	14	14	100.0%	14	100.0%
起き上がる	12	8	66.7%	11	91.7%
全体	1400	1325	94.6%	1369	97.8%

±3フレームのズレを許容した場合、正答率97.8%となり、認識ミスの件数は完全一致の場合で75件だったのが31件となった。このことから、完全一致の場合の認識ミスの半数以上は行動間の境界付近で起きていることが分かる。

個々の行動を見ると、“寝る”の正答率が±3フレームのズレを許容した場合でも他に比べて低い。確認した所、“寝返り”から“寝る”へ移行した際に3フレーム以上、“寝返り”と間違えていることが分かった。

今回の検討で“寝返り”と“寝る”を目視で見分ける際には、前フレームと比較して体幹部の重心移動があったかどうかで判断している。今回行動認識に用いた特徴は体幹部の重心位置のみではないため、認識結果のズレに関しては、体幹部以外の動き特徴などから判断している可能性がある。

7 まとめ

介護・見守り分野への適用を目的として、寝ている人物の姿勢推定と行動認識を試みた。行動認識では通常の睡眠時に発生する行動を姿勢情報に基づく特徴量から認識可能であることを示した。

今後の展開として、行動毎の観測数の差や正解行動の客観的定義の確立といった学習・評価用データへの対応や異常動作の検出、ベッド以外も含めたより複雑な行動の認識といった課題に取り組んでいきたいと考えている。

●出典

本稿は画像センシング技術研究会“第20回画像センシングシンポジウム”論文集の予稿を加筆修正して転載したものである。本稿の著作権は画像センシング技術研究会が有する。

●参考文献

- [1] 東芝情報システム株式会社見守りシステム
<http://www.tjsys.co.jp/page.jsp?id=3072>
- [2] NKワークス見守りシステム
<http://www.nk-w.co.jp/products/medical/mimamori.php>
- [3] 工藤雄太, 指田岳彦, 青木義満: “ベッド上の行動認識に向けた距離画像を用いた姿勢推定”, DIA2014 IS4-3 (2014)
- [4] Leo Breiman, “Random Forests”, Machine Learning 45 (1): 5-32 (2001)