

博士学位論文

輝度勾配自己相関特徴を用いた
起床動作の推定に関する研究

2016年

井上 円

目次

第 1 章	序論	1
第 2 章	研究背景	4
2.1	社会背景.....	4
2.2	看護現場の課題.....	7
2.3	転倒事故の法規的責任.....	10
2.4	問題の解決手段.....	11
2.4.1	ナースコールシステムとの連携.....	11
2.4.2	システムの制約とコストの制約.....	12
2.4.3	センサデバイスの選定.....	13
第 3 章	行動推定と画像の特徴記述	17
3.1	行動推定の変遷.....	17
3.2	行動推定の先行研究に対する考察.....	22
3.3	画像特徴記述と識別手法.....	24
3.3.1	人検出.....	24
3.3.2	HOG 特徴.....	25
3.3.3	フレーム間差分特徴.....	28
3.3.4	立体高次局所自己相関 (CHLAC).....	33
3.3.5	Boosting を用いた識別手法.....	35
第 4 章	実験データ	37
4.1	被検者の活動状況に関する調査.....	37
4.2	実験設備.....	38
4.2.1	撮影施設.....	38
4.2.2	撮影機材.....	39
4.3	データ項目.....	42
4.3.1	カメラ位置.....	42
4.3.2	照明条件.....	48
4.3.3	装飾.....	49
4.3.4	行動パターン.....	51
第 5 章	輝度勾配自己相関特徴を用いた転倒防止のための起床動作検出	54
5.1	まえがき.....	54
5.1.1	提案手法.....	55
5.2	起き上がり動作検出実験.....	60
5.2.1	実験概要.....	60
5.2.2	実験データ.....	60
5.3	低照度環境における評価実験.....	64
5.3.1	実験概要.....	64
5.3.2	実験データ.....	64
5.3.3	実験結果.....	64
5.4	考察.....	65
5.5	むすび.....	67
第 6 章	カメラ設置範囲にロバストな起床動作検出	68

6.1	まえがき	68
6.2	輝度勾配ベクトルの畳み込み	70
6.3	多重解像化	74
6.4	各手法に対する性能比較実験	76
6.4.1	実験概要	76
6.4.2	実験データ	76
6.4.3	カメラ位置に対する検出性能評価	77
6.4.4	動作識別性能評価	78
6.4.5	多重解像に対する検出性能評価	78
6.5	考察	79
6.5.1	カメラ位置評価における検出性能低下	79
6.5.2	カメラ位置に対する汎化性能	80
6.5.3	動作識別性能	81
6.5.4	多重解像の効果	83
6.5.5	誤検出について	86
6.6	むすび	87
第7章	おわりに	88
	謝辞	90
	参考文献	91
	研究業績一覧	103

第1章 序論

2013年9月15日時点の統計によれば，我が国の総人口1億2,726万人に対して高齢者と呼ばれる65歳以上の人口は3,186万人に上り，総人口に占める高齢化率は25.0[%]に達した[1]．更に国別に見ても高齢化率，高齢化速度，平均寿命の全てにおいて日本が1位を占めており[1]-[3]，世界的にも突出した高齢化社会と言われている．

現在のペースで高齢化が進むと十数年後には医療施設の供給能力が需要を上回り，全ての国民が医療や介護を享受する事が出来ない社会になると言われている．この社会変化への対応として，医療にまつわる仕組みを効率化して限られた医療設備で多くの患者の処置を行う事は非常に重要な課題である[4]．

一方，医療処置の過誤や不備が発生した際に，医療従事者に問われる責任は大きい．処置の過失だけでなく看護中の事故や怪我に対しても医療従事者の責任が問われる．そのため医療業務を行う際は事故要因の大きい問題に対して常に重複した確認を求められ，医師や看護師への大きな負担となる．

医療事故報告事例集[5]によれば医療事故の上位3位は，予薬（注射・内服），転倒転落，チューブ・カテーテル類が占める．予薬とチューブ・カテーテル類は事故の主体が看護者にあり，防止策は作業者の事前確認の強化と考えられる．しかし，転倒転落事故については事故の主体が患者側であり，看護者は巡回を増やす事で事故の発生を確率的に低減させる事しか出来ない．この問題の対策として転倒転落の予兆を検知するセンサ類が利用されているが，誤報や失報の頻度が高いため効果を上げていない．加速度センサ，RFID，緊急通報用ボタン等は電池を使用するために定期的なメンテナンスを必要とする等の課題を含んでいる．また，機器の電池切れの際に転倒が発生すると深刻なアクシデントに繋がる．さらに，高齢者はしばしばセンサや通報用ボタンなどの装着を忘れる．転倒ボタンに関しては転倒後に高齢者が意識を失うケースでは役に立たない．また，マットセンサやベッドセンサ等の圧力感知式センサも広く用いられているが，インシデント管理の点から看護介護現場では転倒事故の原因に繋がる配線やケーブル類は望ましくない．更に，ベッドセンサのケーブルは頻繁に断線

が発生する為、各社が断線対策として製品に断線時のアラート機能などを搭載しているが、断線修理のコストを低減する事は出来ない。

画像処理技術の進歩により、前述の問題を改善する手法が提案されている。動画から部分隠れと姿勢の多様性を考慮した転倒検知手法や[6]、単眼カメラを用いて頭部軌跡を3次元で追跡し日常動作から転倒を見分ける[7]等である。しかしながら、これらの手法は転倒自体を検出するアプローチであり、転倒そのものを防止することはできない。

Brown医師の調査によれば、頻尿や夜間頻尿、またはトイレに急いで駆け込むような際に骨折を伴う転倒が発生する[8]。最も頻繁に転倒が発生するシーンは、病室に備えられた簡易トイレからベッドに戻る際や、ベッドからトイレに移動する際であるとされている。つまり、患者がトイレに行く際のベッドからの起き上がり動作を検知し看護者に通知する事で、転倒を効果的に防止する事が出来る。そこで本研究は、カメラで撮影した画像から起床動作を検出する手法を提案する。

人物の行動を推定する処理は、画像の中から人物の動きの変量を抽出する処理と、抽出した変量から動きの種別を推定する処理の組合せにより構成される。人物等の画像から抽出された変量は一般的に特徴量と呼ばれ、動きの種別を推定する処理は、機械学習を用いた識別器で画像特徴を解析する手法が用いられる。とりわけ、特徴量の記述手法と特徴量に対する識別手法は行動の推定精度を大きく左右する。ベースとなる特徴量については、物体認識で効果が確認されている形状特徴や追跡処理で効果が確認されている動き特徴が、行動推定に対しても効果を示している。このように形状を表現する特徴と時間の変量を表現する特徴を組み合わせた特徴記述を用いる事で、時間に対する変位と空間に対する変位を同時に表現する時空間特徴が生成される。近年では、検出精度の向上や行動に対する汎化性、または環境に対する順応性を追求した特徴記述手法に関する研究が良くなされている[9]-[14]。本研究では看護現場の転倒転落という実社会の特定の課題を解決する事を目的とし、固定カメラを用いた実用的な起床動作検出を実現する特徴記述について検討する。

本稿では、第2章で研究背景として社会背景や看護現場の課題、さらに課題を解決する為の最適な運用システムについて述べる。第3章では行動推定の従来研究に対するサーベイを行い、起床動作の推定に対する最適な手法について述べる。第4章では起床動作検出の実利用を予定する運用環境について述べる。設備

の制約条件と検出すべき起床動作等の被写体条件を定義し，それらの設備環境や被写体条件を再現した室内空間で実験データを構築した．なお，4章で構築した実験データは以降の運用評価実験の際に用いる．第5章では起床動作検出の特徴記述手法の提案とその検討結果について述べる．第6章では，起床動作検出の実運用についての議論を行う．5章で構築した検出原理に対する課題と改善策を提案し，実運用面における評価と考察を述べる．第7章は結論であり，本研究についてまとめる．

第2章 研究背景

本章では、起床動作の検出を研究テーマとした背景について述べる。さらに背景にある問題を解決する上で最適な製品形態について述べる。

2.1 社会背景

厚生労働省の統計[15]によれば、2005年(平成17年)を境に国内の人口増加率は減少に転じたが、死亡者数は2011年(平成23年)の120万人から23年後の2038年(平成46年)までに現在比1.5倍の170万人に増加すると試算されている(図2.1)。

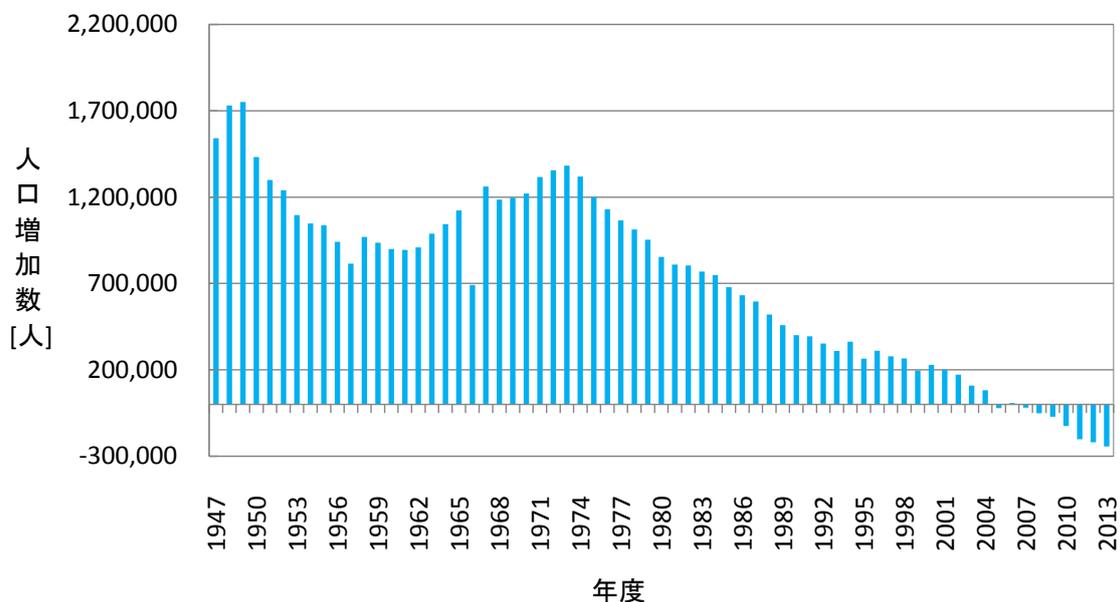


図 2.1 年間人口増減数の推移

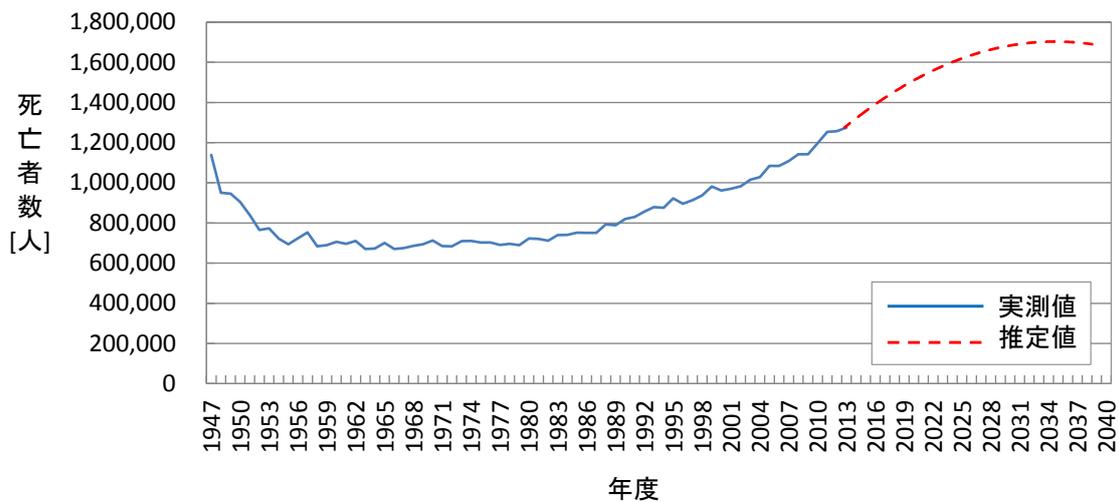


図 2.2 死亡数の年次推移

また、医療機関で死亡する割合は1976年（昭和51年）に自宅で死亡する者の割合を上回り、近年では8割を超える水準に達する（図2.2）[16]。死亡場所と死亡者数の推移を図2.3に示す。

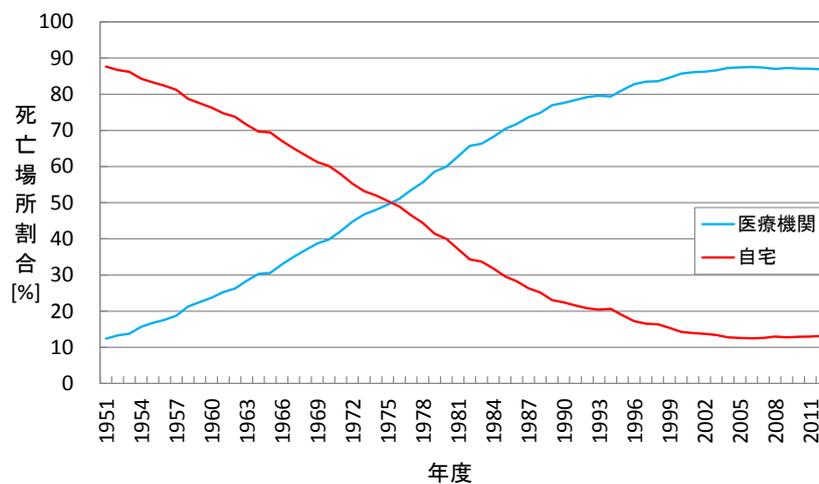


図 2.3 医療機関における死亡割合の年次推移

こうした状況の中、厚生労働省は国による医療保障費の増加は実施せず、医療体制の充実には地域医療制度に依存する方針を打ち出している。OECD(Organisation for Economic Co-operation and Development: 経済開発協力機構)諸国における医療費対GDP比率で諸外国と比較した場合でも、国内の医療保障比率は低い[17]。図2.4に国別の医療費対GDP比率を示す。

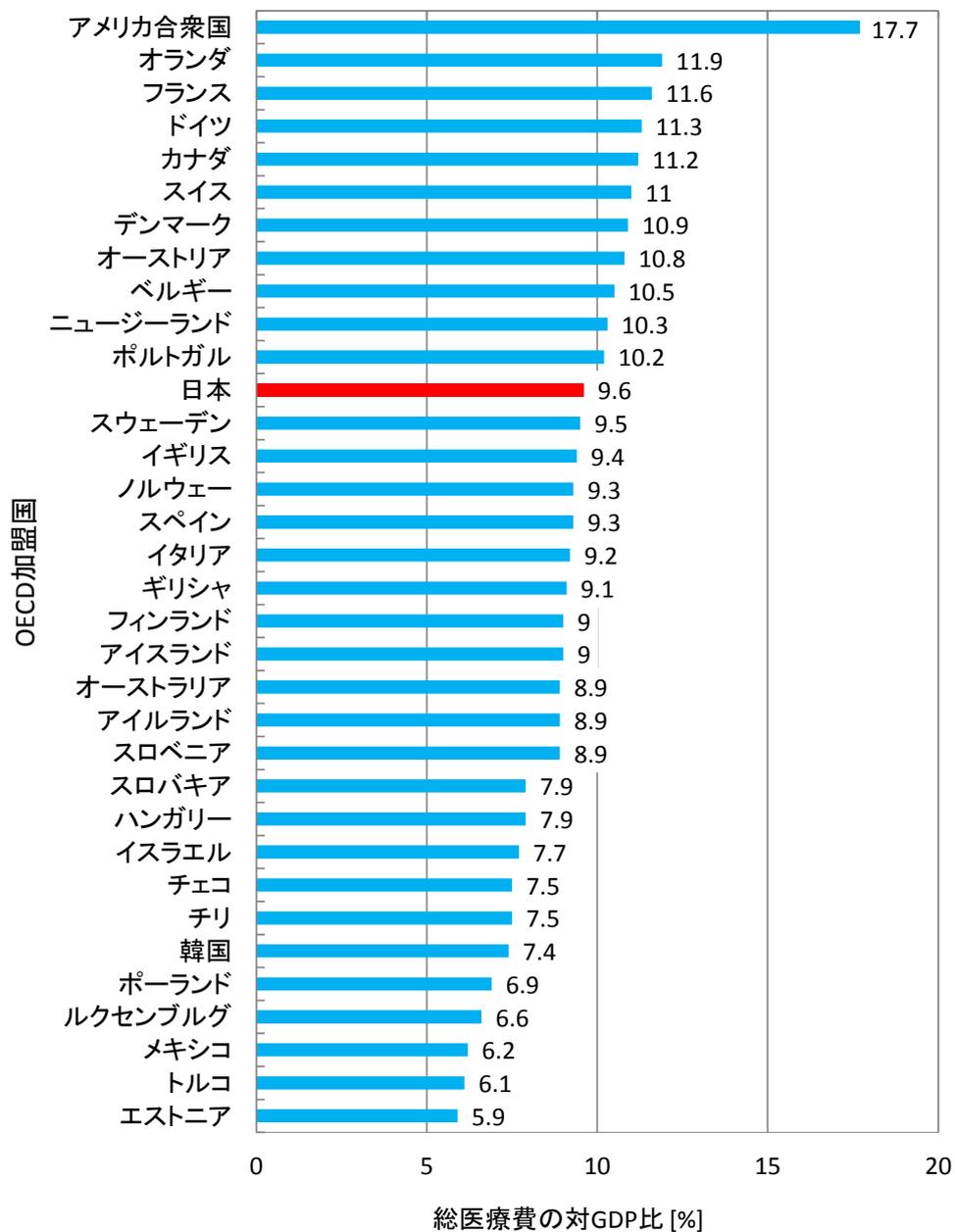


図 2.4 対GDP総医療費（OECD加盟国）

以上の統計を表 2.1 にまとめる。死亡者数と終末期患者は今後増加するが、医療に対する財政支出は減少する事が分かる。病院で死亡する患者数の増加・医療に対する財政支出の削減から、今後国内の地域医療は患者数の増加に伴う効率化のニーズが高まると予想される。とりわけ 2020 年頃迄の今後 7～10 年間に作業効率を向上させる看護・ケア機器の需要が増加すると予想される。

また、2012年に内閣府が発表した医療イノベーション5か年計画[18]では、高齢者や介護現場のニーズに応えるロボット技術の開発及び実用化の促進が謳われている。

表 2.1 背景まとめ

項目	傾向
死亡者数	↗
病院の終末期患者数	↗
医療に対する財政支出	↘

2.2 看護現場の課題

病院や高齢者施設等の研究報告によれば、65歳以上の約30[%]、また老人ホームや高齢者施設では毎年50[%]以上が転倒し、そのうちの半数以上は繰り返し転倒を起こしている。さらに10人に1人は、大腿骨の骨折やその他の骨折、硬膜下血腫などの深刻な怪我を起こす(図2.5)。病院における転倒については多くの機関から報告がなされており[19]-[24]、転倒の発生因子や要因に対する解析も多数報告されている[25]-[36]。

患者の転倒を防止する為に病院では様々なセンサが利用されている。特定の患者の転倒だけを防止する為に、多くのセンサは特定の患者のベッドと組み合わせて運用され、ベッド上の患者の離床を通知する機能を備える。

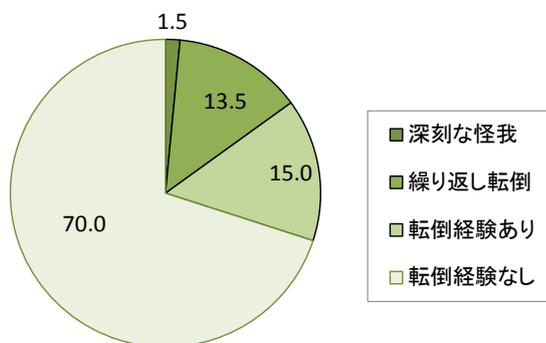


図 2.5 高齢者の転倒発生と症状

転倒防止の主要製品を図 2.6 に示す。ベッドセンサ・マットセンサは加圧センサが内蔵されたシートで、患者の荷重の有無により離床を通知する。しかし、ベッドセンサは患者の寝返りを離床として通知する誤報の発生頻度が高く、マットセンサは看護師への通報を嫌がる患者がマットセンサを避けて離床するケースによる失報が発生する。赤外線センサは、ベッドの側面の移動物体を検知して発報するが、看護師の介助動作や患者の動きによる誤報が発生する。クリップ型センサは、一端をベッドに固定し、残りの一端を患者の衣類に固定し、患者が起き上がる際にクリップが外れる事で、看護師への通知が行われる。しかし、患者の寝返り動作による誤報が発生する。携帯型呼出しボタンは、患者が携帯することで異常の際に通知を行う事が出来るが、呼出しボタンの携帯を忘れるリスクが存在する。パラマウント社が開発した加重センサ式ベッドは、前述のデバイスが課題とした誤検出や未検出を低下した効果的なデバイスである。しかし、一度離床が発生した場合、検出設定を再実行させる必要がある等の設定が複雑という問題と製品価格が高額という問題を持つ。ここ迄に述べたように、看護現場では問題とされている転倒に対して、現在のところ効果的なデバイスは開発されていない。



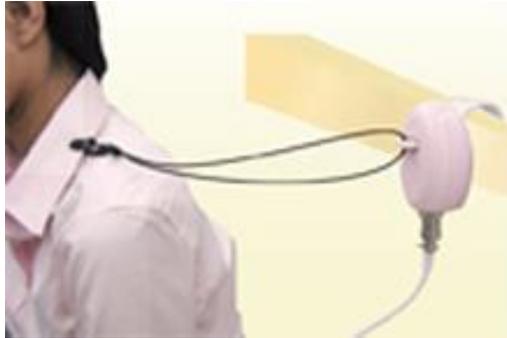
(a) ベッドセンサ[37]



(b) マットセンサ[37]



(c) 赤外線センサ[38]



(d) クリップ型センサ[37]



(e) 携帯型呼び出しボタン[39]



(f) 加圧センサ内蔵ベッド[40]

図 2.6 転倒転落防止の既存製品

2.3 転倒事故の法規的責任

転倒事故が発生した際に、入院中に神経疾患等により運動器不全の状態で行介助等の看護を受けている患者に対して病院側は安全を配慮すべき診療契約上の義務を負う事になる。患者のベッド転落や廊下での転倒により骨折等が発生した場合は、療養上の指示や指導の不十分さが不法行為や債務不履行の問題となる[41][42]。医療行為により患者が被害を受けた場合、病院が責任を問われる法的根拠として、契約の債務不履行による損害賠償請求（民法第 415 条）と不法行為による損害賠償請求（民法第 709 条）が挙げられる。

契約の債務不履行による損害賠償請求について詳しく述べる。診療契約は、患者が疾病の治癒を求め医師が治療を引き受ける事により成立するが、疾病を必ず治癒させることを約束できない医療行為においては、仕事の完成を約束する請負契約（民法第 632 条）を締結する事が出来ない。法律的には診療契約は、医師が医学上の高度な専門知識と技術により患者に代わり、その疾病を治療する事を委任される委任契約（民法第 643 条）であり、法律行為以外の事務の処理を委託する準委任契約（民法 656 条）として考えられている。また、裁判例でも診療契約の内容を「善良な管理者の注意を以て原告の眼病の治療に当たる債務を負担する旨の準委任契約」[43]や「患者と医師が締結する診療契約の法律的性質は、患者の疾病の病的症状を医学的に解明しその症状に従い治療行為を施すことを内容とする準委任契約である」[44]と判示している。従って、医師が診療契約により負担する診療債務の法律的性質は手段債務[45]となり、契約関係に起因する債務不履行が成立すればその責任を負う。判例では責任を任意に主張して損害賠償請求が出来るとしている。

次に不法行為による損害賠償請求について述べる。民法 709 条は「故意又は過失に因りて他人の権利に侵害したる者は之に因りて称した損害を賠償する責に任ず」とし、医療行為に過失（医療水準上要求される注意義務違反）が存在する場合に損害賠償を請求できるとされている。

従って、患者やその遺族は、損害賠償請求の訴えを診療契約の債務不履行に基づき請求する事も、不法行為に基づく損害賠償を請求する事も、またはその両方を択一的に請求する事も出来る。

2.4 問題の解決手段

2章2節で述べたように、これ迄のところ高齢患者の転倒を防止する有効なシステムは提案されていない。そこで、看護現場の設備について検討し有効な転倒防止の手段を検討する。

2.4.1 ナースコールシステムとの連携

看護現場の病室には一般的にナースコールシステムが備えられている。ナースコールシステムとは、病院や介護施設等の病室、浴室、ならびにトイレに設置された呼出ボタン(ナースコール子機)とナースステーションに設置された表示機(ナースコール親機)を接続したシステムである(図 2.7, 図 2.8)。

患者が呼出しボタンを押すことで、ナースステーションに設置されている表示機を呼び出すことができる。表示機(ナースコール親機)には各所の呼出ボタン(ナースコール子機)が設置されている場所の名称が表示されており、ランプ状態でナースコールの呼出元を確認し、通話機能により遠隔で患者の容態を確認する事が出来る。



図 2.7 ナースコール機器

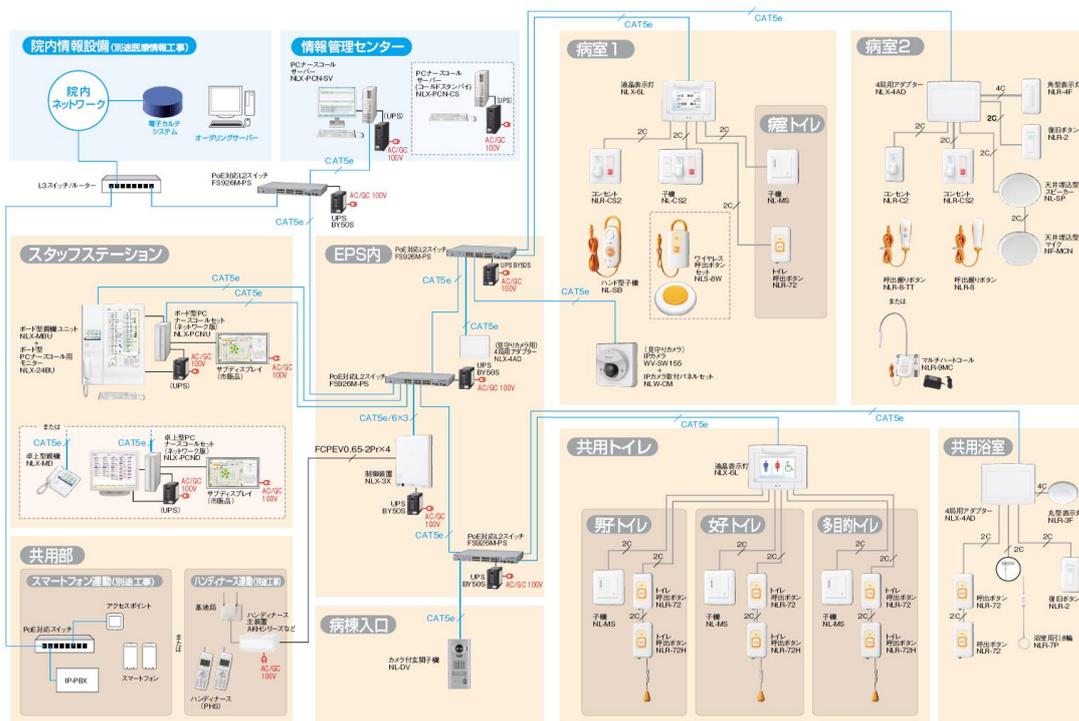


図 2.8 ナースコールシステム接続図

近年のナースコールシステムは幹線がネットワーク化され、機器間で様々なデータを送受信する事が出来る。そのためナースコール子機は呼出しボタンの機能だけではなく、情報表示パネル、生体情報モニタ、カメラ等の付加機能を持つ機器と連動する。従って、このシステムを有効に活用すると、ナースコールシステムの幹線にIPカメラを付加し、カメラからの映像を用いる事で起床動作を検出する事が出来る。

2.4.2 システムの制約とコストの制約

ナースコールシステムに設置されたIPカメラで起床動作を検出する際のシステム上の制約について検討する。はじめにシステム内での起床検知を実行する装置について検討を行う。現在市販されているネットワークナースコールシステムの仕様[46]によれば、一つのシステムでカバーできる病床数は最大240床である。仮に1割の病床の患者が転倒の危険を含む患者である場合、カメラの設置台数は24台となる。起床検知をサーバ上で処理させる場合、常時24台のカメラ映像に対する処理が必要となるが、サーバの処理能力を通常のPC程度と仮

定した場合、サーバ上で画像処理を行い実時間で起床動作を検出することは現実的でない。また、処理能力の高いサーバを用いる場合は費用が高額になり、1台 25 万円程度のナースコール親機で構成されるナースコールシステムのコストを考慮するとユーザに受け入れられるとは考え難い。従って、起床動作の検出装置はネットワークの末端に位置するカメラ側に設置するシステムとなり、カメラから受信した映像に対して逐次検出処理を行い、ネットワークを通じて映像と検出判定結果をサーバに通知する装置となる。(図 2.9)

次に、検出装置のコストから処理能力の上限を検討する。通常ナースコールに付属されるネットワーク上の各機器は 8 万円から 12 万円程度の機材となる。この価格帯で検出処理装置の製造コストを逆算すると、CPU の処理能力上限は 500[MHz]程度と試算される。従って、本研究では 500[MHz]程度の CPU で動作する起床動作検出の研究を行う。

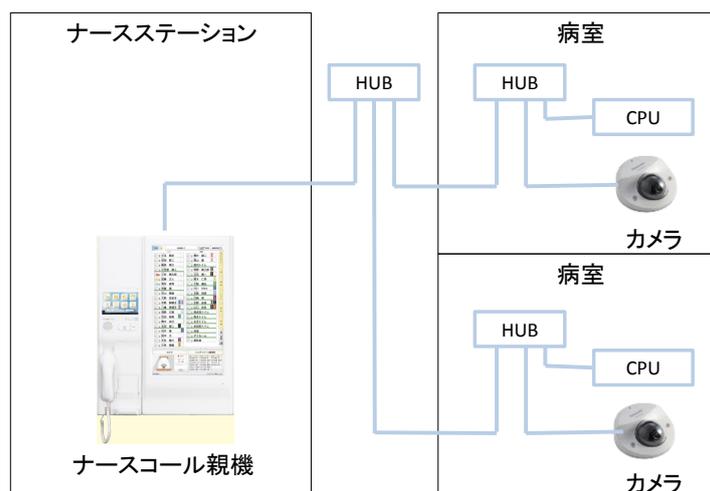


図 2.9 機器構成図

2.4.3 センサデバイスの選定

本研究では、撮像された画像を解析する事で入院患者の起床動作を検出する。撮像デバイスには可視カメラを用いる。可視カメラを用いた場合、室内空間の情報はアレイセンサの平面に透視投影された 2 次元空間の輝度情報となり、画像解析処理にはフォトメトリックな手法を用いる。一方、近年では Microfoft 社

の Kinect[47]に代表されるような距離画像センサが普及している。距離画像センサは、図 2.10 に示す様にセンサから被写体までの距離を画像と同様の XY 平面に記録するセンサである。物体の見えだけでなく、位置関係の様なジオメトリックな手法を組み合わせた処理が行えるため、動画シーケンスから行動を推定するのに効果的な結果をもたらしている[48][49]。ただし、距離画像デバイスはシステムの最終形態を考慮した際に課題を含む。ここでは、その課題について述べる。

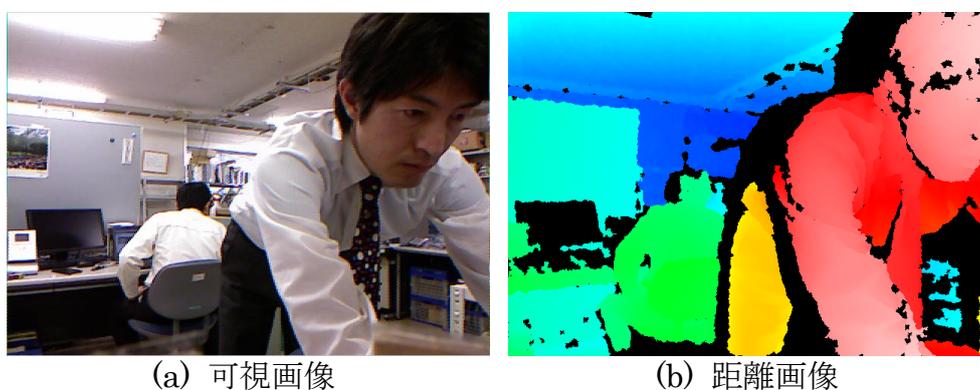


図 2.10 Kinect for Windows による深度センサの撮像例

製品形態を考慮した際の距離画像デバイスの課題は以下の 4 項目である。

- (1) デバイスの原理
- (2) 導入コスト
- (3) 画角
- (4) 製品供給期間

(1) デバイスの原理

現在距離画像デバイスは様々な製品が発売されている(表 2.2)。製品の測距方式は Time of Flight (ToF) 方式, Light Coding 方式, 三角測距方式の三種類に大別される。ToF 方式は距離計測の光レーザ法のひとつであり, 光源から出た光が対象物で反射しセンサに届くまでの時間と, 光の速度($3 \times 10^8[m/s]$)から距離を算出する。light coding 方式はアクティブステレオ法のひとつであるが, パターン化された光を対象物に照射し, センサで得られた画像からパターンのひずみを解析する事で距離を測定する。三角測距方式は複数台のカメラの視差を利用して距離を計測する。上述の測距方式で光レーザ法とアクティブステレオ法を

用いるデバイスは赤外線を照射するが、病院で使用する際に赤外線の照射が問題視されるケースがある。Kinect はセンサ前方に赤外波長近辺の非可視光をランダムな微小ドットパターンで撒布し、投影されたパターンの反射を赤外線カメラで検出する[50]。そのため、赤外波長を照射するデバイス部分は赤くちらついて見える。また、SR4000 は ToF 方式により赤外線を照射するが、Kinect 同様に赤外線照射部が赤く光る。一方、転倒のリスクが高いとされる高齢患者は、せん妄や痴呆、また精神的に不安定な症状を示すことが多い[51]。それらの症状は、不眠症や天井の模様が虫に見えるなどの幻覚を引き起こすことが報告されている[52][53]。本研究を始めるに際し、複数病院へのヒアリング調査を行った結果、発光するデバイスの導入に対しては病院側から押し並べて否定的な意見が寄せられた。また、赤外照射を用いるデバイスは直射日光下における運用の課題があり、窓際の病室等の直射日光が差し込む場所では多くのデバイスが距離情報を取得できない。

(2) 導入コスト

距離画像デバイスの赤外光による課題は既に述べたが、ToF 方式を用いるデバイスでも NipponSignal 社のアンフィニソレイユの様に直射日光下での使用が可能であり、赤外照射部も発光しないデバイスが存在する。また Point Grey 社の Bumblebee の様に三角測距方式のデバイスはパッシブステレオ方式を用いるため、赤外照射の課題を含まない。これらのデバイスは病院での運用に対して非常に親和性が高いが、両者はデバイス単価が数十万となる。2.4.2 項でシステムのコストについて述べたが、センサデバイス 1 台のコストが数十万円に上るとシステム全体のコストに対する不整合が生じる。

(3) 画角

第 4 章にて詳しく述べるが、病室には多くの機材が設置されるため、センサデバイスの設置位置は厳しく制約される。また、設置位置は病院毎に異なる為、病院毎の様々な設置位置に対してベッド周辺の映像を取得する必要がある。その為、センサの水平画角は 100 度程度の広角が必要となるが、現行製品の水平画角は 60 度～70 度である。

(4) 製品供給期間

起床動作検出の搭載を予定するナースコールシステムは想定耐用年数が十数年と長期であり，ナースコールシステムに搭載するセンサデバイスには製品の長期安定供給が要求される．一方，距離画像センサは製品の歴史が浅く，Kinect for Windows や 3D Image Sensor の様に短期間で製品供給が終了するリスクが存在する．

以上の課題を踏まえた場合，現在市販されている距離画像センサは有用な情報を出力するセンサでは有るが，病室内での使用を想定した場合の要求事項を満たすことが出来ない．

表 2.2 測距方式と製品

測距方式	製品	製造元
Time of Flight	Kinect for Xbox One	Microsoft 社
	3D Image Sensor	Panasonic 社
	SR400	Mesa 社
	ZC-1000	Optex 社
	DS311	Softkinetic 社
	アンフィニソレイユ	NipponSignal 社
Light Coding	Kinect for Windows	Microsoft 社
	Prime Sensor	Prime Sense 社
三角測距方式	Bumblebee	Point Grey 社

第3章 行動推定と画像の特徴記述

人物の行動を推定する為には，画像の中から人物の動きを抽出する処理と，抽出した動きの種別を推定する処理の二つの処理が必要となる．人物の動きから抽出された変量は一般的に特徴量と呼ばれる．また，動きの種別を推定する処理は特徴量を用いて識別器を構築する手法が用いられる．本章ではこの二つの処理の理論について述べる．

3.1 行動推定の変遷

Viola らが 2001 年に発表した Haar-like 特徴量と AdaBoost 識別器を用いて画像の中から顔を高速に検出する手法[54]や, Daral らが 2005 年に発表した Hog 特徴量と機械学習を用いて人を検出する手法[55]により，画像から物体を認識する研究が注目を集め，物体認識の研究が増加した．物体認識の研究をさらに時空間に発展させた研究，つまり画像中の人等の時間毎の変化を捉える事により対象の動作を理解する試みは，2000 年代から盛んに行なわれている．動作の理解を大別すると図 3.1 に示す 3 つの研究に分類できる．即ち，集団や群集の動きに対する理解，個人の行動に対する理解，体の部位の動作に対する理解である．

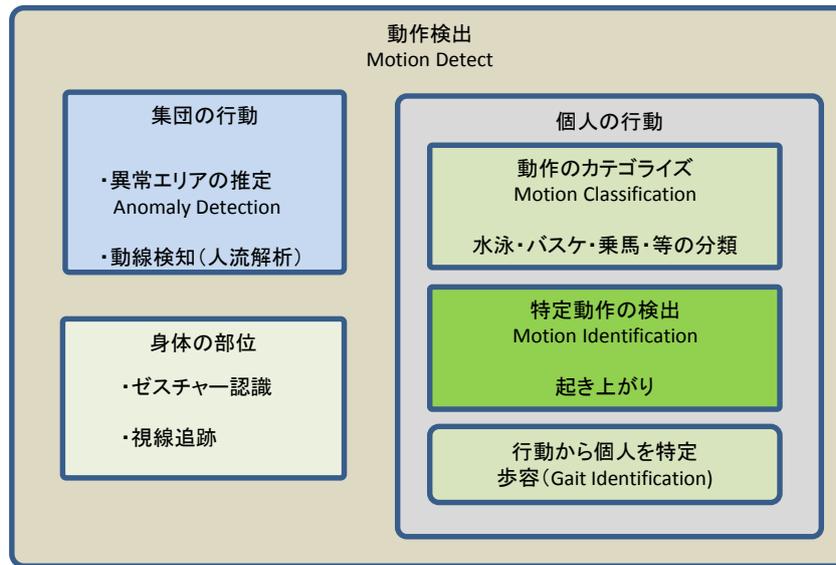


図 3.1 人の行動理解の分類

集団の理解については、異常が発生したエリアの推定(Anomaly detection in crowded scene)や、動線の検出等が研究されている。異常エリアの推定は、駅のコンコースやイベント会場の様な広域エリアを撮影したカメラから、不自然な人の流動や滞留を検出する[56]-[62]。Adam らはオプティカルフローの強度と方向を用いて公共エリアの異常動作検出[56]を行い、南里らは CHLAC と線形判別分析を用いて歩行中の人物の異常動作を検出した[57]。Kratz らは Cuboid と呼ばれる時空間の局所勾配情報と HMM[63]を組み合わせるラッシュ時の地下鉄の改札やコンコースの異常を検出した[60]。Mahadevan らは局所領域に分割されたエリアから時空間特徴を算出する MDT (mixtures of dynamic textures) を用いて混雑シーンの異常を検出した[61]。Kim らは時空間マルコフランダムフィールドを用いて人流の逆行検出を研究[64]した。

行動の推定には様々なアプローチが研究されている。動画の人物から直接行動を推定する方法としては、トラッキングや姿勢推定、時空間の形状テンプレート等が用いられる[65]-[69]。Efros らは 30[pix]程度の低解像度の人物の行動を認識する為にオプティカルフローから得られた時空間特徴を用い、バレー、テニス、サッカーの行動を識別した[70]。Schüldt らは局所時空間特徴と SVM を用いて、Walking, Jogging, Running, Boxing, Hand waving, Hand clapping の 6 種類の動作に対するの識別性能を評価した[71]。Schüldt らが評価に用いたデータセット (KTH Dataset) [71]は屋内、屋外、スケール変化、異なる衣装を組み合わせ

せたシーンで構成されており，行動推定の代表的なデータセットとして 2000 年代に頻繁に性能評価に利用された（図 3.2）.

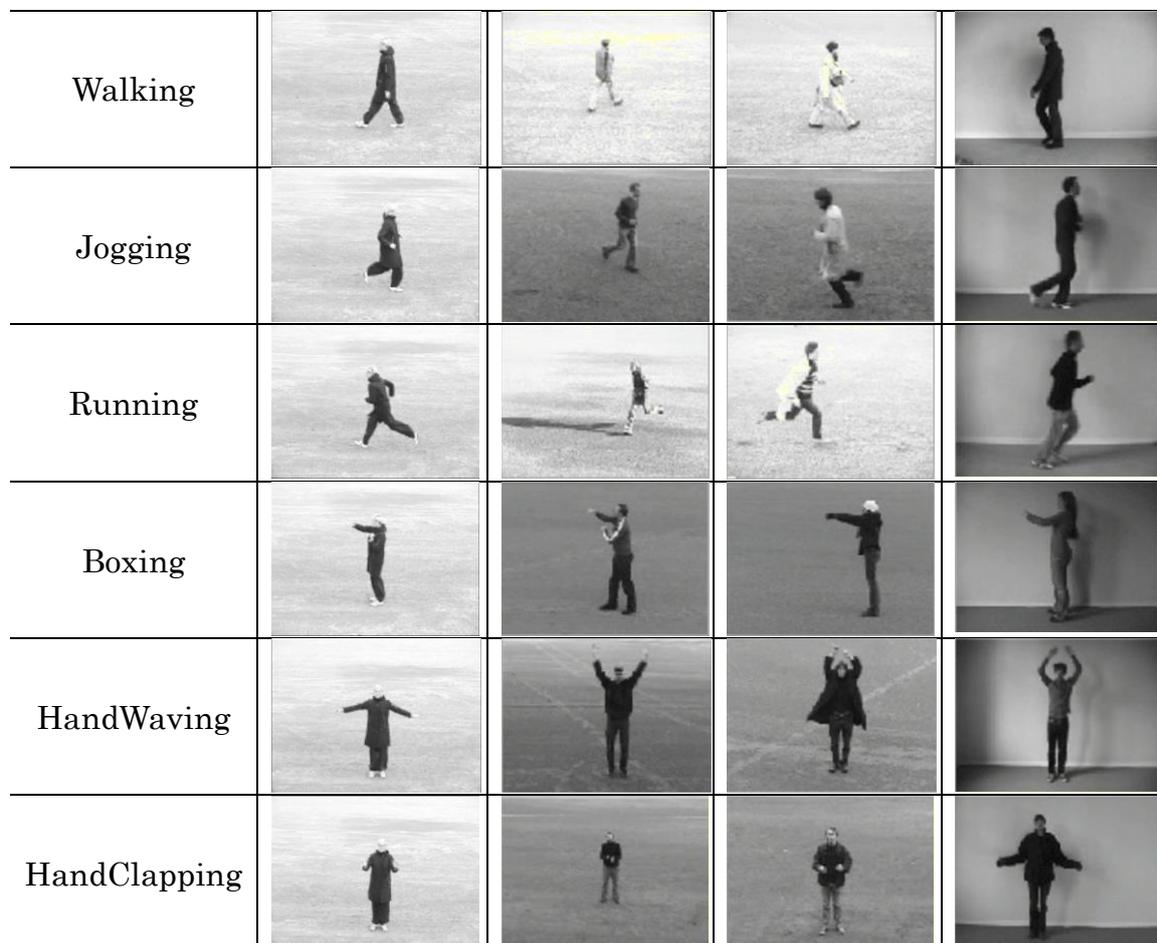


図 3.2 KTH Dataset サンプル [71]

Dollár らは Harris らのコーナ検出 [72] の制約を緩和し空間情報のガウスフィルタと時間情報のガボールフィルタを組み合わせた Cuboid という特徴記述と主成分分析を組み合わせた時空間情報を用いて KTH データセット [71] に対する識別能力を評価した [73]. Kim らは，アフィン変換に対する不変性を備えたテンソル正準相関解析を行動解析に用いた [74]. Nowozin らは，サブシーケンスの時空間特徴を量子化する事で，密な特徴を取り扱う手法を提案した [75]. Gilbert らは時空間データの X-Y 成分，X-t 成分，Y-t 成分に対して検出した Harris コーナ [72] にスケールと方向を組み合わせた特徴記述を行い，マーケットバスケット分析 [76] を用いて識別する手法を提案した [77]. Savarese らは spatial-temporal corre-

lograms と名付けた時間方向に長い時空間上の動き特徴を提案した[78]. Zhang らは、動きの重心からの角度と距離に基づいて算出された、スケールと方向に依存しない MotionContext (MC) 特徴を提案した、更に確率的潜在意味解析法 (pLSA) [79]を MotionContext 特徴に合わせて3次元の潜在変数を扱う処理に拡張した $\omega^3 - pLSA$ を提案した[80]. Bregonzio らは、特徴点の特性に注視されていた従来研究とは異なる切り口として、時空間上の着目点の分布のみに着目し、Dollár らのガボールフィルタを2次元拡張した特徴記述[73]を応用した識別手法を提案した[81]. Willems らはヘッセ行列を用いて時空間特徴を記述し KTH データセットの動作を識別した[82]. Gilbert らは Harris オペレータを用いて算出した着目点と近隣領域の着目点をグループ化した特徴を用いて識別器を構築した[83]. Gillbert らの手法は密な特徴を用いて高速高精度な検出を実現する.

また、動画像から得られた特徴を bag of keypoints[84]の手法を用いて、特徴の位置関係を削減したベクトル記述で表現し動作カテゴリの識別を行う手法が提案されており、近年の行動識別で効果を上げている. Niebles らは Dollár らの特徴[73]を bag of keypoints で記述した教師なし学習の手法を用いて KHT データセット[71], WEIZMANN データセット[68]の動作カテゴリを識別した[85]. しかしながら Niebles らの手法は入力画像データの時空間特徴を明確には記述する事が出来ない. そこで Scovanner らは bag of keypoints を用いて 3DSIFT 特徴をベクトル化する事で詳細な時空間特徴記述を行い、WEIZMANN データセット[68]の動作カテゴリを識別した[86]. Laptev らは隠れや複雑背景シーンにおける歩行者の検出手段として、Harris オペレータ[72]を用いて着目点を算出し、時空間上の着目点同士から HOG やオプティカルフローを計算する事で時空間特徴を抽出し、bag of keypoints でベクトル化した結果を SVM で識別する手法を提案した[87]. Kläser らは HOG を時空間に拡張した特徴を bag of keypoints で記述した[88]. Liu らは、Harris-Laplacian(HAR), Hessian-Laplacian(HES), ならびに MSER Detector[89]で得られた着目点を SIFT で記述し、bag of keypoints でベクトル化した特徴を AdaBoost で識別することで、これまで研究されていないハンドセットカメラ映像を用いた、カメラの視点、背景、隠れ、被写体のアピアランス、照明が著しく変化する映像に対する行動推定を研究した[90]. Marszalek らは HoG, HoF, ならびに SIFT 各々の特徴に対して VisualWord を作成し、SVM 識別器を構築し、行動の推定だけでなく動作からそのシーンを推定する手法を提案した[91]. Han らは着目した人物の特定しがたい行動が含まれた未編集のビデオの行動解析と

いう難しい課題に対して, bag of keypoints を用いて動作に関連したシーンそのものを検出し, 更に Multiple Kernel Gaussian Process Classifier (MKGPC)を用いて低レベルと高レベルのカーネルを結合する事で学習精度を向上させた識別器を提案した[92]. Kovashka らは, HOG3D, HoF, ならびに HOG を組み合わせた VisualWord を階層的に組み合わる特徴を, SVM で識別する手法を提案した[93]. 公開データセットで評価された代表的な研究の検出性能を表 3.1 に纏める.

表 3.1 関連研究の検出性能 [%]

	KTH	HoolyWood2	UCF	YouTube	HMDB51
Schuldt[71]	71.72				
Dollar[73]	81.18				
Niebles[85]	81.50				
Kim[74]	95.33				
Nowozin[75]	84.72				
Wong[94]	86.62				
Fathi[95]	90.50				
Gilbert[76]	89.92				
Klaser[88]	91.40				
Laptev[87]	91.80				
Mikolajczyk[96]	93.20				
Savarese[97]	86.83				
Uemura[98]	93.67				
Willems[82]	84.82				
Zhang[80]	91.33				
Zhao[99]	91.17				
Bregonzio[81]	93.17				
Gilbert[83]	94.45				
Han[92]	94.10	42.12			
Liu[90]	93.80			71.20	
Marszalek[91]		35.10			
Wang[100]		47.70	85.60		

Brendel[101]				77.80	
Ikizler-Cinbis[102]				75.21	
Kovashka[93]	94.53		87.27		
Ullah[103]		55.70			
Bhattacharya[104]				76.5	
Gall[9]	93.50		86.60		
Le[10]	93.90		86.50	75.80	
Yuan[11]	93.70				
Guha[12]			83.80		
Wang[13]	95.30	59.90	89.10	85.40	
Jiang [14]		59.50			
Mathe[105]		61.00			
Kliper-Gross[106]					29.20
Sadanand[107]	98.20		95.00		26.90
Vig[108]		60.00			
Jain[109]		62.50			52.10
Ji[110]	90.20				
Wang[111]		64.30			
Hoai[112]		73.60			65.90
kantorov[113]		58.20			
Sapienza[114]	96.76	57.01		80.06	31.09

3.2 行動推定の先行研究に対する考察

Laptev らが提案したヘッセオペレータを用いる手法[87]は、それ以前の行動推定手法と比較して高い識別能力を示した。また Schüldt らが Laptev らの手法で得た注目点に bag of keypoints を用いた手法を提案した事で、特徴点抽出と bag of keypoints を組み合わせた手法も盛んに研究された。また Liu らの様にヘッセオペレータの代わりに SIFT を用いた特徴点[90]も多く研究されている。この様に

近年の行動推定は、着目点の変位を時空間に適したの画像特徴で表現し、**bag of keypoints** で動作の分布を生成した特徴を **SVM** の様な識別器で分類するという手法がよく研究されている。

しかし、近年の手法は各過程における計算量が非常に大きい。阿倍ら[115]は **SIFT** 特徴の計算負荷を Intel 製 Core2Duo2.66[GHz]のプロセッサを用いた実験により示したが、**QVGA** サイズの画像から 3762 個の特徴点を抽出する際に、全体の処理時間はおよそ 2[s]を要し、その 75[%]は特徴点の方向推定と局所特徴量の抽出が占める事を明らかにした。更に Marszalek ら[91]は、**SIFT** 特徴はシーンの識別能力に優れており、人物の行動推定には **HOG** や **HoF** 特徴量が優れている事を実験により明らかにした。従って、本論文の起床動作検出に関する研究も **SIFT** の様な特徴点ではなく、**HOG** 特徴量を用いる手法が適していると考えられる。また Han ら[92]や Heng ら[116]が述べるように、近年の行動推定の研究はハンディカムで撮影されたホームビデオの様にカメラの視点が移動する映像に対する検出性能の向上に注力しており、固定カメラで撮影した動作の解析に対しては処理が過剰であると考えられる。この過剰な処理とは具体的にはオプティカルフローの算出に当たる。Heng らは移動カメラに対する特徴点の追跡を、階層構造を用いた密なオプティカルフローで表現する事で、行動推定の検出精度を向上させている。しかし固定カメラを用いた場合、動きの発生源は行動を行う人物に限定されるため、高速で処理可能な手法に置き換える事が可能と考える。

ここ迄に述べた行動推定の研究の多くは複数の動作カテゴリを含んだデータセットを用いて動作カテゴリを正確に識別する研究であり、マルチクラス識別に関する問題解決と言える。しかし本研究は起床動作のみをシーンの中から認識する 2 クラス識別の問題に該当する。この二つの違いを静止画での物体認識に置き換えると、先行研究は画像中の物体のカテゴリを問う一般物体認識に相当し、本研究は画像の物体が任意の物体であるか否かを問う特定物体認識に相当する。柳井[117]は **bag of keypoints** の適性について、特定物体認識では数百から数千の **Visual-Word** を用いるのが適切であるのに対し、一般物体認識は数万から百万程度の **Visual-Word** が必要になると述べている。仮に数十万オーダーの **Visual-Word** を用いた識別器を作成した場合、検出処理に要する時間が非常に大きくなると考えられる。従って、先行研究における **bag of keypoints** をクラス分類するという処理は本研究では適さない。しかし、局所的な時空間特徴をヒス

トグラムの分布で表現するという方法は先行研究において効果が示されている [85]-[89]. 一方, 局所的な時空間特徴をヒストグラムの分布で表現した特徴は **bag of keypoints** だけでなく, 同等の手法が小林らにより **CHLAC** 特徴として提案されている [118][119]. **CHLAC** は局所エリアの変位をヒストグラムに投影し, 入力画像全体の時空間の変位を一つのヒストグラムで表現する特徴記述である. 大津らが立体高次局所自己相関と名付けたこの手法は, 時空間特徴として画素の変位を用いる部分が先行研究 [85]-[89] の特徴記述とは異なるが, ヒストグラムの分布を用いた特徴記述は **bag of keypoints** と同等の働きを持つ.

以上を踏まえると, 本論文で研究対象とする固定カメラを用いた 2 クラスの行動識別を高速に実施するには, **HOG** 特徴を基に構成された時空間特徴の自己相関をヒストグラムで記述した手法が有効であると考えられる. そこで **HOG** 特徴の輝度勾配情報から自己相関を算出し画像中の動作を識別する手法として輝度勾配差分自己相関を提案し, 本論文でその効果について述べる. **HOG** 特徴を時空間に展開する手法としてはフレーム間差分特徴 [120] を参考とする.

また本研究では識別器に Liu らが研究で用いた [90] AdaBoost を用いる. AdaBoost は識別境界を線形で分離するが, 学習段階でデータの重みを調整する事により特徴空間上で非線形な境界を生成する. これにより識別時は線形関数の階層的な計算で非線形な境界に対する識別を行う事が出来る.

3.3 画像特徴記述と識別手法

行動推定の基礎となる画像特徴の記述と識別処理について述べる. また画像特徴を体系的に整理するために人検出における画像特徴について述べる.

3.3.1 人検出

一般的に画像の中から人を見つける方法は, 画像から抽出した特徴に対して人らしさの程度を評価する手法が用いられる.

画像の特徴を抽出する方法は, 時刻経過と共に変化する画素の輝度変化を基に算出する手法と, 任意の画素と近隣画素との輝度変化を基に算出する方法に大別される.

時刻経過における画素の輝度変化を抽出する手法としては、隣接フレーム間の輝度の変化を物体とみなして抽出するフレーム間差分、基準となる画像を背景として設定し、入力画像と背景画像の画素の変化を物体とみなして抽出する背景差分、また局所領域の画素の変位を抽出するオプティカルフロー[121]と呼ばれる動きのベクトル等がある。フレーム間差分や背景差分は実装が容易であるが、動き変化を全て対象物体として扱う事や、照明変化を物体として捉える問題がある。また動きベクトルは人の動作を詳細に抽出する手段として適しているが処理時間が大きい。近年ではフレーム間差分特徴[120]の様にフレーム間差分を人の検出に特化させた特徴も提案されている。

近隣の画素との差異から物体の形状を直接的に抽出する手法としては、キーポイントと呼ばれる特徴点を抽出する SIFT[122]や SURF[123]がある。また画像のエッジから物体の形状を特徴として抽出する手法として、HOG (Histogram of Oriented Gradient) [55]や、HOG 特徴を改良した CoHOG[124]や Joint-HOG[125]等がある。算出手段毎の画像特徴を図 3.3 に示す。

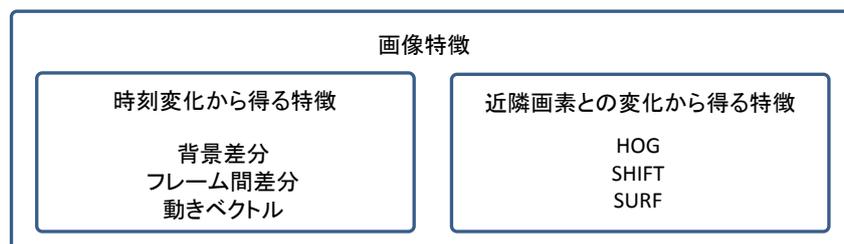


図 3.3 人検出の主な特徴の分類

3.3.2 HOG 特徴

Dalal らにより提案された Histograms of Oriented Gradients (HOG) 特徴は、画像をセルと呼ばれる微小領域に分割し、各セル内の画素から算出した輝度勾配の強度を勾配の方向毎にヒストグラム化した画像特徴である。輝度勾配の角度をヒストグラムの階級に分け、各階級における輝度勾配の強度の総和を階級毎のスコアとする。このとき、画像から算出する勾配の角度は 0 度以上 180 度未満の範囲を取るが、画像特徴を算出する際にはヒストグラムを用いて 20 度毎の 9 方向や 45 度毎の 4 方向等の任意の方向数に集約する。局所領域毎に勾配の強度を算出する構造と、強度を算出する際に近隣の強度に対して行われる正規化

処理により幾何学変化や証明変化の影響を受け難い特性がある．算出手順を以下に示す．

はじめに，各画素 (x, y) に対する勾配の角度 $\theta(x, y)$ と強度 $m(x, y)$ を算出する．

$$m(x, y) = \sqrt{d_x(x, y)^2 + d_y(x, y)^2} \quad (3.1)$$

$$\theta(x, y) = \tan^{-1} \left(\frac{d_y(x, y)}{d_x(x, y)} \right) \quad (3.2)$$

ただし， $d_x(x, y)$ ならびに $d_y(x, y)$ は隣接する画素に対する輝度差であり，画素 (x, y) の輝度値を $I(x, y)$ とすると，輝度差である $d_x(x, y)$ ならびに $d_y(x, y)$ は次の式で算出される．

$$d_x(x, y) = I(x+1, y) - I(x-1, y) \quad (3.3)$$

$$d_y(x, y) = I(x, y+1) - I(x, y-1) \quad (3.4)$$

次に算出した勾配強度 m と勾配方向 θ を用いて，式 (3.5) よりセル領域 c ($p \times p$ ピクセル) における量子化勾配方向 θ' の勾配強度の和を算出する．

$$v_c^H(\theta') = \sum_x \sum_y m(x, y) \delta[\theta', \theta(x, y)] \quad (3.5)$$

δ は Kronecker のデルタ関数であり，勾配方向 $\theta(x, y)$ が量子化勾配方向と同じヒストグラムの要素である場合は 1 を返す．式 (3.5) より算出した量子化勾配方向 θ' における勾配強度の和の集合を， N 方向の勾配方向ヒストグラム $V_c^H = \{v_c^H(1), v_c^H(2), \dots, v_c^H(N)\}$ とする．任意のセルからヒストグラムを算出する過程の概念を図 3.4 に示す．

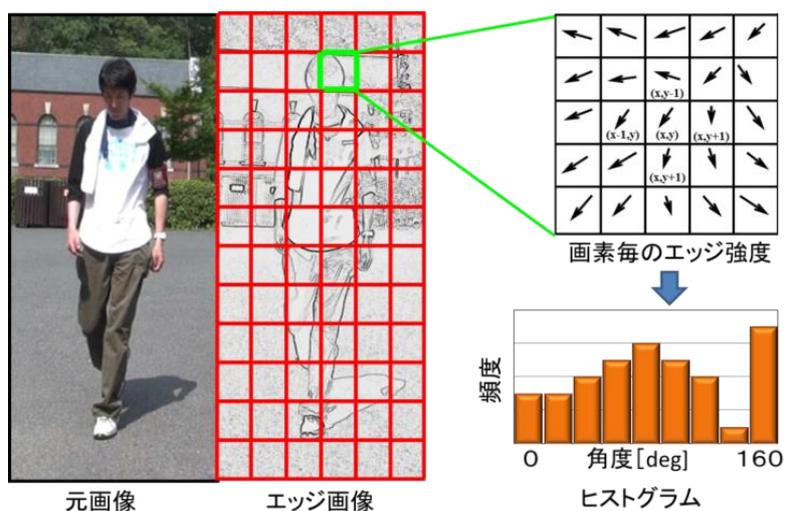


図 3.4 勾配強度の算出過程

最後に、式 (3.6) を用いてブロック領域 ($q \times q$ セル) ごとに正規化し、特徴量を抽出する。

$$v_c^H(n) = \frac{v_c^H(n)}{\sqrt{(\sum_{k=1}^{q^2 N} v_c^H(k)^2) + \varepsilon}} \quad (\varepsilon = 1) \quad (3.6)$$

ブロック領域は 1 セルごとに正規化対象のセル領域が重なるように移動しながら正規化する。正規化後の勾配方向ヒストグラム V_c^H は、 $V_c^H = \{v_c^H(1), v_c^H(2), \dots, v_c^H(B \times N)\}$ となる。ここで B はブロック領域に含まれるセル領域の数である。ブロックとセルの概念を図 3.5 に示す。

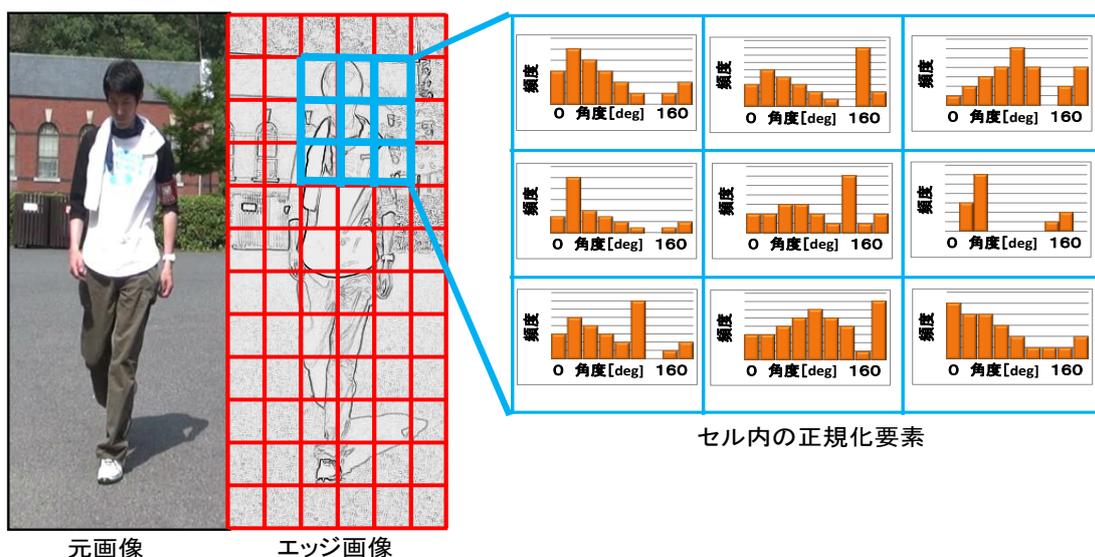


図 3.5 ブロック領域における勾配方向ヒストグラム

3.3.3 フレーム間差分特徴

本研究で用いる時空間特徴の概念を説明する上で参考となる特徴記述手法について述べる。フレーム間差分特徴とは、ピクセル毎の時間経過における輝度変化量を任意の面積毎に抽出してモデル化した特徴量である (図 3.6)。輝度の変化量を捉えるには輝度値の急激な変化 (Motion Trigger) と輝度値の安定度 (Stability Measure) に着目する。はじめに輝度値の急激な変化の検出をおこなう。ここでフレーム k の任意の座標の輝度値を I_k とする。輝度の変化量 T を求めるには、フレーム k から j フレーム前までの各フレームの輝度値 I_{k-j} と I_k との変化量の最大値を算出する。変化量 T は次式により算出する。

$$T = \max\{|I_k - I_{(k-n)}|, \forall n \in [1, j]\} \quad (3.7)$$

急激な輝度値の変化がピクセル上に生じたとき、変化量 T の値は大きくなる。次にピクセルの安定度 S について考える。安定状態の検出には、 k より後の j フレームに着目する。安定度 S は次式により算出する。

$$S = \frac{j(j-1)}{j \sum_{n=0}^j I_{k+n}^2 - (\sum_{n=0}^j I_{k+n})^2} \quad (3.8)$$

安定度はフレーム k から $k+j$ までの輝度値の分散であり，安定した状態では安定度 S は大きくなる．ここで各ピクセルの変化量 T と安定度 S に対して閾値処理を行い，閾値以上の場合動状態，それ以外では静状態とする．人物の歩行シーンに対してフレーム間差分特徴により静状態と動状態を分離した状態を図 3.6 に示す．ここで白色の表示は静状態を表し，黒色の表示は動状態を表す．また同様に動きのある背景に対してフレーム間差分特徴を抽出したデータを図 3.7 に示す．

画素ごとに判定した静動状態をセル領域でヒストグラム化することにより，フレーム間差分特徴となる特徴ベクトルを算出する．

算出した動状態と静状態を用いて 5×5 ピクセルをセルとした領域において，動き状態と静状態の割合を算出する．さらに各セルで作成した動き状態と静状態の割合を特徴量として扱い， 3×3 セルを1ブロックとして正規化をおこなう．

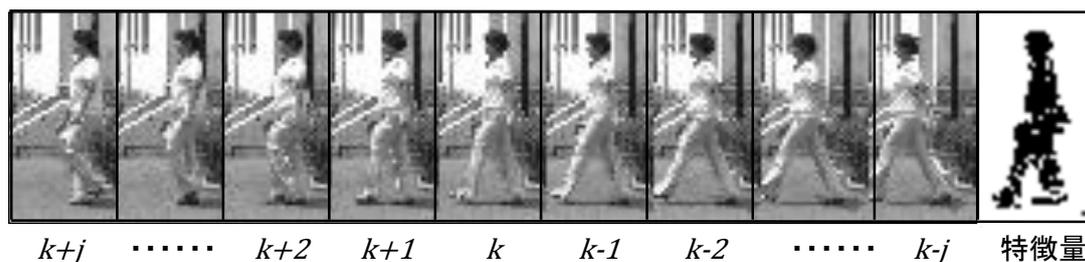


図 3.6 人物のフレーム間差分特徴

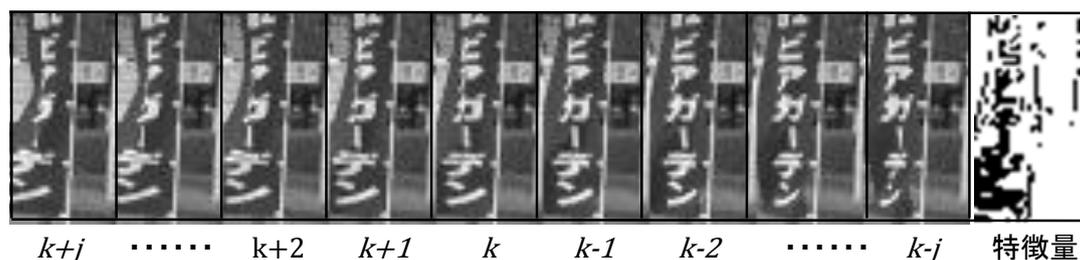


図 3.7 背景のフレーム間差分特徴

i 行 j 列のセル (i, j) の特徴量（動状態と静状態の 2 次元情報）を $F_{ij} = (f_1, f_2)$ とすると， k 番目のブロックの特徴量（18 次元）は $V_k =$

$(F_{ij}, F_{i+1j}, F_{i+2j}, F_{ij+1}, F_{i+1j+1}, F_{i+2j+1}, F_{ij+2}, F_{i+1j+2}, F_{i+2j+2})$ と表すことができる。正規化後の特徴ベクトルを v としたときの正規化は次式となる。

$$v \leftarrow \frac{v}{\sqrt{\|V_k\|_2^2 + \varepsilon^2}} \quad (\varepsilon = 1) \quad (3.9)$$

正規化はブロックを 1 セルずつ移動させることにより行う。入力画像を 30×60 ピクセルとした場合、横方向に 4 ブロック、縦方向に 10 ブロック、合計 40 ブロックに対して正規化を行なう。1 ブロックあたりのセル数は 9 であり 1 セルあたりの特徴ベクトルの次元数は 2 であるため、1 ブロックあたりの特徴次元数は 18 となる。ブロック毎に正規化されたフレーム間差分特徴ベクトルは 40 ブロック \times 18 次元の 720 次元となる。フレーム間差分特徴の抽出概念を図 3.8 に示す。

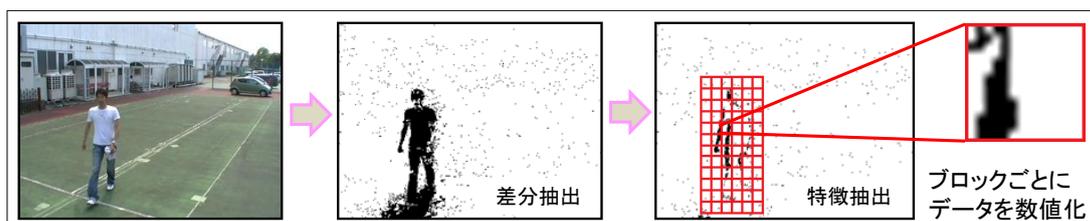


図 3.8 フレーム間差分特徴抽出手順

HOG 特徴とフレーム間差分特徴の抽出結果を比較すると以下の様な差異が見られる。図 3.9 は人物画像に対する特徴抽出結果を示し、図 3.10 は背景画像に対する特徴抽出結果を示す。人物データに対する特徴抽出結果では両者共に人の形状を抽出している事が分かる。しかし背景画像を比較するとフレーム間差分特徴は旗等の動きがあるものに良く反応しており、対して HOG 特徴は背景物体の形状に良く反応しており、両者の抽出する特徴の傾向の差異が確認できる。

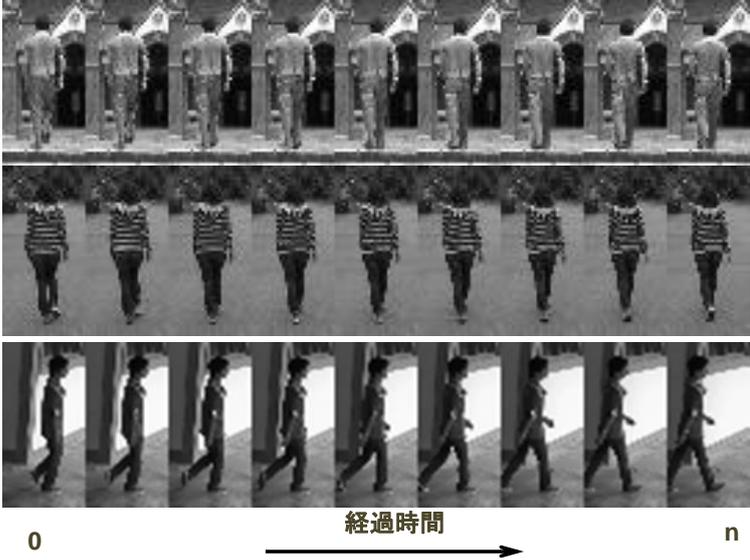
人物データ	フレーム 間差分	HOG
		

図 3.9 人物画像に対する抽出特徴比較

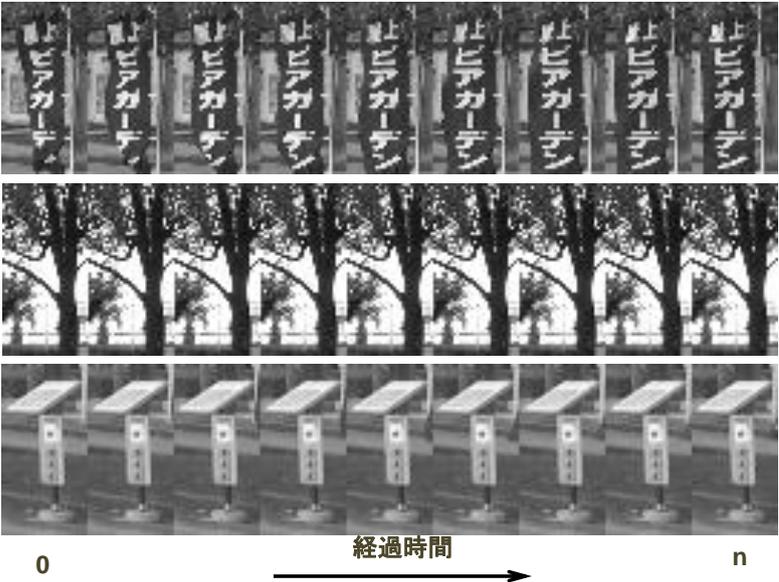
背景データ	フレーム 間差分	HOG
		

図 3.10 背景画像に対する抽出特徴比較

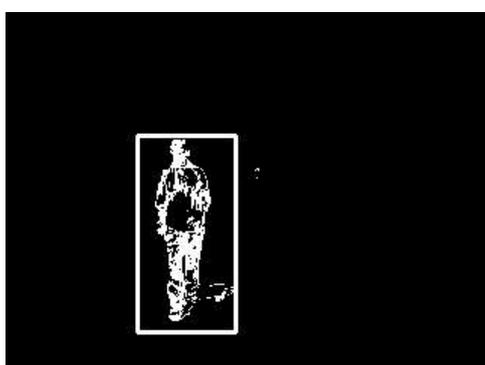
また、フレーム間差分特徴の検出例を図 3.11 に示す。



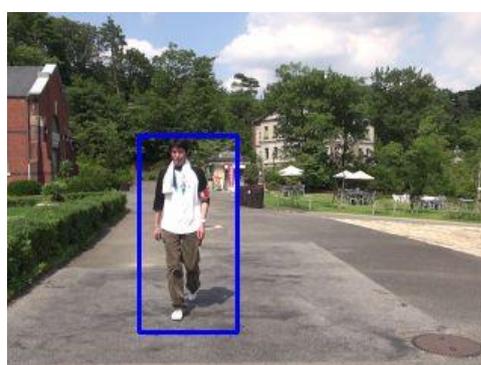
(a_1) 抽出特徴



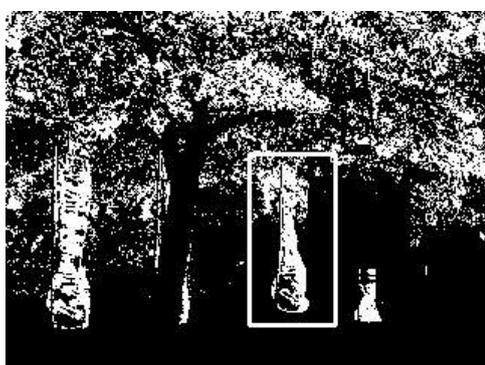
(a_2) 検出結果



(b_1) 抽出特徴



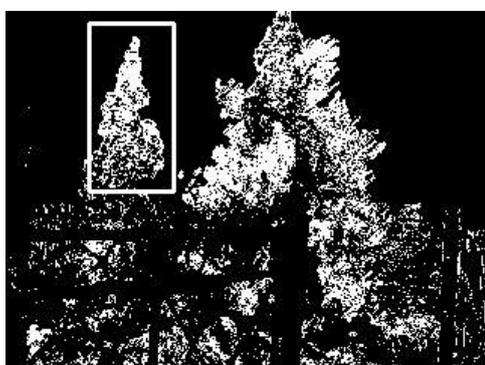
(b_2) 検出結果



(c_1) 抽出特徴



(c_2) 検出結果



(d_1) 抽出特徴



(d_2) 検出結果

図 3.11 フレーム間差分特徴の検出結果

3.3.4 立体高次局所自己相関 (CHLAC)

動画像に対する情報の解析手法として、大津らによる立体高次局所自己相関 (以後 CHLAC とする) [119]が提案されている。CHLAC を用いた研究では、画像中の人物の異常行動検知[126]，歩容を用いた人の認識[127]等が提案されている。CHLAC は単純な特徴記述であるだけでなくシーン中への侵入物体の位置に関わらず同じ記述が可能であり，また複数の同時動作に対応する特徴量の和による表現が可能，さらに対象のモデルを仮定する必要がないという特徴を持つ。

自己相関とは，信号処理において時間または空間のシフトにより得られる自身の信号との整合度合いを測る尺度である。信号のピッチ検出[128]等に用いられる。自己相関 $X(a)$ は式(3.10)であらわされる。

$$X(a) = \int f(r)f(r+a)dr \quad (3.10)$$

- r : 位置
- $f(r)$: 位置 r における出力
- a : r からの変位
- $f(r+a)$: 位置 $r+a$ における出力

高次局所自己相関特徴は図 3.12 に示す様にパターン中央の基準点と黒色で示した変位点との輝度差で構成された特徴であり，変位点の次数を拡張することで特徴パターン数は増加し特徴表現能力は高まる。

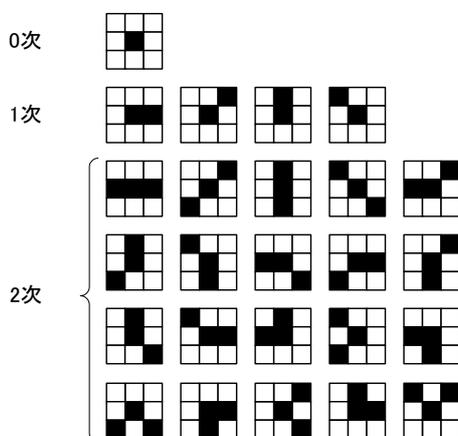


図 3.12 各次元における自己相関特徴

CHLAC は自己相関を時系列の画像データに適応した処理であり，CHLAC 特徴は式(3.11)であらわされる．

$$X(a_1, a_2, \dots, a_N) = \int^{W \times H} f(r) f(r + a_1) \dots f(r + a_N) dr \quad (3.11)$$

- r : 基準点の位置
- $f(r)$: 位置 r における輝度値
- a : 時系列と二次元平面で構成された 3 次元空間における基準点 r からの変位
- $f(r + a)$: 変位点 $r + a$ における輝度値
- N : 変位点の数 (次数)
- W : 画像幅
- H : 画像高

CHLAC は以下の特性を持つことが知られており，人間の動作のように時空間を連続的に変化する動きの性質を捉える上で好ましい定数式とされている．

- ・位置不変性
- ・加法性
- ・モデルフリー

CHLAC の特徴パターン数はパターンサイズ $W \times H$ ，フレーム区間を T ，次数を N とした場合，式(3.12)により求められる．

$$\text{特徴パターン数} = 1 + \sum_{N=0}^{W \times H - 1} W \times H \times T - 1 C_N \quad (3.12)$$

ただし，式(3.12)はパターンの重複を含むマスクパターン数であり，実際の計算に用いられるマスクパターン数は，得られたマスクパターン数から画像走査時に平行移動等価となる組み合わせを除いたものであり，次数 2 の場合を例として算出すると， $N=0$ 次が 1 個， $N=1$ 次が 13 個， $N=2$ 次が 237 個の合計 251 パターンの組合せとなる[129]．

パターンサイズを 3×3 ピクセル，フレーム区間を 3 フレームとした時の 0 次から 9 次までのパターン数（重複を含むパターン数）を図 3.13 に示す

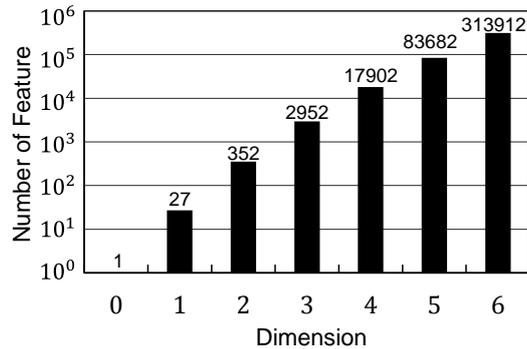


図 3.13 次数と特徴数相関

この結果で示す通り，CHLAC 特徴を用いた画像処理ではマスクパターンの拡張により演算回数が大きく増加する．そのため一般的には次数は 2 を用いることが多い．

3.3.5 Boosting を用いた識別手法

Boosting とは一連の検出精度の高くない単純な分類器を集めてその重み付投票により分類結果を決定する分類器である．分類器で用いられる検出精度の高くない分類器を弱識別器と呼ぶ．

クラス間分離に良く用いられる手法として，超平面分離定理を用いてクラス間を識別する SVM[130]-[135]がある．SVM と比較した場合の利点として Boosting は学習結果が弱識別器の重みとして得られるため，学習後の分類器で特徴量がどのように扱われているかを確認するのが容易である．

Boosting はアルゴリズム的に制限されてはおらず，さまざまな種類の Boosting アルゴリズムが存在する [136]-[139]．最も有名なアルゴリズムとして AdaBoost[140]があり，本研究でも Boosting の手法として AdaBoost を用いる．AdaBoost の具体的なアルゴリズムを以下に示す．

- (1) 初期化： 全てのサンプルの重みを初期化する．

$$\omega_i^1 = 1/m (i = 1, \dots, m) \quad (3.13)$$

i はサンプルのインデックス， ω_i^1 は，学習 1 回目に用いるサンプルの重み，

m はサンプル数を示す.

(2) Do for $t = 1, 2, \dots, T$

ここで, T は学習ラウンドの回数をあらわす.

a) 学習器によりサンプルの重み分布 ω_i^1 に応じて弱識別器 $h_t(t)$ を学習し, 仮説を得る.

$$h_t(x) : x \rightarrow \{-1, +1\} \quad (3.14)$$

b) 学習サンプルに対する誤り率を計算する.

$$\varepsilon_t = \sum_{i=1}^N \omega_i^t [h_t(x) \neq y_i] \quad (3.15)$$

$$\beta_t = \frac{\varepsilon_t}{(1 - \varepsilon_t)}, \quad \alpha_t = \frac{1}{2} \log(1/\beta_t) \quad (3.16)$$

c) サンプルの重みを更新する.

$$\omega_i^{t+1} = \frac{\omega_i^t}{Z_t} \begin{cases} e^{-\alpha_t}, & \text{if } h_t(x_i) = y_i \\ e^{\alpha_t}, & \text{if } h_t(x_i) \neq y_i \end{cases} \quad (3.17)$$

Z_t は重み ω_i^t を正規化する係数である.

(3) 次式より, 最終仮説を得る.

$$H(x) = \text{sign} \left(\sum_{t=1}^T \alpha_t h_t(x) - \lambda \right) \quad (3.18)$$

ただし, λ は閾値であり, 弱識別器の線形和が閾値 λ よりも高ければ正例, 低ければ負例と判定する.

第4章 実験データ

本章では、研究に用いる実験用データについて述べる。画像処理の研究において、顔検出、人検出、行動推定等の成熟した研究分野では多くの評価用の画像データベースが様々な研究機関から公開されている[61][87][141]。しかし、近年研究が盛んになりつつある高齢者のみまもり分野については、現時点では公開データベースが存在しない。特に、本研究が対象とする入院患者の起床動作については、入院患者のプライバシーという観点から施設や病院で収集したデータを公開する事が困難と考えられる。そこで研究を行うに当たり、病室内における撮像対象者の動作と病室環境に対する調査を行い、調査結果から得られた情報に基づいた独自の評価用データベースの構築する。評価用データベースの調査とデータ撮影方法、ならびに評価データについて詳細を述べる。

4.1 被検者の活動状況に関する調査

評価データを作成するに当たり、看護現場で発生する起床動作、ならびに非起床動作について調査を実施した。調査は患者の傾向の異なる3か所で実施した。起床動作については患者の動作を観察し、非起床動作については看護師や見舞客の動作について観察した。調査詳細を表4.1に示す。なお、守秘義務上の理由で病院名は伏せ字とする。また、調査日数の違いは患者の入院都合によるものである。

表 4.1 調査詳細

	診療科	調査日数	性別	年齢
A 病院	整形外科	4	女性	40代
B 病院	脳外科	3	男性	60代
C 病院	脳神経外科	4	男性	40代

4.2 実験設備

4.2.1 撮影施設

病室の条件を再現するために，データ撮影場所は賃貸住宅を使用する．撮影物件の詳細を表 4.2 に示し，撮影場所の間取りを図 4.1 に，室内概要を図 4.2 に示す．

表 4.2 撮影環境

項目	詳細
物件	レオパレスきぼう
所在	名古屋市瑞穂区明前町 3-11
部屋	103 号
間取り	1 K
方角	南向き
採光	南側窓 1600×2300[mm]
期間	2013/08/20 - 2013/08/28

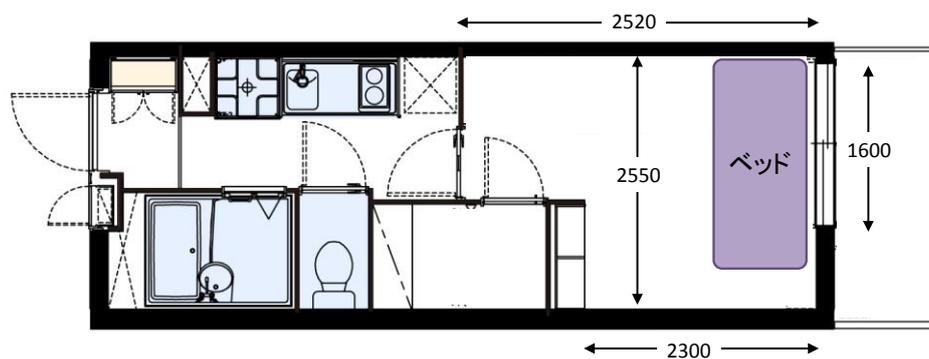
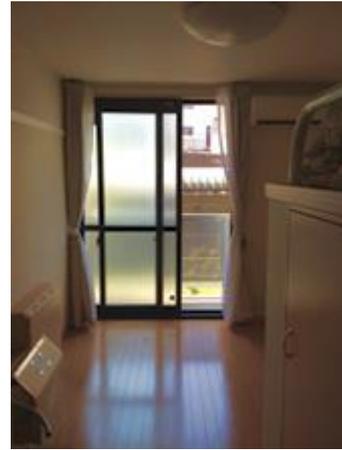


図 4.1 撮影現場間取り



(a) 室内 1



(b) 室内 2

図 4.2 室内

4.2.2 撮影機材

データ撮影に用いた機材について述べる．撮影機材は図 4.3 に示す外観を持つ介護ベッド，三脚で固定した撮影用の IP カメラ（図 4.4），データ記録用のコンピュータ，補助照明を用いて，図 4.5 に記すように室内に配置した．ベッドの仕様を表 4.3 ならびに表 4.4 に，IP カメラの仕様を表 4.5 に示す



図 4.3 介護ベッド外観

表 4.3 ベッド機材

項目	メーカー	型番
ベッド	フランスベッド社	FBN-JJ R30 ANJJ 24JJ
マットレス	フランスベッド社	FM-SH 85×195×8
サイドレール	フランスベッド社	SR-300JJ
ベッドテーブル	フランスベッド社	KT-95HD

表 4.4 ベッド仕様

項目	詳細
高さ	1066 [mm]
幅	972 [mm]
奥行き	2097 [mm]
重量	78 [kg]
駆動	3 モーター
背上げ	0~72 度
脚上げ	0~20 度
上下調整	360~660[mm] ※1

※1 床面からマット上面までの高さ



図 4.4 IPカメラ外観図

表 4.5 IPカメラ仕様

項目	詳細
メーカー	Panasonic
型番	DG-SW155
フォーマット	MJPEG
フレームレート	5[fps]
スーパーダイナミック	OFF
暗部補正	OFF
逆光補正	OFF
光量制御モード	ELC 1/30[s]
ゲイン	On(High)
オートスローシャッタ	OFF 1/30[s]
簡易白黒切り替え	OFF
ホワイトバランス	ATW1
R ボリューム	デフォルト
B ボリューム	デフォルト
デジタル・ノイズリダクション	High
クロマレベル	デフォルト
アパーチャレベル	デフォルト
ペDESTALレベル	デフォルト
アスペクト比	4:3
解像度	VGA
Jepg Quality	5 (0 to 9)



図 4.5 撮影機材

4.3 データ項目

評価データは、病院の環境に対する項目と被写体の行動に対する項目の組合せで作成する必要があると考える。環境はカメラ位置と照明条件について詳細を記載し、被写体については行動パターンの詳細を記載する。

4.3.1 カメラ位置

病院で患者のみまもりを実現する際のカメラの設置についてヒアリングを実施した。ヒアリングは2012年7月18日から7月20日の3日間東京ビッグサイトで開催された国際モダンホスピタルショーの共同研究企業の展示ブースで医療関連の施設担当者や看護担当者に対して実施した。更に後日共同研究企業が病院の担当者と直接打ち合わせを行い、カメラ設置に関する問題点と、カメラの有効設置範囲を設定した。

カメラ設置については既存設備との位置関係が問題となる。病院の設備は診療科目や病院の規模、また建設年代毎に内容や位置が異なる。病室のベッド周辺の主要設備には、テレビや患者の生活用品を格納する床頭台、医療・福祉設備の機能をひとつのユニットに収めたメディカルコンソールが挙げられる。従来のメディカルコンソールは図4.6に示す様に、患者頭部壁面に設置する水平方向に長辺を持つ製品が主流であるが、近年では図4.7に示す様に、ベッドサイドに設置する鉛直方向に長辺を持つ製品の普及が進んでいる。鉛直方向に長辺を持つメディカルコンソールを備えた病院では、コンソール上にカメラを設置する要求がある。また、水平方向に長辺を持つメディカルコンソールを備えた病院では、メディカルコンソールの情報の壁面にカメラを設置する要求や、壁面設置タイプの床頭台（図4.8）にカメラを設置する要求がある。

これらの要求を纏めると、病室内でのカメラ設置位置は図4.9に斜線で示す位置に集約される。従って、データ撮影用のカメラ位置は、図4.9のエリアから均等間隔に選択した6点と、極値判定用にエリア外の3点を加えた合計9点とする。図4.10の位置1については患者に隣接した位置で実運用に設置されることが無い場所である為除外する。なお、実験データは人の動きにより生じる陰など背景の明暗も画像内に記録する。複数台のカメラを並べた場合にはカメ

ラの視野に他のカメラが映り込み、背景の明暗差等を記録する事が出来ない為、カメラを一台ずつ用いて各カメラ位置で実験データの撮影を行う。図 4.10 にカメラ設置位置を示す。また、各カメラ位置の映像を図 4.13 に示す。



図 4.6 従来のメディカルコンソール 写真提供：(株)岡村製作所



図 4.7 近年のメディカルコンソール 写真提供：(株)岡村製作所



図 4.8 床頭台 写真提供：(株)岡村製作所

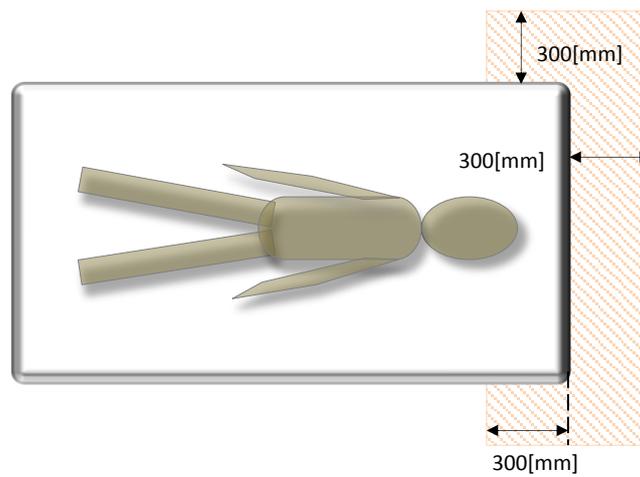


図 4.9 カメラ設置範囲

カメラの設置高さについても、カメラ設置範囲同様にヒアリングによる調査を実施した結果、病室で用いるカメラは患者の状態に応じて不要ならば取り外したいという要望が挙げられた。この理由は患者のプライバシーに対する配慮を明示的に行うという目的によるものである。つまり、入院中の生活をカメラで見られたくない患者については、カメラそのものを取り外したいという要求である。そのため、運用においてはカメラを脱着可能にしておく必要がある。看護現場でカメラの運用を脱着可能にする場合、設置者の身体的特徴が重要な設定基準となる。病棟の看護師は女性も多く、とりわけ小柄な女性でも設置可能なカメラ高さであることが必要となる。そこで160[cm]前後の看護師でも設置出来るように、カメラ高さを170[cm]とした(図4.11)。

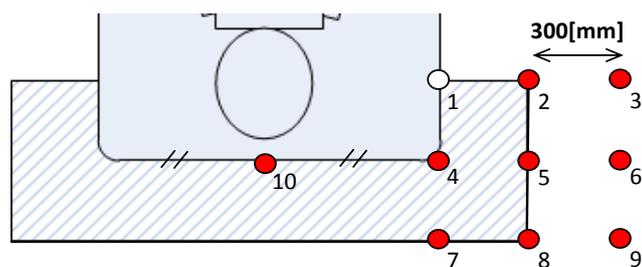


図 4.10 カメラ設置位置

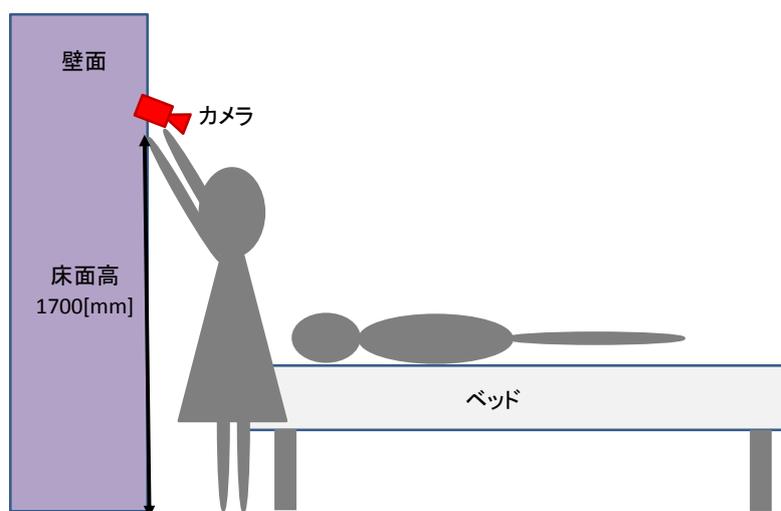


図 4.11 カメラ取り付け高さ

実験データの撮影は、図 4.12(a)に示す様に三脚を用いて高さ 170[cm]の位置にカメラを設置した。また、図 4.10 に示す位置にカメラを設置する為に、ベッド端からの各カメラ位置のマーカを床面に記載し、カメラ位置の垂直上方に撮影用カメラを設置する。マーカを設定する事で、カメラポジションを変更した際に常に同じ条件で実験データを収集する。カメラ位置マーカの設置概要を図 4.12(b)に示す。



(a) カメラ・ベッド位置



(b) カメラ位置マーカ

図 4.12 撮影機器設置概要



(a) 設置位置 1



(b) 設置位置 2



(c) 設置位置 3



(d) 設置位置 4



(e) 設置位置 5



(f) 設置位置 6



(g) 設置位置 7



(h) 設置位置 8



(i) 設置位置 9



(j) 設置位置 10

図 4.13 カメラ位置毎の映像

4.3.2 照明条件

可視光カメラを用いる場合、夜間の低照度時における運用が懸念される。JISの照度基準総則(JISZ9110)[142]によれば深夜の病室内の維持照度は 5[lx]と規定されている。しかし、実際の運用現場においては更に低照度での運用が想定される。そこで、実験データは、照度 0.5[lx]における室内環境の実験データを作成し、評価に用いる。また、昼間に室内照明を点灯した照度 1300[lx]のデータを作成し、夜間と昼間の評価に用いる。室内の照明環境はベッド中央部の照度を 0.5[lx](図 4.14)、ならびに 1300[lx]に調整した(図 4.15)。なお、照度の測定には校正済みの照度計を用い、照度計測位置は図 4.16 に示す様にベッド上面でベッドの対角線の交点にあたるベッド中央で実施した。



図 4.14 室内照度0.5[lx]



図 4.15 室内照度1300[lx]



図 4.16 照度計測位置

4.3.3 装飾

4.1 節で述べた病院施設での調査の結果、病室のカメラで撮影される被写体は3種類に分類できる。

ひとつは患者である。患者は個人の衣類を着用するケースと病院・施設で予め用意された衣類ケースがあり、何れの場合も暖色系の比較的ピンクに近いパジャマ、または寒色系の水色に近いパジャマが着用されている事を確認した。ただし、若年層の入院患者の場合、スウェットの様な衣類を着用する場合は想

定されるため、寒色系、暖色系以外に、スウェット地に良く用いられるグレー、黒の衣類を着用したデータを作成する。実験データに用いた衣類を図 4.17 に示す。

被写体の二つ目は、患者の知人や親族等の見舞客であり、服装については多種多様な組み合わせが想定される。そこで実験データではエキストラの私服(図 4.19)を用いて撮影を実施した。表 4.2 に示す通り撮影は 2013 年 8 月 20 日から 28 日までの土日を除いた 7 日間(夏季)に実施し、5 名のエキストラに対して見舞客の装飾は 35 パターン作成した。

最後の被写体は看護師である。業務における看護師の装飾は看護服が一般的である。看護服には水色やピンクなどの物も使用されるが、病院では白が大半を占める為、データは白の看護服で作成した。実験データに用いた看護服を図 4.18 に示す。



図 4.17 実験データの衣類種別



図 4.18 看護服



(a) sample1



(b) sample2

図 4.19 エキストラの私服サンプル

4.3.4 行動パターン

前項に記載した、患者・看護師・見舞客の3種類の被写体についての行動パターンを規定し、その行動パターンに基づいたデータを作成する。

評価データに用いる行動パターンは起床動作の検出精度を評価する為の起床動作データと、起床以外の動作に対する誤検出精度を評価する為の非起床データに分類する。本研究では起床データを **Positive** データ、非起床データを **Negative** データと定義する。**Positive** データは、仰向け状態からの起き上がり動作、右横を向いた状態からの起き上がり動作、左横を向いた状態からの起き上がり動作をランダムに撮影する。一方、**Negative** データは、看護師と見舞客の動作を模倣した動作と、患者の寝返りを模倣した動作を撮影する。なお、病院施設などでの高齢者の起き上がり動作にかかる時間をおよそ8秒程度と仮定し、作成する評価データのデータ長を8秒とした。**Positive** データのサンプルを図4.20に、**Negative** データのサンプルを図4.21に示す。



(1) sample a



(2) sample b



(3) sample c



(4) sample d



(5) sample e

図 4.20 ポジティブデータサンプル



(1) sample a



(2) sample b



(3) sample c



(4) sample d



(5) sample e

図 4.21 ネガティブデータサンプル

第5章 輝度勾配自己相関特徴を用いた 転倒防止のための起床動作検出

5.1 まえがき

本章では、患者に対する負担が少なく、経済的にも実用が容易な単眼カメラを用いた起き上がり動作の検出手法を提案し、その検出精度を評価する。

提案手法は、エッジの勾配強度と勾配方向から得られた自己相関特徴を基に構成される。動画シーケンスの局所領域から得られた特徴を基にAdaBoost識別器を構築する。自己相関特徴はCubic High order Local Auto-Correlation (以後CHLACとする)特徴[119]の考え方を基礎としている。CHLACは、1)位置不変性(画像の中の物体の位置に結果が依存しない)、2)加法性(画像の中に対象物体が2つ出現した際に、それらの結果を加算した結果が得られる)という特性を持つ。従って、本章で提案する自己相関特徴も位置不変性と加法性という特性を引き継いだ特徴である。病院で用いられるベッドは一般的にキャスターが備え付けられている。従って、検出対象である起床動作が発生する画像座標は様々であり、位置不変性という特性が効果的に働くと考えられる。さらに、加法性という特性は1台のカメラで複数の患者の離床を確認する様な需要が発生した際の発展的な検討も可能であると考えられる。

提案した自己相関特徴はCHLACに比べてプログラム実装が容易である。本章では、提案手法のアルゴリズムの詳細を説明し、その評価実験結果について述べる。評価実験では、実際の環境を模して撮影をした起床動作データベースを用いることで、高齢者の転倒予防に対する有効性を確認する。さらに、単眼カメラを用いる場合、夜間の低照度時における運用が懸念されるが、低照度カメラを用いた実験により、室内照度を0.5[lx]に設定した際の低照度環境の有効性を確認した。

5.1.1 提案手法

提案手法は、動画像の複数フレームから特徴を抽出し、得られた特徴の局所領域毎の傾向をパターンに分類し、分類されたパターン毎のヒストグラムを作成する事で動画像中の動作を記述した後、識別器を用いて任意の動作を検出する。この様に画像からサンプリングされた特徴を、クラスタリングを用いて分類し、ヒストグラムにコーディングする手法は、静止画の一般物体認識手法として bag of keypoints[84]が知られているが、提案手法は bag of keypoints で用いられるコーディング手法を動画像処理に適応する事で、発生する動作に対する位置不変性を獲得している。得られた自己相関特徴を用いて AdaBoost 識別器を構築する。特徴抽出の概略を図 5.1 に示す。

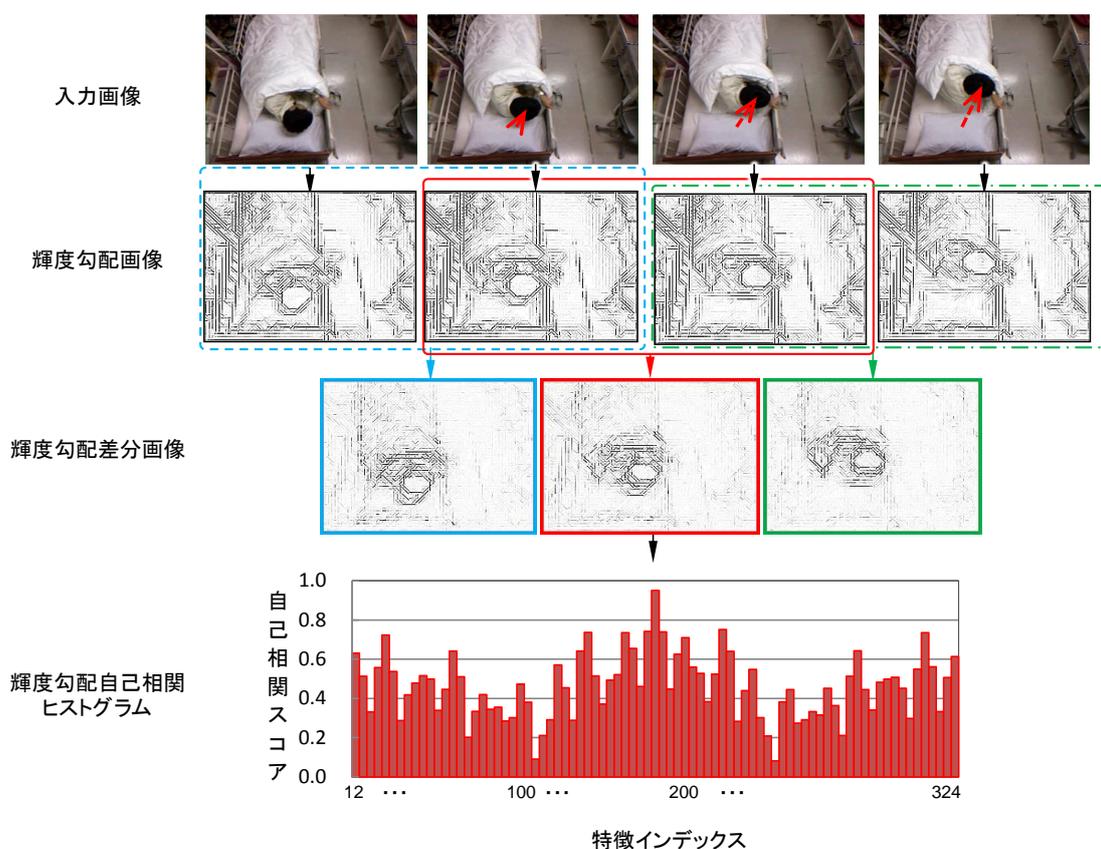


図 5.1 画像特徴抽出の概略

(1) サンプリング

画像特徴には HOG 特徴量[55]でも使用されている輝度勾配ヒストグラムの時間における差分値を用いる。輝度勾配ヒストグラムは、画像をセルと呼ばれる微小領域に分割し、各セル内の画素から算出した輝度勾配の角度と強度を用いて作成したヒストグラムである。輝度勾配の角度をヒストグラムの階級に分け、各階級における輝度勾配の強度の総和を階級毎のスコアとする。このとき、画像から算出する勾配の角度は 0 度以上 180 度未満の範囲を取るが、画像特徴を算出する際にはヒストグラムを用いて 20 度毎の 9 方向や 45 度毎の 4 方向に集約する。本研究ではヒストグラムの階級を予め 4 方向と定義して画像特徴を算出する。輝度勾配ヒストグラムのスコアの算出手順を以下に示す。

はじめに各画素 (x, y) に対する勾配の角度 $\theta(x, y)$ と強度 $m(x, y)$ を算出する。

$$m(x, y) = \sqrt{d_x(x, y)^2 + d_y(x, y)^2} \quad (5.1)$$

$$\theta(x, y) = \tan^{-1} \left(\frac{d_y(x, y)}{d_x(x, y)} \right) \quad (5.2)$$

ただし、 $d_x(x, y)$ ならびに $d_y(x, y)$ は隣接する画素に対する輝度差であり、画素 (x, y) の輝度値を $I(x, y)$ とすると、輝度差である $d_x(x, y)$ ならびに $d_y(x, y)$ は次の式で算出される。

$$d_x = I(x+1, y) - I(x-1, y) \quad (5.3)$$

$$d_y = I(x, y+1) - I(x, y-1) \quad (5.4)$$

続いて、勾配の角度 $\theta(x, y)$ と強度 $m(x, y)$ を用いてヒストグラムを生成する。階級数（以後ビン数とする） B のヒストグラムにおける任意の階級（以後ビンとする）を b として、中心を座標 (x, y) にとる $S \times S$ ピクセルのセルのヒストグラムを算出する際、任意のビン b におけるスコア $h_b(x, y)$ は次式で表される。

$$h_b(x, y) = \sum_{j=-(S-1)/2}^{(S-1)/2} \sum_{i=-(S-1)/2}^{(S-1)/2} (m(x+i, y+j) \times \Delta) \quad (5.5)$$

ただし、 Δ はクロネッカのデルタである。ヒストグラムの各ビンの要素数を c として、勾配の角度 $\theta(x, y)$ をヒストグラムの階級に割り当てた値 $\theta'(x, y)$ を次式で

表わすと,

$$\theta'(x, y) = \theta(x, y) / c \quad (5.6)$$

となる. クロネッカのデルタは次式で表される.

$$\Delta = \delta[b, \theta'(x+i, y+j)] \quad (5.7)$$

仮に b と $\theta'(x, y)$ が等しければ1を, それ以外なら0を返す. 最終的に, 輝度勾配ヒストグラムの差分は, 隣接フレーム間の輝度勾配ヒストグラムの差で求まる. 時刻 t において, 各ビンのスコア値を $h_b(x, y)$ とするビン数 s の輝度勾配ヒストグラムを $H_B(x, y, t)$ とおくと, 各セルの輝度勾配ヒストグラムの差分 $D_B(x, y)$ は次式で得られる.

$$D_B(x, y) = H_B(x, y, t) - H_B(x, y, t-1) \quad (5.8)$$

(2) クラスタリング

5.1.1 項の冒頭で述べた bag of keypoints はサンプリングにより得られた特徴を K-means クラスタリングを用いて分類するが, 本手法では事前に定義したマスクパターンを用いて分類する. ここではマスクパターンの詳細について述べる.

輝度勾配ヒストグラムの差分 (以後, 輝度勾配差分とする) の自己相関は, 現在時刻における着目点 (基準点) とその前後の時刻における着目点 (変位点) から得られた輝度勾配差分の相関を算出することで得られる. また, 基準点と変位点の組合せはマスクパターンと呼ばれる予め定義されたパターンを用いる. 組み合わせを決定する制約条件には自己相関の次数と, 基準点を中心とした領域を用いる.

次数はひとつのマスクパターンに用いる変位点の個数であり, 次数が増加すれば複雑な情報を表現出来るが計算量が増大する. CHLAC では通常 2 を用いており, 本手法でも次数を 2 とする.

マスクパターンは $M \times M \times M$ セルで定義されたブロック (以後, マスクブロック) の中の変位点と勾配の組合せである. CHLAC の変位点はマスクブロックの全てのセルに設定される可能性があるが, 輝度勾配自己相関では変位点の位置は各

フレームで 1 つに限定する．さらに基準点が設定されたフレームには変位点を設定しない．つまり，自己相関の計算対象となるセルは常にフレームに 1 つだけ存在する．次数 2 の変位点の組み合わせ概要を図 5.2 に示す．次数を N とした場合の変位点の組合せ数 P_N を式(5.9)に， P_N に勾配方向を組み合わせたマスクパターン数 P_M を式(5.10)に示す．ただし，マスクパターンを構成するブロック数を $H \times W$ とし，用いる勾配方向数を k とする．

$$P_N = ({}_{W \times H} C_1)^N \quad (5.9)$$

$$P_M = k^{P_N} \quad (5.10)$$

(3) コーディング

サンプリングで述べた画像の全領域から得られた特徴の出力値とクラスタリングで述べたマスクパターンの形状を用いて，マスクパターン全通りに対する自己相関を算出する．本節で得られる出力値はマスクパターン数分の階級数を持つヒストグラムとなる．得られたヒストグラムは学習処理での特徴量として用いる．また，ヒストグラムの各階級は特徴量の特徴インデックスに対応する．処理概略を図 5.3 に示す．変位点を図 5.4 に示す様に基準点からのベクトルで表し，基準点の位置ベクトルを r ，輝度勾配差分 $D_B(x, y)$ を $f(r, k)$ と置くと， N 次の自己相関特徴は次式により得られる．

$$X(a_0, \dots, a_N, k_0, \dots, k_N) = \prod_{n=1}^N \alpha \cdot G(k_0 - k_n) \int_{r \in I} f(r + a_0, k_0) f(r + a_1, k_1) \dots f(r + a_N, k_N) dr \quad (5.11)$$

ただし， a_N は変位ベクトルであり， a_0 は基準点の変位を表すために常に 0， k_N はセル位置 $r + a_N$ における勾配方向， α は係数， $G(\cdot)$ はガウスクアーネル，そして I は入力画像を表す．ガウスクアーネルを用いることで，時間軸上に隣接する勾配方向の強度のみが重みづけされ，角度方向の異なる勾配で構成されるマスクパターンの影響が低減される．例として，勾配方向を 4 とし， $k_0 \neq k_n$ で 0 相当の値を返すガウスクアーネルを用いた場合，得られる特徴ベクトルの次元数は 324 となる．

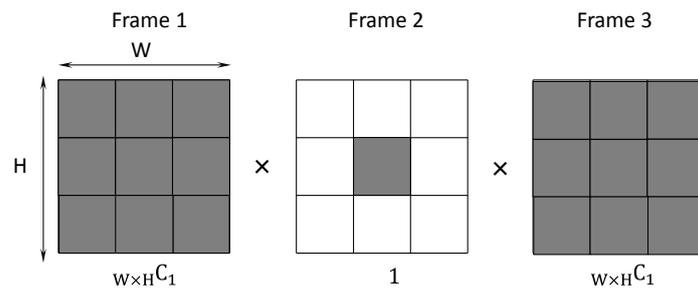


図 5.2 次数2のマスクパターン概要

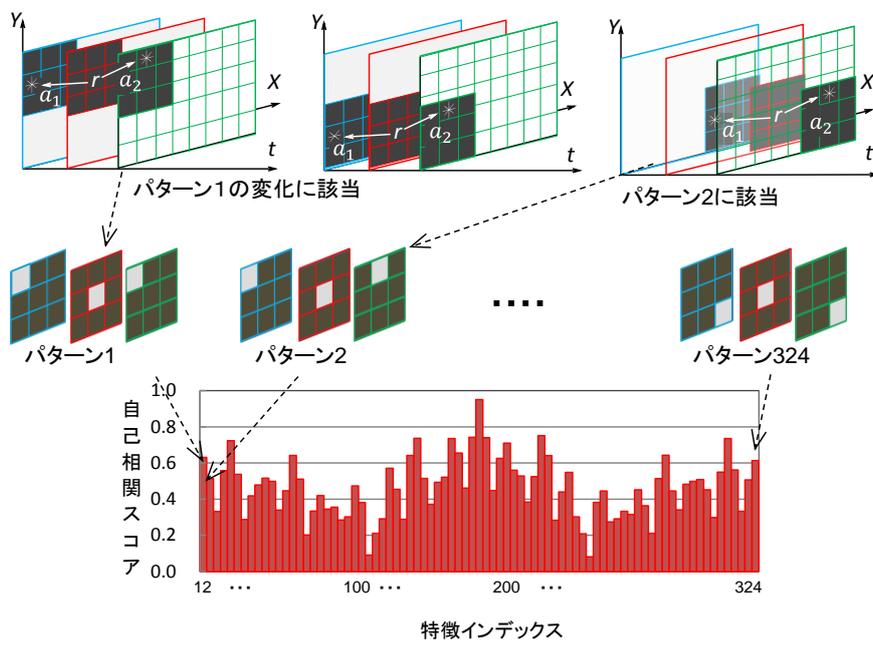


図 5.3 コーディング処理概略

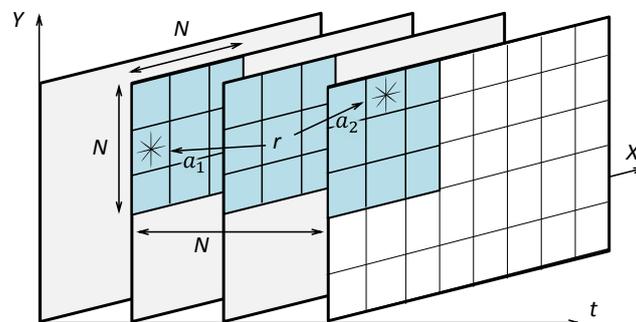


図 5.4 自己相関特徴の抽出概要

5.2 起き上がり動作検出実験

5.2.1 実験概要

CHLAC を用いた従来手法と、提案手法を用いて認識精度の比較実験を行う。実験データは、1)起き上がり動作データ、2)起き上がり以外の動作データの2種類のデータを用いる。動画は 30[fps]で撮影されたものを使用し、自己相関の次数は 2 とする。さらに、実験データの撮影上のノイズや画像センサデバイスのオートアイリスなど自動補正の影響を低減するために幾何学特徴は 8 フレームの移動平均で得られた結果を用いる。また、実験は検出性能を極力維持した状態で処理速度を向上させるという観点からフレームレートを 3[fps]、式 3.17 の T で示した学習回数を 100 回に固定して、解像度を 320×240[pix]、160×120[pix]、80×60[pix]、40×30[pix]の 4 階調に変化させた場合の検出性能を測定した。

5.2.2 実験データ

実験データは起き上がり動作を撮影した 31 データと、起き上がり以外の動作を撮影した 64 データの 2 種類を用いる。起き上がり動作を **Positive** データ、起き上がり以外の動作を **Negative** データと定義する。動作パターンについては、2013 年 3 月に病床数約 600 床の病院の整形外科病棟にて、入院患者ならびに家族の同意のもと、入院患者ならびに見舞客の動作について調査を実施した。その結果、得られた動作傾向に基づき実験データを作成した。

Positiveデータのサンプルを図 5.5に、**Negative**データのサンプルを図 5.6に示す。起き上がり動作は、仰向け状態からの起き上がり動作、右横を向いた状態からの起き上がり動作、左横を向いた状態からの起き上がり動作を均等に用いる。一方、起き上がり以外の動作は、椅子に座した状態からの立ち上がりや、歩行、前かがみ、背伸び等の動作を用いる。病院施設などでの高齢者の起き上がり動作にかかる時間をおよそ8秒程度と仮定し、評価に用いる動画シーケンスのデータ長を8秒とした。

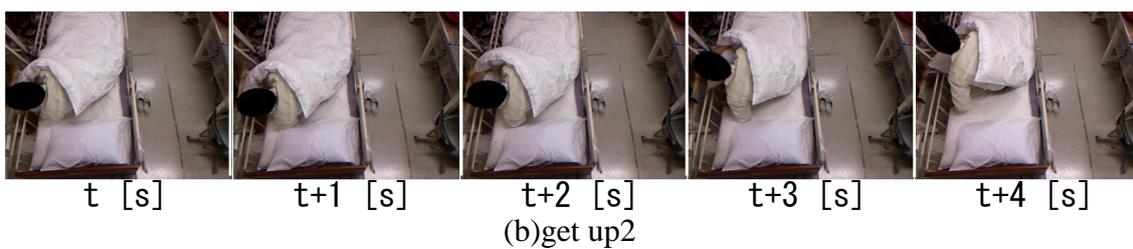
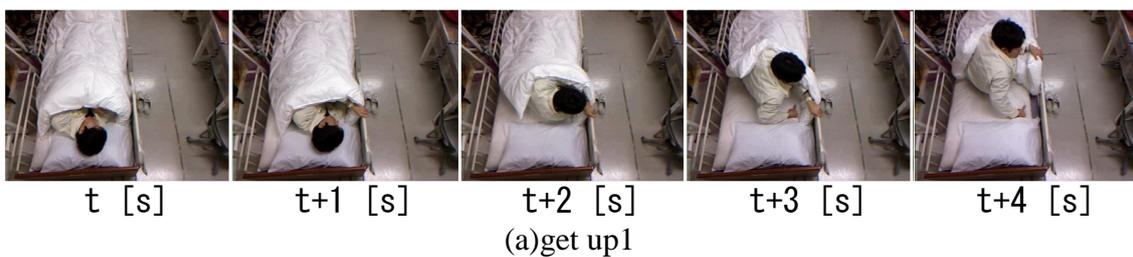


図 5.5 ポジティブデータサンプル

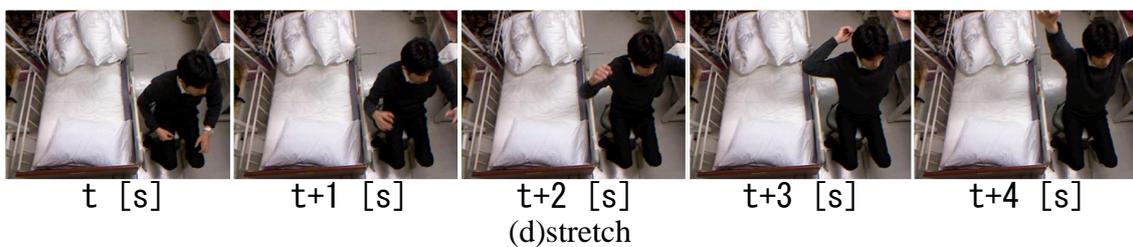
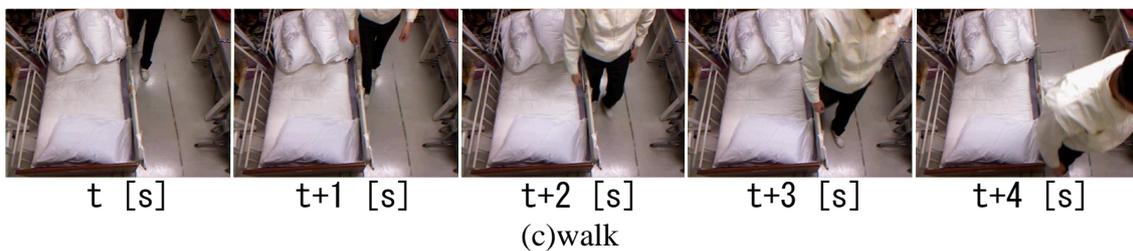
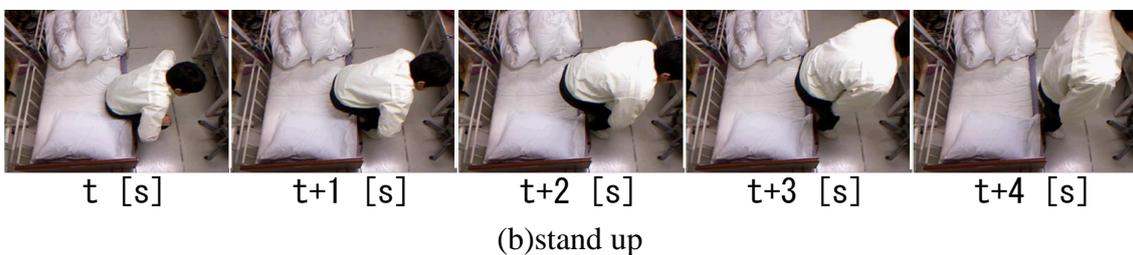
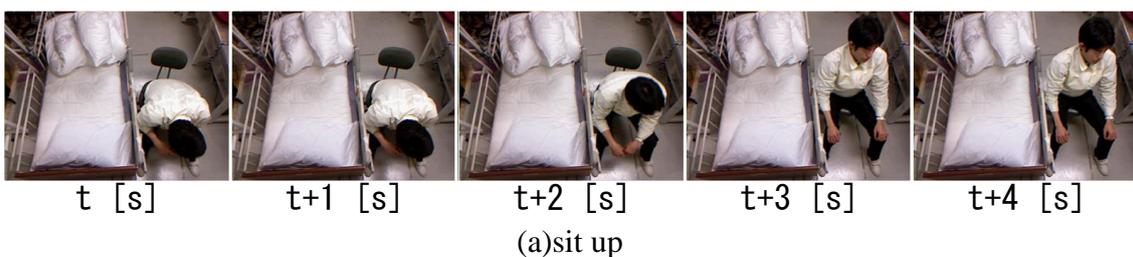


図 5.6 ネガティブデータサンプル

実験結果

交差検定によりエラー率を測定する。ここで、Positive データ数は 31 であり、仰向け、右向き、左向きからの起き上がりがそれぞれ約 10 データ存在する。交差検定の分割数を大きくした場合、評価データに体の向きの何れかが含まれない評価が発生する可能性があり、識別器に対する正確な性能を判断出来ない。そこで体の向きが評価データに一定の割合で含まれることを考慮し、分割数を 4 とした。それぞれの検定において、データの 75[%]を学習データとして使用し、残りの 25[%]を評価データとして用いた。異なる学習データと評価データの組み合わせで、4 通りの試験を各パラメータの組み合わせに対して実施した。評価結果として画像サイズ毎の未検出率を図 5.7 に、画像サイズ毎の誤検出率を図 5.8 に示す。評価の結果、従来手法は画像サイズ 80×60[pix]の時に最も高い性能を示し、未検出率が 14.1[%]発生し、誤検出率が 35.5[%]発生した。提案手法は画像サイズ 80×60[pix]の時に最も高い性能を示し、未検出率が 4.7[%]発生し、誤検出率が 25.8[%]発生した。この結果から提案手法は従来手法を上回る行動推定の性能といえる。

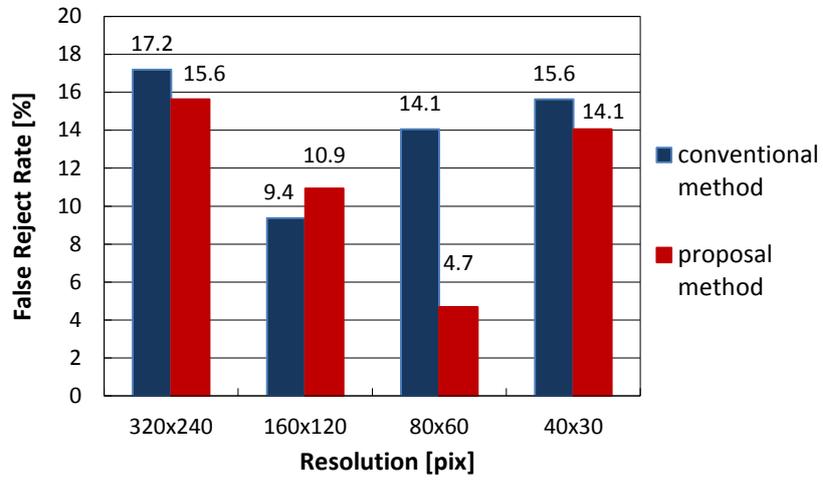


図 5.7 各解像度における未検出率

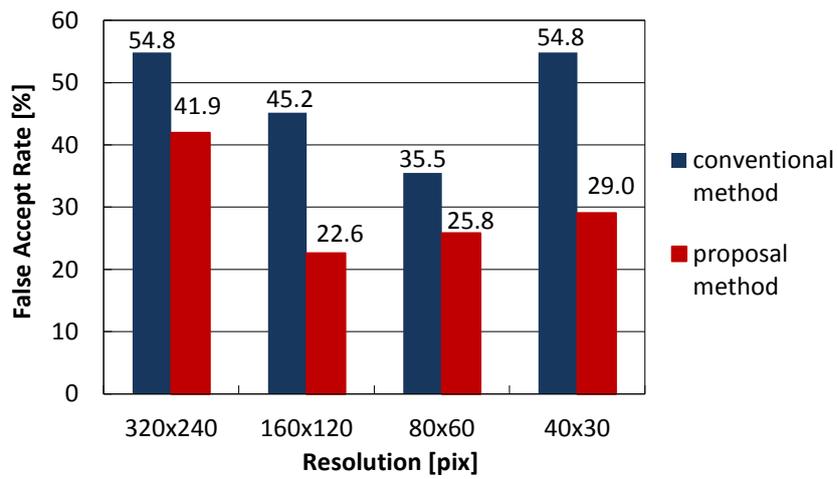


図 5.8 各解像度における誤検出率

5.3 低照度環境における評価実験

5.3.1 実験概要

単眼カメラを用いて、夜間の低照度時における起き上がり動作の検出性能を評価した。4.3.2項で述べたが、JISの照度基準総則(JISZ9110)によれば、深夜の病室内の維持照度は5[lx]と規定されているが、実際の運用現場においては更に低照度での運用が想定されるため、照度0.5[lx]での検出性能を測定し、室内照明を点灯した照度1300[lx]での検出性能と比較することで、低照度時の起き上がり認識精度を調べた。室内の照明環境はベッド中央部の照度を0.5[lx]、ならびに1300[lx]に調整した、

5.3.2 実験データ

実験には、4.3.2項で述べた0.5[lx]ならびに1300[lx]の各照度で撮影した起き上がりデータを用いる。5人が各々20回実施した起き上がり動作、合計100データを用いる。また、低照度時は、通常の照明条件下よりも被写体の明度や彩度が動作の検出に影響する事を考慮し、5人の被験者が4.3.3項に記載した明度や彩度の異なる5種類の服装を着用した。

5.3.3 実験結果

各照度条件における検出率は、0.5[lx]の照明条件で95[%]、1300[lx]の照明条件で91[%]となり、低照度時でも照明点灯時と同等の検出性能であることを確認した。各照度における検出結果を図5.9に示す。

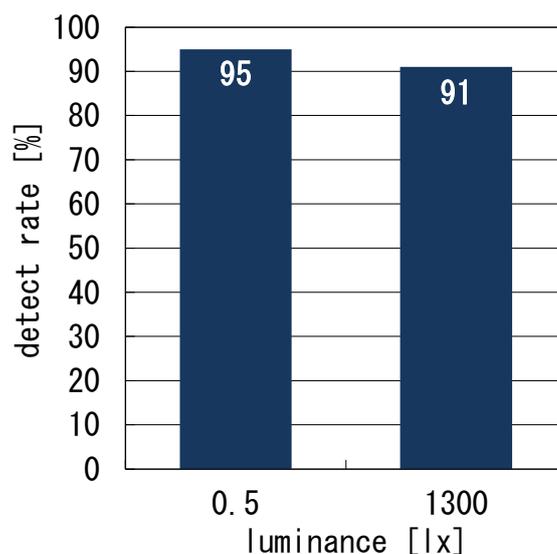


図 5.9 照度毎の検出率

5.4 考察

5.2節の実験で、提案手法の検出性能が従来手法の検出性能に対して優れた要因について考察を行う。

効果が大きいと考えられる要因の一つに、幾何学特徴の形状表現能力が考えられる。CHLACは画像の局所領域の濃淡変化の組み合わせで表現される特徴量であるのに対して、提案手法は局所領域の濃淡変化から輝度勾配を抽出し、その輝度勾配を用いてエッジを表現した。このため提案手法は単なる動きの変化だけでなく物体の形状を表現した特徴であると言える。

物体の形状表現能力が向上することにより、識別性能が向上する例として、視覚的に理解しやすい歩行シーンの例を用いて説明する。図 5.10は画面左から右へ歩行する人物と、さらにその奥で別の人物が画面左方向へ歩行して両者が交差するシーンである。このシーンから従来手法、提案手法それぞれを用いてLow-Levelの幾何学特徴量を抽出した結果が図 5.11の(a), (b)である。

ここで、人物が交差する部分に着目し、その領域を拡大したのが図14の(c)(d)である。丸で囲まれた領域内の横方向の特徴は後ろの人物の特徴であり、縦方向の特徴は手前の人物の特徴である。

この画像からCHLAC特徴を抽出した場合、手前の人物の特徴と奥の人物の特徴とは同一の特徴として処理される。しかし、提案手法を用いた場合、式(12)

の中で示したガウスクーネルにより手前の人物の特徴と奥の人物の特徴は位相差が考慮され、お互いの影響が低減される．性能向上の要因の一つとして、この特徴の表現能力の差があると考えられる．



図 5.10 歩行シーンサンプル画像

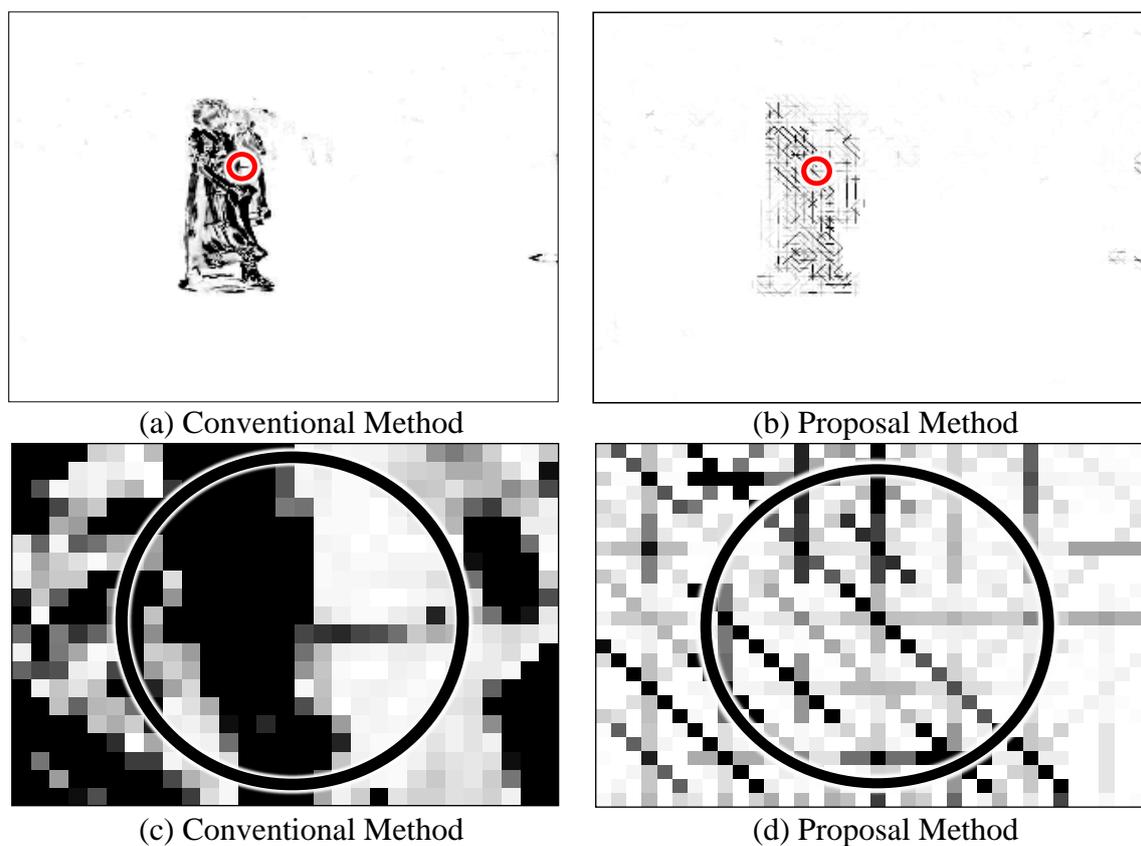


図 5.11 歩行シーンの特徴抽出結果

5.5 むすび

本章では、患者の転倒問題を本質的に解決する為に画像センシングを用いて転倒そのものを未然防止するアプローチを提案した。輝度勾配自己相関を用いた特徴抽出手法から構成された行動推定手法を提案し、従来手法との比較結果を示した。独自に構築した評価データセットを用いた評価の結果、従来手法は未検出率が14.1[%]発生し、誤検出率が35.5[%]発生した。また、提案手法は未検出率が4.7[%]、誤検出率が25.8[%]発生した。提案する輝度勾配差分を用いた自己相関特徴はCHLAC特徴の行動認識性能を上回ることを確認した。

第6章 カメラ設置範囲にロバストな 起床動作検出

6.1 まえがき

5章で述べた起床動作の検出手法は、CHLAC特徴の考え方を基礎としており、1)位置不変性(画像の中の物体の位置に結果が依存しない)、2)加法性(画像の中に対象物体が二つ出現した際に、それらの結果を加算した結果が得られる)という特性を持つが、次に述べる課題を持つ。CHLACと同様に輝度勾配自己相関特徴はXY座標軸の2次元平面の画素情報を時間軸方向に拡張した3次元情報の特徴量であり、任意の時刻での特長量は、平面画像の2次元情報のみで算出される。起床検出においては、頭部付近に設置したカメラとベッド斜め横方向に設置したカメラの映像では、平面画像から得られる人物の姿勢の見えは大きく異なる。更に、検出器の設計当初と運用時でカメラ位置が頭部付近と斜め横方向の様に異なる場合、各々の起床動作から得られた画像特徴は大きな差異が発生し、検出精度低下の原因となる。カメラ位置における画像特徴の差異については6.2節で詳しく述べるが、次に述べる通り、病院では起床検出用のカメラ位置を限定して運用する事は困難であり、これがカメラの設置位置を患者頭部に限定した第5章の手法の問題点と考えられる。

施設や病院に対して実施した、運用時のカメラ設置位置の調査結果では、病院ではコンソールと呼ばれるシリンジポンプやナースコールの設置パネル、または備品棚等、各医院の既設の設備があり、病院毎にカメラの施工予定位置が異なるという結果が得られた。さらに、各病院のカメラ設置予定位置をまとめると、図 6.1の斜線のエリアに集約された。つまり、このエリアの任意の位置に設置したカメラで一定の検出精度を保つ起床動作検出を行う必要がある。しかし、従来手法の輝度勾配自己相関特徴を用いた識別手法は、図 6.1の点Aで示したベッドの中心頭部頭上の定位置に固定したカメラを用いた場合に良好な検出

性能を得られるが、6.4節の実験で述べる結果は、ベッド両側も含めたカメラ位置での評価では検出率が低い。

本章では、輝度勾配自己相関特徴の見えの変化に脆弱であるという問題に対して、特徴量を改善する事で、実運用上のカメラ設置エリアでの起床動作の検出性能を改善する手法を提案する。近年、Deep Convolutional Neural Network[143]やNested Shape Descriptors[144]で画像情報の畳み込みにより回転や平行移動に対する物体検出の汎化性を向上させた例が報告されているが、本研究では動画像の画像特徴に畳み込み処理を適用する事で運動に対する汎化性の向上を検討する。

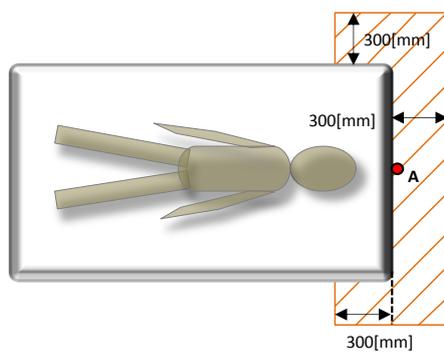


図 6.1 カメラの設置エリア

6.2 輝度勾配ベクトルの畳み込み

カメラ位置を変更する事で、画像中のベッドの見えは変化する。とりわけ、ベッドの角度は大きく変化し、カメラを患者の頭上に設置した映像とベッドサイドに設置した映像ではベッドの見えは45度以上の角度差として現れる。ここで、カメラ設置位置に対するベッドの見えを図に示す。カメラの位置は6.1節で述べた病院の運用環境に従い、図6.1のカメラ設置範囲の端点を重点的に選択し、図6.2の位置とした。また、その際のベッドの見えは図6.3となる。図中にはベッドの角度を明示するためにベッド長手方向の中心に対して赤線を記載する。

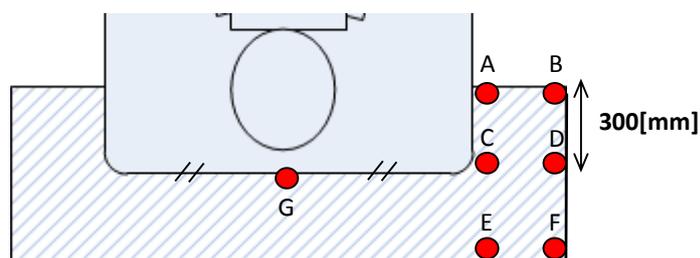


図 6.2 カメラ設置位置



(a) position A



(b) position B



(c) position C



(d) position D



(e) position E



(f) position F



(g) position G

図 6.3 カメラ位置毎の撮影画像

ベッドの角度差が大きくなる事で、起床動作の見えも大きく変化する。この見えの変化に起因して、抽出される画像特徴が大きく異なるという問題が発生する。図 6.3 (g)の様に、ベッドが垂直方向に撮影された画像では、起床の際に上半身は直線運動を行う為、画像から輝度勾配を算出すると比較的画面の垂直方向に向きを持つ特徴が抽出される。しかし、図 6.3 (b)の場合、起床動作における上半身の動きは腰を中心として頭を端点とした回転運動を行う為、画像から輝度勾配を抽出すると斜め方向の画像特徴が最も多く抽出される。角度差による輝度勾配の差異を明示するために、図 6.3(b)ならびに(g)の動画像における人物周辺領域から算出した起床の動作特徴を表す輝度勾配差分特徴を図 6.4 に示す。

式(5.9)が示す通り、勾配の異なる動作に基づく自己相関はそれぞれ異なる次元の特徴として抽出される。つまり、画像水平方向に勾配を持つ物体の動作と画像垂直方向に勾配を持つ物体の動作は、移動方向が同じ場合でも異なる特徴量として区別される。図 6.4 では、同じ起床動作でありながら、見えの変化により異なる輝度勾配が算出されており、この輝度勾配から算出される輝度勾配差分自己相関特徴は、異なる特徴ベクトルとして扱われる。抽出された特徴量は AdaBoost により特徴ベクトル毎の識別が行われるため、識別結果に大きな影響を与える。

そこで画像の見えが異なる起床動作に対して、見えの違いを吸収する特徴抽出手法を提案する。提案手法は輝度勾配を算出した後、自己相関の算出処理において勾配方向を畳み込む処理を行う。ここで、畳み込まれる特徴ベクトルの一例をマスクパターンの形状で図 6.5(a)ならびに(b)に示す。微小時間での画素変化を考えた場合に、図 6.4(a)の様に垂直方向に勾配を多く含む物体の垂直方向運動から得られる特徴は図 6.5(a)に示すマスクパターンの出力が高くなり、同じ物体の動作を側面からとらえた際の傾斜角を持つ輝度勾配の特徴は図 6.5(b)に示すマスクパターンの出力が高くなる。畳み込み処理を用いた場合、この二つの特徴量は同一のベクトルとして扱われ、同じ動作でありながら見えの違いで分離された情報が、同じ動作の情報として統合される。

5.1.1 項で述べた通り、従来手法では異なる勾配方向を組み合わせたマスクパターンの出力値は式(5.9)のガウスクーネルにより低減される。さらにこの出力値は起床動作識別によおける特徴量として用いられるが、ガウスクーネルにより 0 に近い値まで低減される事で、識別器に正例が入力された場合でも、負例が入

力された場合でも出力値が等しくなる。正例と負例に対する判別能力が無い為、この異なる勾配方向が組み合わされたマスクパターンは識別器の要素から除外する事が出来る。従って、マスクパターンの計算処理は同一方向の勾配のみで構成されたマスクパターンだけを扱う問題として置き換えられ、式(5.9)は次のように簡略化できる。

$$X(a_1, \dots, a_N, k) = \int_{r \in I} f(r, k) f(r + a_1, k) \dots f(r + a_N, k) dr \quad (6.1)$$

k は任意の勾配方向を示す。式(6.1)は任意のマスクパターンから得られる出力値を計算する際に、全てのセルの勾配方向が等しい事を表す。次に提案手法の計算処理について述べる。今 K 方向の勾配を P 方向の勾配に畳み込むとすると、 K/P 個の要素が1方向に畳み込まれる。この要素数を Q とすると、方向

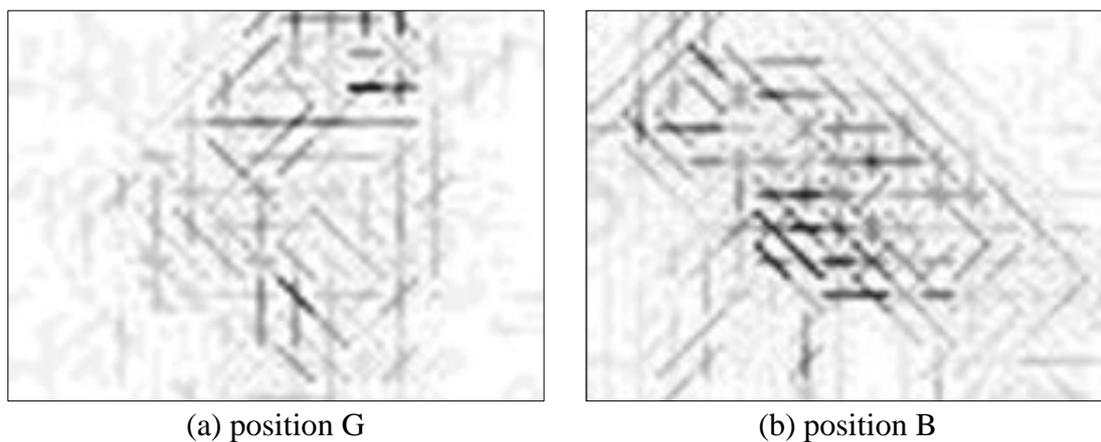


図 6.4 抽出した輝度勾配特徴

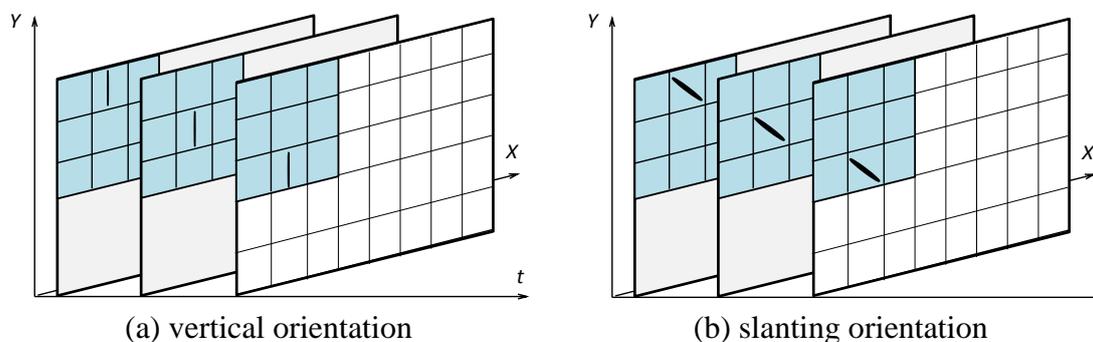


図 6.5 畳み込みマスクパターン

数 K ならびに P と、要素数 Q の関係は次式で表される。

$$Q = K / P \quad (6.2)$$

更に、畳み込まれた勾配の方向を k_p と置くと、提案手法の自己相関特徴は以下の式で得られる。

$$X(a_1, \dots, a_N, k_p) = \int_{q=0}^{Q-1} \int_{r \in I} f(r, k_p \times Q + q) f(r + a_1, k_p \times Q + q) \dots f(r + a_N, k_p \times Q + q) dr dq \quad (6.3)$$

6.3 多重解像化

前節で述べた手法を用いる事でカメラアングルに対して冗長な特徴が取得できるが、勾配の情報を無視して動体の移動経路だけに着目する為に動作の弁別性が低下する懸念が生じる。輝度勾配の畳み込みにより低下した動作の弁別性を向上させる手法として、解像度の異なる画像から患者の動き特徴を抽出する。局所領域の動きに対しては高解像画像を用いて変動を抽出し、大局域の動きに対しては低解像の画像を用いて変動を抽出する。更に本論文で用いる画像特徴は時空間特徴であり、局所大局の概念は空間だけで無く時間軸上にも拡張できる。そこで解像度の多重化と併せてフレーム間隔の多重化を行い時空間に対する多重解像拡張を組み合わせた情報を用いた画像特徴量の抽出を行う。異なる解像度と時間軸の組み合わせで抽出される画像特徴を図 6.6に示す。ただし、単に解像度の多重化処理だけでは、隣り合う解像度間の特徴が疎になる。そこで解像度間の特徴を補完する為に低解像画像に対して図 6.7の様にマスクブロックの間隔にオフセットを加えたパターンを用いた特徴量を追加する。画像特徴を密にする事で検出精度を向上させる効果はWangらのOptical Flowを用いたDense Trajectories[116]でも確認されている。

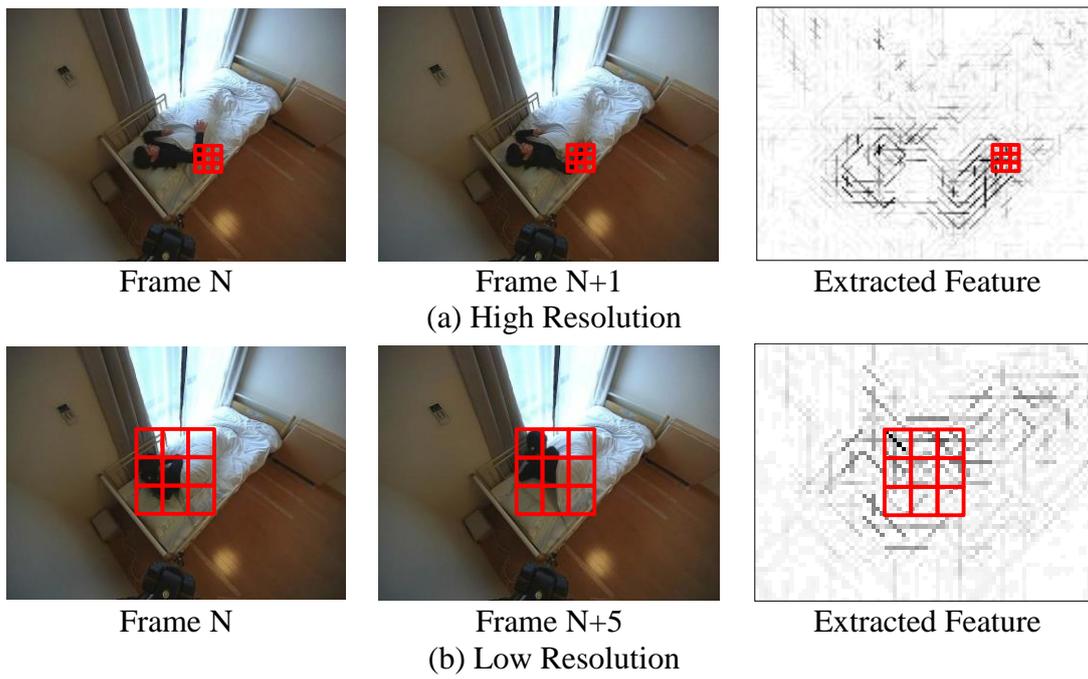


図 6.6 各解像度において抽出された画像特徴

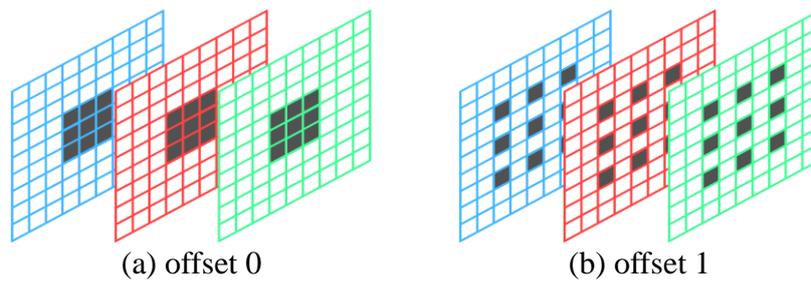


図 6.7 マスクブロックのオフセット

6.4 各手法に対する性能比較実験

6.4.1 実験概要

実験は、本論文で提案する輝度勾配ベクトルの畳み込みと多重解像化の 2 つの実施手法が、従来手法と比較してどの様な効果を示すかを検証する。検証方法は、各手法のカメラ位置に対する検出性能評価と、起き上がり動作と起き上がり動作以外の識別性能評価を行う。実験条件として、動画は 5[fps]で撮影されたものを使用し、式 6.3 の N で示される自己相関の次数は 2 とし、式 6.2 の K で示される輝度勾配は 4 方向を用いる。また提案手法の畳み込みの次元数は 1 とする。実験に用いた識別器の時間と空間の解像度に関するパラメータは、従来手法は時間解像度に 5[fps]、空間の解像度に QVGA を用いた。提案手法は、時間解像度に 5[fps]、2.5[fps]の 2 階調を用い、空間解像度に QVGA と 160×120 ピクセルの 2 階調を用いる。更にマスクブロックのサイズは 1 セル間隔、2 セル間隔、3 セル間隔の 3 種類の組合せを用いる。使用するパラメータの組合せは予備実験により決定した。事前に単独のパラメータで構築した識別器の評価を行い、単独のパラメータで最も高い検出結果を示した解像度を基準として、その周辺の識別器を複数組み合わせさせた多重解像度の識別器の評価を行い、最終的に最も検出率の高い組み合わせを用いた。学習回数は従来手法と提案手法では特徴ベクトルの次元数が異なるため、事前実験の学習曲線から判断し 81 回を用いる。さらに、実験データの撮影上のノイズや画像センサデバイスのオートアイリスなど自動補正の影響を低減するために特徴は 8 フレームの移動平均で得た画像から算出する。

6.4.2 実験データ

実験データには起き上がり動作 2250 データと、起き上がり以外の動作 2940 データの 2 種類を用いる。起き上がり動作を Positive データ、起き上がり以外の動作を Negative データと定義する。起き上がりデータのカメラ位置は、図 4.10 の Pos2 から Pos10 の 9 点で各々 250 データを用いる。Pos1 はカメラが患者の顔の横に位置し、カメラ三脚で検出対象が遮蔽される為に評価から除外した。また、システムの実運用時も Pos1 機材位置により被写体に隠れが発生する様な運用は

制限する。また、実用時の汎用性を確認するために図の斜線部で示した設置仕様境界よりも外側の3点となる Pos3, Pos6 ならびに Pos9 を追加する。起き上がり以外の動作は、看護師と見舞客の動作を模倣した動作 1470 データと、患者の寝返りを模倣した動作 1470 データを用いる。また夜間については、運用を検討する病院では補助光を点灯する事で合意しており、室内の最低照度が 1[Lx]あれば、図 4.14 に記す通り実施機材を用いて十分な輝度が確保できる事を確認している。

6.4.3 カメラ位置に対する検出性能評価

手法毎のカメラ位置の汎化性能に対する評価を実施した。実験では起き上がりデータをカメラ位置毎に分類し、全9か所のカメラ位置をデータの組とする9個のデータセットを作成した。ひとつのカメラ位置データを評価データとして使い、残りの8か所のカメラ位置データを学習データとして用いるLeave-One-Outクロスバリデーションを実施した。従来手法 (Conventional Method) と輝度勾配ベクトルの畳み込み (Method1) ならびに多重解像化 (Method2) の検出率を図 6.8 に示す。更に、各手法のカメラ位置における検出率の平均と標本分散を表 6.1 に示す。評価結果から、Method1を用いる事でカメラ毎の検出率のばらつきを示す標準偏差が4.1低減し平均検出率は6.8[%]向上した。更に多重解像を用いる事で従来手法と比較して検出率のばらつきを示す標準偏差は9.66低減し、平均検出率は13.1[%]向上した。

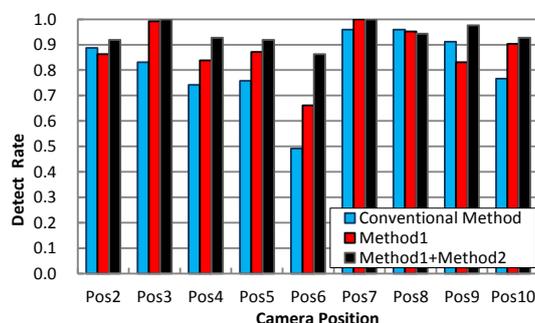


図 6.8 カメラ位置毎の検出率

表 6.1 カメラ位置毎の検出率に対する平均と分散

	Average	STDEV
Conventional	81.1	13.82
Method1	87.9	9.72
Method1+2	94.2	4.16

6.4.4 動作識別性能評価

Positive データの 2250 データと Negative データの 2940 データを用いて、動作識別性能を評価する。Positive データと Negative データ共に半数の 1125 データならびに 1480 データを学習データとして用い、残りの 1125 データならびに 1480 データを評価データとして用いる。実験結果として、式(3.18)の閾値 λ を変化させた際の検出率と誤検出率をプロットした ROC (Receiver Operating Characteristic) カーブを図 6.9 に示す。ここで示す検出率は Positive データを識別し、正しく検出した割合を表し、誤検出率は Negative データを起床動作として誤検知した割合を表す。実験の結果、検出率が 98[%]の際の各手法の誤検出率は、従来手法が 6.7[%], Method1 が 6.5[%], Method1+2 が 0.8[%]を示した。

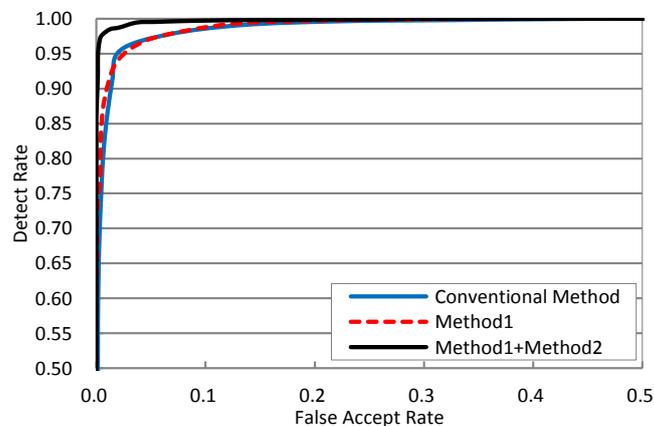


図 6.9 各手法に対する評価結果 (ROC カーブ)

6.4.5 多重解像に対する検出性能評価

多重解像化の効果の詳細を調査する為に、時間と空間の各々の多重解像化に対する検出性能を評価する。6.4.3 項の Method1 に空間の多重解像化のみを加えた手法を Method1+2s, Method1 に時間の多重解像化のみを加えた手法を

Method1+2t として、Positive データの 2250 データと Negative データの 2940 データを用いて、動作識別性能を評価する。Positive データと Negative データ共に半数の 1125 データならびに 1480 データを学習データとして用い、残りの 1125 データならびに 1480 データを評価データとして用いる。実験結果として、式(3.18)の閾値 λ を変化させた際の検出率と誤検出率をプロットした ROC カーブを図 6.10 に示す。

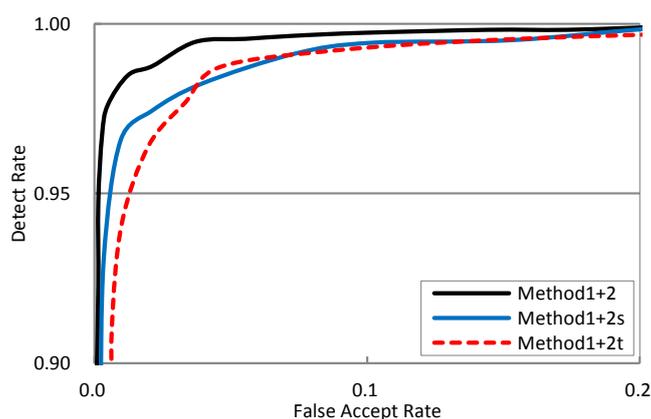


図 6.10 多重解像化に対する評価結果 (ROC カーブ)

6.5 考察

本章では、6.4 節の実験結果を踏まえて、提案手法におけるカメラ設置に対する検出性能の汎化性に関する考察と動作識別性能に関する考察を述べる。ただし、6.4 節の実験結果において、図 6.8 に示す通りカメラ位置 6 の検出精度が突出して低下した為、この問題に関する考察を先に述べる。

6.5.1 カメラ位置評価における検出性能低下

6.4 節の評価実験において、カメラ位置 6 の評価結果は図 6.8 に示す通り他のカメラ位置に比べて明らかな検出性能の低下が見られる。そこで、カメラ位置 6 の評価データを目視で確認したところ、図 6.11 の赤枠部で示す様に評価データの撮影中にベッド上にカーテンを乗せた状態で撮影されたシーンが含まれていた。この状態で起き上がり動作を行うと布団に接触したカーテンの揺れにより、起き上がり本来の特徴とは異なる特徴が抽出されるため検出精度の低下が生じ

た. 実用時にはベッドにカーテンが掛からない等の注意を加えて運用する事でこの問題は回避できるが, 今回の実験では検出性能の低下要因となり得るカーテンの掛かるシーンが評価データに含まれる.



図 6.11 カメラ位置 6 の撮影画像

6.5.2 カメラ位置に対する汎化性能

6.4.3 項の実験において, 他のカメラ位置と比較して起き上がり動作の見えの変化が最大となるのは, 図 4.10 の Pos3 と Pos10 である. 特に Pos10 は他のカメラ位置とは異なり, ベッドの長手方向の中心線が画面に対して鉛直方向となり, 起き上がり動作の見えが大きく異なる. 従来手法の評価結果を確認すると Pos10 の検出率は 76.2[%]であり, 近隣のカメラ位置 Pos7 の検出率 96.0[%]に比べて検出性能の低下が見られる. 一方, 提案手法 Method1 の評価結果においては Pos10 の検出性能は 90.4[%]であり, 従来手法と比較して 14.2[%]の検出率の向上が見られる. そこで提案手法である輝度勾配ベクトルの畳み込みを用いる事で識別器自身の検出傾向がどの様に変化したかを考察する.

AdaBoost は与えられた 2 クラスを分類する為に, 学習データから抽出された特徴に対して識別能力の高い特徴を優先的に弱識別器として採用する. 学習の初期に選択された特徴ほど与えられたデータの性質を良く表す情報であると言える. ここで, 6.4.3 項の実験において, 従来手法と提案手法 Method1 各々の学習初期に選択された特徴を表 6.2, 表 6.3 に記す. 学習開始から 5 順目迄に抽出された特徴のインデックスをカメラ位置毎に記載しており, 同一インデックスについては同色で示した. ここで, 学習回数 1 回目ならびに 2 回目に選択された特徴に注目すると, 提案手法 Method1 では全てのカメラ位置で同一の特徴が選択されているのに対して, 従来手法はカメラ位置 10 で選択された特徴が他の

カメラ位置で選択された特徴とは異なる．つまり従来手法はカメラ位置 10 の起き上がり動作と，他のカメラ位置の起き上がり動作の 2 種類の動作を，より傾向の異なる事象として認識したが，提案手法 Method1 は全てのカメラ位置の起き上がり動作を同一の傾向として認識している．この結果から，提案手法 Method1 は，従来手法と比較してカメラ位置の変動を吸収する特徴量の畳み込みを実現していると言える．

表 6.2 従来手法で選択された特徴インデックス

		Camera Position							
		2	3	4	5	7	8	9	10
Learning Iteration	1	3	3	3	3	3	3	3	36
	2	1	1	1	1	1	1	1	248
	3	285	285	245	285	238	285	238	3
	4	36	40	36	36	141	36	125	1
	5	236	34	248	320	29	320	41	285

表 6.3 提案手法で選択された特徴インデックス

		Camera Position							
		2	3	4	5	7	8	9	10
Learning Iteration	1	0	0	0	0	0	0	0	0
	2	61	61	61	61	61	61	61	61
	3	13	16	7	7	16	16	16	16
	4	32	77	80	80	60	40	60	60
	5	60	20	13	32	20	60	20	20

6.5.3 動作識別性能

提案手法 Method1 は，カメラ位置の変更による動作の見えに対する汎化性向上のみを目的としている為，単独の手法として用いた場合に動作の識別性能の低下を懸念したが，6.4.4 項の識別性能評価の結果では提案手法 Method1 は従来手法と同等の性能を示した．更に提案手法 Method2 と組み合わせる事で，提案手法の検出性能は検出率 98[%]，誤検出率 0.8[%]に向上した．この結果を踏まえて提案手法の実用性について検討する．今回の研究では実証実験の機会が得られない為，他の研究報告の評価結果の数値と本研究の検出性能の比較を試みる．比較対象は実用上有効な手法で，かつ実験データが存在する研究が望ましいが，現在病院で運用されているシステムで最も実績を持つシステムとしては，パラ

マウントベッド社の離床 CATCH が挙げられる。離床 CATCH は力センサを内蔵したベッドで、検出性能の高さから病院で圧力式ベッドセンサの代替品として普及している。初雁らの研究[145]によれば、病院での実証実験において、起き上がり動作の検出率 87[%], 誤検出率 3.1[%]である事が報告されている。ただし、本研究の誤検出率の算出方法が Negative データ数を母数として、Negative データを起き上がりと判定した割合であるのに対して、初雁らの研究の誤検出率の算出方法は、実証実験期間中に起き上がり動作と判定された回数を母数とする。そこで、比較の尺度を統一する為に、初雁らの実験結果を本研究の基準に変換する。以後便宜上、提案手法の誤検出率を机上誤検出率、初雁らの誤検出率を実評価誤検出率とする。

はじめに一定日数における起き上がり動作と、誤検出要因となる寝返り動作の発生頻度を検討する。4.1 節で述べた病院での起き上がり調査では、整形外科病棟の 40 代患者を 4 日間調査した結果、起き上がり動作は 53 回発生した。これより本考察における一日当たりの起き上がり回数を 13.25 回とする。また山下らの報告[146]によれば、成人男性 4 人の平均睡眠時間 335 分における寝返り回数は、4 人の平均で 16 回発生したが、患者は 24 時間仰臥状態にあるため、24 時間換算すると 1 日の寝返り回数は 68.77 回となる。

得られた数値を基に初雁らの評価結果から机上誤検出率を算出する。1 日の起き上がり発生回数 13.25 回、1 日の寝返り回数 68.77 回、初雁らの研究の検出率 87[%], 実評価誤検出率 3.1[%]であり、求める机上誤検出率を x と置くと、机上誤検出率と実評価誤検出率の関係は以下の式で表される。

$$\frac{68.77x}{13.25 \times 0.87 + 68.77x} = 0.031 \quad (6.4)$$

この式を解くと初雁らの手法の机上誤検出率は 0.536[%]となる。つまり、初雁らの手法は検出率 87[%]において机上誤検出が 0.536[%]発生する。図 6.9 に示した提案手法の実験結果では、検出率が 92[%]の際に机上誤検出率が 0.272[%]である。同一条件で比較した結果では無いが、本考察の前提条件下においては、実用実績のある初雁らの手法と比較して、本研究の検出手法は検出性能が高いと言える。

6.5.4 多重解像の効果

多重解像化の効果をもとに6.4.5項の実験結果から考察する。高解像ならびに低解像各々に対して、各データ種別（Negative データ、Positive データ）の画像特徴の平均を算出し、その算出結果に対する比較を行う。図 6.12 から図 6.14 は 6.4.5 項の実験において、時間の多重解像化を用いた際に、各マスクパターンで得られた動きの大きさを数値化した結果である。5.1.1 項で述べた通りマスクパターンの出力は識別器の特徴量であり、ここではマスクパターンの出力値を特徴量のスコアと呼ぶ。図 6.12 は低解像時のデータ種別毎の特徴量のスコアの平均値を示したものであり、図 6.13 は高解像度時のデータ種別毎の特徴量のスコアの平均値を示したものである。破線の値は各条件における、Positive データと Negative データのスコアの差分を示す。図 6.14 は、Positive データと Negative データのスコアの差分を抜き出して高解像度時と低解像度時の傾向を比較したグラフであり、値が高いほど Positive データと Negative データの判別に有効な特徴である事を示す。図 6.14 は高解像と低解像の両方の値を記載したが、先頭から 3 分の 1 に相当する 27 番目迄の特徴は低解像のスコアが高く、後尾から 3 分の 1 に相当する 55 番目以降の特徴は高解像のスコアが高い傾向が現れる。つまり、複数の解像度を用いた場合、どの解像度においても常に同じ動き特徴のパターンが高い識別能力を示すのではなく、識別能力の高い動き特徴のパターンは高解像と低解像で異なる。これは高解像の動き特徴のパターンが局所的な変動を捕捉する特徴であるのに対して、低解像の動き特徴のパターンは、高解像の局所的な動きの変動を包含した大局的な動きの変動を捕捉する為である。図 6.14 ではおよそ 3 分の 1 ずつのスコアが反転しており、提案手法の多重解像化が 6.4.4 節の実験において検出精度を向上させた要因と考えられる。更にスコアの反転が 3 分の 1 である理由をマスクパターンの構造に掘り下げて考察すると、スコアのインデックスとマスクパターンの変位点は図 6.18 の ABC に示す対応関係にある。つまり図 6.18(a)の C に示す後尾のスコアの上昇は、図 6.18(b)の変位点 C に示す画面下方向から上方向への動きを表しており、起床動作の画面下方向から画面上方へと頭部が移動する特性と一致する。空間に対する多重解像化に対しても、時間に対する多重解像化と同様の傾向を確認した。空間に対する多重解像の解析結果を図 6.15 から図 6.17 に記す。

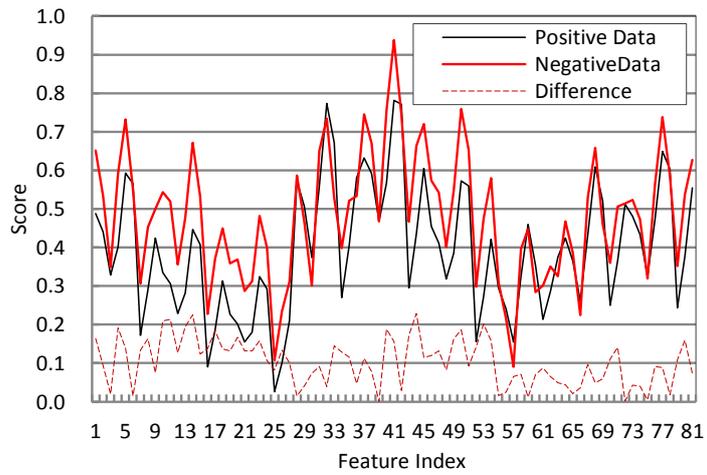


図 6.12 低解像度時のデータ種別毎の平均スコアと差分値（時間多重解像化）

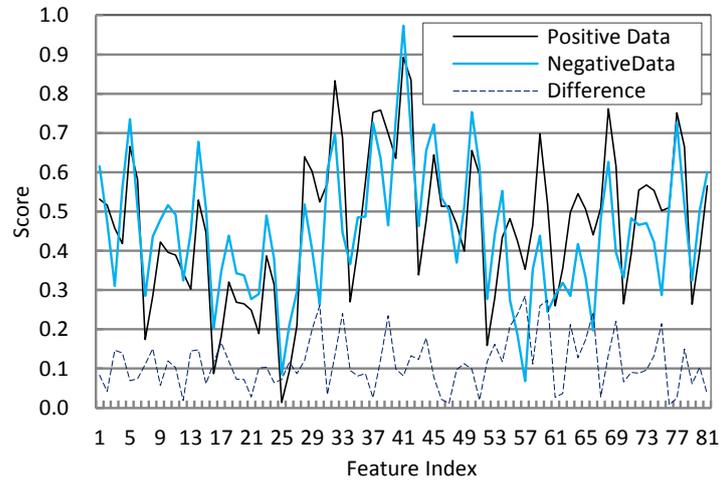


図 6.13 高解像度時のデータ種別ごとの平均スコアと差分値（時間多重解像化）

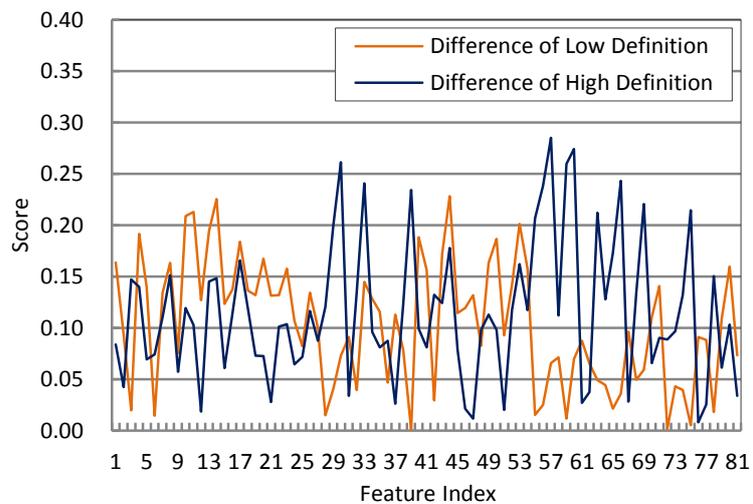


図 6.14 解像度毎の正例と負例のスコア差の比較（時間多重解像化）

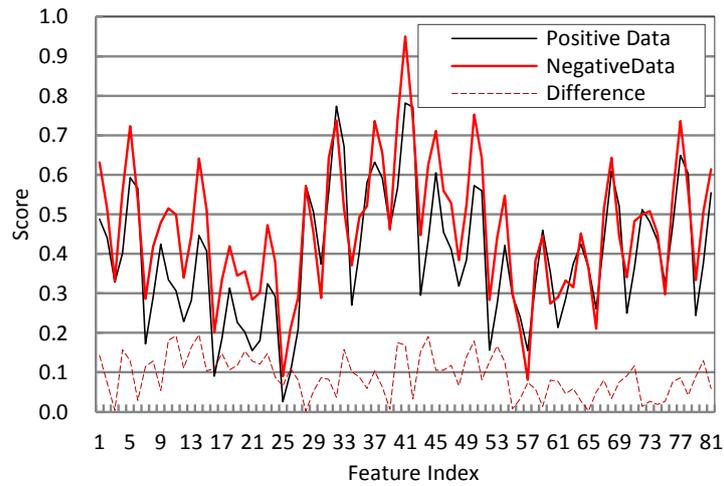


図 6.15 低解像度時のデータ種別毎の平均スコアと差分値（空間多重解像化）

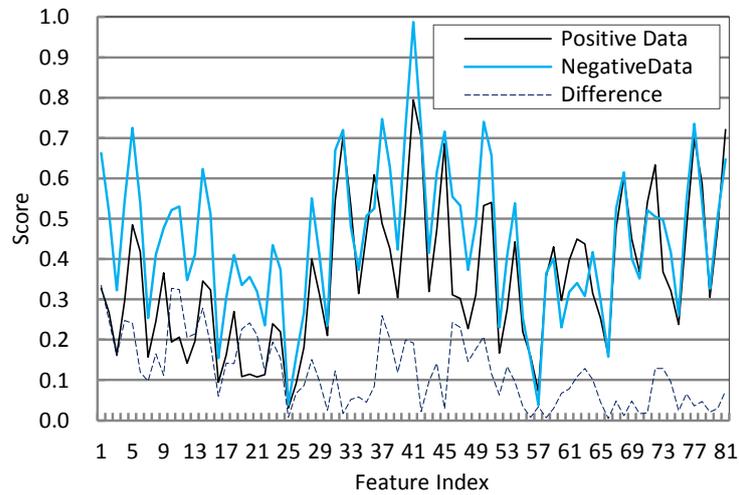


図 6.16 高解像度時のデータ種別ごとの平均スコアと差分値（空間多重解像化）

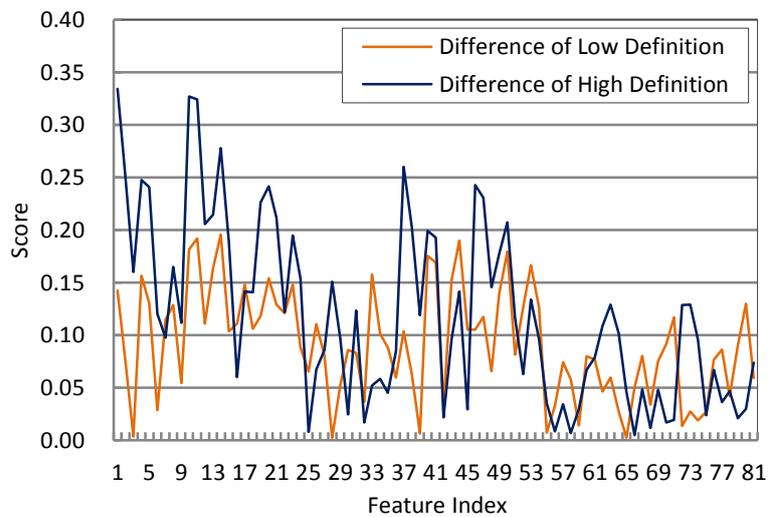


図 6.17 解像度毎の正例と負例のスコア差の比較（空間多重解像化）

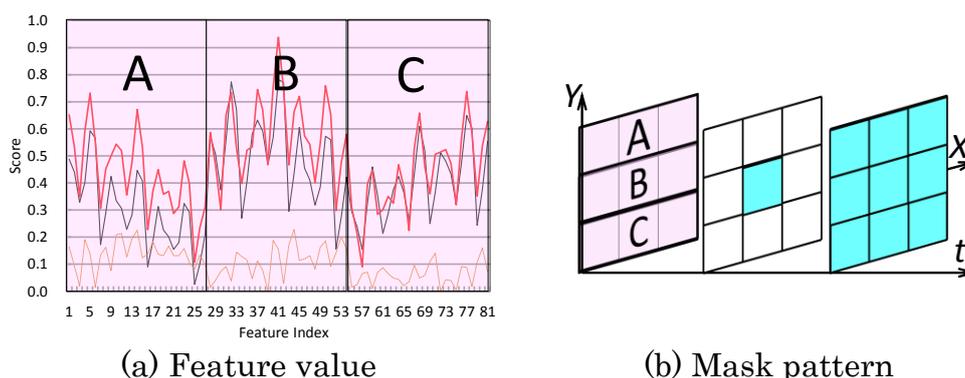


図 6.18 特徴のインデックスとマスクパターンの対応

6.5.5 誤検出について

寝返り動作は、カメラとベッドの位置関係に大きく依存した動き特徴が現れる。従来手法では寝返り動作と起床動作の局所的な特徴が類似してしまうため、誤検出が生じる。一方で、提案手法は動作の局所的な特徴だけでなく、時間ならびに空間的に広い範囲での特徴も同時に捉えることができるため、寝返り動作と起床動作の区別が容易になる。本実験で用いた寝返り動作は1視点のみであるが、実験で想定する45度程度の視点の変更では提案手法の識別性能は低下しないと考えられる、なぜなら両動作は視点に依存して異なる特徴が現れるが、寝返り動作が体の正面に対して左右方向の動きであるのに対し、起床動作は6.5.4項で示した様に上下の動きを伴うため、大局的な特徴を用いることで、どの視点においても両者の区別が可能となるからである。一方で、看護動作については様々な方向の動作が現れるため、ベッドとカメラ位置の関係には大きく依存しない。ここで、6.4.4項の実験において起き上がり動作の検出率を90[%]に揃えた際に発生した看護動作と寝返り動作の誤検出数を表6.4に示す。なお、検出率については6.5.3項で述べた初雁らの研究[145]を上回る値として90[%]と定めた。表6.4が示す様に本データセットで発生した誤検出は看護動作の1件のみである。発生した誤検出は図6.19に示す様に頭部の軌跡が起床動作の頭部の軌跡と類似するシーンである。実際の誤検出率を評価するためには、このような起床動作と類似した動作が実環境でどの程度の頻度で出現するかを考慮しなければならない。今後は、実環境での評価を行い、現場で発生する課題の抽出と改善に取り組む必要があると考える。



図 6.19 誤検出が発生したシーン

表 6.4 各データシーンにおける誤検出数

	Nursing Action	Rolling Over
Conventional	15	6
Method1	9	2
Method1+2	1	0

6.6 むすび

本章では、輝度勾配自己相関特徴のカメラ設置における検出性能の脆弱性を改善するために、輝度勾配ベクトルの統合手法を提案した。実運用の環境をシミュレーションして作成したデータを用いた評価の結果、カメラ毎の検出性能差を大幅に低減した。更に画像特徴の時空間多重解像化を組み合わせる事で検出率98[%]の際に、従来手法に対して誤検出率が5.9[%]低減する事を確認した。提案した二つの手法を組み合わせる事で、カメラ位置にロバストで高い検出精度の起床動作検出を実現した。

第7章 おわりに

本論文では、製品価格を考慮した安価な演算器を用いて、カメラ画像から起床動作を検出する手法について述べた。

高齢化が進行する日本において、現在の労働力の代替となる手段を講じる事は緊急の課題であり、看護業務においても現状に対する改善が必要である。看護業務で問題とされる三つの課題の一つとして挙げられる転倒転落問題を改善するために、画像処理を用いた起床動作検出手法を提案した。また、起床動作検出を行うに当たり、病院の設備を考慮した最適な実施方法ならびに製品を考慮した実用的な運用手段について述べた。

動画から人の行動を推定する手法は近年よく研究されているが、リアルタイムのカメラ画像に対して安価な演算器を用いて実時間で行動推定を行う為には、効率の良い検出手法を検討する事が重要になる。そこで従来研究に対する調査を行い、起床検知に有効な画像特徴を検討した。現在主流となる行動推定は、ハリスオペレータやSIFT等の特徴点を、時間軸と空間軸の双方で関連付けした時空間特徴を更にヒストグラム表現した特徴記述で表し、識別器を用いて行動クラスを判別する手法がある。しかし、この手法は計算処理が複雑になり安価な演算器を用いた処理では処理時間が増大する問題が発生する。そこで、先行研究の各処理過程を調査し、目的とする起床動作検出と従来研究の行動推定手法における3つの差異を明らかにした。一つ目は、画像特徴である。従来研究ではSIFT等の特徴点が用いられるが、SIFTの算出には多くの計算を必要とする。また、人物の行動推定にはSIFTの特徴点よりもHOG特徴が適している。二つ目は、移動カメラと固定カメラの違いである。従来研究はホームビデオの様な移動カメラの映像も考慮した検出手法を用いている。カメラ視点が移動する場合、背景の動きと対象の動きに対する区別が必要となり、動きの方向成分を識別する為にパーティクルフィルタ等の追跡処理を用いる必要がある。一方起床動作は病室の固定カメラを用いる為、カメラ位置は固定となる。固定カメラから対象を抽出する場合は動き成分をダイレクトで特定する事が出来る。三つ目は、動作の同定である。従来研究は検出された動きが複数の動作カテゴリのどのカ

カテゴリに属するかを判定する多クラス分類問題であるのに対し、起床動作検出では動画の動きが起床動作か否かを識別する2クラス分類問題となる。従来手法におけるカテゴリ識別は、特徴量からbag of keypointsを用いてベクトル化した特徴記述をクラスタリングする処理が用いられる。bag of keypointsを用いた静止画のカテゴリ識別では、一般に数百から数千の特徴が用いられるのに対して、特定物体の認識には数十万から百万の特徴量を必要とするため検出処理時間が増大する。ただし、行動推定の従来研究においては特徴のベクトル化が有効であることが示されている。以上の3つを明らかにし、本論文で時空間上の輝度勾配をベクトル表現に変換した新しい特徴量提案した。

先行研究調査の結果から導出した識別手法を実装し、その効果を確認した。画像の時空間特徴記述とベクトル化は輝度勾配自己相関を用い、識別手法にはAdaBoostを用いた。実験の制約事項としてカメラ位置はベッド頭部側の患者頭上の一点に固定した。実験の結果、提案手法が起床動作の検出において有効に機能する事を確認した。

次に、起床動作検出の実運用について検討した。病院や高齢者施設での実運用に対する最大の課題はカメラ設置位置の制約であり、実運用では様々なカメラアングルで患者の起き上がり動作を高精度に検出する必要がある。しかし、異なるカメラ位置で起床動作の検出性能を評価した結果、患者の動きから算出される画像特徴にばらつきが生じ、検出率は大きく低下した。そこで、カメラ位置により生じる画像特徴のばらつきを吸収する為に、特徴算出過程で勾配ベクトルを畳み込む処理を提案した。実験では勾配の畳み込みを用いる事で、どのカメラアングルにおいても優先的に選択される特徴が一致する事を示した。更に、多重解像化によりカメラアングルの大局的な変動を吸収する事で、動きに対して汎化性を向上させる手法を提案した。実験では低解像特徴と高解像特徴の有効成分を比較し、各々が異なる動きに対する識別能力を備えている事を示した。評価では様々なカメラアングルに対する起床動作の検出を行い、勾配の畳み込み処理と多重解像化の効果を確認した。また、現在病院で最も信頼されている転倒防止製品と比較し、本論文で提案した起床動作検出手法の信頼性を示した。

謝辞

本研究を行うにあたり，終始懇切丁寧なるご指導を賜りました名古屋工業大学 梅崎太造教授，岩田彰教授，北村正教授に謹んで感謝いたします。

また，本論文を纏めるにあたり有益なご助言を頂きました名古屋工業大学 田口亮助教，中部大学 保黒政大准教授，服部公央亮助手に深謝いたします。

最後に，研究に対して様々なご協力を頂いた梅崎研究室の皆様に感謝いたします。

参考文献

- [1] 総務省統計局: “高齢者の人口” <http://www.stat.go.jp/data/topics/topi721.htm>
- [2] 内閣府: “高齢化の国際動向”
http://www8.cao.go.jp/kourei/whitepaper/w-2014/zenbun/s1_1_5.html
- [3] 厚生労働省: “平均寿命の国際比較”
<http://www.mhlw.go.jp/toukei/saikin/hw/life/life10/03.html>
- [4] 社会保障審議会医療部会: “医療提供体制の改革に関する意見”
<http://www.mhlw.go.jp/stf/shingi/2r9852000001zh6h-att/2r9852000001zhag.pdf>
- [5] 川村治子: “ヒヤリ・ハット11,000事例によるエラーマップ完全本”. 医学書院, 2003.
- [6] Lv. Qingcong: “A poselet-based approach for fall detection.” In *IEEE International Symposium on IT in Medicine and Education*, Vol.2, pp. 209-212, 2011.
- [7] R. Caroline, J. Meunier, A. Saint-Arnaud, and J. Rousseau: “Monocular 3D head tracking to detect falls of elderly people”, In *IEEE 28th Annual International Conference on Engineering in Medicine and Biology Society*, pp.6384–6387, 2006.
- [8] J. S. Brown, E. Vittinghoff, J. F. Wyman, K. L. Stone, M. C. Nevitt, K. E. Ensrud and D. Grady : “Urinary incontinence: does it increase risk for falls and fractures? Study of Osteoporotic Fractures Research Group.” *Journal of the American Geriatrics Society*, Vol. 48, No. 7, pp. 721-725, 2000.
- [9] J. Gall, A. Yao, N. Razavi, G. L. Van and V. Lempitsky: “Hough Forests for Object Detection, Tracking, and Action Recognition.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 33, No. 11, pp. 2188-2202, 2011.
- [10] Q. V. Le, W. Y. Zou, S. Y. Yeung and A. Y. Ng: “Learning hierarchical invariant spatio-temporal features for action recognition with independent subspace analysis.”, In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3361-3368, 2011.
- [11] J. Yuan, Z. Liu and Y. Wu: “Discriminative video pattern search for efficient action detection.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 33, No. 9, pp. 1728-1743, 2011.

- [12] T. Guha and R. K. Ward: “Learning sparse representations for human action recognition.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 34, No. 8, pp. 1576-1588, 2012.
- [13] H. Wang, A. Kläser, C. Schmid and C. L. Liu: “Dense trajectories and motion boundary descriptors for action recognition.” *International journal of computer vision*, Vol. 103, No. 1, pp. 60-79, 2013.
- [14] Y. G. Jiang, Q. Dai, X. Xue, W. Liu and C. W. Ngo: “Trajectory-based modeling of human actions with motion reference points.” In *European Conference on Computer Vision*, pp. 425-438, 2012.
- [15] 厚生労働省: “平成24年(2012)人口動態統計の年間推計”,
<http://www.mhlw.go.jp/toukei/saikin/hw/jinkou/suikei12/>
- [16] 人口動態統計, 2012年次死亡確定数,
<http://www.e-stat.go.jp/SG1/estat/List.do?lid=000001108739>
- [17] 厚生労働省: “OECD加盟国の医療費の状況 (2011年)”,
http://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryoh/iryohouken/iryohouken11/
- [18] 内閣府: “医療イノベーション5か年計画”,
<http://www8.cao.go.jp/cstp/gaiyo/yusikisha/20120614/siryora-1.pdf>
- [19] 永井新二, 横山良樹, 時岡孝光: “当院における院内転倒の現状”, 日本職業・災害医学会会誌, Vol.53, No.2, pp.88-91, 2004.
- [20] 三重堀建雄, 丸山歩, 山田加代子: “当院における転倒・転落とその背景”, 日本医療薬学会年会講演要旨集, Vol.20, pp.402, 2010.
- [21] 小山照幸, 宮野佐年: “当院における院内転倒の実状”, 日本リハビリテーション医学会誌, Vol.43, pp.253, 2006.
- [22] 福田寛二, 斉藤政克: “当院における転倒・転落の発生状況”, 日本リハビリテーション医学会誌, Vol.43, pp.253, 2006.
- [23] 黒木洋美: “急性期総合病院における転倒・転落事故のまとめおよび防止策に対する取りくみ”, 日本リハビリテーション医学会誌, Vol.43, pp.253, 2006.
- [24] 佐伯覚, 松嶋康之, 蜂須賀研二: “当院における転倒・転落予防の取り組み”, 日本リハビリテーション医学会誌, Vol.45, pp.387, 2008.
- [25] 田代和也, 岡部武史, 太田秀臣, 上出正之, 三森教雄, 石原扶美丈: “院内転倒事故の検討 - より安全な医療のために -”, 慈恵医大誌, Vol.117, pp.91-95, 2002.

- [26] 平松知子, 泉キヨ子, 正源寺美穂: “施設高齢者の転倒予防 : 排泄に関連した転倒者の排泄状況および転倒の実態”, 日本老年看護学会誌, Vol.11, No.1, pp.47-52, 2006.
- [27] 高島幹子, 佐藤京子, 佐藤光子, 浅沼義博: “外科病棟における転倒・転落因子の検討”, 秋田大学学術情報リポジトリ秋田大学医学部保健学科紀要, Vol.15, No.1, pp.22-27, 2007.
- [28] 徳永誠次, 井口茂, 松坂誠應, 平瀬達哉, 武富敦子, 馬場文子: “急性期病院における入院患者の転倒状況とその対応”, 長崎大学学術情報リポジトリ保健学研究, Vol.24, No.1, pp.55-60, 2012.
- [29] 山崎佳奈, 池上直子, 藤本真里子, 松本雅子, 吉田優子, 山村愛子: “転倒転落防止に対する看護婦の意識”, 高知大学学術情報リポジトリ看護研究集録・臨床看護研究集録, Vol.2, pp.102-106, 2002.
- [30] 段重梅, 泉キヨ子, 平松知子: “施設高齢者における転倒予防自己効力感の測定 : 日常生活動作別の分析と関連要因”, 日本老年看護学会誌, Vol.10, No.2, pp.41-50, 2006.
- [31] 新野直明, 中村健一: “老人ホームにおける高齢者の転倒調査 : 転倒の発生状況と関連要因”, 日本老年医学会雑誌, Vol.33, No.1, pp.12-16, 1996.
- [32] 坂本望, 森山英樹, 今北英高, 前島洋, 吉村理, 白濱勲二: “介護老人保健施設痴呆専門棟における転倒の危険因子”, 日本職業・災害医学会会誌, Vol.52, No.3, pp.161-165, 2004.
- [33] 田丸典子, 桶晶子, 篠原裕子, 朱桃亜紀, 片瀬智子, 土田麻奈美, 小野塚恵, 宮中めぐみ, 山内由美子: “転倒を繰り返す脳疾患患者における転倒場面と転倒後の予防ケアの特徴 (第2報) : 3回以上転倒を繰り返した患者4名の分析から”, 金沢大学学術情報リポジトリ看護研究発表論文集録, Vol.37, pp.109-112, 2005.
- [34] 山田楼子, 高島幹子, 佐藤幸美, 伊藤亘, 伊藤登茂子, 浅沼義博: “転倒・転落に伴うインシデント事例の検討と対策 --発生構造に基づいた分類法を用いて--”, 秋田大学大学院医学系研究科保健学専攻紀要, Vol.18, No.2, pp.144-150, 2010.
- [35] 江藤文夫: “高齢者の転倒の原因”, 日医雑誌, Vol.22, No.13, pp.1950-1954, 1999.
- [36] 鈴木隆雄, 杉浦美穂, 古名丈人, 西沢哲, 吉田英世, 石崎達郎, 金憲経, 湯川晴美, 柴田博: “地域高齢者の転倒発生に関する身体的要因の分析的研

- 究：5か年の追跡研究から”，日本老年医学会雑誌，Vol.36，No.7，pp.472-478，1999.
- [37] 株式会社ホトロン：“製品案内” http://www.hotron.co.jp/product05_1.php
- [38] 株式会社マイクロ電子サービス：“製品紹介”
<http://www.mes-care.com/mwsa01.html>
- [39] キヨタ株式会社：“デジタルカタログ”
http://www.kiyota-and.co.jp/digital_catalog/_SWF_Window.html
- [40] パラマウントベッド株式会社：“離床CATCH”
<http://www.paramount.co.jp/contents/949>
- [41] 岡村輝久：“病院における転倒・転落--事故の法的責任”，医療，Vol.60，No. 1，pp.10-14，2006.
- [42] 鹿内清三：“患者のこころ(77). ”，医療のひろば，Vol.44，No.6，pp.8-12，2003.
- [43] 大阪地裁昭和46年4月19日判決，判時646号72項
- [44] 福島地裁会津若松支部昭和46年7月7日判決，下民集，Vol.22，No.7-8，pp.750
- [45] 中野貞一郎：“診療債務の不完全履行と証明責任。”現代損害賠償講座，Vol.4，pp.92，1979.
- [46] アイホン株式会社，“2014-2016総合カタログ”，
<http://www.aihone.co.jp/data/catalog/business/general-catalog/#page=345>
- [47] <http://www.xbox.com/en-US/xbox-one/accessories/kinect-for-xbox-one>
- [48] Shotton, Jamie, et al.: “Real-time human pose recognition in parts from single depth images.” *Communications of the ACM*, Vol.56, No.1, pp.116-124, 2013.
- [49] V. Megavannan, B. Agarwal, and R. V. Babu: “Human action recognition using depth maps”, In *IEEE International Conference on Signal Processing and Communications*, pp.1-5, 2012.
- [50] I. C. Albitar, P. Graebing, and C. Doignon: “Robust structured-light coding for 3D reconstruction computer vision”, In *IEEE International Conference on Computer Vision*, pp. 1-6, 2007.
- [51] S. T. O'Keefe, and J. N. Lavan: “Clinical significance of delirium subtypes in older people”, *Oxford Journals Medicine Age and ageing*, Vol. 28, No. 2, pp. 115-119. 1992.
- [52] J. T. Walker, S. P. Lofton, L. Haynie, and T. Martin: “The Home Health Nurses' Role in Geriatric Assessment of Three Dimensions: DEPRESSION, DELIRIUM, AND DEMENTIA.” *Home Healthcare Nurse*, Vol. 24, No. 9, pp. 572-578. 2006.

- [53] H. Watanuki, T. Kano, K. Kameyama, Y. Tsukui, R. Moroda, Y. Nakano, and K. Kanda: “The Prevalence and Cause of Delirium After Operation of Aged Surgical Patient in Urology Department”, *Gunma University School of Health Sciences*, Vol. 23, pp. 109-116, 2002.
- [54] P. Viola and M. Jones: “Rapid object detection using a boosted cascade of simple features.” In *IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp.511-518, 2001.
- [55] N. Dalal, B. Triggs: “Histograms of Oriented Gradients for Human Detection” In *IEEE Conference on Computer Vision and Pattern Recognition*, vol.1, pp.886-893, 2005.
- [56] A. Adam, E. Rivlin, I. Shimshoni and D. Reinitz: “Robust real-time unusual event detection using multiple fixed-location monitors.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.30, No.3, pp.555-560, 2008.
- [57] T. Nanri and N. Otsu: “Unsupervised Abnormality Detection in Video Surveillance.” In *IAPR International Conference on Machine Vision Applications*, pp.574-577, 2005.
- [58] L Itti P Baldi. “A principled approach to detecting surprising events in video.” In *IEEE Conference on Computer Vision and Pattern Recognition*, Vol.1, pp. 631-637, 2005.
- [59] F. Jiang, J. Yuan, S. Tsafaris, A. Katsaggelos: “Anomalous video event detection using spatiotemporal context.” *Computer Vision and Image Understanding*, Vol.115, No.3, pp.323-333, 2011.
- [60] L Kratz, K Nishino: “Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models.” In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1446-1453, 2009.
- [61] V. Mahadevan, W. Li, V. Bhalodia, and N. Vasconcelos. “Anomaly detection in crowded scenes.” In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1975-1981, 2010.
- [62] Y. Benezeth, P. Jodoin, V. Saligrama, and C. Rosenberger: “Abnormal events detection based on spatio-temporal co-occurrences.” In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2458-2465, 2009.
- [63] L. R. Rabiner and B. H. Juang. “An introduction to hidden Markov models.” *ASSP Magazine*, Vol.3, No.1, pp.4-16, 1986.

- [64] J.Kim and K.Grauman, "Observe locally, infer globally: A spacetime MRF for detecting abnormal activities with incremental updates." In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2921-2928, 2009.
- [65] V. Parameswara and R. Chellappa: "Human action-recognition using mutual invariants." *Computer Vision and Image Understanding*, Vol. 98, No. 2, pp. 294-324, 1998.
- [66] C. Rao and M. Shah: "View-invariance in action recognition." In *IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 316-322, 2001.
- [67] D. Ramanan and D. A. Forsyth: "Automatic annotation of everyday movements." In *Advances in Neural Information Processing Systems*, 2003.
- [68] M Blank, L Gorelick, E Shechtman, M Irani and R Basri: "Actions as space-time shapes." In *IEEE International Conference on Computer Vision*, Vol.2, pp.1395–1402, 2005.
- [69] M. D. Rodriguez, J. Ahmed and M. Shah: "Action mach a spatio-temporal maximum average correlation height filter for action recognition." In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8, 2008.
- [70] A. A. Efros, A. C. Berg, G. Mori and J. Malik: "Recognizing action at a distance." In *IEEE International Conference on Computer Vision*, pp.726-733, 2003.
- [71] C. Schüldt, I. Laptev and B. Caputo: "Recognizing human actions: A local SVM approach." In *IEEE International conference on pattern recognition*, Vol.3 pp.32-36, 2004.
- [72] C. Harris and M. Stephens: "A combined corner and edge detector." *Alvey Vision Conference*, pp. 147–152, 1988.
- [73] P. Dollár, V. Rabaud, G. Cottrell, S. Belongie: "Behavior recognition via sparse spatio-temporal features." In *IEEE International workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, pp. 65-72, 2005.
- [74] Kim, T. K., Wong, S. F., and Cipolla, R: "Tensor Canonical Correlation Analysis for Action Classification." In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8, 2008.
- [75] S. Nowozin, G. Bakir, and K. Tsuda: "Discriminative subsequence mining for action classification." In *IEEE International Conference on Computer Vision*, pp. 1-8, 2007.

- [76] R. Agrawal and R. Srikant: “Fast algorithms for mining association rules in large database.” In *International Conference on Very Large Data Bases*, pp. 487-499, 1994.
- [77] A. Gilbert, J. Illingworth and R. Bowden: “Scale invariant action recognition using compound features mined from dense spatio-temporal corners.” In *European Conference on Computer Vision*, pp. 222-233, 2008.
- [78] S. Savarese, A. DelPozo, J. C. Neibles, and L. Fei-Fei: “Spatio-Temporal correlations for unsupervised action classification.” in *IEEE Workshop on Motion and Video Computing*, pp. 1-8, 2008.
- [79] T. Hofmann: “Unsupervised learning by probabilistic latent semantic analysis.” *Machine Learning*, Vol. 42, No. 1-2, pp. 177-196, 2001.
- [80] Z. Zhang, Y. Hu, S. Chan, and L. T. Chia: “Motion Context: A New Representation for Human Action Recognition.” In *European Conference on Computer Vision*, pp. 817-829, 2009.
- [81] M. Bregonzio, S. Gong and T. Xiang: “Recognising action as clouds of space-time interest points.” In *IEEE Conference on Computer Vision and Pattern Recognition*. pp. 1948-1955, 2009.
- [82] G. Willems, T. Tuytelaars and L. Gool: “An efficient dense and scale-invariant spatio-temporal interest point detector.” In *European Conference on Computer Vision*, pp. 650-663, 2008.
- [83] A. Gilbert, J. Illingworth, and R. Bowden: “Fast realistic multi-action recognition using mined dense spatio-temporal features.” In *IEEE International Conference on Computer Vision*, pp. 925-931, 2009.
- [84] G. Csurka, C. Dance, L. Fan, J. Willamowski, C. Bray: “Visual categorization with bags of keypoints.” In *ECCV Workshop on Statistical Learning in Computer Vision*, Vol.1, No.1-22, pp.1-2, 2004.
- [85] J C Niebles, H Wang, L Fei-Fei: “Unsupervised Learning of Human Action Categories Using Spatial-Temporal Words,” *International journal of computer vision*, Vol.79, No.3, pp.299-318, 2006.
- [86] P. Scovanner, S. Ali and M. Shah: “A 3-dimensional SIFT descriptor and its application to action recognition.” In *ACM International Conference on Multimedia*, pp.357-360, 2007.

- [87] I. Laptev, M. Marszałek, C. Schmid and B. Rozenfeld: “Learning realistic human actions from movies.” In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8, 2008.
- [88] A. Kläser, M. Marszałek, C. Schmid: “A spatio-temporal descriptor based on 3D-gradients.” In *British machine vision conference*, 2008.
- [89] K. Mikolajczyk, et al.: “A comparison of affine region detectors” *International Journal of Computer Vision*, Vol.65, pp.43-72, 2005.
- [90] J. Liu, J. Luo and M. Shah: “Recognizing realistic actions from videos “in the wild”.” In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1996-2003, 2009.
- [91] M. Marszalek, I. Laptev, and C. Schmid: “Actions in context.” In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2929-2936, 2009.
- [92] D. Han, L. Bo, and C. Sminchisescu: “Selection and context for action recognition.” In *IEEE International Conference on Computer Vision*, pp. 1933-1940, 2009.
- [93] A. Kovashka, K. Grauman: “Learning a hierarchy of discriminative space-time neighborhood features for human action recognition.” In *IEEE Conference on Computer Vision and Pattern Recognition*, pp.2046-2053, 2010.
- [94] S. F. Wong and R. Cipolla: “Extracting Spatio Temporal Interest Points using Global Information.” In *IEEE International Conference on Computer Vision*, pp. 1-8, 2007.
- [95] A. Fathi and G. Mori: “Action recognition by learning midlevel motion features.” In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8, 2008.
- [96] K. Mikolajczyk and H. Uemura: “Action recognition with motion-appearance vocabulary forest.” In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8, 2008.
- [97] S. Savarese, A. DelPozo, J. C. Niebles, L. Fei-Fei: “Spatialtemporal correlations for unsupervised action classification.” In *IEEE Workshop on Motion and video Computing*, pp. 1-8, 2008.
- [98] H. Uemura, S. Ishikawa, K. Mikolajczyk: “FeatureTracking and Motion Compensation for Action Recognition.” In *British machine vision conference*, pp. 1-10, 2008.
- [99] Z. Zhao, A. Elgammal: “Information theoretic key frame selection for action recognition.” In *British machine vision conference*, pp. 1-10, 2008.

- [100] H. Wang, M. M. Ullah, A. Kläser, I. Laptev and C. Schmid: “Evaluation of local spatio-temporal features for action recognition.” In *British machine vision conference*, 2009.
- [101] W. Blendel and S. Todorovic: “Activities as time series of human postures.” In *European Conference on Computer Vision*, pp. 721-734, 2010.
- [102] N. Ikizler-Cinbis and S. Sclaroff: “Object, scene and actions: combining multiple features for human action recognition.” In *European Conference on Computer Vision*, pp. 494-507, 2010.
- [103] M. M. Ullah, S. N. Parizi and I. Laptev: “Improving bag-of-features action recognition with non-local cues.” In *British machine vision conference*, pp. 95-1, 2010.
- [104] S. Bhattacharya, R. Sukthankar, R. Jin and M. Shah: “A probabilistic representation for efficient large scale visual recognition tasks.” In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2593-2600, 2011.
- [105] S. Mathe and C. Sminchisescu: “Dynamic eye movement datasets and learnt saliency models for visual action recognition.” In *European Conference on Computer Vision*, pp. 842-856, 2012.
- [106] O. Kliper-Gross, Y. Gurovich, T. Hassner and L. Wolf: “Motion interchange patterns for action recognition in unconstrained videos.” In *European Conference on Computer Vision*, pp. 256-269, 2012.
- [107] S. Sadanand and J. J. Corso: “Action Bank: A high-level representation of activity in video.” In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1234-1241, 2012.
- [108] E. Vig, M. Dorr and D. Cox: “Saliency-based space-variant descriptor sampling for action recognition.” In *European Conference on Computer Vision*, 2012.
- [109] M. Jain, H. Jégou and P. Bouthemy: “Better exploiting motion for better action recognition.” In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2555-2562, 2013.
- [110] S. Ji, W. Xu, M. Yang and K. Yu: “3D Convolutional Neural Networks for Human Action Recognition.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 36, No. 1, pp. 221-231, 2013.
- [111] H. Wang and C. Schmid: “Action recognition with improved trajectories.” In *European Conference on Computer Vision*, pp. 3551-3558, 2013.

- [112] M. Hoai and A. Zisserman: “Improving Human Action Recognition using Score Distribution and Ranking.” In *Asia Conference on Computer Vision*, pp. 3-20, 2014.
- [113] V. Kantorov and I. Laptev: “Efficient feature extraction, encoding and classification for action recognition.” In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2593-2600, 2014.
- [114] M. Sapienza, F. Cuzzolin and P. H. Torr: “Learning discriminative space-time action parts from weakly labelled videos.” *International Journal of Computer Vision*, Vol. 110, No. 1, pp. 30-47, 2014.
- [115] Mitsuru Ambai, Yuichi Yoshida: “CARD: Compact And Real-time Descriptors”, In *IEEE International Conference on Computer Vision*, pp.97-104, 2011.
- [116] H. Wang, A. Klaser, C. Schmid, C. L. Liu: “Action recognition by dense trajectories.” In *IEEE Conference on Computer Vision and Pattern Recognition*, pp.3169-3176, 2011.
- [117] 柳井啓司: “物体認識技術の進歩. ”, 日本ロボット学会誌, Vol.28, No.3, pp.257-260, 2010.
- [118] T. Kobayashi and N. Otsu: “Image feature extraction using gradient local auto-correlations.” In *European Conference on Computer Vision*, pp.346-358, 2008.
- [119] T. Kobayashi, and N. Otsu: “Action and Simultaneous Multiple-Person Identification Using Cubic Higher-Order Local Auto-Correlation,” In *IEEE International Conference on Pattern Recognition*, pp.741–744, 2004.
- [120] 物体検出装置及びインターホンシステム, 特願2012-056086
- [121] M.V.Srinivasan: “Generalized gradient schemes for the measurement of two-dimensional image motion”, *Biological Cybernetics*, Vol.63, pp.421-431, 1990.
- [122] D. G. Lowe: “Distinctive Image Features from Scale-Invariant Keypoints”, *International Journal of Computer Vision*, Vol.60, No.2, pp.91-110, 2004.
- [123] H. Bay, T. Tuytelaars, and L. V. Gool: “SURF:Speed Up Robust Features”, In *European Conference on Computer Vision*, pp.1-15, 2006.
- [124] T. Watanabe, I. Satoshi and Y. Kentaro, “Co-occurrence histograms of oriented gradients for pedestrian detection”, *Advances in Image and Video Technology*, pp.37-47, 2009.

- [125] T. Mitsui and F. Hironobu. "Object detection by joint features based on two-stage boosting." In *IEEE International Conference on Computer Vision Workshops*, pp.1169-1176, 2009.
- [126] T.Nanri, N.Otsu: "Unsupervised Abnormality Detection in Video Surveillance," In *IAPR International Conference on Machine Vision Applications*, pp.574-577, 2005.
- [127] T. Kobayashi, N. Otsu: "A Three-way Correlation-based Method for Human Identification by Gait," In *IEEE Workshop on Visual Surveillance.*, Vol. 1, p. 4, 2006.
- [128] R. Lawrence: "On the use of autocorrelation analysis for pitch detection." *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol.25, No.1, pp.24-33, 1977.
- [129] M. Suzuki: "A Solid Texture Classification Method Based on 3D Mask Patterns of Higher Order Local Autocorrelations." In *International Workshop on Texture analysis and Synthesis*, Vol.48, No.3, pp.1524-1531, 2007.
- [130] Vapnik Vladimir. : "Pattern recognition using generalized portrait method." *Automation and remote control*, Vol24, pp.774-780, 1963.
- [131] Burges. C. J. C: "A tutorial on support vector machines for pattern recognition." *Knowledge Discovery and Data Mining*, Vol.2, No.2, pp.121-167, 1998.
- [132] Cristianini, N. and J.Shawe-Taylor. : "Support vector machines and other kernel-based learning methods." *Cambridge University Press*.
- [133] Müller, Klaus-Robert, et al.: "An introduction to kernel-based learning algorithms." *IEEE Transactions on Neural Networks*, Vol.12, No.2, pp.181-202, 2001.
- [134] Schölkopf, Bernhard, et al.: "New support vector algorithms." *Neural computation*, Vol.12, No.5, pp.1207-1245, 2000.
- [135] Herbrich Ralf.: "Learning kernel classifiers." *MIT Press*, 2002.
- [136] R.E.Schapire: "Using output codes to boost multiclass learning problems." In *International Conference on Machine Learning*, Vol. 97, pp.313-321, 1997.
- [137] R.E.Schapire, Y.Singer: "Improved boosting algorithms using confidence-rated predictions." *Machine learning*, Vol.37, No.3, pp.297-336, 1999.
- [138] R.E.Schapire, Y.Freund, P.Bartlett, W.S.Lee: "Boosting the margin: A new explanation for the effectiveness of voting methods." *Annals of statistics*, pp.1651-1686, 1998.
- [139] V. Koltchinskii, D. Panchenko: "Empirical margin distributions and bounding the generalization error of combined classifiers." *Annals of Statistics*, pp.1-50, 2002.

- [140] Y.Freund, R.E.Schapire: “A desicion-theoretic generalization of on-line learning and an application to boosting.” *Computational learning theory*, pp.23-37, 1995.
- [141] INRIA Person Dataset, <http://pascal.inrialpes.fr/data/human/>
- [142] JISC 日本工業標準調査会, <http://www.jisc.go.jp/app/pager?id=2789157>
- [143] A. Krizhevsky, I. Sutskever, and G. E. Hinton. : “Imagenet classification with deep convoluntional neural networks.”, In *Advances in neural information processing systems*, pp. 1097-1105, 2012
- [144] J. Byrne and J. Shi: “Nested Shape Descriptors.” In *IEEE International Conference on Computer Vision*, pp. 1201-1208, 2013.
- [145] 初雁卓郎, 椎野俊秀, 村井真也: “ベッド上の患者行動を推定・通知する「離床CATCH」の提案”, *労働科学*, Vol.88, No.3, pp. 94-102, 2012.
- [146] 山下浩史, 佐藤慎, 徳増来斗, 兼子光治, 鈴木祐介, 一之瀬大資, 大城昌平: “睡眠時の体動と嚔下の関係性について”, 第47回理学療法学会大会, Ae0050-Ae0050, 2012.

研究業績一覧

学術論文（査読あり） 主：2，副：0

- [1] 井上円, 田口亮, 梅崎太造, “輝度勾配自己相関特徴を用いた転倒防止のための起床動作検出”, 電気学会論文誌C, Vol.134, No.2, pp.242-249, 2014.
- [2] 井上円, 田口亮, 服部公央亮, 梅崎太造, “輝度勾配ベクトルの畳み込みを用いたカメラ設置範囲にロバストな起床動作検出”, 電気学会論文誌C, Vol136, No3 Accepted

国際会議（査読あり） 主：1，副：0

- [3] **Madoka Inoue**, Ryo Taguchi, Taizo Umezaki, “Vision-Based Detection of Getting-up Behavior for Fall Prevention”, WMSCI 2013 World Multi-Conference on Systemics, Cybernetics, and Informatics, pp.224-228, 2013.

研究会・口頭発表（査読あり） 主：3，副：0

- [4] 井上円, 田口亮, 梅崎太造, “自己相関特徴とコヒーレンスに基づく推定手法を用いた転倒防止のための起床動作検出”, FIT(情報科学技術フォーラム), RI-001, 第3分冊pp.1-6, 2013.
- [5] 井上円, 田口亮, 服部公央亮, 梅崎太造, “輝度勾配ベクトルの畳み込みを用いた起床動作の検出精度向上”, 第54回日本生体医工学会大会, P1-4-2-F, pp. P1-4-2-F-177, 2015.
- [6] 井上円, 田口亮, 服部公央亮, 梅崎太造, “実利用の制約条件下における起床動作検出システムの検討”, SSII (Symposium on Sensing via Image In-formation), IS3-20, pp.IS3-20-1-IS20-5, 2015.