

動的人間行動モデリングによる一人暮らしの高齢者の生活支援システムに関する研究

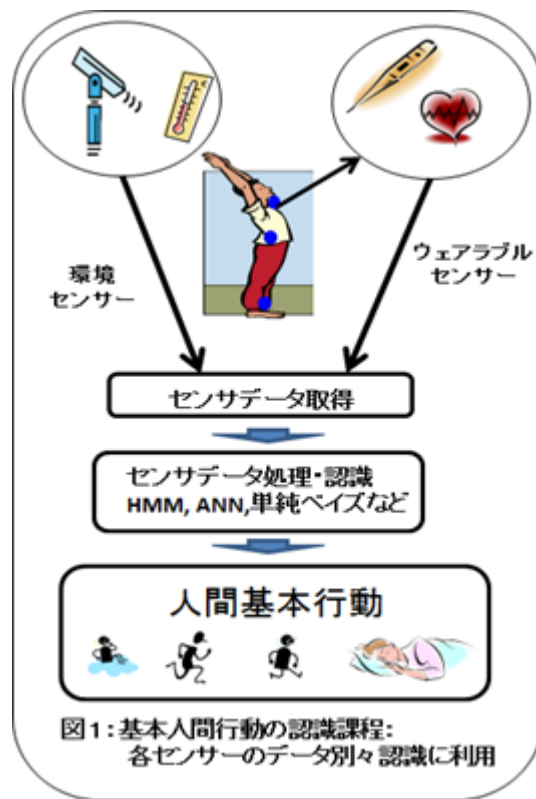
代表研究者 Stephen Karungaru 徳島大学大学院ソシオテクノサイエンス研究部

1 はじめに

近年、高齢者に占める単身者数の割合が増加している。総務省「国勢調査」、国立社会保障・人口問題研究所「日本の将来推計人口（2013年1月）」に基づき推計した、65歳以上に占める単身者数割合は、2年後で31.62%なる。また2030年には50代の3人に1人が一人暮らししているという報告がある（藤森克彦「2010単身急増社会の衝撃」）。その結果、高齢者の介護が困難となり、老老介護も増加すると予想されている。家庭でも高齢者の介護や看護、見守りシステムが必要とされている。しかし、現在、一般家庭にはそのような見守りシステムは存在しない。そのため、孤独死も度々発生しており（「2012 高齢者社会白書」）、高齢者介護や見守りシステムの開発は喫緊の課題である。したがって、高齢者が1人でも健康で安全な生活を送れるような見守り・サポートシステムの開発が必須である。

一人暮らしの高齢者が、家庭での自立した生活を継続できる生活支援システムを実現するためには、人間行動モデリングの実現が不可欠である。本研究は、ウェアラブルセンサー（心拍、加速等）や家庭にあるセンサー（カメラ等）のデータをニューラル・ネットワーク（特に追加学習機能）を使って生活支援システムの構築を目的に、人間行動モデリングの実現を図ることを目指す。本研究ではカメラとソフトコンピューティング手法を組合せ、人間行動に加えその動作を行った**時間・場所を観察できる動的人間行動モデリング**を実現する。モデルの例は「毎朝台所で新聞を読む」や「お昼居間でテレビを見る」などの日常生活における人の行動のことである。このシステムを、生活リズムを崩す可能性がある一人暮らしの老人の安心・見守りに適用することが目的であり、将来的には介護生活に陥ることのない安心・安全な日常生活が過ごせる生活のアドバイスをを行うことを目指す。現在、人間行動の研究は、カメラによる視覚と他種センサーに基づいた「**基本行動認識**」が多く行われている（河口 et. al, 2011）。それらの提案されている手法はサポートベクタマシン、ベイズ・クラシファイヤー、k-近傍法、線形判別分析、隠れマルコフモデル、単純ベイズ分類器、決定木及び人工ニューラルネットワーク（以後、ANN と略記）を用いている。基本行動は「座る、歩く、寝る」などである。しかし、基本行動認識だけでは時間と場所が含まれず、人間の生活リズムを把握しにくいことが判った。そこで行動認識と同時に**動作を行った時間と場所を加えて学習させる**と人間の生活リズムを把握できると考えられる。これを**動的人間行動モデリング**と呼び、ひとり暮らしの老人が安全・安心に暮らせる見守りシステムに応用することを考える。この動的人間行動モデリングの実現には、基本行動の認識をベースに複合センサーの使用と同時に基本動作認識と行動時間・場所情報を考慮した ANN（追加学習）、遺伝的アルゴリズムや決定木等の手法の組合せが有効と考えられる。申請者は、これまでの研究によって「動的人間行動モデリング」を実現するために必要な基本的な成果を得ている。

申請者はカメラを利用した室外内での人間行動を研究してきたが、カメラの視野や画質に依存した限界を克服するために複合センサーの使用・行動時間・場所情報を含む人間行動モデリングの実現が必要となった。そこで他のセンサーデータ（ウェアラブルセンサーや室内のセンサー）を追加することにより、動的人間行動モデリングを実現する本研究では、追加学習・決定木・遺伝的アルゴリズムを導入することにより動的人間行動モデリングを実現する。基本行動は「座る、歩く、寝る」などである。その認識過程を図1に示す。



購入した監視カメラシステムから8台のカメラは研究室に設置しました。人物領域は背景差分を用いた候補領域から、HOG 特徴量を用いて検出を行う。特徴量を評価するデータはAdaBoost を用いて作成されたものを使用している。複数のカメラで同一の人物を検出している場合は射影変換を用いた座標系の変換によりマッチングを行う。図1に示すような複数のカメラの映像に対し、実験環境を模した図1(左下)に示す画像と対応点を取得、変換パラメータを算出する。

2 提案手法

本研究は、身につける（ウェアラブル）センサー（心拍、加速等）や家庭にあるセンサー（カメラ等）のデータに対し、決定木と遺伝的アルゴリズムを利用して入力ベクトルを作成する。その後ニューラルネットワーク（以後、ANN と略記する）（追加学習）、で学習させ、人間行動モデリングを認識する。

研究計画として

1. 環境・ウェアラブルセンサーデータと ANN を利用して人間の基本行動を認識する。そのときセンサーデータの重要性を判断するため遺伝的アルゴリズムを利用する。
2. 行動時間・場所のデータを加えて決定木により特徴ベクトルを作成する方法を開発する。

3. 候補領域抽出

閾値の自動設定には判別分析法を用いる。判別分析法とはクラス内分散とクラス間分散との比から分離度を求め、分離度が最大となる閾値を求める手法である。

図2.1のように閾値 t で二値化した場合、閾値よりも低レベル側の画素数を ω_{-1} 、平均を m_{-1} 、分散を σ_{-1} 、閾値よりも高レベル側の画素数を ω_{+1} 、平均を m_{+1} 、分散を σ_{+1} 、画像全体の画素数を ω_t 、平均を m_t 、分散を σ_t とする。ここでクラス内分散 σ_w^2 の計算式を式(3.1)、クラス間分散 σ_b^2 の計算式を式(3.2)に示す。

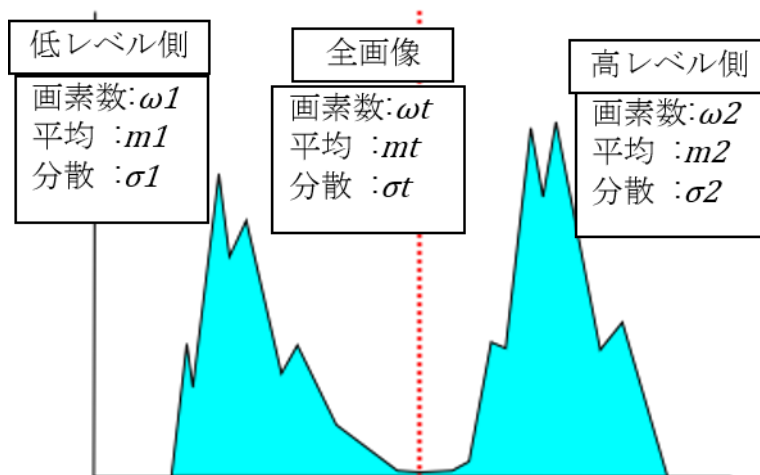


図 3.1: ヒストグラム

$$\sigma_w^2 = \frac{\omega_1 \sigma_1^2 + \omega_2 \sigma_2^2}{\omega_1 + \omega_2} \quad (3.1)$$

$$\sigma_b^2 = \frac{\omega_1 \omega_2 (m_1 - m_2)^2}{(\omega_1 + \omega_2)^2} \quad (3.2)$$

全分散 σ_t の計算式を式(2.3)に示す. よって分離度を求める式は式(2.4)となる. ここで, 全分散 σ_t は閾値に関係なく一定であるため, クラス間分散 σ_b^2 が最大となる閾値 t を求める. また, 式(2.2)の分母は一定であるため式(2.5)の式を用いて計算を行う.

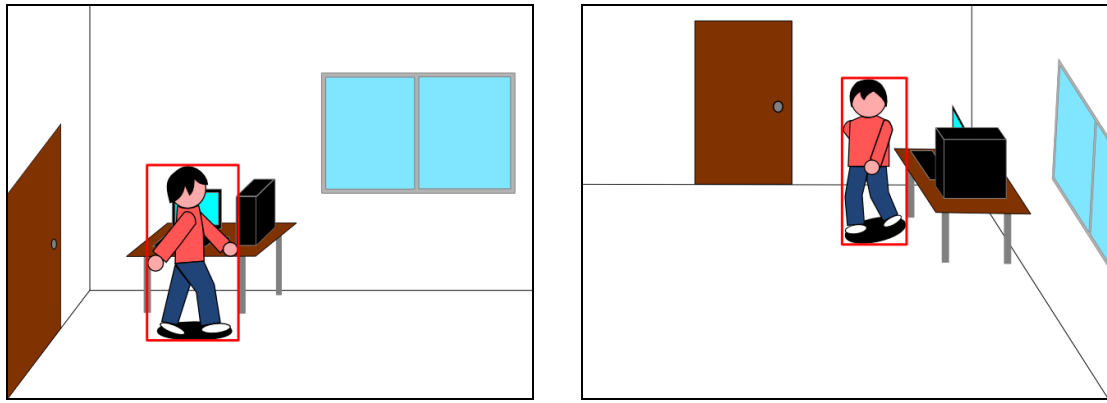
$$\sigma_t^2 = \sigma_b^2 + \sigma_w^2 \quad (3.3)$$

$$\frac{\sigma_b^2}{\sigma_w^2} = \frac{\sigma_b^2}{\sigma_t^2 - \sigma_b^2} \quad (3.4)$$

$$\omega_1 \omega_2 (m_1 - m_2)^2 \quad (3.5)$$

4. 候補領域の統合

図 4.1(a), (b)に示すような任意の位置に設置されたカメラから取得する画像から同一の物体を検出しているか判断する必要がある. 色相や形状を用いたマッチング手法では, 人物を検出している場合に服装の色相が様々であることや撮影する角度により形状が大きく変化するという問題が生じる. そこで, それぞれの映像における床の座標を合わせることで検出位置から同一物体のマッチングが可能になる.



(a) 一方のカメラ

(b) 他方のカメラ

図 4.1 候補領域検出

本システムでは、あらかじめ用意された実験環境を模したマップを利用し、各カメラの映像から床領域の座標をマップに合わせるように変換することで、それぞれのカメラの検出を統合する。本節では床領域の座標変換に用いる射影変換について説明し、次に候補領域のマップ座標への変換について説明する。候補領域の座標 (x, y) をマップにおける座標 (u, v) に変換する式を式(3.1)に示す。

$$u = \frac{ax + by + c}{gx + hy + 1}$$

$$v = \frac{dx + ey + f}{gx + hy + 1}$$

(4.1)

式(4.1)における (a, b, c, d, e, f, g, h) は射影変換係数であり、式(3.2)に示す方程式を解くことで求められる。

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix} = \begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -X_1x_1 & -X_1y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -X_2x_2 & -X_2y_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -X_3x_3 & -X_3y_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -X_4x_4 & -X_4y_4 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -Y_1x_1 & -Y_1y_1 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -Y_2x_2 & -Y_2y_2 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -Y_3x_3 & -Y_3y_3 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -Y_4x_4 & -Y_4y_4 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{pmatrix}$$

(4.2)

5. 人物検出

人物を検出するための特徴量として HOG(Histograms of Oriented Gradients)特徴量を利用する。HOG 特徴量から人物を特定するためのデータを AdaBoost により作成する手法を利用した。また、複数のカメラで候補領域が検出されている場合、各カメラにおいて算出された強識別器から総合の強識別器を用いて認識を行った。

6. 行動認識

検出された人物に対して基本行動として“歩行”，“走行”，“停止”，“着席”，“デスク作業”，“転倒”，“会話”行動の認識を行う．認識には検出領域の大きさや画像から得られる特徴を利用している．行動認識のフローチャートを図3に示す．さらに“荷物の放置”，“うろつき”行動を不審行動として認識を行う．“荷物の放置”行動は人物領域から切り離された人物ではない領域が存在する場合に認識される．また，“うろつき”行動は一つのカメラにおいて長時間“歩行”認識されている人物が存在している場合に認識される．

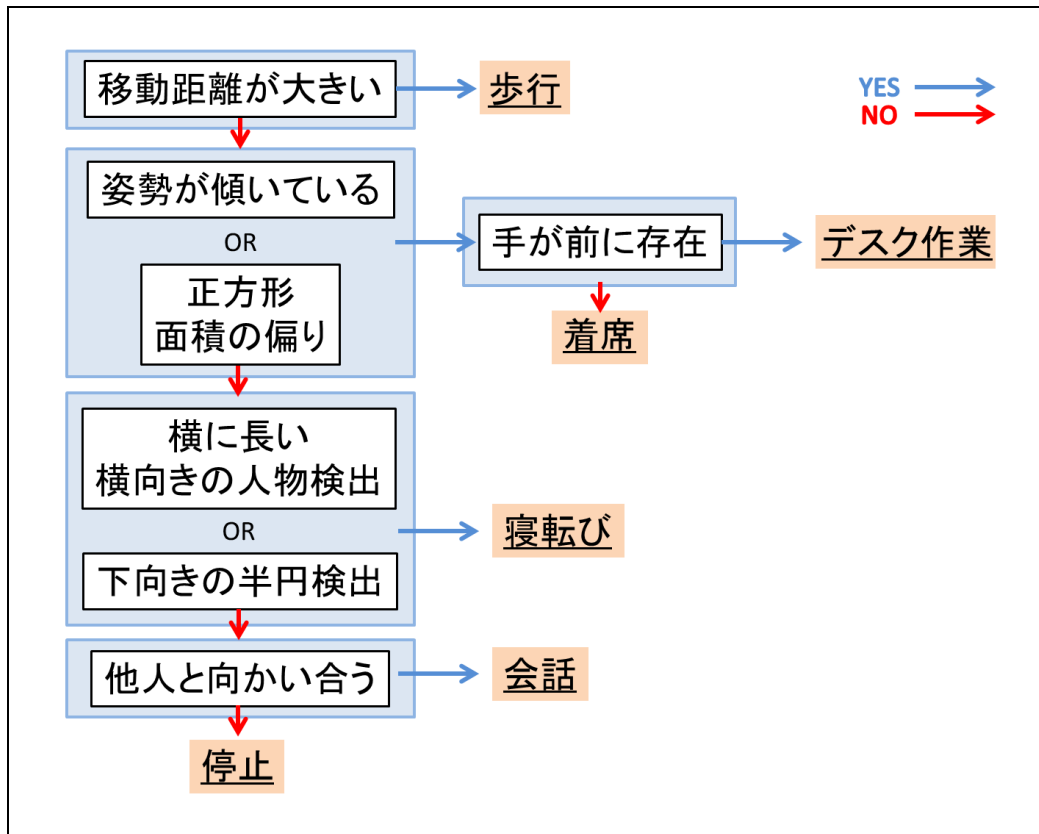


図6：行動認識フローチャート

7. 実験

実験環境として情報センター8階に8台のカメラを設置し，人物の行動認識を実施し，804号室の稼働率を算出する．行動認識結果及び稼働率を表示するためのマップを図7.1に示す．このマップは情報センター8階804号室および廊下を簡易的に描いたものである．

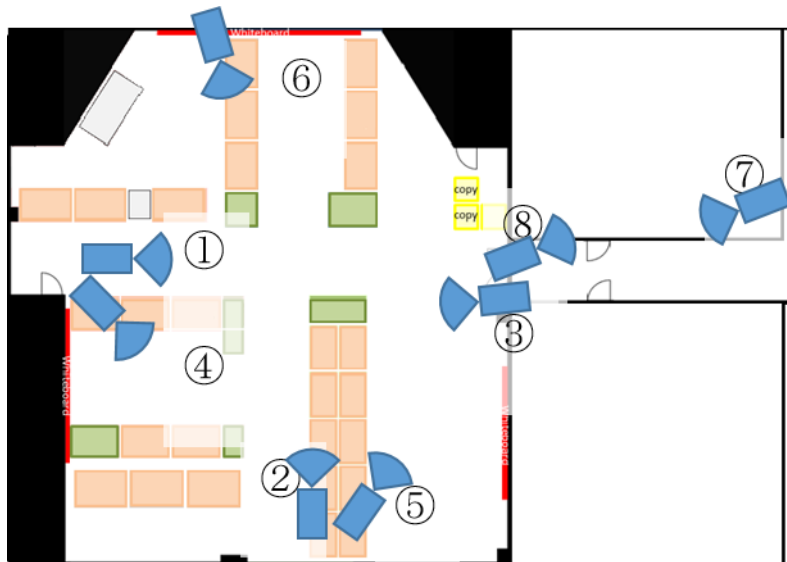


図 7. 1:実験に用いるマップと監視カメラの設置位置

すべてのカメラにおける人物検出領域の足元の座標を変換パラメータにより図 7. 2 の画像に反映させることで、近傍に存在する他カメラの検出が同一の人物であると判断することができる。設置するカメラはアナログカメラを用いた監視カメラシステムである H.264 Digital Video Recorder (ZMOD0 社)を用いる。

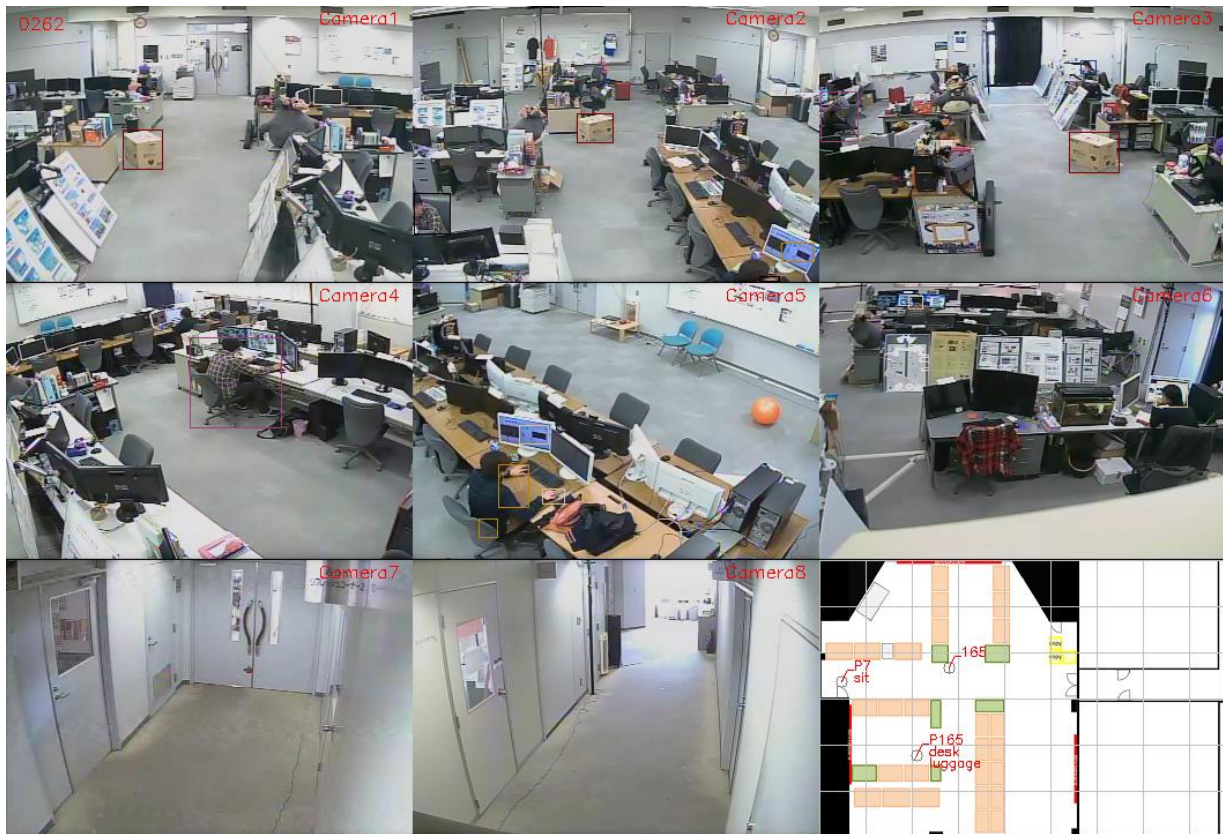


図 7. 2:取得画像と環境を模したマップ

Table 1. Human detection Accuracy

Correct	Detected	Accuracy
4608	4565	99.0

Table 2. Recognition accuracy

	Walk	Run	Stop	Sit	Desk	Fall	Talk	Leave	Accuracy
Walk	1305		157	20	6				87.7%
Run	28	33							45.9%
Stop			685	64					91.5%
Sit	3			1382	67	4			94.9%
Desk				196	1566				88.9%
Fall	2			78		821			91.1%
Talk			321				407		55.9%
Leave								4	100%

おわりに

本研究では複数のカメラを用いて人物の行動認識を行い、ログとして記録するシステムを作成し、出力されたログを用いて撮影されている部屋の使用率や稼働率を算出した。まず、序論として行動認識技術の利用分野の拡大と監視カメラを使用する利点について述べた。第2章では、ブロックを用いた判別分析法により、入力画像から背景差分を用いて候補領域を抽出する方法について述べた。第3章ではそれぞれのカメラにおいて同一の候補領域をマッチングする手法について述べ、さらに足元が隠れている候補領域への対策について述べた。第4章では人物検出に用いるためのHOG特徴量について説明し、HOG特徴量を評価するデータをAdaBoostにより作成する方法について述べた。また、複数のカメラで検出されている場合の処理についても述べた。第5章では認識する行動である“歩行”行動、“停止”行動、“着席”行動、“デスク作業”行動、“寝転び”行動、“会話”行動、“荷物の持ち込み”行動の特徴と認識方法について述べた。第6章ではシステムの最終処理として人物の検出および認識された行動から監視している部屋の使用率および稼働率の算出方法について述べた。第7章では本システムの有効性を証明するために複数のシーンに対して実験を行い、認識精度を示した。全ての行動に対して高い認識精度を得られたことから本システムの有効性を確認することができた。

今後の課題として、変動する撮影環境への対応や認識できる行動の追加が挙げられる。

【参考文献】

- [1] F. Lv, X. Song, B. Wu, V. Kumar Singh, and R. Nevatia, Left-Luggage Detection using Bayesian Inference, 9th Intl. Workshop on Performance Evaluation of Tracking and Surveillance (PETS-CVPR'06), 2006.
- [2] Rubner, Y., Tomasi, C.; Guibas, L. J. The Earth Mover's Distance as a metric for image retrieval. Int. J. Comput. Vision, pp. 40-99, 2000.
- [3] E. Auvinet, E. Grossmann, C. Rougier, M. Dahmane and J. Meunier, Left-luggage detection using homographies and simple heuristics, IEEE international workshop on performance evaluation in tracking and surveillance(PETS), 2006.
- [4] 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.
- [5] A. Yilmaz and M. Shah. Actions Sketch: A Novel Action Representation, In IEEE Conference on Computer Vision and Pattern Recognition, pp.984-989, 2005.
- [6] I. Laptev and T. Lindeberg. Space-time interest points, In IEEE International Conference on Computer Vision, pp.432-439, 2003.

- [7] C. Rao, A. Yilmaz and M. Shah. View-Invariant Representation and Recognition of Actions, In International Journal of Computer Vision 50(2), pp.203-22, 2002.
- [8] F. Lv and R. Nevatia. Recognition and Segmentation of 3-D Human Action using HMM and Multi-Class AdaBoost, In European Conference on Computer Vision, pp. 359-372, 2006.
- [9] S. Karungaru, K. Terada and M. Fukumi , Detection of Abandoned Luggage and Owner Tracking at Sensitive Public Areas, IEEJ Transactions on Electronics, Information and Systems, Vol.133-C, No.1, 2012.
- [10] H. Hung Liao, J. Ying Chang, and L. Gee Chen, A Localized Approach to Abandoned Luggage Detection with Foreground-Mask Sampling, IEEE Fifth International Conference on Advanced Video and Signal Based Surveillance, pp.133-139, 2008.
- [11] G. Gomez, and C. Orrite-Uruuela, Automatic Left Luggage Detection and Tracking Using Multi-Camera UKF, 9th PETS,CVPR, IEEE, pp. 59-66, 2006.
- [12] L. Li, R. Luo, R. Ma, W. Huang, K. Leman, Evaluation of AnIVS System for Abandoned Object Detection on PETS 2006 Datasets, 9th PETS, CVPR, IEEE, pp. 91-98, 2006.
- [13] N. Dalal and B. Triggs, Histograms of Oriented Gradients for Human Detection In Proceedings of IEEE Conference Computer Vision and Pattern Recognition, San Diego, USA, pp. 886-893, June 2005.
- [14] Y. Freund and R. t E Schapire, A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting, Journal of Computer and System Sciences, Vol. 55, No. 1, Pp. 119- 139 1997.
- [15] M. Arie, A. Moro, Y. Hoshikawa, T. Ubukata, K. Terabayashi and K. Umeda, Human Detection Using Subtraction Stereo with HOG Features, The Robotics and Mechatronics Conference 2010, 2P1-D17.

〈発 表 資 料〉

題 名	掲載誌・学会名等	発表年月
Fuzzy Rules based Indoor Human Action Recognition using Multi Cameras	Proceedings of the IEEE Symposium on Computational Intelligence for Multimedia, Signal and Vision Processing	2014年12月
Feature Extraction from Wearable Sensors for Human Action Recognition	Proc. of 2015 RISP International Workshop on Nonlinear Circuits and Signal Processing	2015年3月
Elderly Monitoring System using Surveillance Cameras	Proc. of 2015 RISP International Workshop on Nonlinear Circuits and Signal Processing	2015年3月
Human Action Recognition using Wearable Sensors and Neural Networks	Proceedings of 2011 Proc. of 2015 ASCC Emerging Control Techniques for a Sustainable World	2015年5月