非日常性の認識に向けた多数センサの関連付け

服部 傑 亀田 能成 大田 友一

† 筑波大学大学院システム情報工学研究科 〒 305-8573 つくば市天王台 1-1-1

E-mail: †{hattori,kameda,ohta}@image.esys.tsukuba.ac.jp

あらまし 日常生活環境におけるセンサ群を用いた人間行動認識とそれに基づく非日常性認識に向けた研究動向について述べると共に、我々の取り組みの第一段階として多種類のセンサの関連性の解析について報告する。本稿では、ネットワークカメラ、マイクロフォン、近接センサの三種類のセンサを数個ずつ配した環境で人間行動がどのように観測されるか実験を行い、その出力データの解析・考察を行った。その結果、出力データに対する特徴量が比較的単純なものであっても、複数のセンサを組み合わせれば人間行動認識が可能であるとの示唆が得られた。キーワード センサフュージョン、環境認識、行動認識、ユビキタス

A Study of Multi-Sensor Correlation Toward Abnormality Recognition

Takashi HATTORI[†], Yoshinari KAMEDA[†], and Yuichi OHTA[†]

† Graduate School of Systems and Information Engineering, University of Tsukuba 1-1-1 Tenoudai, Tsukuba, 305–8573 Japan

E-mail: †{hattori,kameda,ohta}@image.esys.tsukuba.ac.jp

Abstract In this paper, recent researches of human action recognition and abnormality recognition in a daily scene are surveyed and our preliminary results and discussions are presented. The main concern of this paper is to investigate correlation of multi-sensor outputs of three kinds of sensors; i.e., networked video cameras, microphones, and near-range sensors for recognizing human actions in a scene. The results of our preliminary experiment indicate that basic features are sufficiently useful for the purpose if multiple sensors are adequately utilized.

Key words Sensor fusion, Scene recognition, Action recognition, ubiquitous

1. はじめに

インテリジェントルームやアウェアホーム、インテリジェント在宅福祉、ビデオサーベイランスなど、人間が住む環境自体を付加価値の高いものにし、より住みやすい社会にしていこうという動きが社会的に認められてきている。このような高付加価値型居住環境システムでは、人間生活を邪魔することなく人間の生活を暖かく見守る機能が必須である。言い換えれば、人間が日常生活を送っている間はシステムからの積極的な干渉を控える一方で、非日常的な問題が発生した状況ではその問題解決のための支援を行う機能が求められている。

非日常性の認識はその実現のための技術前提であり、これからの現代社会において特に実現が望まれている技術であるといえる。我々の研究グループでは現在、非日常性の認識を目標とする研究に取り組んでいる。本稿では、日常生活環境におけるセンサ群を用いた人間行動認識とそれに基づく非日常性認識に向けた研究動向を紹介すると共に、我々の取り組みの第一段階として多種類のセンサの関連性の解析について述べる。

2. 関連研究

これまで、日常生活環境中の人間行動認識という目的のために、様々なセンサシステムや手法が提案されてきているが、それらはある場所でのある動き・行動の認識というように、特定の目的のために設計されていて、センサの設置場所や設置形態などは、限られた数のセンサを有効利用できるように事前に決定されていた。これは、これまでは認識したい内容・対象に比べて獲得できる入力情報量が比較的少なかったため、各センサの入力を最大限利用する必要があったからである。しかし、人間の生活空間が多岐に渡り、かつ生活の自由度は束縛されるべきではないことから、これからの居住環境システムでは様々なセンサの設置位置や形態は設計者の都合で決定されるべきではなく、与えられたものとして考えるべきである。それに加えて、近年の電子技術の発展に伴い、映像センサや音像センサなどのマルチメディアパッシブセンサはネットワーク結合され、小さくなり、安価・大量に入手・設置できるようになりつつある。

我々は、人間の日常生活空間内に設置された映像センサ(ビ

デオカメラ)・音センサ(マイクロフォン)・近接センサ(PSDセンサ)というマルチメディアパッシブセンサを中心に、多数のセンサが非均一非較正に設置された状態から、ネットワークを介した互いのセンサの同定・通信・協調による人間行動の認識のための理論構築を目指している。そのアプリケーションとして非日常性の認識を取り上げている。

多数のセンサを人間の活動環境に設置して行動認識や生活補助を行う取り組みは、ユビキタスコンピューティングの一環として、様々なところで研究が成されている。例えば、講義室やオフィスを対象とする Classroom2000 [1], ORL Active Floor [2], The Intelligent Room [3], Intelligent Workspace [4] [5], 自動講義アーカイブシステム (CARMUL) [6] [7] [8] や、一般家庭を想定した Awarehome [9], MAV home [10], Ubicare [11], Sensing room [12], ゆかりプロジェクト [13] などがそうである。

このような環境埋め込み型センサ群を用いた人間の行動認識の研究では、フォトダイオード・圧電素子の出力や RFID タグのコード情報、さらには Smartdust [14] に将来みられるような小型センシングデバイスが利用されている。これらのセンサは、小型単機能であるだけに特定の行動認識や検出には好適であるが、環境に設置するセンサを選択した時点で認識可能な事象が限定される。

それに対して、最近、カメラ群・マイクロフォン群という大量のデータ処理が要求されるセンサ群を利用し、人間の行動認識、特に非日常性の認識をより柔軟に行おうとする研究動向が見られる。

環境埋め込み型カメラ群については、数が増えることを想定するべきであるので、全てのカメラについてキャリブレーションを前提とするのは現実的ではない。複数の映像センサからの自動較正および対象認識という研究[15]も行われているが、視体積の交差が少ない場合には対応が難しい。このため、三次元空間的に明確な位置関係がわからないままでも人間行動認識可能な技術が望まれる。このことは、マイク群についても言える。

人間行動の非日常性の認識には、環境埋め込みカメラによる 人間行動の認識が必要となる。これまでの研究では、人物行動 を何らかのシンボルで表し、それを観測結果と照合する手法が 多く採用されている[16][17][18][19]。

環境埋め込みカメラに限らずに様々なセンサー記録から人間 行動や軌跡を認識する研究としては、単一カメラに基づく取り 組み [23] [24] [25] [26] [27] [28] [22] や、そこからを多人数認識に拡張した研究など [20] [21] が挙げられる。

これらの人間行動認識結果を元に、異常検出を行おうという 取り組みがすでに発表されている [29] [30] が、その適用範囲は まだ限定的である。

3. マッシブセンシングシステムの構成

本節では人間行動認識のために、多数センサの出力間の関連性がどのように利用できるかについて調べるため、我々が構築しているマッシブセンシングシステムについて述べる。

本研究では、システムが利用するセンサからの情報として、 画像(映像)情報・音響情報・距離情報、の三種を考えている。 図 1 は、本研究で構築している環境センシングシステムの概略である。システムは、基本的にセンサとその情報を取り込み、処理するための PC から構成されている。それぞれ三種の情報に対応して、

- 距離情報:距離センサ+距離センサデータ処理用 PC
- 音響情報:マイクロフォン群+音声データ処理用 PC
- 画像情報:ネットワークカメラ群+画像処理用 PC 群となっている。

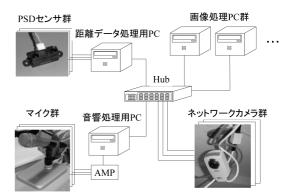


図 1 システム構成

カメラセンサとしてネットワークカメラを用いたのは、

- PC とセットで設置する必要性が無いため、設置の自由 度が高く、ある程度の量を設置することが可能
- PC ー台あたりが扱えるカメラの数の自由度が高い などの利点があるためである。逆に、欠点としては、
 - 各カメラで撮影した映像の同期が取れない
- ネットワークの状態によってはフレームレートが安定しない

といったことがあげられる。

マイクロフォン群としては、ファンタム給電可能な8本の有線マイクとアンプ、信号処理用の PC と AD 変換ボードを1セットとして用いる。

また、距離センサとしては、扱いが容易であること、比較的 小型であること、センサとしての特性が優れていること、安価 であること等の利点をもつ PSD センサを用いる。このような 単純な機能のセンサは、将来的に人間の生活環境中に遍在する ようになると考えられている。本研究でも当センサは環境中に比較的密に配置することを今後予定している。

4. 実 験

人間行動認識のために、多数センサの出力間の関連性がどのように利用できるかについて述べる。本稿では、環境埋め込みセンサの位置は不明と仮定する。

4.1 実験環境

本実験では、ネットワークカメラとして AXIS 205 (AXIS) を用いた。また、マイクロフォンとして F-720 (SONY)、距離 センサとして GP2D12 (SHARP) を用いた。音声信号と距離 センサからの出力は、AD 変換ボードを用いて PC に取り込んだ。音声は、 $26.0~{\rm kHz}$, $12~{\rm bit}$, $\pm~5~{\rm V}$ の範囲でサンプリングをした。マイクのゲインは、アンプからの出力が最大となると

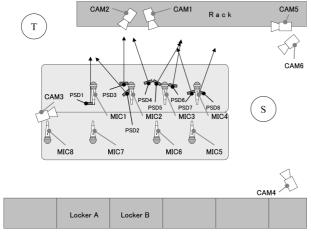


図 2 センサ類配置

きに \pm 5 V の範囲に収まるように事前に調整を行った。また、距離センサ の出力は、 $10.0~\mathrm{kHz}$, $12~\mathrm{bit}$, \pm 5 V の範囲でサンプリングを行い、 $100~\mathrm{thz}$ サンプルの平均値をデータとして記録した。距離センサは、測定可能範囲が $10 \sim 100~\mathrm{cm}$ 程度で、近距離に物体があるほど高い電圧を出し、 $2.4 \sim 0.3~\mathrm{V}$ の範囲で変化する。

ネットワークカメラからのデータの取得には PC を用い、一台の PC で 2台のネットワークカメラの映像を処理した。また、ネットワークカメラとデータ取得用 PC 群は 100BASE-TX の LAN で接続した。取得した画像の解像度は 320×240 pixel で、カメラの画角は水平画角 48°、垂直画角 54°である。ネットワークカメラから得る 1 フレームの画像サイズは約 30 KB 程度で、映像取得の平均フレームレートは 15fps 程度であった。

各センサを図 2 のように設置した。図 2 は対象空間を真上から見下ろした図で、縦に 3 m、横に 5m 程度の広さである。中央に PC やモニタなどが設置された机があり、片側には高さ 2.1 m のロッカーが並んでいる。ネットワークカメラ (CAM) は天井に設置し、マイク (MIC) は机上に設置した。距離センサ (PSD) は図の位置のオブジェクト (モニタなど) に設置した。図中の矢印は距離センサの最良感度方向を表している。距離センサは、他の二種類のセンサと比較して対象センシング範囲が狭いことやセンシング環境中に大量に設置されている事を想定していることから、本実験では比較的狭い範囲に密に設置した。実験環境の外観と各カメラからの映像例を図 3 に示す。

観測対象として、図 4 のように、対象空間を反時計回りに移動する動作(動作 1),時計回りに歩行する動作(動作 2)と、ロッカー A ないし B の前まで移動し、その開閉を行い、さらにもとの位置まで戻る動作(動作 3A, 3B)を反復した。全ての動作は図 2 の S の地点から開始した。また、実験環境中で所定の動作を行う人物以外に、図 2 の T の位置にもう一人の人物が存在し、所定の動作を行う人物と不定期に会話を行った。各動作は数十秒程度で完結し、それぞれの動作について約 10 分間ずつ動作を反復したデータを取得した。

実験では、各センサから得られる一次情報として以下のようなものを用いた。

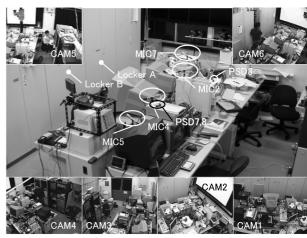


図3 実験環境外観と各カメラからの映像

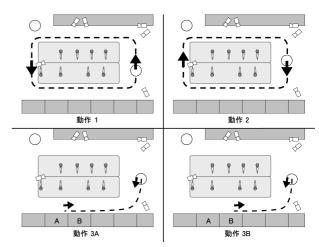


図 4 各動作の移動軌跡

- 映像特徴量:カメラから得られた映像のフレーム間差分の輝度値を5秒毎に全画素で平均した値(図7・図8・図9)
- 音響特徴量:マイクから得られた音声データの振幅の 100 msec 毎の平均値(図 10・図 11・図 12)
- PSD 特徴量: 距離センサから得られたデータの 10 msec 毎の出力電圧値(図 13・図 14)

なお、動作 3 は距離センサの配置位置周辺を移動していないため、グラフは割愛した。

図 7 から図 14 までは取得したデータの一部をグラフにプロットしたものである。動作 1 のグラフでは、 $220 \sim 253$ 秒・ $257 \sim 290$ 秒・ $290 \sim 320$ 秒・ $324 \sim 358$ 秒、動作 2 のグラフでは、 $24 \sim 65$ 秒・ $74 \sim 114$ 秒・ $119 \sim 155$ 秒・ $169 \sim 206$ 秒、動作 3 のグラフでは動作 3A が $175 \sim 203$ 秒・ $226 \sim 248$ 秒、動作 3B が・ $279 \sim 295$ 秒・ $311 \sim 330$ 秒の間に行われている。なお、図 10 から図 12 についてはグラフ間のスケールが異なることに注意する。

4.2 考 察

本節ではまず、各センサでのデータの傾向について述べる。 映像特徴量としては、動作 1 , 動作 2 について考えると CAM1 , 2 , 6 が、毎回の動作に対して同じ特徴を持ったデータを出力している。これは、該当するカメラが動作を行う人物のみを撮影可能な位置にあるためである。逆に CAM5 では、シーン中の

もう一人の人物が映像中に映り込むために出力されるデータに 規則性はみられない.

また、CAM4 において動作 1 では規則性のあるデータが得られているが、動作 2 では規則性が見られない。これは、CAM5 と同様に動作 2 の過程でシーン中のもう一人の人物が映像中に映り込んだ影響である。また CAM3 は、行動の起点となる地点を撮影しているために動作を行う人物のみを撮影可能ではあるが、規則性の無いデータを出力している。

これらのことから、逆に CAM1,2,4,6のデータの相関から同じ動作が繰り返されていることが推定可能だと考えられる。

動作 3A, 3B では、CAM4 が動作 3A, 3B に関わらず同一の特徴を持つ特徴量を出力している。また、動作 3A のみを撮影可能な CAM2 は動作 3A に対してのみ規則性のある特徴量を示している。CAM3 では、動作の開始地点が撮影されるために規則性が乱れてしまっていることが分かる。一方 CAM1, 6 では、3A, 3B それぞれの二回では互いに類似性を示すが、3A, 3B 間でみると異なる特徴を示している。

これらから、データの相関を見ていくことで動作 3A, 3Bを 判別することが可能だと考えられる。

以上のような傾向から、カメラの台数を増やしていけば、今 回用いたような基本的な特徴量であってもクラスタリングをす ることによって、行動が分類可能であると推測できる。

マイクセンサでは、動作 1,2 に関して殆んどデータに規則性が見られない。これは、動作に伴う発生音であっても一般的な物音程度であれば会話や環境ノイズにまぎれてしまうことが原因と考えられる。

しかし、音の種類によっては有意な情報として利用可能なものもある。例えば、オブジェクトを移動させる際に発生するラップ音があり、動作 1 の 240, 275 秒、動作 2 の 128, 188 秒の時点で MIC8 からそのような音が観測されていることが分かる。

動作 3 では、ロッカーを閉める際に発生する音が毎回観測されており、その信号レベルは一般の音声に比べて大きく、際立った特徴を持っているが、これだけではどちらのロッカーで開閉が行われたかは判別不能である。そこで、この二種の動作が同じかどうか判別する方法の一つとして、映像データと相関を取ることが考えられる。例えば、CAM2 の観測結果は動作3A にのみ影響を受けることから、そのデータを併用すれば両者の動作が判別可能である。

また、それ以外の方法としてマイクセンサ系全体でデータ間の相関関係を調べることも有効である。そのための方法として、本稿では音源の定位を行う手法の一つである CSP 法 [31] [6] を利用する。まず、CSP 法についてその原理を簡単に説明する.入力信号を $x_1(n),x_2(n)$ とし、その離散フーリエ変換をそれぞれ $X_1(k),X_2(k)$ と表す。式 2 で振幅によって正規化した相互相関を求めて、逆離散フーリエ変換することによって白色化相互相関係数を得る。この白色化相互相関係数 CSP(n) が最大値となるような n (式 3) が、マイクロフォン間での信号の到来時間差となる。また音源からの距離の差 d は、サンプリング周波数を F_s 、音速を c としたとき、式 4 で表される。

$$X(k) = \sum_{n=n_0}^{n_0+N-1} x(n) \exp(-2\pi j \frac{k}{N} n)$$
 (1)

$$CSP(n) = IDFT \frac{X_1(k)X_2(k)^*}{|X_1(k)||X_2(k)|}$$
 (2)

$$m = argmax_n(CSP(n)) \tag{3}$$

$$d = c \frac{m}{F_c} \tag{4}$$

図 5 は、MIC6 と 7 の間で CSP 係数を計算したものである。 上側が動作 3A のデータに対して計算したもので、MIC6 への信号が MIC7 への信号と比較して 10 サンプル (0.38 msec)遅れて到達していることが分かる。また、下側が動作 3B のデータに対して計算したもので、逆に MIC7 への信号が 16 サンプル (0.61 msec)遅れて到達していることが分かる。このことから、動作 3A では MIC7 に近い位置のロッカーの開閉を行い、逆に動作 3B では MIC6 に近い位置のロッカーの開閉を行っていることが推定できる。

このように、マイクからのデータだけではロッカーの開閉した位置に関わらず同じデータに見えていたデータが、マイク系の情報を統合することで別の事象として判別できる。

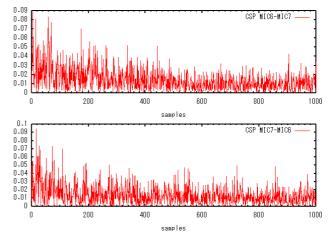


図 5 MIC6, MIC7 の間の CSP 係数

距離センサデータでは、PSD2, 3, 5, 6, 8 が各動作に対して規則的な反応を返している。あるセンサが反応している間は、そのほかの PSD センサが基本的に無反応であるが、これは距離センサの有効範囲が 1 m 程度と狭いためである。

有意な反応を返しているセンサからのデータを拡大した例を 図 6 に示す。図 6 は、上段が動作 1 の観測例で、下段が動作 2 の観測例である。動作 1 , 動作 2 でセンサの反応順が逆転して おり、このことからこれら二つの動作が動作順序的にみて逆の 関係にあることが推定可能である。

また、異種のセンサ間のデータを統合することで、より正確な状況認識が可能になる。例えば、動作3の観測結果を調べたとき、映像特徴量だけでは、ロッカーの前に移動し元の位置に戻った、程度の情報しか得られない。このとき、映像だけでロッカーの開閉を行わなかった場合のデータと判別するために

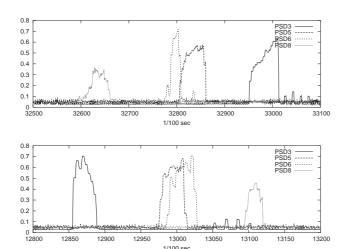


図 6 上:動作1,下:動作2距離センサデータ(拡大)

は、より詳細な画像解析が必要であると考えられる。しかし、 マイクのデータと合わせてデータ間の相関を調べることで、今 回用いた観測データの特徴量程度の情報でも十分に判別可能だ といえる。

5. おわりに

本稿では、日常生活環境におけるセンサ群を用いた人間行動 認識とそれに基づく非日常性認識に向けた研究動向を紹介する と共に、我々の取り組みの第一段階としての多種類複数センサ の関連性解析について述べた。

本稿では、ネットワークカメラ、マイクロフォン、近接センサの三種類のセンサを数個ずつ配した環境で人間行動がどのように観測されるか実験を行い、その出力の利用方法について考察を行った。その結果、複数種類の複数のセンサからの情報を統合していくことで、それぞれのセンサからの特徴量が比較的単純であっても、人間行動認識のためには役立つという示唆が得られた。

今後は、本稿で得られた知見を元にマッシブセンシングシステムを利用し、非較正なセンサ群の利用による人間行動の自動認識と、それに基づく非日常性の認識方法について研究を進めていく予定である。特に、今回用いているようなマルチメディアパッシブセンサはリアルタイムで継続的に出力が出続けるので、そのようなデータ構造に適した非日常性認識理論の構築を目指していきたい。

文 献

- Gregory D. Abowd, "Classroom 2000: An Experiment with theInstrumentation of a Living Educational Environment," IBM Systems Journal Special issue on Pervasivecomputing, vol.38, no.4, pp.508-530, 1999.
- [2] Michael D. Addlesee, Alan Jones, Finnbar Livesey, and Ferdinando Samaria, "ORL Active Floor," IEEE Personal Communications, vol.4, no.5, pp.35-41. 1997.
- [3] Rodney. A. Brooks, "The Intelligent Room Project," Proceedings of the Second International Cognitive Technology Conference (CT'97),pp.271-277, 1997.
- [4] Brad Johanson, Armando Fox, and Terry Winograd, "The Interactive Workspaces Project: Experiences with Ubiquitous Computing Rooms," IEEE Pervasive Computing (Spe-

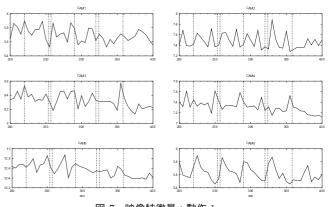
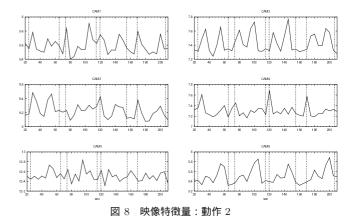
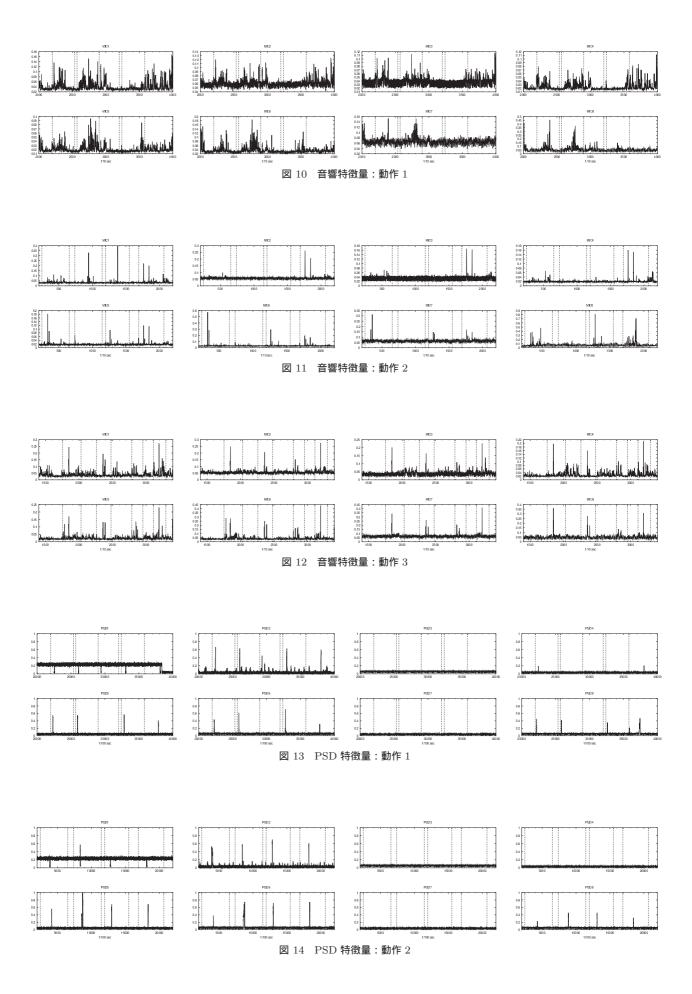


図 7 映像特徴量:動作 1



— 5 —



— 6 —

- cial Issue on Overviews of Real-World Ubiquitous Computing Environments), vol.1, no.2, pp.67-74, 2002.
- [5] Jan Borchers, Meredith Ringel, Joshua Tyler, and Armando Fox, "Stanford Interactive Workspaces: A Framework for Physical and Graphical User Interface Prototyping," IEEE Wireless Communications (Special Issue on Smart Homes), vol.8, no.6, pp.64-69, 2002.
- [6] 西口 敏司, 東 和秀, 亀田 能成, 角所 考, 美濃 導彦, "講義自動撮影における話者位置推定のための視聴覚情報の統合," 電気学会論文誌 C, vol.124, no.3, pp.729-739, 2004.
- [7] Yoshinari Kameda, Satoshi Nishiguchi, Michihiko Minoh, "CARMUL: Concurrent Automatic Recording for Multimedia Lecture," IEEE International Conference on Multimedia and Expo (ICME2003), Vol.1, pp.129-132, 2003.
- [8] Satoshi Nishiguchi, Kazuhide Higashi, Yoshinari Kameda, Michihiko Minoh, "A Sensor-fusion Method of Detecting A Speaking Student," IEEE International Conference on Multimedia and Expo (ICME2003), Vol.2, pp.677-680, 2003.
- [9] Cory D. Kidd, Robert Orr, Gregory D. Abowd, Christopher G. Atkeson, Irfan A. Essa, Blair MacIntyre, Elizabeth Mynatt, Thad E. Starner and Wendy Newstetter, "The Aware Home: A Living Laboratory for Ubiquitous Computing Research," Proceedings of the Second International Workshop on Cooperative Buildings (CoBuild'99), 1999.
- [10] Diane J. Cook, Michael Youngblood, Edwin O. Heierman, III, Karthik Gopalratnam, Sira Rao, Andrey Litvin, and Farhan Khawaja, "MavHome: An Agent-Based Smart Home," IEEE International Conference on Pervasive Computing and Communications (PerCom'03), pp. 521-524, 2003.
- [11] "Ubiquitous Computing for Healthcare in the Community," http://www.ubicare.org
- [12] Taketoshi Mori, Hiroshi Noguchi, Aritoki Takada, Tomomasa Sato, "Sensing Room: Distributed Sensor Environment for Measurement of Human Daily Behavior," First International Workshop on Networked Sensing Systems(INSS2004), pp.40-43, 6 2004.
- [13] 美濃 導彦, "ゆかりプロジェクトの目的と概要 -UKARI プロジェクト報告 No1-," 情報処理学会 第 66 回全国大会, vol.5, no.3, TTC-2, pp.5-8, 2004.
- [14] Jason Hill, Robert Szewczyk, Alec Woo, Seth Hollar, David Culler, and Kristofer Pister, "System Architecture Directions for Networked Sonsers," Proc. 9th International Conferenceon Architectural Support for Programming Languages and Operating Systems, vol. 3, no. 11, pp. 93-104, 2000.
- [15] Chris Stauffer, Kinh Tieu, "Automated multi-camera planar tracking correspondence modeling," Proc. Computer Vision and Pattern Recognition, I, pp.259-266, 2003.
- [16] 青木 茂樹, 大西 正輝, 小島 篤博, 福永 邦雄, "HMM による 行動パターンの認識," 信学会論文誌 D-II, vol.J85-D2, no.7, pp.1265-1270, 2002.
- [17] Junji Yamato, Jun Ohya, and Kenichiro Ishii, "Recognition Human Action in Time-Sequential Images Using Hidden Markov Model," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp.379-385, 1992.
- [18] Aaron F. Bobick and James W. Davis, "The Recognition of Human Movement Using Temporal Templates," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 23, no. 3, pp. 257-267, 2001.
- [19] 西村 拓一, 十河 卓司, 小木 しのぶ, 岡 隆一, 石黒 浩, "動き変化 に基づく View-based Aspect Model による動作認識," 信学会 論文誌 D-II, vol.J84-D2, no.10, pp.2212-2223, 2001.
- [20] 和田 俊和, 佐藤 正行, 松山 隆司, "選択的注視に基づく複数対象 の動作認識," 信学会論文誌 D-II, vol.J82-D2, no.6, pp.1031-1041, 1999.
- [21] 佐竹 純二, 尺長 健, "階層的注視点制御による動画像上での複数人物追跡、"信学会論文誌 D-II, vol.J86-D2, no.8, pp.1212-1221,

- 2003.
- [22] Chris Stauffer, "Minimally-Supervised Classification using Multiple Observation Sets," Proc. International Conference on Computer Vision (ICCV2003), pp.297-304, 2003.
- [23] Edwin O. Heierman III, G. Michael Youngblood, and Diane J. Cook, "Mining Temporal Sequences to Discover Interesting Patterns," KDD Workshop on Mining Temporal and Sequential Data, 2004.
- [24] Namrata Vaswani, Amit Roy Chowdhury, and Rama Chellappa, "Activity Recognition Using the Dynamics of the Configuration of Interacting Objects," IEEE Computer Vision and Pattern Recognition (CVPR), 2003.
- [25] Patrick Peursum, Hung H. Bui, Svetha Venkatesh, and Geoff West, "Human action segmentation via controlled use of missing data in HMMs," 17th International Conference on Pattern Recognition (ICPR2004), 2004.
- [26] Nan T. Nguyen, Svetha Venkatesh, Geoff West, and Hung H. Bui, "Learning people movement model from multiple cameras for behaviour recognition," Joint IAPR International Workshops on Structural and Syntactical Pattern Recognition and Statistical Techniques in Pattern Recognition (SSPR2004), p315-324, 2004.
- [27] Nan T. Nguyen, Hung H. Bui, Svetha Venkatesh, and Geoff West, "Recognising and monitoring high-level behaviors in complex spatial environments," IEEE CVPR, 2003.
- [28] Sebastian Luh, Hung H. Bui, Svetha Venkatesh, and Geoff West, "Recognition of human activity through hierarchical stochastic learning," IEEE International Conference on Pervasive Computing and Communications, p416-422, 2003.
- [29] Lukman Sasmita, Wanquan Liu, and Svetha Venkatesh, "An Immunological Approach to Raising Alarms in Video Surveillance," Pacific Rim Conference on Multimedia (PCM2004), 2004.
- [30] 青木 茂樹, 岩井 嘉男, 大西 正輝, 小島 篤博, 福永 邦雄, "人物の位置・姿勢に注目した行動パターンの学習・認識と非日常状態検出への応用," 信学会論文誌 D-II, vol.J87-D2, no.5, pp.1083-1093, 2004.
- [31] 西浦 敬信, 山田 武志, 中村 哲, 鹿野 清宏, "マイクロホンアレー を用いた CSP 法に基づく複数音源位置推定," 信学会論文誌 D-II, vol.J83-D2, no.8, pp.1713-1721, 2000.