

分散視覚情報に基づいた人物追跡技術に関する研究 —生活活動を見守る知的空間の構築に関する研究—

渡辺 博己 棚橋 英樹 田中 等幸
赤塚 久修 窪田 直樹* 清水 早苗*

Study on Distributed Visual Information for Human Tracking - Construction of Smart Environment for Supporting Human Activities -

Hiroki WATANABE Hideki TANAHASHI Tomoyuki TANAKA
Hisanobu AKATSUKA Naoki KUBOTA* Sanae SHIMIZU*

あらまし 治安の悪化や高齢社会の進展に伴い、ユビキタスコンピューティングを中心とした知的空間構築技術を用いて生活の安全性、快適性、利便性を向上させることが重要な課題となっている。そこで本研究では、知的空間を構築する上で重要となる人物追跡技術に着目し、複数の分散視覚システムから得られた映像に基づいて人物を追跡する手法について検討する。また、人物を追跡することにより得られた情報から、動作の種別を判定する方法についても検討する。

キーワード 知的空間, 分散視覚, 視体積交差法, 人物追跡, 動作推定

1. はじめに

治安の悪化や高齢社会の進展に伴い、個人生活の安全性、快適性、利便性向上を支援するサービスの需要が増加している。また、コンピュータやセンシング・通信機器の高機能化・小型化により、日常生活のあらゆる場所で利用者の状況に応じた情報支援が受けられるユビキタスコンピューティング技術への期待と要望が年々高まっている。

そのような中で、家庭にユビキタスコンピューティング技術を取り入れ、家庭環境を知的化することにより、安心、安全、快適な生活空間を構築する研究が数多く行われている^{[1]~[6]}。例えば、米国ジョージア工科大学のAware Homeでは、センサとコンピュータによって家族の状況を把握してくれる家と、その機能を利用したアプリケーションが研究されている^[1]。また、(独)情報通信機構のユビキタスホームでは、家電製品だけでなく様々な機器の相互接続性の実証実験、家全体のロボット化、その中で人の行動把握とそれに基づく新たなサービスが研究されている^[4]。

様々な生活シーンにおいて、状況に応じたサービスを人に提供するためには、人とその周りの環境とを同時に

センシングし状況を把握することで、その人の要望を理解する技術が重要であり、画像センシング技術は解決すべき課題の一つである。画像センシング技術は、人にセンサを装着させる必要が無いため、日常生活を束縛することなく生活行動を取得できると共に、人とその周りの環境の状況を同時に取得できるため、人と人、人と物とのインタラクションの理解などに有効である。しかし、画像センシング技術を用いた人の状況把握には、カメラの前に立つなどの制約を与えない限り、常に人の位置と姿勢を取得する必要がある。

そこで本研究では、連続的に入力される時系列画像間で人の位置と姿勢の計測を繰り返す人物追跡技術について検討する。また、得られた追跡情報から動作種別の判定を試みる。

2. システムの構築

人物追跡技術は、人の位置と姿勢を取得することにより、人の動線や行動パターンを取得できるため、不審者の検出・監視だけでなく、居住者の動作推定などに有効である。一般に、人物追跡で動線を取得するだけであれば、画像上の二次元平面で人の位置を取得することにより、移動軌跡を求めることが可能である。しかし、三次元で変化する人の姿勢を推定するためには、三次元空間

*財団法人ソフトピアジャパン IT研究センター



図1 ステレオユニット



図2 1/8球型全方向ステレオシステム

上で人を追跡する必要がある。

三次元空間上で人を追跡するためのカメラシステムとして、マルチカメラ（対象物を囲むように何台ものカメラを配置し撮影するカメラシステム。多視点カメラとも称す。）とステレオカメラがある。検出した領域をボリュームデータとして処理することで、画像上で人と人が重なり合っている場合でも、精度よく分離することができるため、個々の行動パターンを容易に取得することが可能となる。

そこで本研究では、図1に示すステレオカメラをベー

スとしたステレオユニットにより、図2のような分散ステレオ視覚システム（マルチカメラが一点を集中して撮影するのに対して、分散視覚は多数台のカメラの視野を補うように配置する^[7]。本稿でいう分散ステレオ視覚とは分散視覚で使用するカメラをステレオカメラで構成したカメラシステムを意味する。）を構成し、それを複数利用することでマルチカメラのように配置した撮像系を用いる。使用した分散ステレオ視覚システムは1/8球型全方向ステレオシステム^[8]（以下、1/8-SOS）である。図3に人物追跡のために構築したプロトタイプシステム（以下、システム）の構成を示す。本システムでは、空間の一辺が480cmとなるようにパーティションで区切り、その天井の四隅それぞれに1/8-SOSを設置した。天井までの高さは270cmである。

各1/8-SOSからは3枚のカラー画像と6枚のモノクロ画像が取得され、これらの画像は光変換機能を有するデータ処理ユニットに転送される。データ処理ユニットは四つの1/8-SOSから転送された36枚の画像データを統合し、光ケーブルにより1.2Gbps×2の光信号として画像データ制御部へ出力する。画像データ制御部は高速シリアル信号として送られてきた光信号を画像データに復調し、システム制御部のメモリへ高速に転送する。システム制御部では、ステレオユニット毎にキャリブレーションデータを管理し、ステレオビジョンソフトウェアライブラリ^[9]を用いて、補正画像、距離画像を生成し表示する。1/8-SOSの補正には一つの基準座標系を用いて導出した回転、並進パラメータを利用する。1/8-SOS上のステレオユニットは高い精度で組み付け加工されており、この配置パラメータをステレオユニット間のキャリブレーションに利用する。個々のステレオユニットに関しても、それぞれユニット単位でキャリブレーションを行っており、レンズ歪み及びカメラ間のアライメントずれの影響をソフトウェアによって除去する。

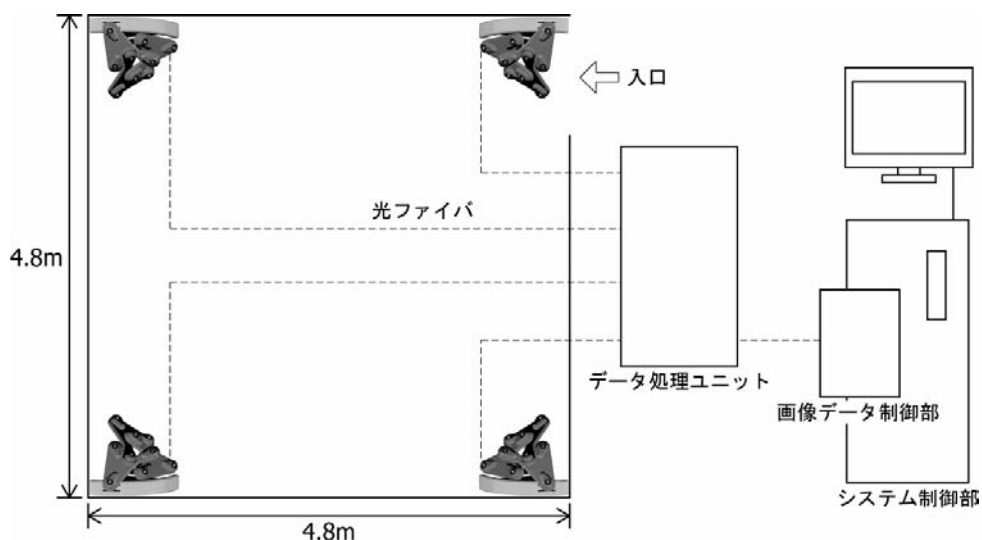


図3 プロトタイプシステムの構成

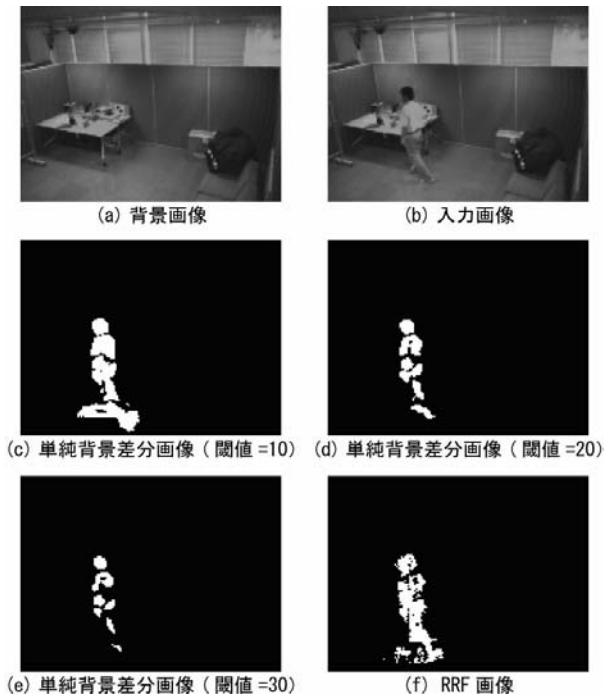


図4 前景領域の抽出例

3. 人物の追跡

室内で人物を追跡するために、複数の1/8-SOSで同時刻に取得したカラー画像からシルエット画像を生成し、視体積交差法により人物領域を協調して検出する。1/8-SOSからはカラー画像だけでなく距離画像も取得できるが、本研究では、環境に存在する人の数を一人として追跡処理を実行するため、カラー画像のみを使用する。以下、それぞれの処理について詳細を述べる。

3. 1 シルエット画像の生成

シルエット画像は、各ステレオユニットのセンタカメラから得られたカラー画像に濃淡処理を行い、Radial Reach Filter^[10]（以下、RRF）により前景領域を抽出し、膨張・収縮処理を加えて生成する。

単純背景差分を用いる場合、出現物体の明度分布と背景の明度が近い領域において未検出が生じるだけでなく、影などの明度変化による過検出の影響も大きい。また、カメラを設置した環境により、閾値を手動で設定する必要がある。一方、RRFは濃淡情報をベースに、ピクセル単位の分解能で局所的なテクスチャを評価し、変化領域の明度分布に依存せず、しかも影などの明度変化の影響を低減したロバストな検出が可能である。また、RRFで用いる閾値は統計的に決定することが可能であるため、複数の画像を与えるだけで自動的に設定することが可能である。図4に単純背景差分とRRFによる前景領域の抽出例を示す。

3. 2 人物領域の検出

得られたシルエット画像とキャリブレーションデータに基づいて、視体積交差法により投票を行うことで、ボ

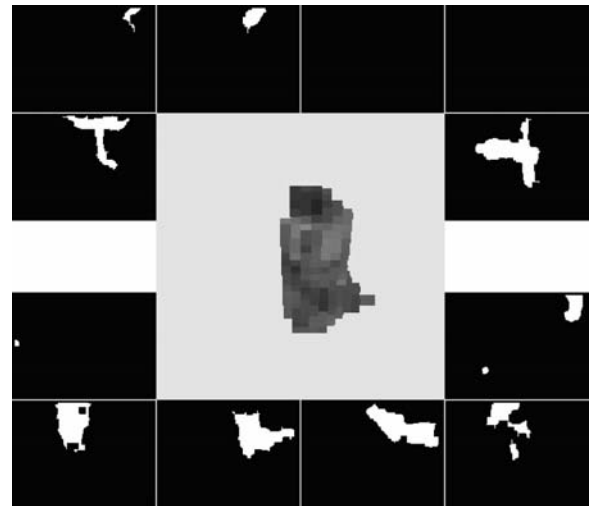


図5 ボクセル領域の投票結果例（空間分解能:10cm）

クセル空間上で人物領域を検出する^[11]。

視体積交差法では、カメラ位置とシルエット画像の各画素とを結ぶ直線上でボクセル投票を行い、各カメラからの投票が重なる部分をボクセル領域として取得する。通常、視体積交差法は3次元形状復元に多く用いられ、形状復元の精度を高めるために、高い空間分解能をもつボクセル空間を用いている。しかし、高分解能のボクセル空間への投票は処理速度の低下を招き、人物検出処理の負担となる。本研究では、人物検出を目的としているため、形状の精度はあまり重要ではない。そこで、ボクセル空間の分解能を低く設定することにより、処理の高速化を図る。図5にボクセル空間解像度を10cmに設定した投票結果の例を示す。なお、各シルエット画像の解像度は160×120pixelである。

3. 3 人物領域の追跡と動作の推定

視体積交差法を用いることで、床面上の位置だけでなく、人物の三次元的な像が検出できるため、簡単な動作の推定が可能である。

人物領域の追跡を実現するためには、各フレームで人物の空間上の位置を検出する必要がある。そこで、人物の床面上の位置は、床面を基準として高さ方向のボクセル数を加算することにより三次元ヒストグラムを生成し、最もボクセル数が多い位置を人物位置とする。また、人物の高さは、ボクセル領域内で最も高い位置に存在するボクセルの高さとする。

次に、人物領域を追跡することにより得られた床面上の位置と領域の高さ情報から動作の推定を行う。推定する動作は「歩く」、「止まる」、「座る」の基本的な3種類の動作である。「歩く」、「止まる」の動作は、床面上の位置の変化の大きさから推定する。「座る」の動作は、人物が歩いているときの高さ情報を基準に、人物の高さが75%以下となる止まった状態として推定する。図6に追跡結果と動作の推定例を示す。図中の白線は人物の移動軌跡を表している。

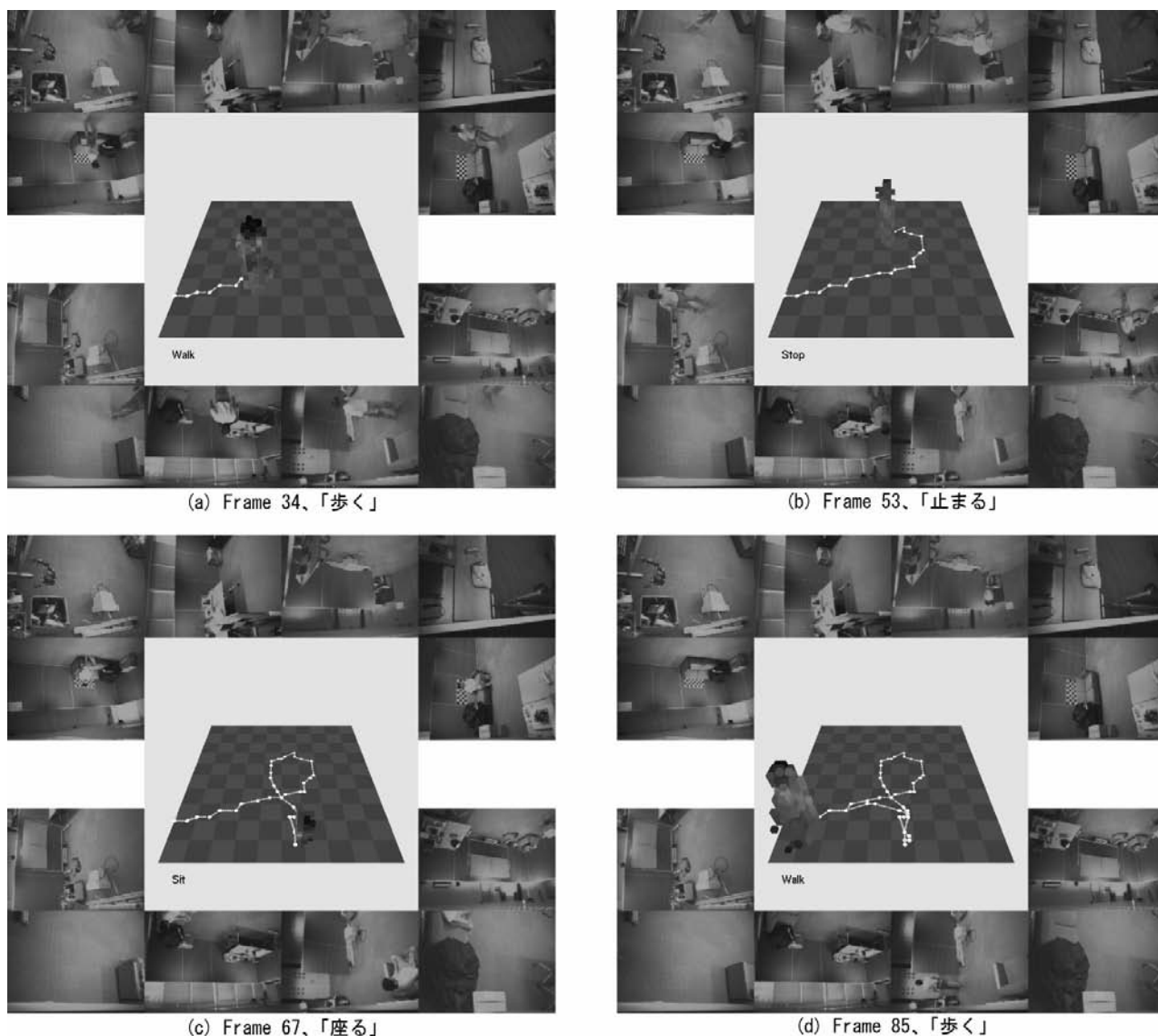


図6 人物領域の追跡と動作の推定例

4. まとめ

安心、安全、快適な生活環境を実現する知的空間を構築するために、日常生活を束縛することなく生活行動を取得できる画像センシング技術に着目し、特に人物の動線や行動パターンを取得できる人物追跡技術について検討した。天井の四隅に1/8球範囲の視野を持つ1/8-SOSを設置し、それらの映像から視体積交差法により環境内の人物像を復元することで、分散視覚システムとマルチカメラシステムとを統合した人物追跡を行った。

人物追跡では、三次元空間上で人物領域のボリュームを求めることで、位置と姿勢の推定が可能となり、「歩く」、「止まる」、「座る」の三つの動作に分類することができた。これらの動作に環境内の物体の情報を付加することにより、より多くの動作について推定が可能になると考えられる。

構築したプロトタイプシステムでは、一人に対して追跡、動作推定を行ったが、今後は複数人に対して稼動するようアルゴリズムを検討していく予定である。そのためには、前報告で行った視体積交差法とステレオ法とを利用したハイブリッド処理^[8]による人物追跡の実現すると共に、プログラムの高速化によるリアルタイムシステムを構築する必要がある。

謝 辞

本研究の遂行にあたり、1/8球型全方向ステレオシステムを貸与して頂いた(財)ソフピアジャパンIT研究センターの丹羽義典センター長をはじめ、関係者の皆様に感謝する。また、有益な助言を頂いた岐阜大学工学部の山本和彦教授、早稲田大学総合研究機構WABOT-HOUSE研究所の富永将史氏に感謝する。

文 献

- [1] Aware Home, <http://www.awarehome.gatech.edu/>
- [2] Easy Living, <http://research.microsoft.com/easyliving/>
- [3] Smart Rooms,
<http://vismod.media.mit.edu/vismod/demos/smarteroom/>
- [4] ユビキタスホーム,
http://www2.nict.go.jp/x/x164/past_pj/disco/research/ubiquitous_home.html
- [5] 人を見守るデジタルヒューマン,
<http://www.dh.aist.go.jp/research/enabling/>
- [6] T. Mori, H. Noguchi, A. Takada, T. Sato, “*Sensing Room: Distributed Sensor Environment for Measurement of Human Daily Behavior*,” First International Workshop on Networked Sensing Systems (INSS2004), pp.40-43, 2004.
- [7] 富田文明, “3次元視覚システムの技術動向と展望”, 人間主体の知的情報技術に関する調査研究V, (財)日本情報処理開発協会先端情報技術研究所編, pp.218-228, 2002.
- [8] 渡辺博己, 棚橋英樹, 浅井博次, 平湯秀和, 清水早苗, 山本和彦, “協調3次元センシング技術を用いた人物検出手法に関する研究—生活活動を見守る知的空間の構築に関する研究—”, 岐阜県生産情報技術研究所研究報告, 第7号, pp.8-13, 2006.
- [9] Point Gray Research Inc., “User’s Guide and Command Reference - TRICLOPS Stereo Vision Software Development Kit (SDK).”
- [10] 佐藤雄隆, 金子俊一, 丹羽義典, 山本和彦, “Radial Reach Filter (RRF)によるロバストな物体検出”, 信学論(D-II), Vol.J86-D-II, No.5, pp.616-624, 2003.
- [11] 富永将史, 本郷仁志, 興水大和, 丹羽義典, 山本和彦, “人物行動把握のための複数人物追跡システムの構築”, 動的画像処理実利用化ワークショップ (DIA2004), pp.79-84, 2004.