

ステレオカメラとモーションセンサを用いた 室内用位置追跡手法の提案

Detection of Human Location by Integrating Different Kinds of Sensors

新野毅^{*1} 田島潤一^{*1} 橋詰和範^{*1} 新谷公朗^{*2} 芳賀博英^{*1} 金田重郎^{*1}
Tsuyoshi Shinno Junichi Tajima Kazunori Hashizume Kimio Shintani Hirohide Haga Shigeo Kaneda

^{*1}同志社大学大学院工学研究科

Graduate School of Engineering, Doshisha University

^{*2}常磐会短期大学幼児教育科

Department of Early Childhood Education, Tokiwakai College

In this paper, we propose a detection method of indoor multiple human location and direction by using a stereovision camera and motion sensors. The proposed method features the integration of two kinds of sensing devices based on direction information. Using direction information instead of conventional location information to integrate sensors, we can build a practical detection system because it can detect location even when humans move in way other than specific walking. The experimental evaluation result shows the success rate of detection was 90 percent. So, direction information is effective for integrating sensors. We think the accuracy of the location and direction information that are obtained by the proposed system are sufficient to be used by a ubiquitous service.

1. はじめに

ユビキタスコンピューティング環境では、人の位置は非常に価値のある情報となる。特に、多くのユビキタスサービスは室内を想定して提案されており、室内での位置検出は重要な研究課題である。UWBなどの電波を用いた位置検出手法は、端末が人体に隠れてしまうと電波が受信機に到達せず検出がおこなえない[日立 07]。また超音波を用いた位置検出手法は、端末が傾くと正確な検出が困難になる[古河]。

コミュニケーションの検出や、情報家電の操作をおこなう場合、人の位置情報に加えて方向情報も必要である。2人の人間が近くにいる、お互いに向かい合っているとわかれば会話していると推測でき、コミュニケーションが検出できる。また複数の家電が近くにあり非接触で操作をおこなう場合、操作する人がどの方向を向いているかわからないと、操作対象の特定が不可能である。

そこで本稿では、室内で利用することができ、複数の人間の位置と方向の検出が可能な手法を提案する。提案手法ではステレオカメラとモーションセンサを用いてデータを測定する。測定したデータはオフラインで回収して分析されるため、ノイズの影響を受けない。さらに人の姿勢から受ける影響も低く、ロバストな手法である。プロトタイプシステムを構築し、同時に検出する人数を変化させながら検出をおこなったところ、検出成功率は90%であった。

2. 従来のアプローチ

2.1 概要

我々は、これまでに単眼カメラとモーションセンサを用いた位置検出手法を提案してきた[KAWAI 05]。単眼カメラを用いれば高精度な位置検出が可能であるが、個人を特定することはできない。モーションセンサを用いれば位置検出の精度は低いが、センサ装着者を特定することができる。そこで単眼カメラとモーションセンサから得られる情報を、それぞれが検出した

位置情報をもとに統合することにより、個人特定可能な高精度位置検出がおこなえることを示した。

2.2 問題点

しかし従来のアプローチには、次のような問題が存在する。

- センサデータ処理
 - － 歩数で位置を計算しているため、歩幅一定で前進しているとき以外は位置検出ができない
- 動画処理
 - － フレーム間差分方式を用いたため、検出対象が停止しているときは検出できない
 - － 単眼カメラなので奥行きが検出できず、検出範囲全体を撮影可能な比較的高い位置に設置する必要がある
- 照合処理
 - － モーションセンサで求められる位置情報は、検出誤差が蓄積し時間とともに精度が低くなるので、カメラデータとの照合ミスが多発する
 - － 単純に相関が高いセンサデータとカメラデータを結び付けるので、1つのセンサデータに複数のカメラデータが結びついて対処されない

3. 提案手法

3.1 概要

今回提案するシステムはステレオカメラと、地磁気を検出可能なモーションセンサ(方位センサ)、小型データロガーから構成される。図1に提案システムの全体イメージを示す。検出時には方位センサを検出対象の腰に装着し、ロガーを携帯させる。ステレオカメラと方位センサから得られたデータは、計測後に回収し照合処理をおこなう。このシステムはステレオカメラを用いているため、奥行きが検出できる。したがって単眼カメラと違い、床と平行にカメラを設置しても撮影が可能である。さらにカメラとセンサから得られた2種類のデータを、位置情報の代わりに方向情報に基づき統合する。これによりセンサから位置情報を求める必要がなくなり、歩幅が一定の前進をしていない場合も検出がおこなえる。方向情報は地磁気より検出可能であり、誤差が蓄積しないので、照合ミスも減少する。

連絡先: 新野毅, 同志社大学大学院工学研究科,
京都府京田辺市多々羅都谷 1-3, 0774-65-6979,
tshinno@ishss10.doshisha.ac.jp

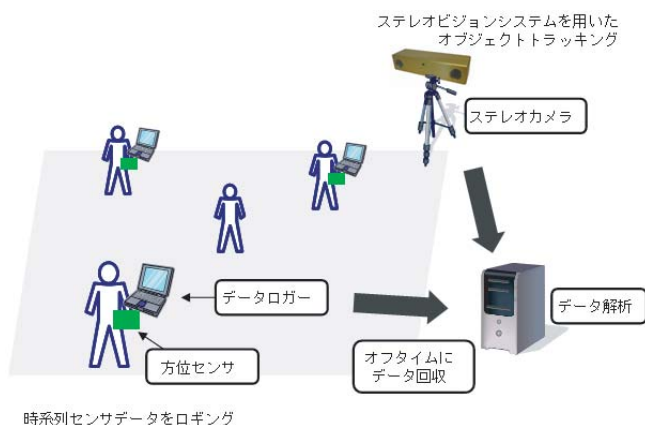


図 1: 提案システムのイメージ

処理の順序を図 2 に示す．センサデータ処理と動画像処理に関する詳しい説明は後述する．ここで動画像処理によって得られる経路候補の数を M ，方位センサの数を N ，サンプリングの時間間隔を t とする．方位センサからは検出対象の ID と時系列方向データ $D_N(t)$ が得られる．ステレオカメラからは追跡経路候補が得られる．経路候補からは対象の時系列位置データ $X_M(t)$ がわかる．さらに移動方向から対象の時系列方向データ $R_M(t)$ がわかる．この 2 つの時系列方向データを基にして ID と経路候補を結びつける．

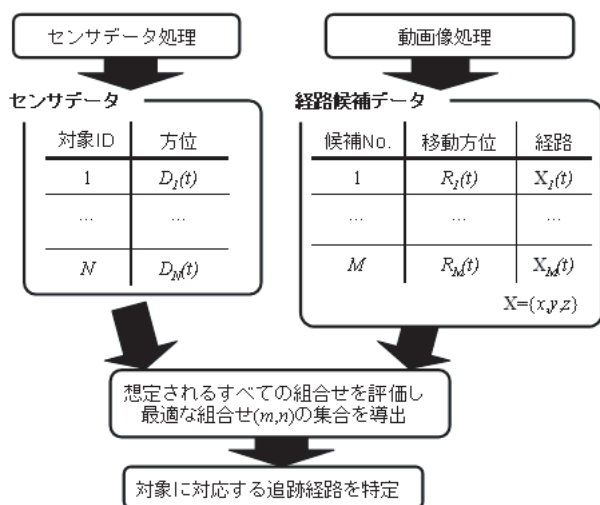


図 2: 処理順序

3.2 照合アルゴリズム

まず，方位センサによって得られるデータとステレオカメラによって得られるデータの，すべての組み合わせを考える．ステレオカメラの動画像処理によって得られる経路候補の数 M と，方位センサの数 N は，必ずしも等しくない．後述する動画像処理で交差や分岐という状態が生じると，実際に検出対象が移動していない経路も正しい可能性がある経路候補として保存されるので， M が N より大きくなることもある．1 つの方位センサに複数の経路候補は結びつかないので，架空の方位センサを用意して N を M より大きくする．新たに定義した経

路候補の数 m と方位センサの数 n は，式 (1) の範囲をとる．

$$\begin{aligned} m &= 1, \dots, M \\ n &= 1, \dots, N, \dots, M + N \end{aligned} \quad (1)$$

経路候補と方位センサの組み合わせの正しさを判断するために，評価関数 $f(m, n)$ を用いる．経路候補と実際の方位センサが結びついているときは，それぞれの方向データである $R_m(t)$ と $D_n(t)$ の計測時間内の誤差の平均をとり，正規化したものが $f(m, n)$ となる．また経路候補と架空の方位センサが結びついているときは， $f(m, n)$ は経路候補が正しい経路である確率 p_m の余事象となる．この評価関数を式 (2) に示す．

$$f(m, n) = \begin{cases} g(m, n) & (m \leq M, n \leq N) \\ 1 - p_m & (m \leq M, n > N) \end{cases} \quad (2)$$

$$g(m, n) = 1 - \frac{\sum_{t=1}^T |R_m(t) - D_n(t)|}{\pi T}$$

$$p_m = \frac{\text{追跡終了フレーム} - \text{開始フレーム} + 1}{\text{総フレーム数}}$$

経路候補と方位センサが結びつくとき 1，結びつかないとき 0 とする割り当て変数を $a(m, n)$ とすると，すべての経路候補が正しい方位センサと結びついているとき，式 (3) に示す評価規範 J が最大になる．以上のことから，すべての経路候補と方位センサの正しい組み合わせが判明する．

$$J(\mathbf{a}) = \sum_{m=1}^M \sum_{n=1}^{M+N} a(m, n) f(m, n) \quad (3)$$

3.3 センサデータ処理

3.3.1 解くべき課題

検出対象の方向は，測定した地磁気によって計算することができる．対象の北から時計回りの方位は式 (4) により計算できる．

$$\arctan\left(\frac{MAG_{left}}{MAG_{front}}\right) \quad (4)$$

方向を求めるとき，方位センサの軸が対象の正面と左方向に固定されていることが理想である．しかしセンサは対象の体に装着されるので，対象が動くとき軸も動く．そのため実際に対象が向いている方向と，計算により求まる方向にずれが生じる．

3.3.2 補正方法

加速度情報を基に検出対象の移動状態と姿勢を判定し，判定結果に応じた補正をかける．ここで移動状態とは「静止」「歩行」「走行」と定義し，姿勢とは「直立」「仰向け」「うつ伏せ」と定義する．

移動状態は 3 軸の合成加速度から判定する．合成加速度が $1.1G$ 以下なら「静止」， $2.0G$ 以上なら「走行」，それ以外なら「歩行」と判定する．「静止」のときはセンサがほとんど動いていないと考えられるので，過去 $1.2sec$ の移動平均をとるにとどめる．「歩行」と「走行」のときは過去 $0.6sec$ 間の最大値と最小値を探し，その平均をとる． $0.6sec$ はおよそ 1 歩分の時間に相当する．その時間の最大値と最小値は，それぞれ足を踏み出したタイミングと着地したタイミングであり，実際の対象の向きはその中間だと考えられるからだ．しかし「歩行」と「走行」のときには，この処理をおこなっても大きな誤差が発生する恐れがある．そのため直前の方向と加重移動平均を取

ること、大きな誤差の発生を抑える。動きが少ないときのほうが計算結果の信頼性が高いので、重みは「静止」:「歩行」:「走行」=3:2:1 とつける。

姿勢はピッチ角から判定する。ピッチ角は各軸にかかる重力加速度の大きさから計算することができる。対象が体をそらせればピッチ角が小さくなり、体を前に傾ければピッチ角が大きくなるよう、センサが装着されているとする。ピッチ角が $-\pi/4rad$ 以下であれば「仰向け」、 $\pi/4rad$ 以上なら「うつ伏せ」、それ以外なら「直立」と判定する。この判定は移動状態が「静止」のときのみおこなわれる。地磁気データの座標系変換をおこなうことで計算結果の精度向上を目指す。回転角が $\pi/2rad$ に近いとセンサの測定誤差の影響が大きくなってしまふ。そのため「直立」のときと「仰向け」「うつ伏せ」のときとは、異なるワールド座標系に変換し、回転角が $\pi/2rad$ に近づかないようにする。「直立」のときに用いる変換行列を式(5)に、「仰向け」「うつ伏せ」のときに用いる変換行列 R を式(6)に示す。

$$R = \begin{bmatrix} \cos\theta & -\sin\theta\cos\phi & \sin\theta\sin\phi \\ \sin\theta & \cos\theta\cos\phi & -\cos\theta\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} \quad (5)$$

$$R = \begin{bmatrix} \cos\theta\cos\phi & -\sin\theta & \cos\theta\sin\phi \\ \sin\theta\cos\phi & \cos\theta & \sin\theta\sin\phi \\ -\sin\phi & 0 & \cos\phi \end{bmatrix} \quad (6)$$

3.4 動画処理

3.4.1 検出対象の追跡

ステレオカメラを用いて撮影した映像をフレーム単位に分割し、フレームごとにオブジェクトが検出されたとして、そのオブジェクトに対する追跡処理を述べる。今、フレーム f において、オブジェクト $Object(f)$ が検出されたとする。このとき、過去のオブジェクト $Object(f-1)$ の予測移動範囲内に $Object(f)$ が存在した場合、同じオブジェクトが移動したとして追跡する。また、 $Object(f-1)$ が存在しない場合には、フレームをさかのぼり予測移動範囲を広げる。最大で $LOSTframe$ 分さかのぼり、処理を続ける。

予想移動範囲は1辺が $MoveRangem$ の正方形とした。正方形にしたのは、コンピュータの計算量を軽減させるためである。 i フレームさかのぼったときの予想移動範囲を式(7)に定義する。

$$MoveRange = \frac{3 \cdot i}{LOST} + 1 \quad (7)$$

3.4.2 対象が複数の場合の対応

検出対象が複数である場合、次に示す5つのパターンが発生する。

- (a): $Object(f-1)$ の予想範囲外に $Object(f)$ が存在する
- (b): 予想範囲内に $Object(f)$ が存在しない
- (c): 予想範囲内に $Object(f)$ が複数存在する
- (d): $Object(f)$ が複数の予想移動範囲内に存在する
- (e): (c) と (d) が混在する

(a) の状態は新しいオブジェクトが現れたのだと考え、新しく追跡を始める。(b) の状態はオブジェクトが検出範囲内から出て行ったのだと考え、追跡を終了する。(c) の状態は2つの経路候補が発生し、どちらが正しい経路なのか分からない。よってこの状態を分岐と定義し、経路を複製して追跡を続ける。(d) の状態は統合と定義し、2つの経路候補を残したまま追跡を続ける。しかしこのような処理をすると、状態(e)のときに経路

候補が急速に増加してしまう(図3左参照)。したがってこのような状態を交差と定義し、一定範囲内にある複数オブジェクトを1つにまとめる処理をおこなう(図3右参照)。この処理により、複数のオブジェクトが接近していても、分岐と統合の繰り返しによる経路候補の増大を防ぐことができる。

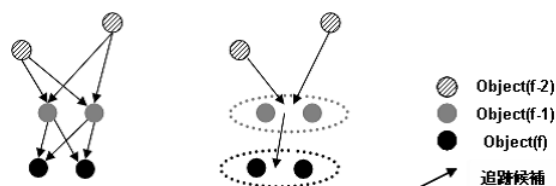


図3: 交差状態の処理

3.4.3 方向検出

方位センサとの照合に必要な方向情報は、オブジェクトの移動ベクトルから求める。しかし動画処理により得られる位置情報には、誤差が含まれる。よってオブジェクトの位置情報の移動平均をとることで、スムージングをおこなう。このスムージング処理の後、移動方向を求める。

4. プロトタイプシステム

4.1 モーションセンサ

市販の複合センサと小型PCを用いて、センサ処理部を実装した。センサについては、パイテック製のセンサモジュールTDS01Vと、NECトーキン製のモーションセンサMDP-A3U9Sの2種を用いて実装した。2種のセンサを用いるのは、提案した方向補正法の汎用性を検証するためである。前者のTDS01Vは3軸地磁気、3軸加速度、絶対気圧センサを内蔵している。後者のMDP-A3U9Sは3軸加速度、3軸角速度、3軸地磁気センサを有するモーションセンサである。TDS01Vについては、小型のモジュールとなっており、サンプリングレートが最大10Hzと低く安価である。一方のMDP-A3U9Sは、サンプリングレートが最大125Hzと高く、前者に比べ高価である。

また小型ロガーに関してはソニー製VAIO type Uを利用する。この小型PCに2種類のセンサをUSB接続し、センサデータの処理をおこなう。

4.2 ステレオカメラ

ステレオカメラはPoint Grey Research社のBumblebeeを用いる。オブジェクト抽出には同社のCensys3Dシステムを用いる。ステレオカメラをIEEE1394ケーブルにてコンピュータに接続し、制御及び映像の取得をおこなう。検出情報の座標については、x-y平面を床と同じ平面とし、原点はカメラ設置点の直下とした。

5. 評価実験

5.1 実験環境

鉄筋コンクリート建ての室内にて、 $4 \times 5m$ の検出範囲を設定した。ステレオカメラは検出範囲全域が撮影可能となるよう、高さ約2.6mの位置に設置した。照明は一般的な蛍光灯を使用する。太陽光の影響は少ない。また、センサを所持する被験者は、腰に2種のセンサノードを装着し、USBにて接続された小型ロガーを携帯しながら移動する。

5.2 実験内容

2人の被験者のうち、1人の被験者にセンサを装着する。センサを装着した被験者には予め設定した経路を動いてもらい、もう1人の被験者には自由に動いてもらう。さらに被験者の数を増やし、3人にする。3人の被験者のうち2人にセンサを装着し、実験をおこなう。被験者2人の試行を5回、被験者3人の試行を5回、合計10回の試行をおこなう。

5.3 実験結果

予め設定した経路と検出された経路の誤差を表1に示す。誤差がもっとも大きいところで、X軸方向に0.32m、Y軸方向に0.40mという高い精度を示した。参考のために、被験者2人の試行で検出されたセンサ所持者の経路を図4に示す。この経路はあらかじめ指定した経路と一致する。

表1: 設定経路と検出経路の誤差

最大誤差 [m]		誤差の平均 [m]	
X 軸方向	Y 軸方向	X 軸方向	Y 軸方向
0.32	0.40	0.08	0.11

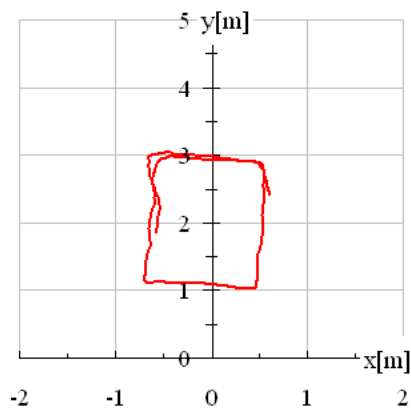


図4: 検出経路

また方向情報の精度も同時に検証したところ、TDS01V センサで最大誤差が1.09rad、誤差の平均は0.29rad。MDP-A3U9S センサで最大誤差が1.46rad、誤差の平均は0.39radであった。

照合アルゴリズムによって求めたセンサと経路候補の組み合わせが、正しい組み合わせなのか精度を検証した結果を表2に示す。すべての組み合わせが正しいときに、照合が成功したとみなす。被験者2人のときには100%、被験者3人のときには80%、全体で90%の成功率であった。また、センサデータの補正効果を確認するために、補正前の値とも照合をおこなった。補正後の照合成功率の上昇から、提案した補正法の効果が確認できた。

5.4 考察

表1より、検出された位置情報の最大誤差は0.4mであった。Censys3Dによって得られる位置情報が、頭部の右側面のものであるときと左側面のものであるときがあるため、頭部の大きさ分の誤差が生じる。しかし0.4m以内に複数の人間が接近している状況はまれであり、ユビキタスサービスを提供するには十分な位置検出精度だと考える。

表2: 照合結果

被験者 [人]	センサ [個]	試 行 [回]	経路 候補 [本]	センサデータ補正			
				TDS01V		MDP-A3U9S	
				補正前	補正後	補正前	補正後
2	1	1	6	成功	成功	成功	成功
		2	6	成功	成功	成功	成功
		3	3	成功	成功	成功	成功
		4	4	成功	成功	成功	成功
		5	2	成功	成功	成功	成功
3	2	1	7	失敗	成功	成功	成功
		2	7	失敗	失敗	失敗	失敗
		3	5	成功	成功	成功	成功
		4	6	成功	成功	成功	成功
		5	5	成功	成功	成功	成功

表2をみると、被験者が2人である場合には完全に照合が成功しているが、被験者が3人である場合には1つの試行で照合が失敗している。ステレオカメラの映像を確認すると、このときの試行でセンサ装着者の1人が「後歩き」や「横歩き」をおこなっていた。動画像処理から得られる方向情報は移動方向であるため、このような動きをすると計算結果が体の方向と異なる。よって、照合アルゴリズムによる対応付けが失敗した。このような動きに対する対応が今後の課題となる。

6. まとめ

本稿では、ステレオカメラとモーションセンサを用いた個人識別可能な位置検出手法について提案し、方向による照合手法の有用性を検証した。著者らの研究室では、動画像処理とセンサ情報処理を組み合わせた手法について研究開発が行われてきた。しかし従来の手法は複数の問題が存在し、実用的な手法ではなかった。本著で提案する手法も同じアプローチをとっているが、新たに提案した方向情報による照合の導入により、従来手法の問題が改善された。提案した手法に対する評価実験をおこない、方向による照合が有効であると証明できた。

参考文献

- [日立07] 株式会社日立製作所: UWB 無線を用いた高精度位置計測システム「日立 AirSenseTMUWB エントリーモデル」を発売, <http://www.hitachi.co.jp/New/cnews/month/2007/12/1210a.pdf>
- [古河] 古河機械金属株式会社: 超音波3次元位置測位システム, <http://www.furukawakk.jp/products/ZPS1.htm>
- [KAWAI05] Jun KAWAI, Kimio SHINTANI, Hirohide HAGA and Shigeo KANEDA: IDENTIFICATION AND POSITIONING BASED ON MOTION SENSORS AND A VIDEO CAMERA, The 4th IASTED International Conference on WEB-BASED EDUCATION, No.461-809.