

# 複数センサ融合にもとづく最適特徴の動的選択

## Dynamic Fusion of Visual Features Based on Multisensor Data

村北 卓也<sup>\*1</sup>  
Takuya MURAKITA

池田徹志<sup>\*1</sup>  
Tetsushi IKEDA

石黒 浩<sup>\*1</sup>  
Hiroshi ISHIGURO

<sup>\*1</sup> 大阪大学大学院 知能・機能創成工学専攻  
Department of Adaptive Machine Systems, Osaka University

### 1. はじめに

近年、コンテキスト・アウェア・コンピューティング[1]、ユビキタス・コンピューティング[2]などと呼ばれる分野において、行動認識システムを利用して、人びとの生活を支援しようとする試みが盛んである。ここでいう行動とは、手話などのジェスチャから、雑踏における人びとの群行動まで幅広い意味を含んでいるが、人びとの居所を表す位置は、これらの中でも恣意性がほとんどなく、数学的にも扱いやすい。また、人びとの移動は環境構造によって制約されるため、位置は環境との相互作用に基づく行動(テレビを見る、冷蔵庫をあけるなど)を調べる上でも興味深い特徴であるといえる。

GPS (Global Positioning System) は、位置認識の代表的なシステムであるが、利用者は電波受信装置を身につける必要がある。日常において円滑な生活支援を実現するためには、システムがトランスペアレント(transparent)であることが不可欠である。つまり利用者がシステムの存在に気付かないほどに支障がないものでなければならない。

我々はこのようなシステムを実現するために有効なセンサ情報は二種類に絞られると考える。すなわちビデオカメラなどが観測する赤外光から可視光領域の情報、およびマット状の床圧力センサが得る圧力情報である。我々は各々の情報に基づく位置認識(人間追跡)システムを既に構築している。



(a) Omnidirectional camera



(b) Floor pressure sensor

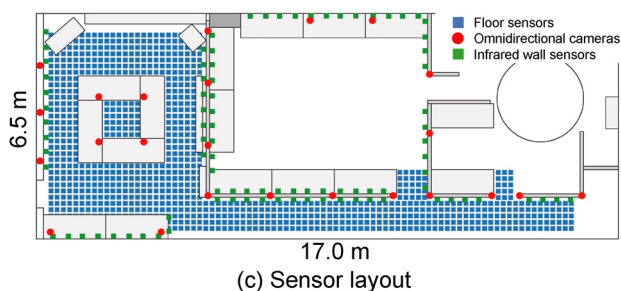


図 1. センサネットワークシステム

連絡先: { murakita, ikeda } @ed.ams.eng.osaka-u.ac.jp  
ishiguro@ams.eng.osaka-u.ac.jp

### 2. センサネットワークシステムを用いた先行研究

#### 2.1 システム構成

図 1 は、全方位カメラ、赤外線壁センサ、床圧力センサからなるセンサネットワークシステムを示す。研究室内に、24 台の全方位カメラ、100 個の焦電センサ、約 37 m<sup>2</sup>の床センサが配置されている。

#### 2.2 全方位カメラを用いた人間追跡システム[3]

全方位画像をパノラマ画像に展開し、背景画像との差分を動径方向に積分することで、人間の存在を示すヒストグラムを得ることができる(図 2)。ヒストグラムから、人間のみかけの方位を算出し、三角測量の原理にもとづいて位置を測定した。このシステムは、リアルタイムかつ、数 cm の比較的小さい誤差(標準偏差)で人間を追跡することができた。しかし、人間を示す領域を画像から安定にセグメントするために、背景差分をもちいたため、追跡誤差が照明条件に左右されやすいという欠点があった。

#### 2.3 床圧力センサを用いた人間追跡システム[4]

人間は歩行状態にあるとき、両足によって概ね複数の床センサ領域を反応させる。したがって床センサ情報から人間の位置を求めることは多対一の対応問題に相当する。この問題を解決するために 4 状態スイッチングモデル(図 3)を適用した。床センサは体重を直接検知するため、外乱に対して安定であるが、人びとが混雑した状況においては、多対一の対応付けが困難となった。90%以上の確からしきで人間を区別するためには 70cm 以上の間隔を保つ必要があった。

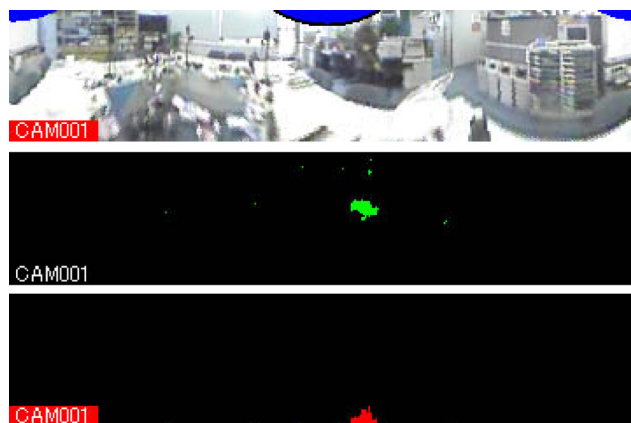


図 2 (上) パノラマ画像 (中) 背景差分 (下) ヒストグラム

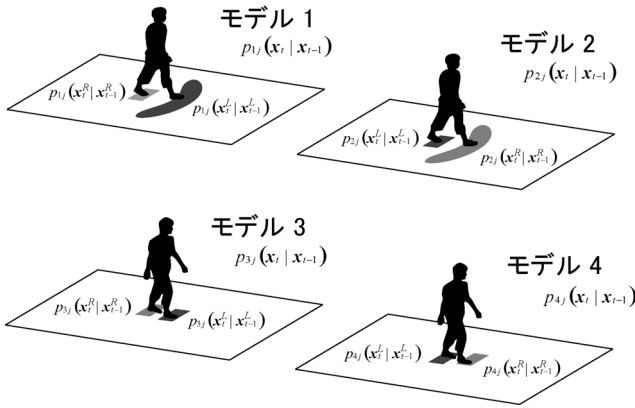


図 3. 4 状態スイッチング予測モデル

### 3. センサ情報の融合

#### 3.1 センサ情報の融合における利点と欠点

展開パノラマ画像における人間領域のセグメンテーションは照明条件に影響されやすいが、床センサの位置情報を利用することで、人間に関係のある差分を重点的に選択することができる。また、カメラによって推定された、比較的誤差の小さい位置情報を利用して床センサの多対一対応の確からしさの改善も期待できる。

逆に、不確かなセグメンテーションにもとづく位置情報を床センサ情報と融合したとき、多対一対応の誤りが増大し、センサ融合の処理が悪循環に陥ることも考えられる。したがって、センサ情報の確からしさを確率的に評価し、有効な情報のみを選択できるように推定理論を構築する必要がある。

#### 3.2 ベイズ理論にもとづく人間追跡

時間  $t$  における、 $N_h$  人の人間の位置を表す集合を

$$X_t = \{\mathbf{x}_{1,t}, \dots, \mathbf{x}_{i,t}, \dots, \mathbf{x}_{N_h,t}\}, \quad (1)$$

$N_s$  個のセンサによって得られた観測の集合を

$$Z_t = \{z_{1,t}, \dots, z_{j,t}, \dots, z_{N_s,t}\} \quad (2)$$

とする。観測  $z_{j,t}$  は  $N_f$  個の特徴を含む集合である;

$$z_{j,t} = \{d_{1,j,t}, \dots, d_{k,j,t}, \dots, d_{N_f,j,t}\} \quad (3)$$

時間  $t$  における、処理の目的は、

$$p(X_t | Z_t) = \{p(\mathbf{x}_{i,t} | Z_t^i); i = 1, \dots, N_h\} \quad (4)$$

を求め、 $\hat{\mathbf{x}}_{i,t} = E[p(\mathbf{x}_{i,t} | Z_t^i)]$  を推定することである。集合  $Z_t^i$  は、観測  $Z_t$  のうち、人間  $i$  に対応付けられた特徴の集合である;

$$Z_t^i = \{d_{1,t}^i, \dots, d_{k,t}^i, \dots, d_{m,t}^i\}. \quad (5)$$

(4)式を求めるためにベイズの定理を利用する;

$$p(\mathbf{x}_{i,t} | Z_t^i) = \text{const} \cdot p(Z_t^i | \mathbf{x}_{i,t}) \cdot p(\mathbf{x}_{i,t}). \quad (6)$$

ここで尤度関数の、特徴に対する独立性を仮定する;

$$p(Z_t^i | \mathbf{x}_{i,t}) = \prod_{k=1}^m p(d_{k,t}^i | \mathbf{x}_{i,t}) \quad (7)$$

(6)式の事前確率分布として、時間  $t-1$  における、位置の推定値を与えることで、次の差分方程式を得る。

$$p(\mathbf{x}_{i,t} | Z_t^i) = c \cdot \prod_{k=1}^m p(d_{k,t}^i | \mathbf{x}_{i,t}) \cdot p(\mathbf{x}_{i,t} | \hat{\mathbf{x}}_{i,t-1}) \quad (8)$$

### 4. 議論

システムのふるまいは、観測モデル;  $p(d_{k,t}^i | \mathbf{x}_{i,t})$  および、予測モデル  $p(\mathbf{x}_{i,t} | \hat{\mathbf{x}}_{i,t-1})$  によって決定される。観測モデルは、単峰性関数になることが望ましいが、背景モデルやノイズ、クラッターなどの要因を考慮すると、観測モデルを一意に決定することは極めて困難である。

ノイズやクラッターの影響を排除するための方法として、観測された分布にガウシアンをフィッティングすることが考えられる。ガウシアン分散に閾値を設け、不確かな分布と考えられる場合は、観測モデルを一様分布にすればよい。一様分布は(8)式の尤度に影響を与えないため、結果としてこの観測は無視される。

一方、リアルタイムに近い周波数で位置を推定することができれば、時間的に連続する画像における人間の運動はごくわずかなものとなる。したがって、予測モデルはガウシアンなどで若干の推定誤差を表現する程度でよいものと考えられる。

### 5. 参考文献

- [1] B. Schilit, N. Adams, and R. Want, "Context -Aware Computing Applications", Proceedings Workshop on Mobile Computing Systems and Applications. IEEE, December 1994.
- [2] J. Hightower and G. Borriello, "Location Systems for Ubiquitous Computing", Computer, 34(8), 2001, 57-66.
- [3] T. Sogo, H. Ishiguro, and M. Trivedi, "N-ocular stereo for real-time human tracking", In Panoramic Vision: Sensors, Theory and Applications (R. Benosman and S.B. Kang, Eds.), Springer Verlag, 2001.
- [4] T. Murakita, T. Ikeda, and H. Ishiguro, Human Tracking using Floor Sensors based on the Markov Chain Monte Carlo Method, Seventeenth International Conference on Pattern Recognition (ICPR), pp.917-920, Aug. 2004.