

I-088

# ステレオ視による閉空間人物動作解析

## Stereo Camera Surveillance Sensor for Closed Space

林 健太郎<sup>†</sup>  
Kentaro Hayashi

橋本 学<sup>†</sup>  
Manabu Hashimoto

鷲見 和彦<sup>†</sup>  
Kazuhiko Sumi

### 1. まえがき

人間の居住空間には、個室やエレベータかご等、不特定多数の人物が利用できるさまざまな閉空間が存在する。しかし閉ざされた空間という特殊な状況が、利用者にいわずらや犯罪を誘発させる原因となったり、何らかの原因により倒れたまま他人に発見されず放置されたりという問題がある。こうした異常事象を自動検出できれば、防犯や安全につながる。

画像による挙動解析は単眼カメラが主流であり、ステレオカメラは価格と計算コストの面で実用化困難と考えられている。しかし、現実の場面では距離情報をうまく利用することで単眼にはできないことを低コストで実現できる可能性がある。

ここでは、ステレオカメラを用いた異常事象(本論文では倒れおよび暴れと定義)検出手法と、それを用いたりリアルタイム監視システムを提案する。実験の結果から、ステレオカメラによる挙動解析が有効であることを示す。

### 2. 異常事象検出手法

本論文では人物の倒れと暴れを閉空間内における異常事象と定義し、これらをステレオカメラで検出する手法を提案する。倒れとは、床面上に低い物体(高さは閾値により決定される)が長時間横たわっている状態とする。また平時であれば閉空間内の大きな動きはほぼあり得ないという理由から、単に画像上に大きな動きを生じたときを暴れと定義する。ただし、ペットや子供の動きは含まない。図1に本システムの認識アルゴリズムの流れを示す。

最初に異常事象検出に共通となるステレオ距離計算方法について述べる。我々は高速ステレオ対応付け手法[1]を採用した。得られた視差  $d[\text{pixel}]$  から奥行き  $z[\text{mm}]$  を以下の式で求める。

$$z = \frac{q_x D}{2 \tan(a_x/2)} \cdot \frac{1}{d} \quad (1)$$

ここで  $q_x$  は水平画像解像度(=640)[pixel]、 $D$  はカメラ間距離 [mm]、 $a_x$  は水平画角 [rad] である。

#### 2.1 倒れ検出

最初にカメラ設置時のパラメータを用いて、カメラから見たときの床面の状態(距離、傾き)をあらかじめ計算しておく。そして、閉空間内を上下二つの3次元ボックス領域に分割し、床面から近いボックス領域に含まれる計測点数が、床面から遠いボックス領域の計測点数よりも多い状態が一定時間以上継続すると倒れが発生していると判断する。

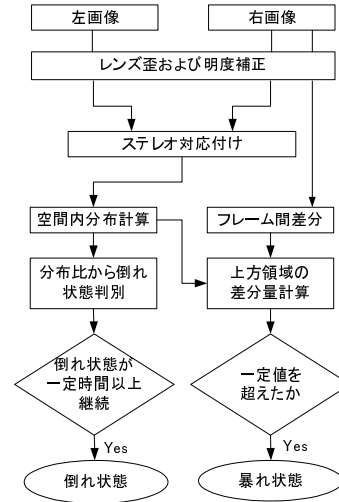


図1: 検出アルゴリズムの流れ

ただし信頼性の低いステレオ距離観測値が3次元領域中に定常的に混入すると、認識率が著しく低下する。そこで床面領域を画像上の部分領域で指定し、部分領域以外の計測値については無視する。こうすることで下のボックス領域に含まれる計測値が床面上にあることがほぼ確実に保証される。上のボックス領域についても同様である。検出の一例を図2に示す。図中(b)(d)のグレー領域は上ボックス領域、白領域は下ボックス領域に含まれるステレオ計測点を示す。また、同色の矩形領域は上記部分画像を示す。画像の右端のバーはシステムが持つ倒れ確信度を示しており(d)では確信度が高くなっている。このように、ステレオ距離情報が得られれば、単眼では必須のシルエット解析など複雑な手法を用いることなく、人物の倒れを簡単に検出することができる。

#### 2.2 暴れ検出

ビデオシーケンスから暴れ等の事象を背景差分から検出する手法[2]が提案されているが、真横からの映像を仮定しており閉空間内のカメラ設置条件には合わない。

閉空間における動き量検出は暴れ検出にほぼ直結しており、動き量検出として最もシンプルかつ高速なフレーム間差分の面積を用いれば十分である。ただしこのままでは、子供等の暴れと犯罪時の暴れを区別できない。そこで倒れ検出時の上下ボックス領域を応用し、下領域での動きを無視することで両者を区別する。背景差分からだけではこのような区別が非常に困難であるが、ステレオ距離情報があれば容易である。

図3(b)に暴れ時の変化領域を示す。暴れ時は明らかに面積が大きく、この場合であれば閾値1200~2000(全

<sup>†</sup> (社)電子情報通信学会, IEICE

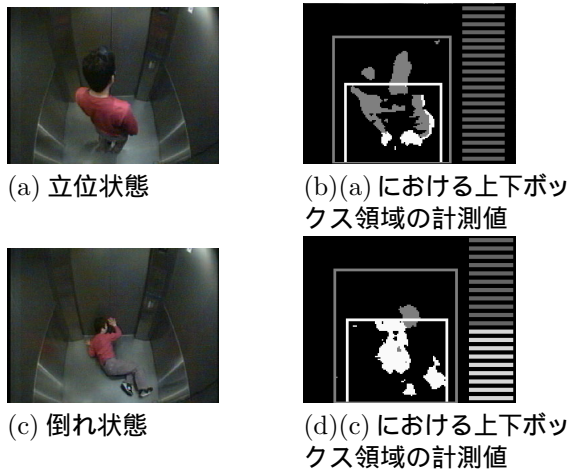


図 2: 倒れ検出例

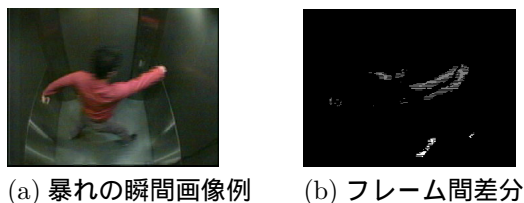


図 3: 暴れ例

面積の約 7.6 ~ 10% [pixel] の範囲内で検出できる。

### 3. 実験結果

今回は代表的な閉空間としてエレベータを対象とする。エレベータ内で一人の人物が倒れと暴れの挙動を模擬し、それをデジタルテープに記録してオフライン検出実験を行う。

図 4(a) にシステム全体の構成イメージを示す。閉空間内上方に設置されたステレオカメラヘッド (図 4(b)) からの映像を 2ch 同時入力可能なキャプチャ I/F (iEi 社製 IVC-200G を利用) を介して PC に入力、画像処理/認識は全て PC 上でを行い、認識結果を画面へ出力する。ステレオカメラヘッドの取り付け位置は、床面からの高さ約 2050 [mm]、水平俯角約  $-50$  [度] (下向き)、またカメラ間距離 150 [mm] である。

ステレオカメラには、CIS 社製 DCC-2030N (NTSC 映像) を 2 個、レンズは水平画角 109 [deg] の物を使用した。これらをステンレス製の台座にバネとネジで取り付け、画角を調整して固定した。

倒れ挙動では、手前や奥の位置から床面への倒れに至るシーケンスを 5 通り模擬した。倒れ状態が 1 分以上継続すると倒れと判断する。システムはこれら 5 通りの倒れの全てを検出することができた。

暴れ挙動では、数分以内にパンチやキック等の挙動を、カメラに対する向きを変えて模擬した。本実験で用いた暴れシーケンスには、パンチやキックなどの挙動が 10 通り含まれている。これを面積の閾値を変化させながら

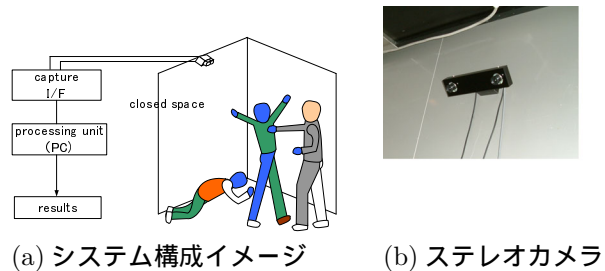


図 4: システム概略

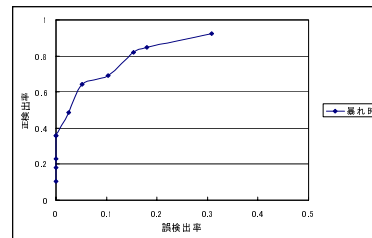


図 5: 暴れ検出における ROC 曲線

検出したときの ROC 曲線を図 5 に示す。通常時の面積最大値は 1175 であり、この場合閾値が 1175 よりも大きければ通常時の挙動を誤検出することはない。閾値 1175 以上での正検出率の最大は 0.69 であり、検出率 69% とみなせる。

上記倒れ暴れをリアルタイムに検出するシステムを PC (Linux) 上に実装した。Pentium4 3.06GHz の PC を用いて、X ウィンドウ上への処理結果表示時 15.2 [fps]、表示なし時 30 [fps] を達成した。なお、画像処理時間の内訳は歪み補正 5.72 [ms]、明度補正 1.72 [ms]、フレーム間差分 0.73 [ms]、ステレオ対応付け 1.50 [ms]、挙動解析部 1.7 ~ 2.7 [ms] となっている。画像処理時間計 12.4 < 33 [ms] であり、今後より複雑な挙動解析も可能である。

### 4. あとがき

閉空間内の異常事象検出のためのリアルタイムステレオ画像センサを提案した。本論文では倒れと暴れのみを扱ったが、我々の実験において倒れ 100% 検出、暴れ 69% 検出を確認した。暴れ検出率に課題が残るものの、ステレオカメラによる挙動解析が有効であることを確認できた。

### 参考文献

- [1] 岡田慧, 加賀美聡, 稲葉雅幸, 井上博允. PC による高速対応点探索に基づくロボット搭載可能な実時間視差画像・フロー生成法と実現. 日本ロボット学会誌, Vol. 18, No. 6, pp. 896-901, 2000.
- [2] Ankur Datta, Mubarak Shah, and Niels Da Vitoria Lobo. Person-on-person violence detection in video data. In *Intl. Conf. on Pattern Recognition*, pp. 433-438, 2002.