

人工神経回路網による人検出システムの開発

尾崎裕章^{*1} 服部公央亮^{*1} 田口亮^{*1} 保黒政大^{*2} 梅崎太造^{*1}
^{*1}名古屋工業大学 ^{*2}中部大学

1. まえがき

近年、住宅や会社等における防犯システムが広く普及してきている。しかし、不法侵入検知システムは監視場所の環境変化や小動物の移動に対して敏感に反応することが多く、侵入者か否かを判断する場合に支障をきたすことがある。また、重要施設などにおいては監視カメラを使用したシステムも用いられているが、監視員が常に侵入者の有無を確認する必要があり、人件費が問題となる。さらに、人がカメラ画像を監視する場合には、監視員の疲労やチェックミスによる見落としが存在するという危険性を伴う。

従来からカメラ画像における人検出として背景差分法を用いた手法や人工神経回路網を用いた手法が提案されている^{[1][2]}が、背景差分法では環境変化や小動物の移動に弱く、人工神経回路網では環境に適した人画像や背景画像を選定する必要がある。そこで本研究では、人であるか否かを判定する人工神経回路網に監視カメラの画像を入力して、得られた出力から背景との差分により人を検出する手法を提案する。これにより、背景に対する誤検出を軽減して、人のみの検出が可能となる。

2. 提案手法

2.1 人工神経回路網

人工神経回路網とは、脳機能の特性をコンピュータ上のシミュレーションにより表現することを目指した数学モデルである。人工神経回路網において教師あり学習と教師なし学習という2つの訓練手法が存在する。本稿では、教師信号の入力により問題に最適化されていく教師あり学習を使用する。多層ニューラルネットワークを訓練する場合には、誤差逆伝播法を使用する。

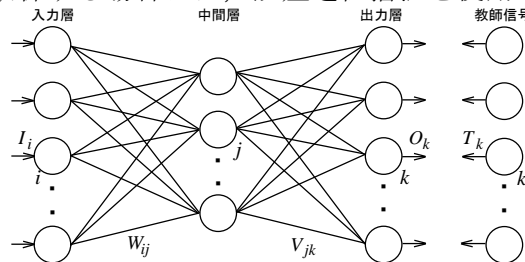


図1 人工神経回路網の構成

人工神経回路網の構成は、図1に示す入力層、中間層、出力層からなる階層型人工神経回路網である。本手法では入力層 20×40 、出力層 20×40 、中間層 256 ユニットのシグモイド係数 $U = 6.0, 7.0, 8.0, 9.0$ の三階層人工神経回路網を構築した。

2.2 学習データ

学習データには、MIT CBCL PEDESTRIAN DATABASE から 924 枚の人の全身画像を用いて、この画像数を1セットとした。抑制データには、6411 枚の背景画像を用いた。画像は人工神経回路網に入力する前に元画像に対して前処理を行う。処理手順および処理画像は図2に示す。

処理後の画像に対して、拡大縮小、平行移動、回転、濃度変換、左右反転のランダム化を行う。ランダム化により入力画像の学習パターンを増やすことで、認識パターンの位置ずれ、大きさ、変形等の変動に対する認識能力が高くなり、ノイズの影響を受けにくい人工神経回路網が構築される。また、人の領域を1、背景の領域を0としてガウス分布で変化する濃淡画像を教師信号とする。教師信号には入力画像と同様の拡大縮小、平行移動、回転、左右反転のランダム化が行われる。入力画像に対して使用する教師信号の例を図3に示す。

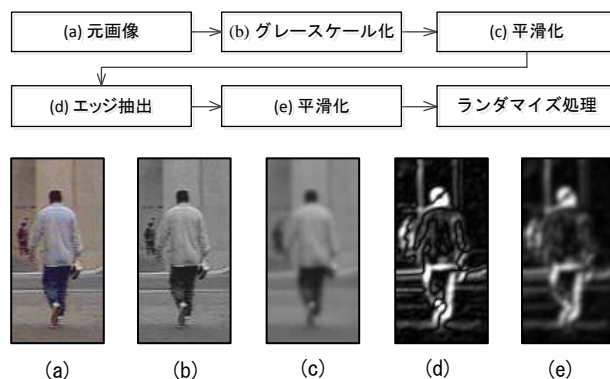
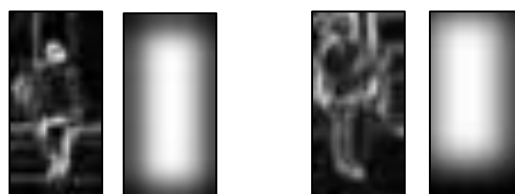


図2 処理手順および処理画像



(a) ランダム化なし (b) ランダム化あり

図3 入力画像に対する教師信号

Human Detection System with Neural Network
 Hiroaki OZAKI^{*1}, Koosuke HATTORI^{*1}, Ryo TAGUCHI^{*1},
 Masahiro HOGURO^{*2}, Taizo UMEZAKI^{*1}
^{*1} Nagoya Institute of Technology, ^{*2} Chubu University,

3. 実験結果

3.1 学習実験

前記の人工神経回路網を用いて学習した。10セット毎に得られた人工神経回路網からクロズドデータの誤差を計算して、式(1)において学習誤差 e を RMSE で評価した。 $out_k(x_i, y_j)$ は各画素値の出力値、 $teach_k(x_i, y_j)$ は同画素の教師信号の値、 k は入力画像の画像番号、 W および H は画像の幅と高さ、 N は(ランダムイズのパターン数) \times (1セットの画像数)である。学習結果は、シグモイド係数 9.0 のとき誤差が最小で収束した。

$$e = \sqrt{\frac{1}{N} \sum_{k=0}^{N-1} \left(\frac{1}{WH} \sum_{j=0}^{H-1} \sum_{i=0}^{W-1} \{out_k(x_i, y_j) - teach_k(x_i, y_j)\}^2 \right)} \quad (1)$$

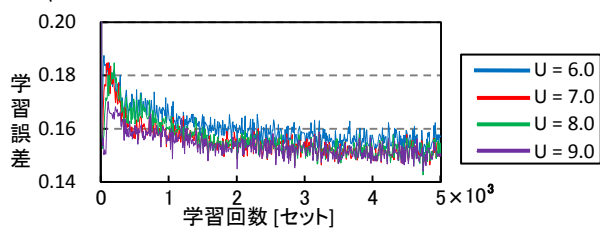


図4 学習誤差曲線

3.2 検出実験

学習した人工神経回路網より人の検出を行う。検出用データには、CAVIAR DATASETを用いた。対象画像および背景画像に対して、学習時と同様の処理を行い、設定した大きさで画像から切り出し、人工神経回路網に入力する。これを複数の領域で行い、得られた出力を統合することにより図 5(a2)(b2)に示すような出力が得られる。得られた 2 つの出力の差分を取ることで図 5(c2)に示すように人に対する反応が得られた。

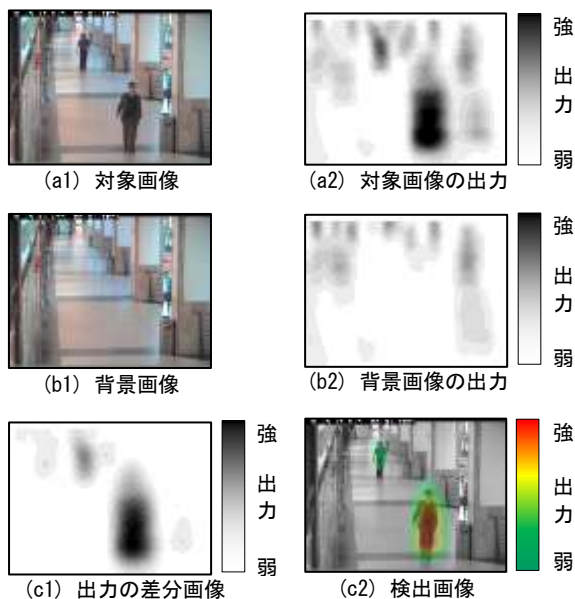


図5 人工神経回路網の検出結果

評価実験として、従来手法である背景差分法 (B.S.), 人工神経回路網(N.N.)を用いた手法、背景差分法と人工神経回路網を組み合わせた提案手法の精度比較を行う。比較には、Detection Error Tradeoff(DET) カーブを用いる。DETカーブは、手法の閾値を変化させたときの未検出率と誤検出率で表され、原点に近づくほど精度が良いことになる。結果は誤検出率が 10.0[%]の場合に、未検出率が背景差分法で 2.2[%], 人工神経回路網を用いた手法で 25.4[%], 提案手法で 15.9[%]であることが確認された。各手法における評価結果の DET カーブを図 6 に示す。

提案手法では人が横を向いている場合や体が見切れている場合、人が小さい場合において検出率が低いが、正面もしくは背を向いた人に対する検出率が高い傾向にあることが確認された。未検出の例を図 7 に赤枠で示す。

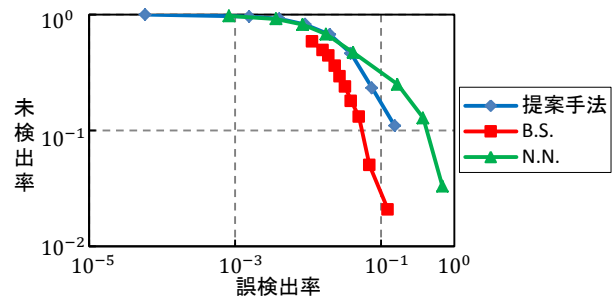


図6 DETカーブ

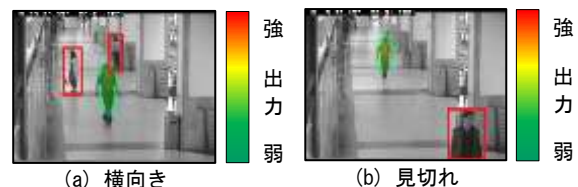


図7 未検出例

4. まとめ

本稿では、人であるか否かを判定する人工神経回路網から得た出力を用いて背景差分を行い、人を検出する手法を提案した。今後の課題は、姿勢変化や人の服装をより多様にして人工神経回路網を構築することで、多様な環境における人検出の精度の向上を目指す。

参考文献

- [1] 岩田健司, 佐藤雄隆, 尾崎竜史, 坂上勝彦, “統計的リーチ特徴法に基づくロバスト背景差分”, 電子情報通信学会論文誌, Vol.92, No.8, pp.1251-1259, 2009.
- [2] 舘俊太, 武藤佳恭, “3層ニューラルネットワークと変形デンプレート法による動画の人物検出”, 情報処理学会論文誌, Vol.44, No.14, pp.48-56, 2003.