

人物検出技術の高度化に関する研究

—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 田中 靖哲* 長谷部 斉志* 菊池 春秀*
 下中 智* 川合 亘* 橋本 周司** 富永 将史**
 清水 早苗*** 飛谷 謙介***

Study on an Advanced Method of Human Detection

- Construction of Smart Environment for Supporting Human Activities -

Hiroki WATANABE Yasunori TANAKA* Hitoshi HASEBE* Haruhide KIKUCHI*
 Satoshi SHIMONAKA* Wataru KAWAI* Shuji HASHIMOTO** Masafumi TOMINAGA**
 Sanae SHIMIZU*** Kensuke TOBITANI***

あらまし 近年、監視カメラシステムには、映像を的確に捉え、状況を分析・判断できる機能を具備したインテリジェント化が求められている。そこで、本研究では、監視カメラとして用いられるネットワークカメラの映像を対象に、出現した物体が人であるか判断するための前処理として重要となる物体の検出精度の向上について検討する。また、検出した物体の動き情報を用いることで、人であるか判断する手法について検討する。

キーワード 監視カメラシステム、インテリジェント化、物体検出、人物判定

1. はじめに

近年、監視カメラシステムには、映像を撮り続けるシステムから、映像を的確に捉え、状況を分析・判断できるシステムへのインテリジェント化が求められている。そのため、不審者や物体の動きの把握、管理者負担の低減などの機能を向上させることは、次世代監視カメラシステムとしての市場優位性の確保に大きく貢献すると考えられる。

そのような中で、監視対象となる画像内に現れた物体が人であるか判断する機能は、監視カメラシステムにおいて高い有用性を持つ。例えば、人が荷物を運んで置いたシーンでの人と荷物とを識別する技術、あるいは、人が静止したまま動かないシーンでの人を検知する技術などは、様々なアプリケーションに応用できるため、次世代監視カメラシステムにおいて重要な要素技術と成り得る。

現在の監視カメラシステムには、人であるか判断する

技術に類似した機能として、モーション検知がある。しかしながら、モーション検知機能は対象となる画像内で、色の变化や動きで判断するため、前述のような人と荷物との識別や静止した人の検知は困難である。

そこで本研究では、近年、監視カメラとして注目されているネットワークカメラの映像を対象に、出現した物体が人であるか判断するための前処理として重要となる物体の検出精度の向上について検討する。また、検出した物体の動き情報を用いることで、人であるか判断する手法について検討する。

2. 物体検出精度の向上

監視カメラシステムにおいて、監視対象となるエリアに侵入した物体が人であるか判断するためには、画像中から出現物体のみを精度よく検出および分離することが重要である。しかしながら、想定するカメラがネットワークカメラである場合、取得される画像はJPEG圧縮された画像であるため、ブロックノイズやモスキートノイズが発生し、通常の画像パターン検出では精度よく出現物体を検出することが困難である。

一方、我々はこれまで、背景から分離するための画像

* 株式会社電算システム

** 早稲田大学総合研究機構WABOT-HOUSE研究所

*** 財団法人ソフトピアジャパン IT研究センター

パターン検出技術としてRadial Reach Filter^[1](以下, RRF)を用いてきた. 従来の背景差分法^[2]やフレーム間差分法^[3]など画像の明度差に基づく手法では, 演算量が少ないというメリットを持つ反面, 照明条件の変化の影響を受けやすく, 更に物体表面の明度と背景の明度とが近い部位は検出できず, 検出領域にいわゆる「虫食い」が生ずるなどの問題があった. RRFは, 明度変動の影響を抑えながら, ピクセル単位の分解能で局所的なテクスチャを評価するため, 物体表面の明度分布, あるいは対象シーンの明度変化に鈍感でありながらも出現物体をロバストに抽出することが可能であり, 次の二つの特徴的な要素で構成されている.

- ・ 着目点から放射状 (8 方向) にしきい値 (T_p) 以上の明度差を持つ点を探索し, ペア (8 組) を組む.
- ・ 各ペアにおける明度の大小関係 (差分の正負) を二値符号 (8 組分=計 8bit) とし, 着目画素の評価値とする.

しかしながら, RRFもJPEG画像中のノイズに対しては影響を受けるため, 一定以上の明度差を持つように明度差のしきい値 T_p を決定しなければノイズに対する耐性を高めることができない.

そこで, 本章では, JPEG画像中のノイズを低減した物体検出を実現するために, RRF画像の生成時における明度差しきい値 T_p の決定方法について検討するとともに, 画像ノイズの影響を抑えた物体検出手法について検討する. また, 影や照明変動による誤検出の問題や, システムの実装時に必要となる背景更新の問題についても検討する.

2. 1 明度差しきい値の決定

明度差のしきい値 T_p の目的^[1]は, ペアを結ぶ条件として一定以上の明度差を持つことを規定することで, ノイズに対する耐性を高めることである. 一方で, 必要以上に大きな T_p は, 大きなリーチを生み出す可能性があり, 空間解像度の低下が起きる可能性がある.

図1は, 実際に T_p の値を変化させたRRF画像である. T_p の値が小さい図1(c), 図1(d)はノイズの影響を強く受けて誤検出が発生していることがわかる. 一方, 図1(e)ではテクスチャの状況によって適切なリーチ長が設定されることで, 頑健なペアが形成され, テクスチャの弱い面においてもノイズの影響を受けにくくなっていることがわかる. これに対し, 図1(f)は大きなリーチが生み出されたことにより, 空間解像度が低下し, 未検出領域が発生していることがわかる. そのため, T_p は撮像系のノイズ特性を測定することにより予め合理的に決定する必要がある. 通常は, 測定されたノイズ分布の 2σ 程度に T_p を設定することで, ノイズ成分の大部分に影響されない最小の T_p を決定することができる.

本研究では, 適切なリーチ長を設定するために, 監視対象となるシーンにおいて, 物体を出現させずに撮影した二枚の画像の差を取るという簡易的な方法で, 撮像系

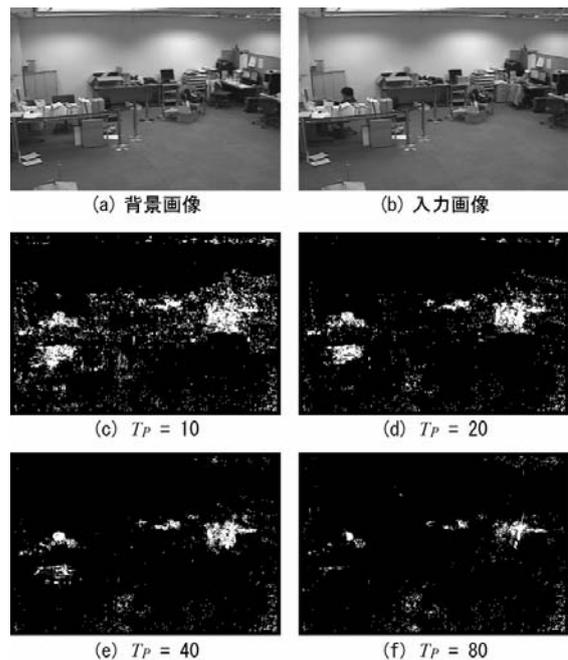


図1 明度差しきい値の違いによるRRF画像の生成例

のノイズ分布の標準偏差 $\sigma=22.27$ を得た. 約 2σ となる図1(e)の例では, ノイズの大部分の影響を受けない出現物体の検出ができていことがわかる. システムに実装する場合は, 背景画像を取得する際, 連続して同一背景画像を複数枚読み込み, ノイズ分布を測定することで明度差しきい値 T_p を決定する構成とすることが必要である.

2. 2 画像ノイズの低減

JPEG圧縮ではブロック単位で変換するため, 圧縮率を上げるとブロックの境界にブロックノイズと呼ばれるノイズが生じる. また, 色の境界など急激な変化をする部分にはモスキートノイズと呼ばれるノイズが生じる.

そこで, これらのノイズを低減するために, 取得した画像にガウシアンフィルタ処理を加える. ガウシアンフィルタは, ガウス関数にしたがって画素値を平均化することにより, もとの画像より輪郭がぼやけた画像が生成できる. そのため, ガウス関数を適用する範囲が大きくなるほど全体がぼやけた画像となる.

図2は図1の背景画像, 入力画像に対し, ガウス関数の適用範囲の大きさを 3×3 , 5×5 , 7×7 と変化させてガウシアンフィルタを施した場合のRRF画像の生成例である. 図1(e)に比べ図2(b), 図2(d), 図2(f)はノイズの影響が減少し, 出現物体が精度よく検出できていることがわかる. 適用範囲の大きさは状況に応じて使い分ける必要があるが, 図2(e)のようにかなりぼやけた画像に対しても適切に物体を検出することが可能である. なお, 明度差しきい値 T_p は上述と同様に, 予め測定したノイズ分布より σ を求め, 約 2σ となるように設定した. ガウス関数の変数 σ は

$$\sigma = (n/2 - 1) * 0.3 + 0.8 \dots\dots\dots (1)$$

により与えられ, n は適用範囲の大きさである^[4].

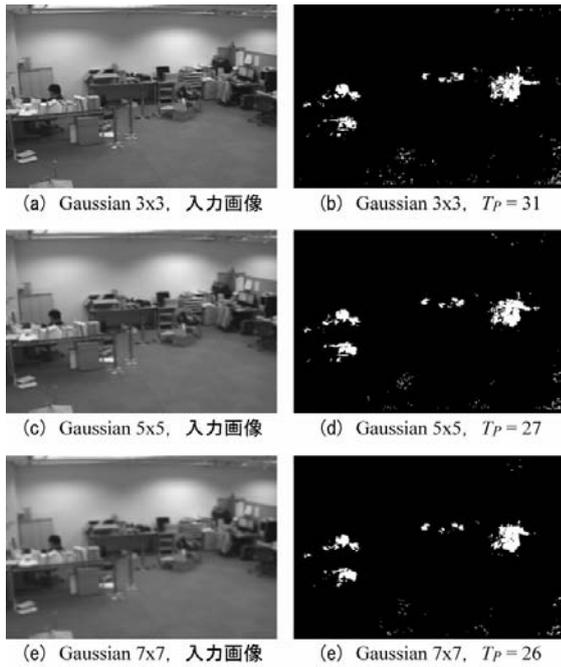


図2 ガウシアンフィルタ違いによるRRF画像の生成例

2. 3 誤検出の低減

前節で述べたとおり，RRFは局所的なテクスチャを評価することにより物体表面の明度分布，あるいは対象シーンの明度変化に鈍感でありながら出現物体をロバストに検出することが可能である．しかし，テクスチャの少ない領域が撮像サイズに対して大きな割合を占める（例えば，壁，天井，床面等）場合は，着目点に対し明度の大小関係を計算する点（評価点）の位置が遠くなる可能性が高く，局所的なテクスチャを評価することが困難である．そのため，このような領域に対して明度変化（例えば，日中屋外の太陽光の変化，移動物体による影，移動する人工照明の照射，出現物体からの反射等）が生じた場合，評価値の正負が反転することがあり，結果として図3，4に示すような誤検出が生まれやすくなる．そこで，相関係数により出現物体として検出された領域が誤検出であるか判定する．

相関係数とは二つのデータ列間の相関（類似性の度合い）を示す統計的指標であり，

$$\rho = \cos \theta = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \|\vec{y}\|} \dots\dots\dots (2)$$

により表される．画像における相関係数を考える場合， \vec{x} は出現物体として検出された領域における画素全ての明度値を成分とするベクトルで， \vec{y} は背景画像内の同領域の明度値を成分とするベクトルである．また， θ はベクトル空間内で \vec{x} ， \vec{y} ベクトルがなす角度である．

計算された相関係数は二つのベクトルの大きさの積で割っているため，検出領域内の画素全てに同程度の明度変化が加わる場合は背景画像の同領域と高い相関値を示



図3 明るさの変化による誤検出例

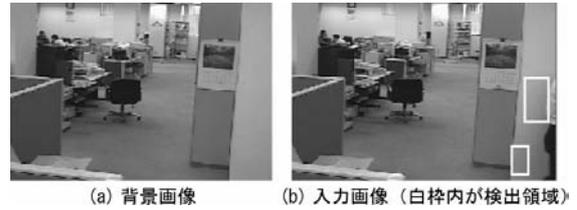


図4 人の影による誤検出例

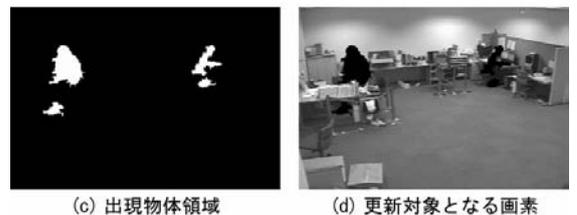


図5 背景の更新例

す．そのため，図3に示す明るさの変化による検出領域や，図4に示す人の影による検出領域は，誤検出領域として判断することが可能である．実験では，相関係数が0.98以上のものを誤検出領域として判定することにより，図3，4における誤検出を防ぐことができた．

2. 4 背景の更新

頻繁に自動車が横行する道路や人通りが多い繁華街などでは，出現物体を含まない背景画像を取得することが困難である．仮に，車や歩行人を含んだ画像を背景画像としても誤検出の要因となり，安定して物体を検出することができない．そのため，常に出現物体を含んだ画像から背景画像を推定および生成する処理が必要となる．

そこで，あらかじめ取得した背景画像（図5(a)）に対して，現在の入力画像（図5(b)）と比較し，出現物体領域（図5(c)）を抽出する．その後，出現物体領域以外の画素値全て（図5(d)の黒抜き以外）を背景画像における同座標の画素値と置き換え，これを順次繰り返すことにより背景画像を更新する．

3. 人物の判定

前章では、RRFを用いた物体検出手法に、ノイズの影響を低減する前処理と誤検出を低減する後処理を加えた高精度な物体検出処理について述べた。これにより検出された物体領域には、監視対象エリアに入ってきた人と、その人が持ち込んだ物とが含まれており、この二つを区別することが必要である。

本研究では、物体領域を人と物とに区別するため、動き情報を用いたアプローチをとる。まず、動きのある領域を人、静止している領域を物とした場合、静止している人を物と判定してしまう問題がある。そこで、人は自律的に動き、物は人により動かされ自律的に動かないと仮定し、物体領域を動きのある「動人物」と、かつて動いていたが静止している「静人物」、そして静止している「物体」の三つに区別する。

3. 1 「動人物」の判定

「動人物」はフレーム間差分で動きのある領域(以下、動領域)を検出することにより判定する。まず、物体領域のうち、フレーム間差分により動領域を検出する。ここで、フレーム間差分は現時刻の入力画像と一つ前の入力画像より求める。このとき、一つ前の入力画像の物体領域が動領域として検出されるが、背景差分との積を求めることにより、現時刻の動領域のみを求めることが可能である。次に、検出した動領域が背景差分で検出した物体領域と重なりがあるか判定し、物体領域が動領域を含んでいれば「動人物」とする。図6に「動人物」と判定される例を示す。

これに対し、物体領域において動領域がない場合、その領域は「静人物」、または「物体」の二つの可能性がある。これらの判定には、オプティカルフローを利用することにする。

3. 2 「静人物」の判定

「静人物」領域は、「動人物」領域にオプティカルフローを計算するためのフロー点を配置することにより判定する。ここで、オプティカルフローの計算にはLucas & Kanadeの手法^[5]を用いる。

「動人物」が前フレームから引き続き動いている場合、前フレームの「動人物」と判定された領域の周辺に動領域が検出されるため、フロー点はその動領域に移動する。ここで、検出された動領域にフロー点の移動が検出された時、その領域を引き続き「動人物」と判断し、前フレームのフロー点を消去した後、新たに検出された「動人物」領域にフロー点を配置する。

「動人物」領域の周辺に動領域が検出されない場合、フロー点は移動する動領域がなく、その場に留まる。この物体領域のうち動領域がなくフロー点が留まっている領域を「静人物」と判定する。図7に「動人物」が静止し「静人物」となるシーンの判定例を示す。時刻 $t=1, 2$ においては、人が動いているため、フレーム間差分により

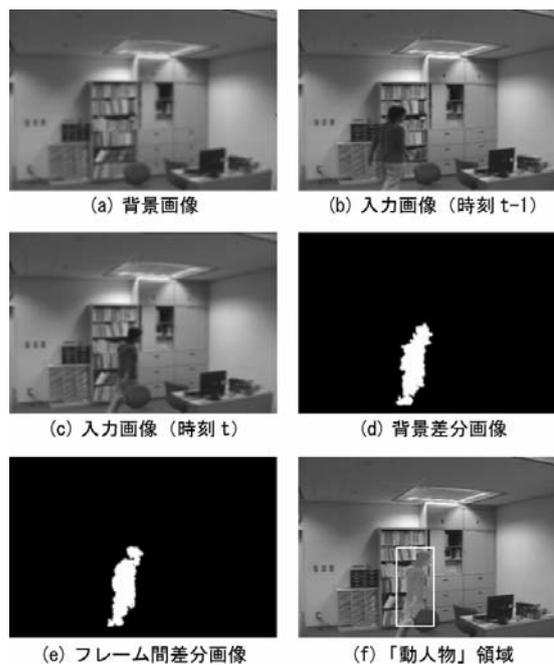


図6 「動人物」領域の判定例

動領域が検出されており、フロー点もその領域へ移動する。時刻 $t=3$ において人が静止したため、周囲に動領域が検出されずフロー点が留まり、「静人物」と判定されている。

3. 3 「物体」の判定

本研究では、物は人により動かされると仮定しているため、人と一緒に動いている物は「動人物」領域に含まれる。しかし、人が物から離れると同時に静止状態の「物体」となる。

そこで、人が物を置いて離れる時の人の動きに注目する。つまり、物体領域が人と物とに分裂する際、動いている領域を「動人物」、動きのない領域を「物体」と判定する。図8に人が物の位置を変え、離れるシーンにおいて、物が「物体」と判定される例を示す。 $t=1, 2$ までは、人と物は共に動いており「動人物」と判定されている。 $t=3$ において領域が分裂する際、動きのある領域は「動人物」、動きのない領域は「物体」と判定されている。フロー点が移動する動領域がある場合、前フレームでのフロー点を消去して、新たに検出された動領域にフロー点を配置するため、フロー点は動領域に移り、「静人物」とはならず、「物体」と判定することができる。

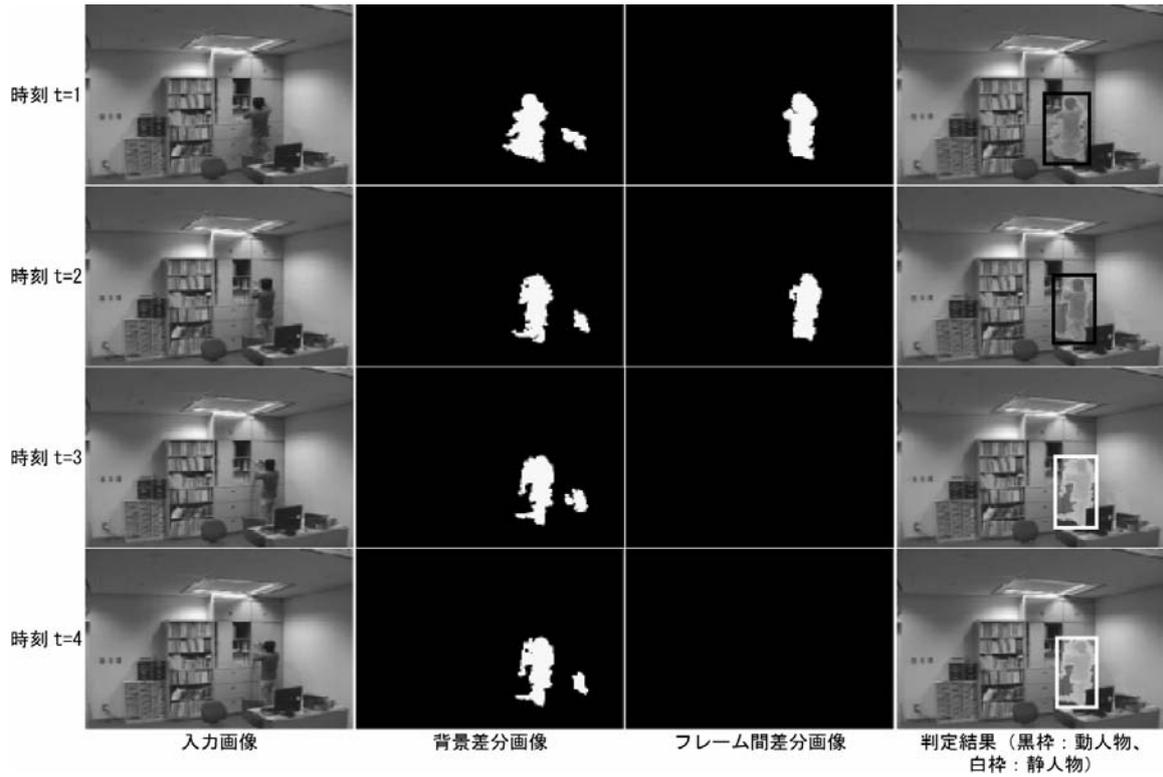


図7 「静人物」領域の判定例

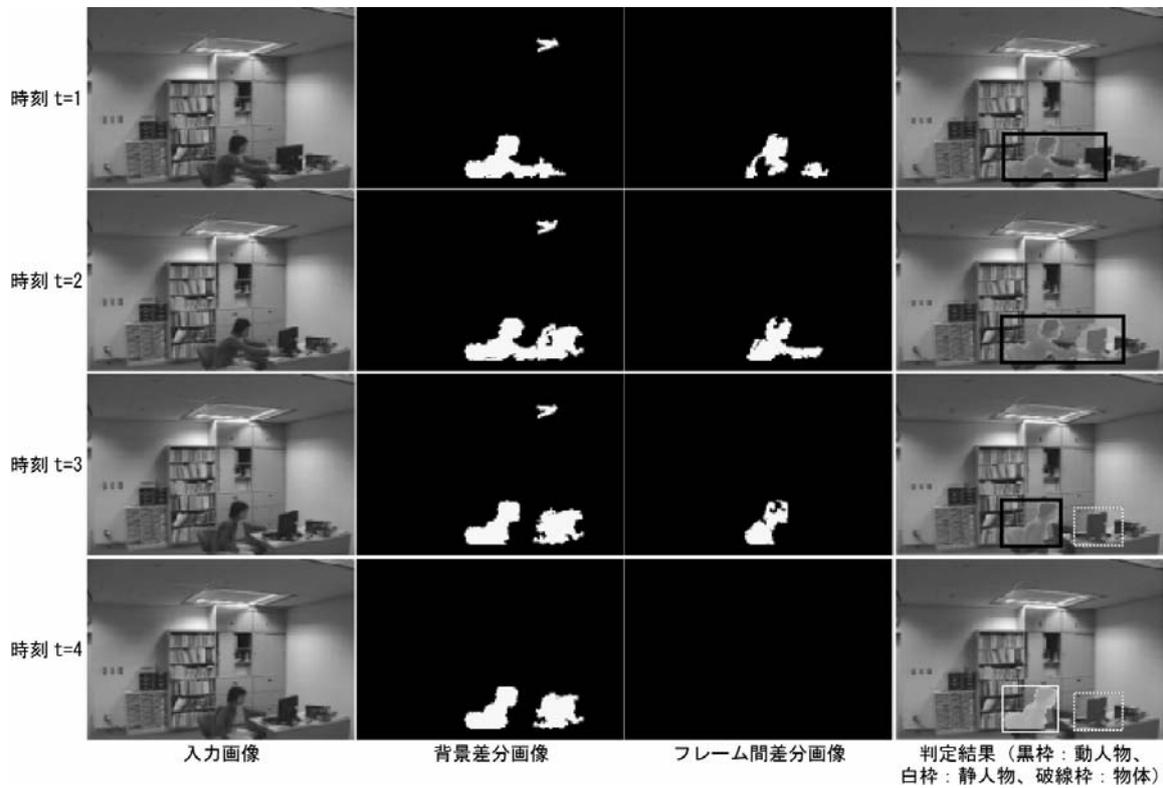


図8 「物体」領域の判定例

4. まとめ

ネットワークカメラを用いた監視カメラシステムにおける人と物との判定機能を実現するために、物体の検出精度を向上させ、誤検出を低減することが人と物とを精度よく判定するための出発点になるとの考えから、高精度な物体検出手法について検討した。さらに、精度よく検出された物体に対し動き情報を用いることで、人と物とを判定する手法を提案した。これにより、人と人が運んだ（動かした）物とを区別することや、静止した人を検出することが可能となった。

本研究では、人と物を、人は自律的に動き、物は人により動かされ、通常は動かないと仮定した。そのため、ボールやカーテンなど監視対象エリア内で動いている物は人と判定される。このようなシーンにおいても人と物とを判定するためには、人らしい動き、形状などについて検討する必要がある。

謝 辞

本研究は財団法人ソフトピアジャパン「平成18年度共同研究開発事業(共同研究制度)」の一環としてなされた。ここに記して謝辞とする。

文 献

- [1] 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.
- [2] 山田浩正, 伊藤渡, 上田博唯, “背景差分法における波の誤検出抑制法の検討”, 信学技報, PRMU98-109, 1998.
- [3] 影広達彦, 太田友一, “動画像からの背景画像の自動生成と適応的更新”, 画像の認識・理解シンポジウム(MIRU'94)講演論文集II, pp.263-270, 1994.
- [4] Intel Corporation, “Open Source Computer Vision Library Reference Manuals.”
- [5] Bruce D. Lucas and Takeo Kanade, “An Iterative Image Registration Technique with an Application to Stereo Vision,” Proc. of 7th International Joint Conference on Artificial Intelligence (IJCAI), pp.674-679.