

照明変化に頑健な背景差分

松山 隆司[†] 和田 俊和[†] 波部 斉^{†*} 棚橋 和也^{†**}

Background Subtraction under Varying Illumination

Takashi MATSUYAMA[†], Toshikazu WADA[†], Hitoshi HABE^{†*}, and Kazuya TANAHASHI^{†**}

あらまし 背景差分は、画像中の移動対象を検出する有効な手法として広く利用されている。しかし、背景差分を行なうには、背景部分での画像の変化が観測されないという前提条件が必要であるため、その適用範囲は限られている。本論文では、照明変化による背景部分の画像の変化が起きた場合にも適用可能な背景差分による移動対象検出法を提案する。本手法は、異なった考え方に基づく二つの対象検出法に基づいている。一方は、照明に不変な特徴を用いて背景画像と観測画像の比較を行なう手法である。他方は、背景差分を行なう前に観測画像の照明条件を推定し、輝度の正規化を行なう手法である。これら二つの手法は互いに相補的であり、最終的に双方の検出結果を統合することで精度の高い検出結果を得ることができる。実験の結果、実画像に対する本手法の有効性が示された。

キーワード 背景差分, 対象検出, 照明変化, 正規化距離, 固有画像解析

1. はじめに

背景差分は、画像中の移動対象を検出する有効な手法として利用されている。しかし、背景部分での画像の変化が観測されないという前提条件が必要であるため、その適用範囲は限られている。

一般に、画像中の背景シーンの変化は (1) カメラの動き (2) 照明変化 (3) 背景中の物体の動き (木の揺れや CRT のフリッカなど) によって引き起こされる。

カメラの動きに対しては、我々は視点固定型パン・チルト・ズームカメラ [1] [2] を提案した。このカメラを用いると、回転・ズームによって視点の移動が起きないため、視線・ズームのコントロールを行ないながら背景差分が行える。

照明変化や背景中の物体の動きに対しては、多くの手法が提案されている。[3]–[5] は、背景画像の画素値の変化を確率分布によってモデル化し、移動対象に相当する画素を検出している。[7]–[8] は、背景中の物体の見え方の変化や照明変化に対して適応的に背景画

像を再生成している。近年では、Toyama らがマルチレベルの背景差分法を提案している [9]。この手法では、空間方向と時間方向の拘束条件を統合し、Wiener フィルタによる画素毎の背景差分の精度向上を実現している。

本論文では、照明変化に対して頑健な背景差分法を提案する。前提条件として、背景シーン中の物体はすべて静止しているものとするが [9] で示されているような、空間的、時間的連続性を利用した背景差分法と組み合わせることによって、幅広い背景変化に対応した手法が実現できるものと考えられる。

本論文では、まず、異なった考え方に基づく二つの検出法を導入する。一方は、照明に不変な特徴を用いて背景画像と観測画像の比較を行なうものである。他方は、背景差分を行なう前に観測画像の照明条件を推定し、輝度の正規化を行なうものである。これら二つの手法は互いに相補的であり、最終的に双方の検出結果を統合することにより精度の高い検出結果を得る方法を提案する。最後に性能評価実験により、照明変化に対する本手法の有効性を示す。

2. 照明に不変な特徴を用いた背景差分

2.1 正規化距離

正規化距離 [6] (Normalized Vector Distance; N-

[†] 京都大学大学院情報学研究所, 京都市左京区
Graduate School of Informatics, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto, 606-8501 Japan

* 三菱電機 (株)

** (株)NTT データ

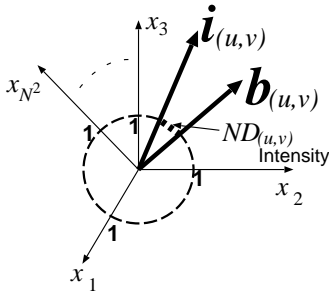


図1 正規化距離

Fig.1 Normalized Vector Distance

VD)は、照明変化の影響を受けにくい特徴である。正規化距離を計算するために、まず、画像を $N \times N$ 画素のブロックに分割し、そのブロックを N^2 次元のベクトルで表現する。ここで、ベクトルの要素はそれぞれブロック中の輝度値に対応している。 $i(u,v)$ と $b(u,v)$ をそれぞれ観測画像と背景画像中の同一位置のブロックに対応するベクトルとする。以下、このベクトルを画像ベクトルと呼ぶ。すると、正規化距離は次式で表される(図1)。

$$ND(i(u,v)) = \left| \frac{i(u,v)}{|i(u,v)|} - \frac{b(u,v)}{|b(u,v)|} \right|, \quad (1)$$

但し、 $|\cdot|$ はベクトルの大きさを表している。

定義から明らかなように、 $ND(i(u,v))$ は入力画像ベクトルの大きさ、すなわちブロック内での明度の一様な変化の影響を受けない。実際、正規化距離と正規化相互相関は以下の関係式を満たす。

$$ND(i(u,v)) = \sqrt{2(1 - \cos \theta)}, \quad (2)$$

但し、 θ はベクトル $i(u,v)$ と $b(u,v)$ の間の角度を表す。ブロック間で求められる正規化相互相関は $\cos \theta$ に他ならない。

しかし、一般に $i(u,v)$ にはノイズが含まれている。正規化距離は比の計算を行なうため、入力画像ベクトルの大きさが小さいときにはノイズの影響が大きくなり、正規化距離の値は不安定になる。

この問題に対して、本手法では以下のように対処する。

(1) 正規化距離の統計的性質の解析に基づき、移動対象検出時の閾値を明度変化に対して適応的に変化させる。

(2) ブロック内変化の空間的な性質を評価することで正規化距離を拡張する。

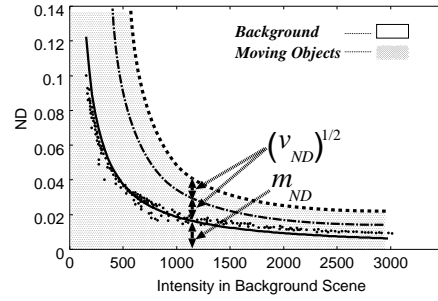


図2 正規化距離の統計的性質に基づいた閾値の決定

Fig.2 Adaptive threshold determination based on the statistical properties of NVD.

2.2 適応的な閾値の決定

観測画像中で、背景シーンに対応する画像ベクトルを $i(u,v)$ と表し、ノイズ n の影響は以下のように加法的であると仮定する。

$$i(u,v) = \widetilde{i(u,v)} + n, \quad (3)$$

但し、 $\widetilde{i(u,v)} = \alpha b(u,v)$ である。 α は照明変化によるブロック内での一様な明度変化を表すパラメータである。 n の各要素は、平均 0、標準偏差 σ の互いに独立な正規分布に従うとする。

すると、正規化距離に関して、以下に示す定理を導出することができる(付録1. 参照)。

[定理 1]

$ND(i(u,v))$ の平均 m_{ND} 、分散 v_{ND} は、 $|i(u,v)|$ 、 σ 、 N を用いて、以下のように近似できる。

$$\begin{cases} m_{ND}(i(u,v)) = \frac{\Gamma(\frac{N^2}{2})}{\Gamma(\frac{N^2-1}{2})} \frac{\sqrt{2}\sigma}{|i(u,v)|}, \\ v_{ND}(i(u,v)) = \left[\frac{\Gamma(\frac{N^2+1}{2})}{\Gamma(\frac{N^2-1}{2})} - \frac{\Gamma(\frac{N^2}{2})^2}{\Gamma(\frac{N^2-1}{2})^2} \right] \frac{2\sigma^2}{|i(u,v)|^2}, \end{cases} \quad (4)$$

但し、 $\Gamma(\cdot)$ はガンマ関数である。

この定理は、理想的な照明変化とノイズが観測される条件では「正規化距離に対するノイズの影響は明度 $|i(u,v)|$ のみによって決まる」ということを示している。したがって、以降に述べるように、背景の明度 $|i(u,v)|$ が推定できれば、ノイズの影響を考慮した精度の高い検出を実現できる。この明度の推定手法については次節で述べる。以降の議論では、推定によって観測画像の各ブロックにおける $m_{ND}(i(u,v))$ と $v_{ND}(i(u,v))$ が既知であるとする。

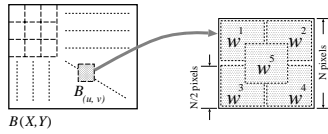


図3 画像ブロック中の小ブロック
Fig. 3 Small windows in an image block

図2はこの定理を用いた閾値の決定法を示している．図中で、横軸は $|\widehat{i}_{(u,v)}^B| = |\alpha b_{(u,v)}|$ を表し、縦軸は $ND(i_{(u,v)})$ を示す．図中下部の実線は式(4)による平均値 m_{ND} を示し、その付近の点は実画像から計算された平均値を示している．計算では、ブロックの大きさは経験的に決め、 $N = 16$ としている．この結果から定理の妥当性が確認できる．中央の一点鎖線は $m_{ND} + \sqrt{v_{ND}}$ を示し、太い点線は $m_{ND} + 2\sqrt{v_{ND}}$ を示す．図は $m_{ND} + 2\sqrt{v_{ND}}$ を閾値として移動対象を検出している様子を示している．このように、各ブロックにおける閾値をブロックの明度 $|\widehat{i}_{(u,v)}^B|$ に応じて変化させることにより、適応的な背景差分が実現できる^(注1)．

2.3 空間的特徴との統合

図2に示したように、適応的に閾値を変化させることによってノイズによる誤検出を低減させることができる．しかし、この処理は暗い領域における感度を低下させていることに他ならない．これはブロック相関だけでなく、色度差など、比の計算を用いた画像処理において一般的な問題であり、この解決のためには輝度値以外の情報を加味しなければならない．

本研究では、ブロック内の輝度値変化の空間的な特徴を考慮することで、正規化距離による検出精度の向上を図る．すなわち、ブロック内での輝度値変化の空間的な特徴に基づいて、ノイズによる変化と移動対象による変化の識別を行なう．

ブロック内での輝度値変化の空間的な構造を特徴づけるため、以下の仮定を導入する．

[仮定1] ノイズによる輝度変化はブロック内に独立かつ一様に分布する．一方、移動対象による輝度変化はブロック内の特定の領域に集中して分布する．

この仮定に基づき、以下のような評価尺度を定義する．まず、図3に示すように、背景画像と観測画像の同一位置のブロック $B_{(u,v)}$, $I_{(u,v)}$ をそれぞれ小ブロックに分割する． m を一つのブロック中の小ブロックの

(注1): ノイズ成分の標準偏差 σ は撮像系ごとにあらかじめ求めておく．

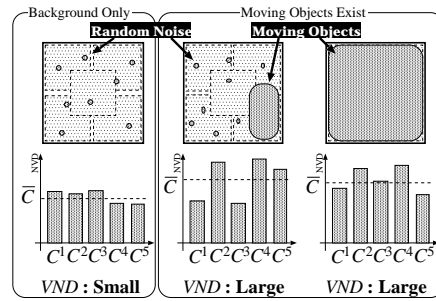


図4 VND の値とブロック内の空間的特徴との関係
Fig. 4 Spatial configurations in a block and corresponding VND values

数とする(図3では $m = 5$ としている)． $w_{B_{(u,v)}}^j$ と $w_{I_{(u,v)}}^j$ はそれぞれ $B_{(u,v)}$, $I_{(u,v)}$ 内の j 番目の小ブロックを表すものとする．

このとき、ブロック内の小ブロックについて計算された正規化距離の分散は次式で表される．

$$VND(i_{(u,v)}) = \frac{1}{m} \sum_{j=1}^m \left(C_{(u,v)}^j - \overline{C_{(u,v)}} \right)^2, \quad (5)$$

但し、 $C_{(u,v)}^j$ を小ブロック $w_{B_{(u,v)}}^j$ と $w_{I_{(u,v)}}^j$ の間の正規化距離とし、 $\overline{C_{(u,v)}} = \frac{1}{m} \sum_{j=1}^m C_{(u,v)}^j$ をブロック内の平均値とする．

$VND(i_{(u,v)})$ を用いることにより、ブロック内の変化の空間的な特徴を解析することができる．図4に示すように、ブロック内に現れる変化の空間的な特徴は3種類に分類できる．

(1) 背景: ノイズが含まれない理想的な状況では、それぞれの $C_{(u,v)}^j$ が0となり、 $VND(i_{(u,v)}) = 0$ となる．照明が暗く、観測画像のノイズの影響が大きいとき、 $C_{(u,v)}^j$ は、非零の値をとる．しかし、ノイズは一樣に分布しているため、その値は互いにほぼ等しくなる．したがって、その分散である $VND(i_{(u,v)})$ は照明条件によらず小さな値となる．

(2) 背景と移動対象が混在: 移動対象が存在する小ブロックに対応する $C_{(u,v)}^j$ は大きな値となり、それ以外は小さな値となる．よって、その分散 $VND(i_{(u,v)})$ は大きな値となる．

(3) 移動対象: 移動対象が背景と同じテクスチャを持っていない限り、それぞれの $C_{(u,v)}^j$ はランダムな大きな値をとる．よって、 $VND(i_{(u,v)})$ は大きくなる．

以上のように、 $VND(i_{(u,v)})$ の大小によって移動対象の有無を識別できる．

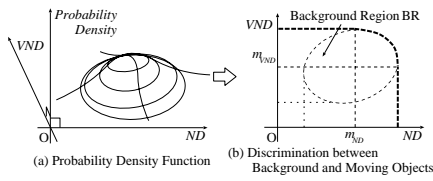


図 5 正規化距離に基づく背景差分
Fig. 5 Object detection based on NVD

2.4 正規化距離に基づく背景差分

ここまでの議論で、照明変化に不変な二つの特徴量、 $ND(i_{(u,v)})$ と $VND(i_{(u,v)})$ が得られた。これらの特徴量を統合し、移動対象検出を行なうために以下の仮定を導入する。

[仮定 2] $ND(i_{(u,v)}^B)$ と $VND(i_{(u,v)}^B)$ はそれぞれ正規分布 (平均: $m_{ND}(i_{(u,v)}^B)$, 分散: $v_{ND}(i_{(u,v)}^B)$) および、正規分布 (平均: $m_{VND}(i_{(u,v)}^B)$, 分散: $v_{VND}(i_{(u,v)}^B)$) に従う。

この仮定によると、2次元ベクトル $(ND(i_{(u,v)}^B), VND(i_{(u,v)}^B))$ は、2次元の正規分布に従う (図 5 (a))。その正規分布は $|i_{(u,v)}^B|$ 、すなわち、背景シーンの照明の強さによって決定される。

図 5 (b) は移動対象と背景の識別境界を示している。ブロックについて計算された $(ND(i_{(u,v)}^B), VND(i_{(u,v)}^B))$ が網かけ領域 BR の中にあれば、背景であり、それ以外の場合は、ブロックには移動対象が含まれていると判断される。 BR は以下のように定義される。

$$\begin{aligned} lk(ND(i_{(u,v)}^B), VND(i_{(u,v)}^B)) &> TH1 \quad (6) \\ \text{or } ND(i_{(u,v)}^B) &< m_{ND}(i_{(u,v)}^B) \quad (7) \\ \text{or } VND(i_{(u,v)}^B) &< m_{VND}(i_{(u,v)}^B) \quad (8) \end{aligned}$$

ここで、関数 lk は 2次元の正規分布に基づく尤度を示し、 $TH1$ は $\sqrt{v_{ND}(i_{(u,v)}^B)}$ によって決まる閾値を示している。

2.5 性能評価

ここまでに述べてきた手法の有効性を検証するため、計算機実験を行なった。結果を図 6 に示す。実験で用いた画像は、照度、点灯する組合せを制御できる蛍光灯が設置された屋内で、固定カメラを用いて撮影した濃淡画像である。ブロックサイズ 16×16 とした。

実験では、下記の 3 つの手法による対象検出結果の比較を行なった。

・ R_{ND} : ノイズによる正規化距離の変動を考慮せず

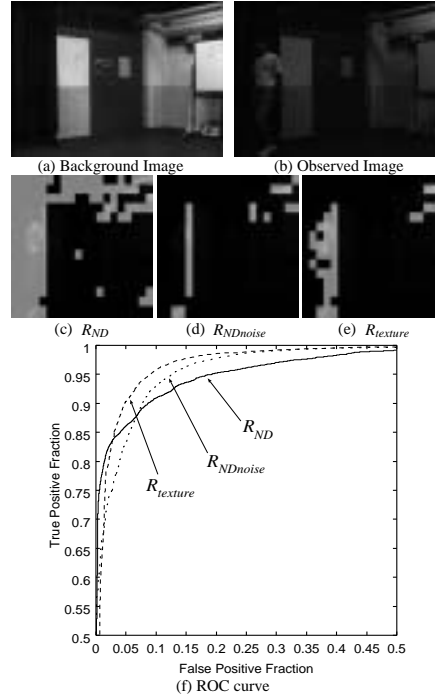


図 6 正規化距離に基づく手法の性能評価

Fig. 6 Performance evaluation of the NVD-based method

$ND(i_{(u,v)})$ に対する固定閾値を用いた結果。

・ $R_{NDnoise}$: 照明変化は簡略線形モデル^(注2)に従うとして背景シーンの照明変化の推定処理を行ない、ノイズによる正規化距離の変動を考慮に入れて $ND(i_{(u,v)})$ に対する閾値を決定する。

・ $R_{texture}$: 2.4節で提案した手法。ここでは上記と同じ簡略線形モデルを用いて $|i_{(u,v)}^B|$ を推定した結果を用いる。

図 6 の中で、(a) が背景画像、(b) が暗い照明下での観測画像の例、(c),(d),(e) がそれぞれ (b) を処理した結果 R_{ND} , $R_{NDnoise}$, $R_{texture}$ を示している。これらの検出結果の差が顕著に現れているのは、黒い天井やカーテンなどの暗い領域である。 R_{ND} ではノイズの影響で誤検出が目立っており、 $R_{NDnoise}$ では検出感度を落すように閾値を設定している (図 2) ため検出漏れが目立っている。これに対して、 $R_{texture}$ では空間的特徴を加味した評価を行なっているため、検出精度が向上している。

(注 2): $|i_{(u,v)}^B| = \alpha_{(u,v)} |b_{(u,v)}^B|$ 中の照明変化は $\alpha_{(u,v)} = k_1 u + k_2 v + k_3$ で表され、観測画像と推定された画像との間で輝度の差が最小になるように k_i ($i = 1, 2, 3$) を決定する。

これら 3 つの結果を定量的に比較するために次のようにして得た ROC 曲線を図 6 (f) に示す．まず，評価に先立って，映像中の移動対象領域を正確に与えておく．ある閾値において，各フレームでの検出率を映像全体にわたって平均したものを記録する．その上で，閾値をさまざまに変化させながら検出率の変化をプロットしていくことで ROC 曲線が得られる．ROC 曲線では縦軸が対象を正しく検出した割合 (True Positive)，横軸が背景を誤って対象として検出した割合 (False Positive) を示す．図 6 (f) の ROC 曲線は照明変動下における平均検出率の変化を示しており，これをみても提案手法の検出精度が他の手法よりも高いことが分かる．

他の研究 [9] においても， $R_{NDnoise}$ を単純化したものが提案されている．そこでは，照明変化に対する頑健性が示されている．このことは，図 6 に示されている結果と一致する．

3. 照明条件の推定に基づく背景差分

2. 章で述べた検出手法を用いれば，照明変化に不変な対象検出が行なえるが，

- ノイズの影響に適応的に検出を行なうためには，背景シーンの照明の強さ $|i_{(u,v)}^B|$ を推定しなければならない．

- 背景と対象の両者が同じテクスチャをもっている場合や，ともにテクスチャのない様な輝度分布をもっている場合に移動対象を検出することができないという問題点がある．

この問題に対し，本章で述べる手法では，

- (1) 観測画像の照明条件を推定する．
- (2) 推定結果を用いて，輝度値を照明変化に対して正規化し，背景差分を行なう．

の順序で移動対象の検出処理を行なう．ここで，(1) の推定処理によって，前章で述べた正規化距離に基づく移動対象検出に必要な情報を得ることができる．さらに，正規化距離を用いた手法では，背景と対象が同一テクスチャの場合に対象の存在と照明変化を識別できないが，本章で述べる手法では背景シーンの明度を推定するため，そのような場合でも対象検出が可能である．

一方，照明条件の推定を行なうためには，以下に述べるように，正規化距離に基づく対象検出が重要な位置を占める．このように，本論文で述べる二つの手法は相補的な働きをし，互いの精度を高めている．



図 7 照明変動下で撮影した画像

Fig. 7 Images taken under varying illumination.

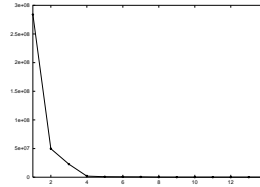


図 8 固有値 (縦軸: 大きさ, 横軸: 固有値の次数)

Fig. 8 Eigen values (vertical axis: magnitude, horizontal axis: index of eigen values)



図 9 固有画像

Fig. 9 Eigen images

3.1 照明変動下における輝度変化モデル

[10] では，照明変化による輝度値の変化を表す Illumination Cone モデルが提案されている．そこでは以下の仮定が導入されている．

[仮定 3]

- 物体の表面は完全拡散面である．
- 各物体は凸であり，影が生じない．
- すべての光源は無限遠点にある．

ベクトル $\{i_1, i_2, \dots, i_n\}$ が異なる照明下で撮影した画像を表すとす．Illumination Cone モデルによると，これらのベクトルは， M^2 次元空間中の高々 3 次元のコーン状の部分空間に分布する．この部分空間は 3 つの固有ベクトル $i_{eigen1}, i_{eigen2}, i_{eigen3}$ によって定義される．換言すれば，任意の照明条件下で撮影された画像に相当するベクトル i_{any} は次式で表される．

$$i_{any} = a_1 i_{eigen1} + a_2 i_{eigen2} + a_3 i_{eigen3}, \quad (9)$$

但し， $a_k = i_{any} \cdot i_{eigenk}$ ($k = 1, 2, 3$) である．

この照明変化モデルの有効性を確かめるために，屋内を広い画角で撮影した画像を用いた実験を行なった．これは，先に述べた仮定は，人の顔のような狭い領域に対しては成立するが，広い画角で撮影した画像の中

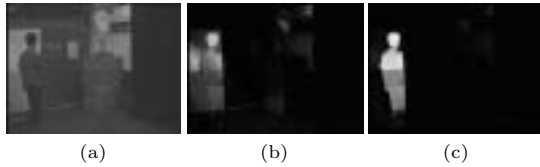


図 10 照明条件推定の安定性。(a) 移動対象を含む観測画像 (b) 観測画像の全画面を用いた推定の残差 $(a_1 \ a_2 \ a_3) = (9303.7 \ 1013.0 \ -1363.3)$ 平均残差 2.49 (c) 移動対象を除去して推定したときの残差 $(a_1 \ a_2 \ a_3) = (7961.1 \ 540.4 \ -434.9)$ 平均残差 1.23
Fig. 10 Stability of the illumination estimation

では必ずしも仮定が満たされるとは限らないからである。

室内のパノラマ画像を得るため、視点固定型パン・チルト・ズームカメラ [1] [2] を用いた。このカメラでは異なるパン・チルト角で撮影された画像を切れ目なく合成し、パノラマ画像を得ることができる。図 7 は照明の強さや点灯パターンを変えて撮影したパノラマ画像の例である。このシーンには表面が反射するホワイトボードや、マネキンや椅子の影、近くにある光源に照らされている物体などが存在し、仮定 1 が必ずしも成立しない環境である。

観測された画像ベクトル $\{i_1, i_2, \dots, i_n\}$ から主成分分析を用いて固有値、固有ベクトルを求める。その際には、まず平均ベクトル i_{avr} を各ベクトルから引き、 $\{i_1 - i_{avr}, i_2 - i_{avr}, \dots, i_n - i_{avr}\}$ の共分散行列の固有値、固有ベクトルを算出する。これは、実画像では画像ベクトルの分布の主成分が原点から離れていることが多いためである。

図 8 は、図 7 中の画像から計算された固有値を表している。必ずしも仮定が成立しない中で、3 つの固有値が際だって大きな値をとっている。図 9 はそれぞれの固有値に対応する固有画像を示す。これらの実験結果は実世界の屋内シーンに対しても Illumination Cone モデルが有効であることを示している。6. 章ではこのモデルの有効性について、より詳細な評価を行っている。

3.2 移動対象検出手法

これらの実験をふまえ、観測画像の照明条件を推定し、照明変化の影響を受けない背景差分を実現する。

(1) 背景モデルを構成するために、さまざまな照明条件のもとで背景画像を撮影する。

(2) それらの背景画像に対して主成分分析を行ない、固有ベクトル $i_{eigen1}, i_{eigen2}, i_{eigen3}$ を得る。

(3) 処理対象の観測画像 i に対して、観測時点での背景画像を得るための係数ベクトル $a = (a_1 \ a_2 \ a_3)^t$ は一般化逆行列を用いて、

$$a = (E^t E)^{-1} E^t (i - i_{avr}) \quad (10)$$

と求める。但し、行列 E は 3 つの固有ベクトルを用いて $E = [i_{eigen1} \ i_{eigen2} \ i_{eigen3}]$ で定義される。

(4) 推定された照明条件での画像ベクトルを以下のようにして得る。

$$\tilde{i} = a_1 i_{eigen1} + a_2 i_{eigen2} + a_3 i_{eigen3} + i_{avr}. \quad (11)$$

(5) 画素毎の差分を \tilde{i} と i の間で行ない、移動対象を検出する。

このアルゴリズムは、移動対象が画像中で占める大きさが十分小さいことを前提としている。同様の手法を用いた研究 [11] においても、屋外で小さな移動物体の検出を実現している。しかし、屋内では移動対象が画像中で大きな面積を占めることが多く、そのような場合には観測画像の照明条件の推定、すなわち a_k ($k = 1, 2, 3$) の導出は移動対象の影響を大きく受けてしまう。図 10 がこのことを裏付けている。図中では、(a) が移動対象 (人) を含む観測画像を示しており、(b) が観測画像全体を用いて照明条件を推定した結果の残差を示している。(c) は観測画像から移動対象を除いて照明条件の推定を行った結果の残差を示している (ここでの残差はそれぞれ 4 倍してある)。図の下に示している値は推定された係数 a_k ($k = 1, 2, 3$) と背景領域における観測画像と推定された画像との平均残差を示している。(b) では平均残差はそれほど大きくないが、局所的には 0 から 255 までの画素値で数十の差が生じており、これらが誤検出の原因となる。このように、精度よく照明条件の推定を行うためには移動対象を除外する必要がある。本手法ではそのために前節で述べた正規化距離に基づく検出手法を用いる。

4. 二つの手法を統合した背景差分法

ここまでに二つの検出手法が示したが、それらはともに問題を持っている。

正規化距離による検出：背景シーンの画像ブロックの明るさ $|i_{(u,v)}^B|$ が既知でなければならない。観測画像ブロック $I_{(u,v)}$ が移動対象で占められていても、そのブロック内の照明強度が事前に得ていた値よりも変化している可能性があるため、観測時点での (u, v) での照明変化を推定しなければならない。

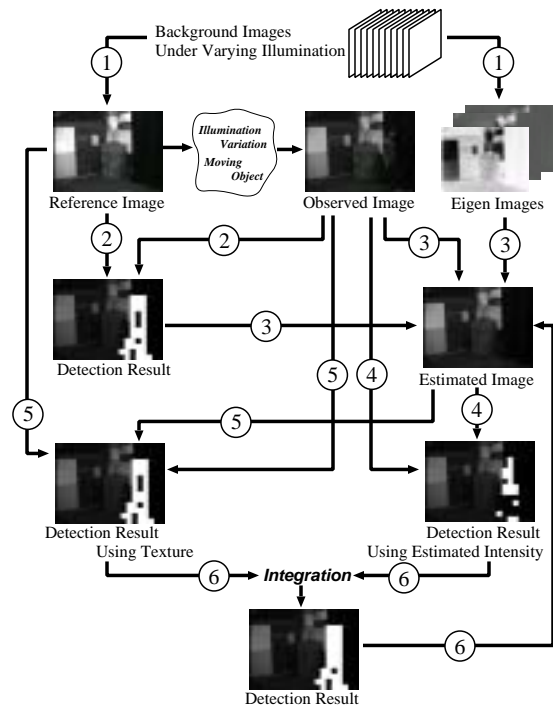


図 11 提案手法の概要

Fig. 11 Overview of Proposed Method

照明条件の推定による検出：より正確な照明条件推定を行うためには、移動対象に相当する領域を観測画像から除外した後で推定を行う必要がある。

これらの問題を解決するため、以下に示すように、二つの検出手法を反復的に実行する(図 11)。

ステップ(1) さまざまな照明条件のもとで背景画像を撮影し、固有画像 i_{eigen1} , i_{eigen2} , i_{eigen3} を得る。また、明るい照明下の背景画像から各画素のメディアンを計算し、メディアン画像 i_{median} を得る。

ステップ(2) 処理対象の観測画像を i とする。まず、 i_{median} を背景画像として正規化距離に基づく対象検出を行う。ここでは、 i の照明条件が未知であるので、 $|i_{(u,v)}^B| = |i_{(u,v)}|$ とする。ここで、 $|i_{(u,v)}|$ はブロック (u, v) での観測画像ベクトルの大きさを指す。 $|i_{(u,v)}^B|$ は閾値の決定のみに用いられるので、このように近似しても対象を大まかに検出することができる。

ステップ(3) i , i_{eigen1} , i_{eigen2} , i_{eigen3} から移動対象とされたブロックに含まれる画素を除く。その後、残された部分ベクトルに対して照明条件推定を行い、係数 a_1 , a_2 , a_3 を得る。この係数を用いて、観測時点の照明条件に正規化された背景画像 \tilde{i} を式(11)を用



図 12 背景画像

Fig. 12 Reference background image

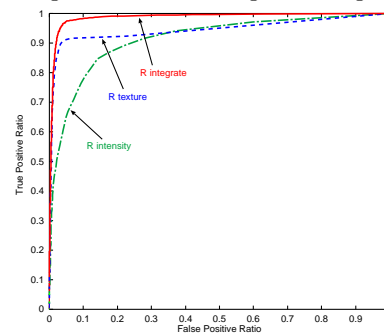


図 13 明るい照明下での検出性能

Fig. 13 Performance under light illumination

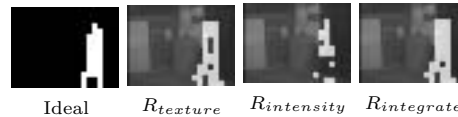


図 14 明るい照明下での検出結果

Fig. 14 Detected foreground objects under light illumination

いて算出する。

ステップ(4) 観測画像と推定結果の背景画像の間で輝度の差分をとり、ブロック単位で移動対象の検出を行う。ここで、ブロック (u, v) は条件 $||\tilde{i}_{(u,v)}| - |i_{(u,v)}|| > TH2$ を満たすときに移動対象が存在するとされる。ここで、 $TH2$ は推定処理に含まれる誤差に基づいて決定される閾値である。

ステップ(5) 一方、推定結果を用いて、 $|i_{(u,v)}^B| = |i_{(u,v)}|$ とすることで、正規化距離に基づく対象検出を再び行うことができる。

ステップ(6) ステップ(4)と(5)で得られた二つの検出結果に対し、ブロック毎に OR 演算を施し、最終的な結果を得る。この結果を用いて、ステップ(3)の推定処理を再度実行する。推定結果の変化が少なくなり、検出結果が安定するまで、ここまでの処理を繰り返していく。

5. 実験

図 12 は背景画像として用いるメディアン画像 i_{median} を示している。実験では、サイズ 320×240 , 256 階調の濃淡画像を用いている。ブロックの大きさは 16×16

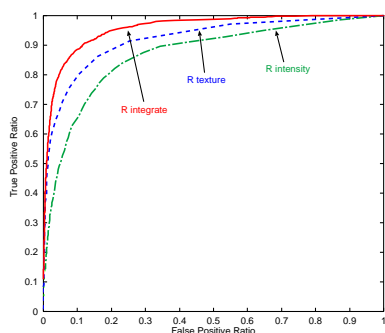


図 15 暗い照明下での検出性能

Fig. 15 Performance under dark illumination

Ideal $R_{texture}$ $R_{intensity}$ $R_{integrate}$

図 16 暗い照明下での検出結果

Fig. 16 Detected foreground objects under dark illumination

とした．実験環境では，光量が連続的に変化でき，点灯パターンも選択できる天井照明が設置されている．その中で撮影した，移動対象として人を含む画像を用いて実験を行った．

性能評価のための比較した手法は以下の 3 つである．

$R_{texture}$: 4 章のステップ 2 で述べた正規化距離に基づく検出結果である．

$R_{intensity}$: 3.2 章で述べた照明条件推定に基づく検出結果である．ここでは画像全体を推定に用いている．

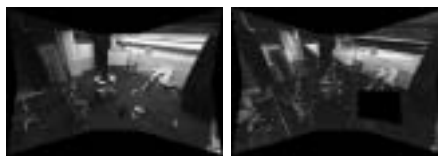
$R_{integrate}$: 4 章で述べた 2 つの手法の統合による検出結果である．ここではステップ (3) - (6) を 1 回のみ行っている^(注 3)．

なお， $R_{intensity}$ ， $R_{integrate}$ を得るために必要な固有画像は，異なった照明条件下での 13 枚の背景画像から算出した．また， $R_{integrate}$ を得るための処理速度は約 3 枚/秒である^(注 4)．

図 13 は照明が明るい下での映像に対してえられた ROC 曲線を示している．図 14 はあるフレームにおける検出結果を示している．ここでは，(a) が移動対象の正解領域，(b) が $R_{texture}$ ，(c) が $R_{intensity}$ ，(d) が $R_{integrate}$ を示す．それぞれの結果は“最適”な閾値によるものである．ここで“最適”とは，ROC 曲線

(注 3): この手法では 2 つの閾値 $TH1, TH2$ が存在するため，まず $TH2$ 一定で $TH1$ を変化させて対象を正しく検出する率が最も高くなる点を求める．ついで， $TH2$ を変化させて再び検出率が最高となる点を求める．このような点をつなぐことで，ROC 曲線を得る．

(注 4): PentiumII400MHz × 2 の PC



Observed image Error image

図 17 局所的な照明条件の変化

Fig. 17 Variations of local illumination conditions

(True Positive, False Positive)=(1, 0) からの距離を重み付けして計算し，それが最小となった閾値を指す．ここで，重みとして True Positive:False Positive = 3:1 としている^(注 5)．図 15 と図 16 は暗い照明下での実験結果を同様に示している．

二つの結果ともに，提案手法による結果 $R_{integrate}$ が他の結果 $R_{texture}, R_{intensity}$ に比べて際立って良いことが分かる． $R_{intensity}$ の検出率が悪いのは，移動対象が大きな領域を占め，推定の精度が低下しているためである．照明が明るい場合と，暗い場合の結果を比べると，前者が良いことが分かる．これは，照明が明るい方が観測画像の SN 比が高いためである．

6. おわりに

本論文では，照明変化に対して頑健な背景差分を用いた移動対象検出法を提案した．まず，二つの検出手法について述べた．一方は入力画像と背景画像の対応する画像ブロックに対して定義される正規化距離に基づくものであり，他方は固有画像解析を用いた照明条件の推定に基づくものである．これらの手法を統合することにより，照明変化に対して頑健な移動対象検出を実現し，実験により提案手法の有効性を示した．

本手法を拡張し，アクティブカメラを用いてより広いシーンでの対象検出を行なう場合には，照明条件の推定方法を変更する必要がある．これは，実世界では物体の幾何的な構造や照明条件によって，さまざまな局所的な照明変化が観測されるからである．このことを確かめるため，以下の 2 つの実験を行った．

図 17(a) はある照明条件におけるパノラマ画像を示している．この画像の中から，4 つの折れ線で角を示した四角形領域をとりだし，照明変化を推定して係数 a_k ($k = 1, 2, 3$) を得る．これから，式 (11) を用いて推定した照明条件での画像全体を生成する．この画像

(注 5): 背景を誤って対象としてしまう割合を対象の検出漏れの 1/3 にするという意図による



図 18 各ブロックにおいて支配的な固有値の数
Fig. 18 Number of dominant eigen values for a block

と観測画像 (a) の差を (b) に示す。四角形領域内では差が少なく、局所的な照明変化は正しく推定できているが、それ以外の領域では誤差が大きくなっている。

次に、ブロック単位での固有画像解析の結果を示す。ここでは、パノラマ画像中央の長方形領域を 15×7 のブロックに分け、固有ベクトルを求めた。図 18 は各ブロックの次元、つまり、支配的な固有値の個数を示している。これを見ると、次元がブロック内に存在する物体表面の幾何的特徴をよく表していることが分かる。つまり、1次元は平面（床や壁）に相当し、2次元は円柱状（椅子の脚や壁の端）、3次元は局面（マネキンなどの複雑な表面）に相当する。この結果はブロック毎の固有画像解析を用いて、3次元空間での物体表面の種類を分類し、照明変化のもとでの局所的な輝度値変化をモデル化できることを示している。

現在、我々は視点固定型カメラを用いて動的に背景差分を行なうシステムの研究を行なっている。そこでは、ブロック毎の固有画像解析を行なって局所的な照明条件変化をモデル化し、同時に物体の3次元特徴を得ることができる。

また、色情報を用いることで、提案手法の精度向上が期待できる。また [9] のように、空間的、時間的な拘束条件を用いることで、動的に変化する物体が背景に存在する場合でもそれらの変化を解析することができる。具体例として、時間的な相関関係に基づき、背景物体の変化をモデル化する手法 [12] を提案している。

文 献

- [1] T. Wada and T. Matsuyama: Appearance Sphere: Background Model for Pan-Tilt-Zoom Camera, Proc. of ICPR, Vol. A, pp. 718-722, 1996.
- [2] T. Matsuyama: Cooperative Distributed Vision – Dynamic Integration of Visual Perception, Action, and Communication –, Proc. of Image Understanding Workshop, pp. 365-384, 1998
- [3] 中井 宏章: “事後確率を用いた移動物体検出手法”, 情処研報, SIG-CV90-1, 1994
- [4] L.Davis: Visual Surveillance and Monitoring, Proc.

of VSAM Workshop, 1997.11

- [5] W.E.L. Grimson, C. Stauffer, R. Romano and L. Lee: Using adaptive tracking to classify and monitor activities in site, Proc. CVPR, pp. 22-29, 1998
- [6] 長屋 茂喜, 宮武 孝文, 藤田 武洋, 伊藤 渡, 上田 博唯: “時間相関型背景判定法による移動物体検出”, 信学論 D-II, Vol. J79-D-II, No. 4, pp. 568-576, 1996
- [7] 影広 達彦, 大田 友一: “動画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, Vol. II, pp. 263-270, 1994
- [8] 高藤 政雄, 北村 忠明, 小林 芳樹: “空間微分および差分処理を用いた車両抽出法”, 信学論 D-II, Vol. J80-D-II, No. 11, pp.2976-2985, 1997
- [9] K. Toyama et al: Wallflower: Principles and Practice of Background Maintenance, Proc. of ICCV, pp. 255-261, 1999
- [10] P. N. Belhumeur and D. J. Kriegman: What is the set of images of an object under all possible lighting conditions?, Proc. CVPR, 1996, pp. 270-277.
- [11] N. Oliver, B. Rosario, and A. Pentland: A Bayesian Computer Vision System for Modeling Human Interactions, Proc. of Int'l Conf. on Vision Systems, Gran Canaria, Spain, 1999
- [12] T. Matsuyama, T. Ohya, and H. Habe: Background Subtraction for Non-Stationary Scenes, Proc. of ACCV, 1999.

付 録

1. 定理1の導出

雑音 n_i が加わった背景シーンのみの観測画像ブロック (大きさ $N \times N$) は次式で表せると仮定する。

$$\{i_{(u,v)}^B\}_i = \widetilde{i_{(u,v)}^B} + n_i \quad (\text{A.1})$$

以降、観測データの番号を表す i は省略する。ここで、 n の各要素 n_k は互いに独立に平均 0、標準偏差 σ の正規分布に従うと仮定しているため、 n が従う確率密度関数 $f_n(n)$ は、次式で表される。

$$f_n(n) = \frac{1}{(2\pi\sigma^2)^{\frac{N^2}{2}}} \exp\left(-\frac{1}{2\sigma^2} \sum_{k=1}^{N^2} n_k^2\right) \quad (\text{A.2})$$

ここで、図 A.1 (注6) に示す $ND_{(u,v)}$ の近似を用いる。すなわち、 $\widetilde{i_{(u,v)}^B}$ に垂直な平面 P を考え、雑音ベクトル n を P に正射影し、ベクトル n' を得る。P に平行な原点からの距離 1 がある平面 Q を考える。この平面 Q 上で、ベクトル $\widetilde{i_{(u,v)}^B} + n'$ の交点と、 $\widetilde{i_{(u,v)}^B}$ の交点との間の距離 ND' は

(注6): 実際には $N \times N$ 次元の空間であるが、簡単のため図では 2次元平面で記述している。

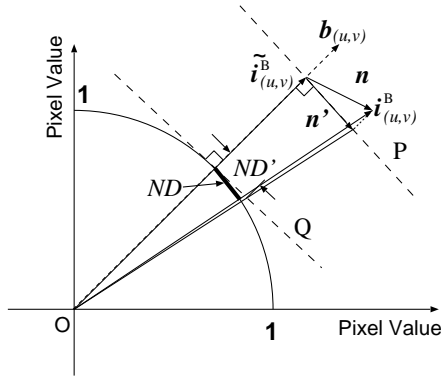


図 A-1 正規化距離 $ND_{(u,v)}$ に対する雑音の影響を求め
るための近似

$$ND' = \frac{1}{|\tilde{i}_{(u,v)}^B|} |n'| \quad (\text{A-3})$$

となる．これを正規化距離 ND の近似とし，以下の
手順で ND' が従う確率密度関数を求める．

- (1) $|n'|$ の従う確率密度関数の導出．
- (2) 正規化距離を近似する ND' への置換．

(1) 雑音ベクトルは等方性を持つことから， n を
 $n_{N^2} = 0$ の平面に射影したベクトルを n'' とし， n''
が従う確率密度関数を以下のように求める．

$$f_{n''}(n'') = \frac{1}{(2\pi\sigma^2)^{\frac{N^2-1}{2}}} \exp\left(-\frac{1}{2\sigma^2} \sum_{k=1}^{N^2-1} n_k^2\right).$$

$$|n''|^2 = \sum_{k=1}^{N^2-1} n_k^2 \text{ および } N^2 \text{ 次元空間での半径 } r \text{ の}$$

$N^2 - 1$ 次元球面の表面積が $\frac{2\pi^{\frac{N^2-1}{2}}}{\Gamma(\frac{N^2-1}{2})} r^{N^2-2}$ になる

ことを用いると， $|n''| = r$ として， $r (r > 0)$ が従う確
率密度関数は

$$f_r(r) = \frac{2r^{N^2-2}}{(2\sigma^2)^{\frac{N^2-1}{2}} \Gamma(\frac{N^2-1}{2})} \exp\left(-\frac{r^2}{2\sigma^2}\right)$$

となる．よって， $|n'|$ が従う確率密度関数は，

$$f_{n'}(|n'|) = \frac{2|n'|^{N^2-2}}{(2\sigma^2)^{\frac{N^2-1}{2}} \Gamma(\frac{N^2-1}{2})} \exp\left(-\frac{|n'|^2}{2\sigma^2}\right)$$

となる．

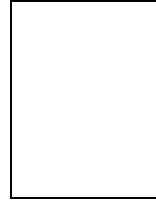
次に，求められた $f_{n'}(|n'|)$ に対して，式 A-3 の変
数変換を施すことによって， $d = ND' (d > 0)$ が従う
確率密度関数は，

$$f_{ND}(d) = \frac{2|\tilde{i}_{(u,v)}^B|^{N^2-1} d^{N^2-2}}{(2\sigma^2)^{\frac{N^2-1}{2}} \Gamma(\frac{N^2-1}{2})} \exp\left(-\frac{|\tilde{i}_{(u,v)}^B|^2 d^2}{2\sigma^2}\right)$$

となる．さらに， $|\tilde{i}_{(u,v)}^B| \gg |n|$ の範囲では， $|\tilde{i}_{(u,v)}^B| \simeq$
 $|\tilde{i}_{(u,v)}^B|$ と仮定することができる．この分布から，平均
値，分散を計算すると，定理 1 が得られる．

(平成 x 年 xx 月 xx 日受付)

松山 隆司 (正員)



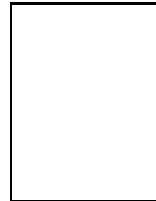
1976 年京大大学院修士課程修了．京大助
手，東北大助教授，岡山大教授を経て 1995
年より京大大学院電子通信工学専攻教授．
現在同大学院情報学研究科知能情報学専攻
教授．工博．画像理解，人工知能，分散協
調視覚の研究に従事．1980 年情報処理学
会創立 20 周年記念論文賞，1990 年人工知能学会論文賞，1993
年情報処理学会論文賞，1994 年電子情報通信学会論文賞，1995
年第 5 回国際コンピュータビジョン会議 Marr Prize，1996 年
国際パターン認識連合 Fellow，1999 年電子情報通信学会論文
賞，2000 年画像センシングシンポジウム優秀論文賞．人工知
能学会評議員，情報処理学会理事，同学会フロンティア領域委
員会委員長．

和田 俊和 (正員)



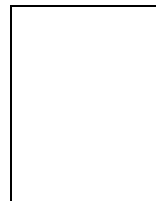
平成 2 年東工大大学院博士課程修了．同
年岡山大学工学部助手．平成 9 年京都大
学大学院工学研究科助教授．工学博士．画像
理解，パターン認識の研究に従事．平成 7
年 David Marr 賞．平成 9 年 情報処理学
会山下記念研究賞．平成 11 年 電子情報通
信学会論文賞各受賞．

波部 斉



1999 年京大大学院工学研究科修士課程
修了．同年三菱電機(株)入社．在学中に画
像理解の研究に従事．

棚橋 和也



2000 年京大大学院情報学研究科修士課
程修了．同年(株)NTT データ入社．在学
中に画像理解の研究に従事．