

室内における背景画像の推定と影の除去

森 田 順 也[†] 岩 井 儀 雄[†] 谷 内 田 正 彦[†]

画像中の物体抽出手法の1つである背景差分法では、対象物体の影や照明変化等による背景変化への対応が問題となることが多い。これらの問題に対する先行研究では、その多くが単一光源あるいは単一色光源を暗に仮定しているが、室内では様々な光源色が存在するため、複数光源色に対応した手法が必要となる。本研究では、画素値を各光源からのエネルギーの線形和で表す背景画像のモデル化を行い、室内の複数照明条件下における影の影響を抑えた物体抽出手法を提案する。また、光源の明るさ変化に対応するために、入力画像と背景画像の輝度の比率を最初に推定する。実画像を用いた実験により本手法の有効性を検証した。

Adaptive Background Estimation and Shadow Removal in Indoor Scenes

JUNYA MORITA,[†] YOSHIO IWAI[†] and MASAHICO YACHIDA[†]

Background subtraction algorithm is susceptible to both global and local illumination changes such as shadows, sunlight and reflection. These changes are sometimes the cause of failure of object tracking, gesture recognition and posture estimation. In this paper, we propose a method for detecting object casting a shadow in an indoor scene by modeling a pixel value as the total energy received from each lightsource. Our proposed method can also estimate a rate of illumination change in order to improve background subtraction. The usefulness of the method is demonstrated by experimental results.

1. はじめに

画像処理によって画像中の移動物体を検出する手法の1つに、あらかじめ撮影しておいた背景画像と入力画像との差分を計算し、差分値の大きい領域を物体として抽出する背景差分法がある^{1),2)}。この手法は、背景が変化しない場合には単純な処理で良い結果が得られることから広く利用されている。しかし、背景画像と異なる領域をすべて抽出するため、対象物と影を1つの領域と見なして抽出する場合が多く見られる。人物の姿勢推定や追跡、ジェスチャ認識等を行う際にその影が影響を及ぼすことがあり、前処理として対象物と影を分離することが重要となる。

また、照明の点灯/消灯等照明条件の変化により背景画像の変化が起きた場合には、背景差分法は適用できないという問題点もある。実画像に適用する場合には、これらの照明変化に対応することも重要である。

これらの問題に対して、従来から多くの手法が提案されている。屋外においては、対象物体の3次元形

状モデルに基づいた手法^{3),4)}や、天文学の知識から推定した太陽と対象物体の位置関係を利用して影を分離する手法^{5),6)}等がある。これに対して、そのような幾何学情報を利用しない手法として、Gordonら⁷⁾はステレオ視による距離情報と色情報を統合し、それぞれのアプローチの欠点を補完した物体検出手法を提案している。しかし、複数のカメラが必要であることや、低いコントラスト領域では距離情報が安定に求まらないという問題は残る。また、フレーム間差分を用いた手法⁸⁾もあるが、ヒューリスティックなルールを用いているため、そのルールの前提条件が成り立たない場合には対応できないと思われる。松山ら⁹⁾は、入力画像と背景画像の対応する画像ブロックごとに正規化距離を求め、照明変動環境下において安定に物体抽出を行っている。しかしながら、ブロック単位での処理のため物体抽出の空間分解能は低く、またブロック内での非一様な明度変化には対応できないという欠点がある。

一方、統計的なアプローチも数多く提案されている^{10)~13)}。Staufferら¹⁰⁾は、背景画像の画素値の変化を混合ガウス分布によってモデル化し、物体に相当する画素を検出している。物体表面の反射と画素値の

[†] 大阪大学大学院基礎工学研究科
Graduate School of Engineering Science, Osaka University

関係はガウス分布を用いて統計的にモデル化されることも多い。Ohta は、背景画素の分布と物体画素の分布をそれぞれガウス分布でモデル化し、物体検出に χ^2 検定を用いている¹¹⁾。また、色変動のモデルに基づいた手法も提案されている^{13),14)}。しかしながら、これらの手法は、単一光源、あるいは単一色光源を暗に仮定しており、太陽だけが唯一の光源である屋外環境では安定に動作するが、複数光源、複数色光源では仮定が崩れて安定に動作できない。たとえば室内では、タングステン電球 (3,500 K) や昼光色蛍光灯 (6,500 K) 等様々な光源色が存在する。そのため、室内において背景差分法を適用するには、様々な光源色に対応した手法が必要となる。

複数光源環境下では、画素値は各光源の光量に起因するエネルギーの線形和で表す背景画像のモデル化が行われる^{15),16)}。新宮ら¹⁶⁾は、そのようなモデルを用いて照明条件に対応した背景画像を合成する手法を提案している。しかし、照明装置の放つ光量を示すパラメータをシステム側で毎回取得する必要がある。本研究においても、室内環境における物体抽出を目的とするため、照明は調光可能か、点灯/消灯を制御できることを前提とするが、入力画像の照明条件は画像から毎回推定を行うため、最初に 1 回だけ、照明条件を既知として背景成分を推定するだけでよい。これは、舞台照明等のように、撮影中に照明条件が変化するような環境においても、照明条件をシステムに入力することなく、同一の背景画像を用いて連続して処理が行えるようにするためである。本研究では、まず同様に背景画像のモデル化を行い、あらかじめ照明制御下での背景画像を撮影し、各照明の光量変化による画素値の変化を測定しておく。そして、入力画像が与えられたときにその照明条件を推定する手法を提案し、合成した背景画像を用いて物体領域を抽出する。また、背景画像モデルに基づいて、影を含む物体領域から影領域を分離する手法を提案する。

以下、2 章では背景画像のモデル化を行い、3 章で照明変化と影を考慮した背景差分法の具体的処理について説明する。続いて、4 章で本手法を実画像に適用した実験例を示し、その有効性を検証する。

2. 背景画像モデル

本章では、まず光源からの光のエネルギー量とカメラで観測される画素値との関係をモデル化し、影を考慮するためのパラメータを導入することで、任意の明るさを表現できる背景画像を生成する手法を示す。

画像上のある画素 j の各 RGB 成分についてその物

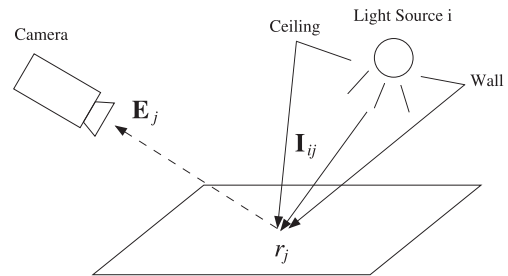


図 1 光源からの光量と画素値の関係
Fig. 1 Relation between energy from light sources and pixel values.

体表面の反射係数を r_j とし、光源 i から j に届く光量を I_{ij} 、観測される画素値を E_j とすると、その関係は式 (1) のようになる (図 1)。ただし、ここではカメラへの入射光量と画素値の関係が線形であると仮定している。

$$E_j = r_j I_{ij}, \quad (1)$$

$$E_j = (R_j, G_j, B_j)^T. \quad (2)$$

ここで r_j は、対角要素に RGB 各成分に対する物体表面の反射係数 (拡散反射率、および鏡面反射率を含む) r_j を持つ対角行列であり、 I_{ij} には天井や壁からの反射光も含まれる。

次に、影を考慮に入れるために、パラメータ S_{ij} を導入すると、画素値 E_j は次式のようになる。

$$E_j = r_j S_{ij} I_{ij}. \quad (3)$$

ここで、 S_{ij} は光源 i から画素 j へ届く光の、シーン中への侵入物体による遮蔽関係を表し、0~1 の値をとる。つまり、 $S = 0$ のときは、遮蔽物体により照明光が完全に隠されて影になっている状態か、照明が消灯されている状態を表している。また、 $S = 1$ は照明光がすべて届いている状態を表す。

複数光源の場合の画素値は、各光源からの光量の和になることから次式のようになる。

$$E_j = r_j \sum_i S_{ij} I_{ij}. \quad (4)$$

ここで、 $L_{ij} = r_j I_{ij}$ とおくと、

$$E_j = \sum_i S_{ij} L_{ij}. \quad (5)$$

となる。

このモデル化により、室内の複数照明条件下における背景差分法を扱うことができる。ここで、 L_{ij} を光源 i による画素 j の背景成分と呼ぶことにする。この背景成分 L_{ij} は、式 (5) より前もって推定しておくことができる。これは、室内における光源は調光器や点灯/消灯スイッチにより制御可能であることが多く、 S_{ij} を既知とすることができるためである。一度 L_{ij}

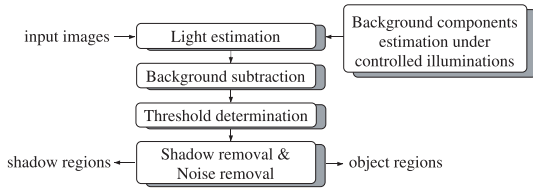


図 2 処理の流れ

Fig. 2 Outline of process.

を求めておくと、式 (5) より入力画像の S_{ij} を推定することができる。具体的な推定処理については、次章で詳しく述べる。

3. 照明変化と影を考慮した背景差分法

本章では、照明変化と影への対策の具体的処理について述べる。2 章で述べたように、各光源による背景成分 L_{ij} を各画素であらかじめ推定しておき、それを用いてカラー化した背景差分処理を行う。

3.1 処理の流れ

本手法の処理の流れを図 2 に示し、各処理の詳細を次節以降で述べていく。なお、ここでは簡単化のため主に 2 光源の場合について述べるが、任意の光源数の場合も本手法を拡張することで、同様に照明変化および影を考慮した背景差分が行えると考えられる。

3.2 背景成分の推定

簡単化のため 2 光源の場合を考える。このとき、画像上のある画素 j における画素値 E_j は、式 (5) より

$$E_j = S_{1j}L_{1j} + S_{2j}L_{2j}. \quad (6)$$

となる。2 章で述べたように、影や照明変化に対応する任意の背景を合成するためには、各光源による背景成分 L_{1j}, L_{2j} が必要となる。式 (6) より、 (S_{1j}, S_{2j}) が異なるときの背景が 2 組以上あれば、各画素ごとに連立方程式を解くことで背景成分 $\tilde{L}_{1j}, \tilde{L}_{2j}$ が推定できる。

ここで、光源の遮蔽関係を表すパラメータ S は、照明の明るさの変化を表すとも考えることもできるので、照明の明るさを変化させた背景画像を 2 枚以上撮影しておいて背景成分の推定を行う。なお、光源数が N 個の場合は、照明の明るさを変化させた背景画像が N 枚以上あれば、同様に連立方程式を解くことで N 個の背景成分を推定できる。

3.3 照明の明るさ推定

実環境では、照明の点灯/消灯により、入力画像の照明条件が背景画像での照明条件とは異なることが考えられる。そこで、実画像で背景差分を行う際には照明変化に対応できることが重要となる。照明変化には明るさ変動と色変動の 2 種類あるが、本研究では照明

の明るさ変動についてのみ扱い、色変動については扱わない。

シーン中への侵入物体がない場合、つまり背景領域においては、照明の明るさ変化は式 (6) において S_{1j}, S_{2j} が全画素で一定に変化していると考えられる。そこで、

$$S_1 = S_{1j}, S_2 = S_{2j} \quad \forall j. \quad (7)$$

とおくと、画像中から 2 点以上選べば、式 (6) より、3.2 節で推定した背景成分 $\tilde{L}_{1j}, \tilde{L}_{2j}$ を用いた連立方程式を解くことで $\tilde{S}_1, \tilde{S}_2$ を推定することができる。しかし、シーン中に侵入物体が存在する場合は、入力画像中のどの点を用いて推定を行うかが問題となる。これは、物体領域およびその影領域では式 (7) が成り立たないためであり、画像全体を用いて照明推定を行った場合は、対象物体の画像中での面積が小さい場合は問題ないが、室内では画像中の対象物体は大きな面積を占めると考えられるため、対象物体およびその影の影響を大きく受けてしまう。精度良く照明条件の推定を行うためには、何らかの手法で対象物体およびその影を除外した領域で推定を行う必要がある。そこで、画像中の対象物体が大きい場合にも、その影響を低減できる照明条件推定方法について述べる。

まず一様乱数により、画像上のランダムな 2 点の組 (画素番号 $j = 1, 2$) を数多く選んでくる。このとき、式 (7) が成り立つと仮定すると、式 (6) より、

$$E_1 = S_1\tilde{L}_{11} + S_2\tilde{L}_{21}, \quad (8)$$

$$E_2 = S_1\tilde{L}_{12} + S_2\tilde{L}_{22}.$$

の連立方程式が成立する。いま、輝度値 T を RGB 成分の和

$$T = R + G + B. \quad (9)$$

で表す。このとき、画素値 E_j の輝度値を T_j 、画素 j の背景成分 $\tilde{L}_{1j}, \tilde{L}_{2j}$ の輝度値をそれぞれ $\tilde{T}_{1j}, \tilde{T}_{2j}$ とすると、式 (8) は以下の輝度値の連立方程式に書きかえることができる。

$$T_1 = S_1\tilde{T}_{11} + S_2\tilde{T}_{21}, \quad (10)$$

$$T_2 = S_1\tilde{T}_{12} + S_2\tilde{T}_{22}.$$

背景領域から 2 点を選んできた場合、この連立方程式を解くことで $(\tilde{S}_1, \tilde{S}_2)$ を推定することができる。しかし、物体やその影領域を含む 2 点を選んだ場合は、そもそも式 (7) の仮定が成り立たないため、式 (10) も成立しない。そこで、そのような outlier を除去するために、式 (10) の連立方程式を解いた結果が、

$$0 \leq S_1, S_2 \leq 1. \quad (11)$$

の条件を満たすもののみを採択することにする。そうすることで、ランダムに選んだ 2 点の組合せのうち、(背景, 背景) および (影, 影) の組合せ以外のもの

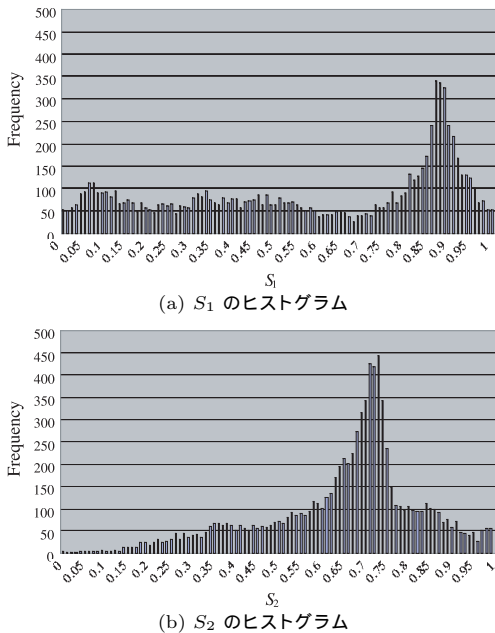


図 3 照明条件推定における S_1 , S_2 のヒストグラムの例

Fig. 3 Histograms of S_1 , S_2 constructed in light estimation process.

をかなり排除することができる．よって， $\tilde{S}_1$, $\tilde{S}_2$ それぞれでヒストグラムを作成すると図 3 のようになり，ヒストグラムのピークとなる値をその入力画像における照明条件 ($\tilde{S}_1$, $\tilde{S}_2$) とすればよい．

図 3 の例ではヒストグラムの間隔を 0.01 とし， $(\tilde{S}_1, \tilde{S}_2) = (0.87, 0.74)$ と推定できる．なお，N 光源の場合も，ランダムに N 点を選んでくることで同様に照明条件推定ができると考えられる．この結果より，入力画像の照明条件に対応した背景成分を

$$\tilde{\mathbf{L}}'_{1j} = \tilde{S}_1 \tilde{\mathbf{L}}_{1j}, \quad \tilde{\mathbf{L}}'_{2j} = \tilde{S}_2 \tilde{\mathbf{L}}_{2j}. \quad (12)$$

のように推定する．また，

$$\tilde{\mathbf{E}}_j = \tilde{\mathbf{L}}'_{1j} + \tilde{\mathbf{L}}'_{2j}. \quad (13)$$

として背景画像を合成することができる．この推定背景を用いれば，入力画像との差分値がしきい値 T_a 以上ならば物体とすることで，照明の明るさ変化に対応した背景差分が行える．しかし，影領域も同時に検出してしまうため，次節では影除去の方法について述べる．

3.4 影を考慮した背景差分

3.4.1 影の除去

入力画像中に移動物体が存在する場合に，3.3 節で述べた方法で影を含む物体抽出を行った後，影を除去する方法について述べる．

光源が L_1 , L_2 の 2 つの場合，画像は以下の 4 つの領域に分割できると考えられる．

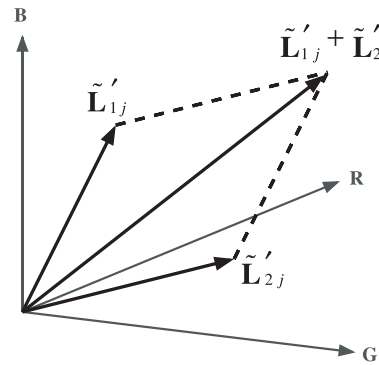


図 4 背景と影を表す 3 次元ベクトル

Fig. 4 Three vectors expressing the background and shadows.

- (1) 2 つの光が当たっている背景
- (2) L_1 の光のみが当たっている影
- (3) L_2 の光のみが当たっている影
- (4) 物体

このとき，3.2 節と 3.3 節で推定した背景成分ベクトル $\tilde{\mathbf{L}}'_{1j}$, $\tilde{\mathbf{L}}'_{2j}$ を用いると，RGB 空間において，図 4 に示す 3 つのベクトルが背景および影に相当する，物体を表すベクトルは，この 3 つのベクトルとは異なる方向を向くと考えられる．

ただし，実環境においては，光源に大きさがあるために，光源の一部の光のみが届く半影 (soft shadow) と呼ばれるぼやけた影が生成されることが多く，上記のような 4 つの領域に分けるのは難しい．そこで，この半影への対応が必要となる．半影領域内の画素値は，図 4 での $\tilde{\mathbf{L}}'_{1j}$ と $\tilde{\mathbf{L}}'_{2j}$ の 2 つの背景成分ベクトルで作られる平面 (平行四辺形) 内に分布する．よって，RGB 空間において観測値とその平面との最短距離を特徴量とすることで，影と物体が分離できると考えられるが，影と認識する領域が大きくなることで半影に対応できるようになると同時に，物体を影であると誤検出する割合も大幅に増すことになる．また， $\tilde{\mathbf{L}}'_{1j}$, $\tilde{\mathbf{L}}'_{2j}$ のなす角が小さくなるほど，平面との距離を計算する過程で誤差が大きくなると考えられる．

ここで，RGB 空間中の背景成分ベクトルに対する半影の分布を調べるために，照明変化のない状態で撮影した画像から，手で背景，本影，半影の領域を切り出し，その各画素ごとに RGB 値から (S_1 , S_2) の推定を行った．図 5 にその結果を示す．図 5 の四角形は，図 4 での平行四辺形に相当する．図 5 より，半影は，背景成分ベクトルで作られる平行四辺形の内部よりも 4 辺周辺に多く分布しているのが分かる．これは，半影が影のエッジ付近の領域に相当するため，複数光源

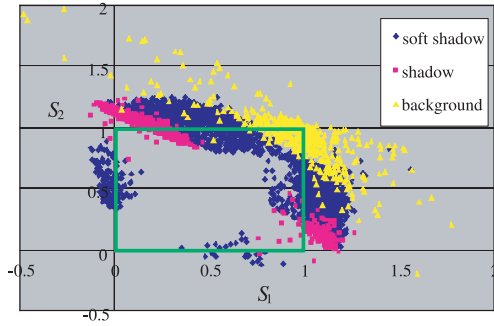


図 5 半影の分布

Fig. 5 The distribution of soft shadow.

の場合，一方の光源による影のエッジ付近の領域は，そのエッジどうしが重なり合う領域に比べてかなり大きいことを意味している．したがって，入力画素値と4辺との最短距離 d_{min} を用いて影と物体の分離を行う．もちろん，半影どうしが重なり合う領域は，平行四辺形の内部にくるために誤識別の要因となるが，影領域全体に対してその割合は非常に小さいといえる．

以上のことから，特徴量 f を次のように求める．

$$d_{min} = \min_{0 \leq \alpha \leq 1, 0 \leq \beta \leq 1} \left[\|\mathbf{E}_j - \alpha \tilde{\mathbf{L}}'_{1j}\|, \|\mathbf{E}_j - \beta \tilde{\mathbf{L}}'_{2j}\|, \|\mathbf{E}_j - (\alpha \tilde{\mathbf{L}}'_{1j} + \beta \tilde{\mathbf{L}}'_{2j})\|, \|\mathbf{E}_j - (\tilde{\mathbf{L}}'_{1j} + \beta \tilde{\mathbf{L}}'_{2j})\| \right] \quad (14)$$

$$f = \frac{d_{min}}{\|\mathbf{E}_j\|}. \quad (15)$$

ここで， $\mathbf{E}_j$ は入力画素値， $\|\mathbf{E}_j\|$ は RGB 空間における原点からの距離を表している．この特徴量 f があるしきい値 T_b より小さければ影，大きければ物体として識別する．

ただし，本手法では RGB 空間の原点付近の値であっても影として判別することがあるため，黒い物体も影として判別されることがありうる．実際，人物の髪領域が影として判別されることが多い．そこで，これを防ぐために，明るさがあるしきい値 T_c 以下の場合にはすべて物体であると判別する．これは，実環境においては環境光の影響により影領域であっても真っ暗になることは稀であり，暗い領域は物体表面の色が元々黒いと考えられるためである．

以上まとめると，背景領域，影領域，物体領域は以下のように判別される．

- (1) $\|\mathbf{E}_j - (\tilde{\mathbf{L}}'_{1j} + \tilde{\mathbf{L}}'_{2j})\| \leq T_a$ のとき背景
- (2) $\|\mathbf{E}_j - (\tilde{\mathbf{L}}'_{1j} + \tilde{\mathbf{L}}'_{2j})\| > T_a$ かつ $f \leq T_b$ かつ $\|\mathbf{E}_j\| > T_c$ のとき影
- (3) それ以外のとき物体

なお， N 光源の場合は， N 個の背景成分ベクトルで

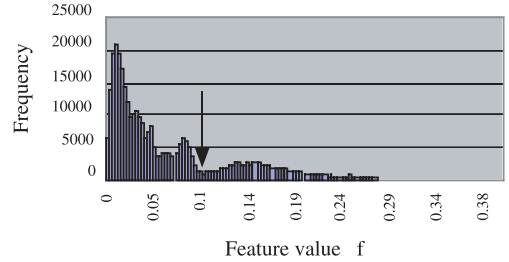


図 6 ある画像における本手法の特徴量のヒストグラム

Fig. 6 An example of histogram of the proposed image features.

張られる空間の各辺との最短距離を用いることで同様に考えることができる．ただし，光源数が増えるほど計算量が増すため，計算量削減のためには何らかの工夫が必要である．

3.4.2 しきい値の自動決定

3.4.1 項で求めた特徴量から背景（影を含む）と物体を正確に分離するためには，適切なしきい値 T_a, T_b, T_c を選ぶ必要がある．本項では，しきい値 T_b の自動決定法について述べる．特徴量のヒストグラムを作成すると，ヒストグラムは物体と背景（影）の2クラスに対応し双峰となり，その間の谷部を最適しきい値と決定することになる．しかし，観測シーンによってこの谷部の位置，つまり適切なしきい値が異なるため，これを最適に決定する必要がある．

本手法では，特徴量のヒストグラムから大津の方法と呼ばれる自動しきい値選定法¹⁷⁾によりしきい値 T_b の決定を行った．これは濃淡画像の2値化を行うためのものであるが，本手法の特徴量を濃淡画像の濃度レベルと同様にして考えることで適用できる．

このしきい値決定法を図6に示すヒストグラムに適用すると，最適しきい値が0.096となり，正確にヒストグラムの谷底を検出できているのが確認できる．

なお， T_a, T_c については，現在のところ実験による経験値を用いている．

4. 実画像における評価実験

本手法の有効性を検証するために評価実験を行った．光源として，3本の蛍光灯をそれぞれ点灯/消灯できる照明装置を用いた．図7が実験に用いた照明である．これらの蛍光灯を点灯/消灯することで，照明の明るさ変化および照明条件の変化とした．照明およびカメラの位置，向きは固定である．なお，上記の照明光以外の光源の影響を受けないようにするため，夜間に室内で実験を行った．実験では，サイズ 720×486 画素，8bit カラー，対象として人物を含めた画像を用



図 7 実験で用いた照明

Fig. 7 Lights used for experiments.



図 8 照明条件を変化させて撮影した背景画像

Fig. 8 Background images captured under various light conditions.

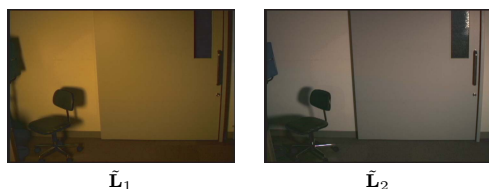


図 9 推定された背景成分画像

Fig. 9 Estimated background components.

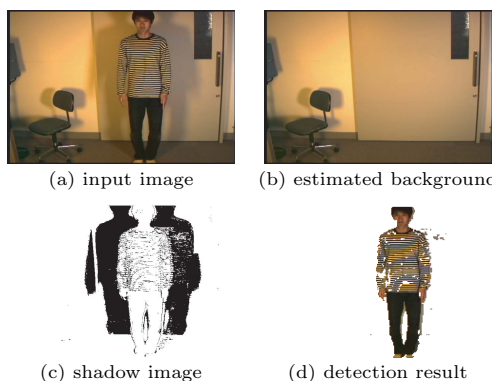


図 10 実験結果 (2 光源)

Fig. 10 Result images (2 lights).

いて処理を行う。

4.1 実験結果

図 8 に、照明色の異なる 2 つの光源を用いたときに、照明条件を変化させて撮影した背景画像を示す。図 9 は図 8 の背景画像から推定した背景成分画像である。図 10 は図 9 の背景成分画像を用いた本手法による実験結果であり、(a) 入力画像、(b) 入力画像から照明条件を推定して合成した背景画像、(c) 影の分離結果、(d) 物体抽出結果、を示している。また、2 つの照明色が似ている場合の実験結果も図 11 に示す。

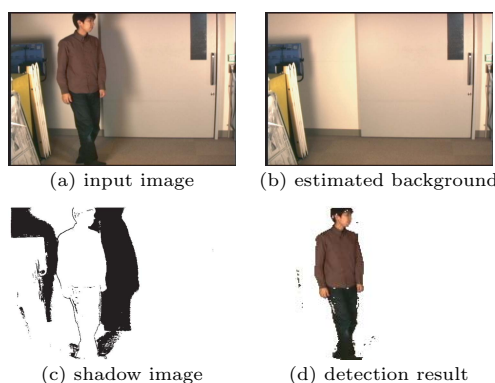


図 11 実験結果 (照明色の似ている 2 光源)

Fig. 11 Result images (similar colored 2 lights).

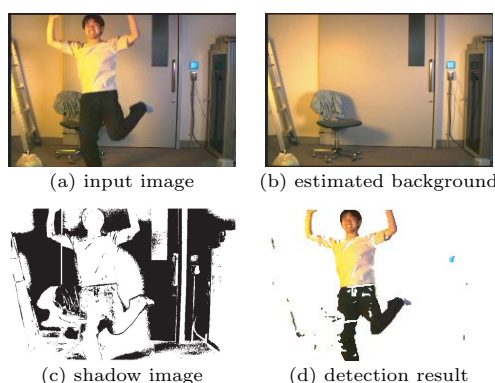


図 12 実験結果 (3 光源)

Fig. 12 Result images (3 lights).

これらの結果より、影の影響を大幅に低減し、良好な物体抽出結果が得られているのが分かる。2 つの照明色の違いの大きさによらず、正確な物体抽出が行えることも確認できた。

さらに、照明色の異なる 3 光源を用いた実験も行った。図 12 にその結果を示す。なお、図 10～図 12 の実験における照明の配置は、それぞれ異なるものである。図 12 の結果より、光源数が増えて影領域がより広範囲になった場合でも、正確に物体抽出が行えているのが分かる。

また、精度評価を行うために、2 光源と 3 光源のそれぞれに対して、図 10～図 12 とは対象人物、撮影場所、照明の明るさ、照明色等が異なる環境で実験を行い、各 10 枚の実験結果を得た。

4.2 照明の明るさ推定の有効性

照明の明るさ変化がおこったときの本手法による照明の明るさ推定の有効性を検証する。3.3 節で述べたように、2 光源での照明の明るさ推定では、画像中からランダムな 2 点を選ぶことで $(\hat{S}_1, \hat{S}_2)$ を推定し、0～1 の値を採択することを繰り返す。画像領域を背

表 1 組合せのパターン

Table 1 Patterns of pair selection.

	BG	SH	SS	OB
BG	A ₁	B ₁	B ₂	B ₃
SH	B ₁	A ₂	C ₁	B ₄
SS	B ₂	C ₁	C ₂	B ₅
OB	B ₃	B ₄	B ₅	B ₆

BG: background SH: shadow
SS: soft shadow OB: object

景, 本影, 半影, 物体に分けると, この 2 点の組合せは表 1 のようになる.

A_i は推定する S の値が正しい値となる組合せを表し, B_i と C_i はノイズとなりうる正しくない組合せを表している. それぞれの組合せから推定された $(\tilde{S}_1, \tilde{S}_2)$ は, 式 (11) によりある割合で棄却される. ここで, A_i, B_i, C_i のそれぞれの組合せにおいて, 棄却されずに採択される割合を順に R_1, R_2, R_3 とする. このとき, 提案手法がうまく働く条件は, 正しい組合せが全体の過半数を占めることで, 以下の式で表すことができる.

$$\sum_i R_1 P(A_i) \geq \sum_i R_2 P(B_i) + \sum_i R_3 P(C_i). \quad (16)$$

画像中における背景, 本影, 半影, 物体の割合をそれぞれ P_1, P_2, P_3, P_4 とすると, 式 (16) は次のように書き直せる.

$$R_1(P_1^2 + P_2^2) \geq R_2\{P_1(1 - P_1) + P_2(P_1 + P_4) + P_3(P_1 + P_4) + P_4\} + R_3(2P_2P_3 + P_3^2). \quad (17)$$

$P_4 = 1 - \sum_{i=1}^3 P_i$ を用いて整理すると,

$$(R_1 + R_2)(P_1^2 + P_2^2) \geq (R_3 - R_2)(2P_2P_3 + P_3^2) + R_2. \quad (18)$$

となる. ここで, 4.1 節で述べた精度評価用の 10 枚の画像に対して R_i をそれぞれ求め, その平均をとると,

$$R_1 = 0.408, R_2 = 0.109, R_3 = 0.274$$

であった. よって, 照明の明るさ推定がうまく働くための条件は,

$$0.517(P_1^2 + P_2^2) \geq 0.165(2P_2P_3 + P_3^2) + 0.109. \quad (19)$$

となる. もちろん, 実験で用いた 10 枚の画像はこの条件を満たしている.

なお, 本手法は, 照明が局所照明である場合や, 背景画像中に影領域が存在する場合等, 照明の影響を受けない領域があるときには安定に動作しない可能性がある. 問題となるのは, そのような領域中から 2 点を選んできた場合であり, このとき推定する $(\tilde{S}_1, \tilde{S}_2)$ の

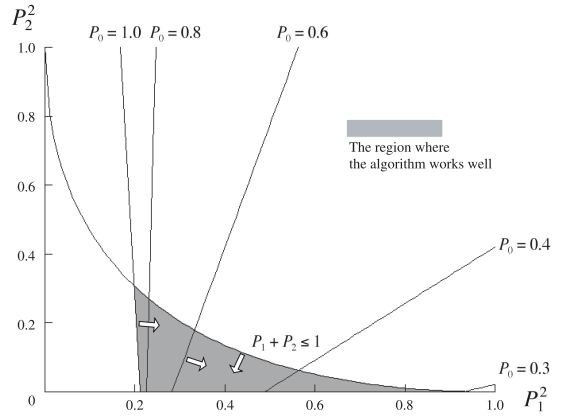


図 13 本手法の適用可能領域
Fig. 13 Feasible solution region.

値は不安定になると考えられる. そこで, 画像中において, すべての照明の影響を受ける領域の割合を P_0 とすると, 本手法がうまく働く条件は以下の式で表すことができる.

$$\begin{aligned} \{1 - (1 - P_0)^2\} R_1(P_1^2 + P_2^2) \geq & R_2(1 - P_1^2 - P_2^2 - P_3^2 - 2P_2P_3) \\ & + (1 - P_0)^2 R_1(P_1^2 + P_2^2) \\ & + R_3(2P_2P_3 + P_3^2). \end{aligned} \quad (20)$$

R_i を代入して整理すると,

$$(-0.816P_0^2 + 1.632P_0 - 0.299)(P_1^2 + P_2^2) \geq 0.165(2P_2P_3 + P_3^2) + 0.109. \quad (21)$$

ここで, 解析を容易にするために, $P_2 > P_3$, つまり本影領域が半影領域よりも大きいと仮定する. 半影領域は影領域のエッジ付近に相当することから, この仮定は妥当であるといえる. このとき,

$$(\text{右辺}) \leq 0.495P_2^2 + 0.109.$$

となり, 条件は以下の式で表される.

$$\begin{aligned} (-0.816P_0^2 + 1.632P_0 - 0.299)P_1^2 \geq & 0.109 \\ & + (0.816P_0^2 - 1.632P_0 + 0.794)P_2^2. \end{aligned} \quad (22)$$

よって, P_0 を様々に変化させたとき, 本手法がうまく働く領域は図 13 のようになる.

図 13 より, $P_0 = 1.0$ の場合は $P_1^2 = 0.2$, つまり画像中の背景領域の割合が約 45% 以上ならば, 照明の明るさ推定は成功するといえる. これは, ノイズとなる物体領域や影領域が画像中の半分を占めていても背景画像を正しく推定できることになり, 本手法の有効性を示している. なお, P_0 が小さくなるほど本手法の適用可能領域は狭くなり, $P_0 = 0.3$ あたりが限界であることが確認できる.

4.3 物体抽出の精度評価

4.3.1 精度評価実験

本手法による物体抽出の精度評価を行うため, 4.1 節

表 2 物体抽出の精度 (2 光源)

Table 2 The accuracy of object extraction (2 lights).

image	$R_D \times 100(\%)$	$R_I \times 100(\%)$
(a)	96.2	1.49
(b)	89.1	1.39
(c)	92.4	1.64
(d)	94.1	3.16
(e)	89.5	2.12
(f)	90.8	1.08
(g)	90.1	1.94
(h)	89.0	7.80
(i)	87.1	4.99
(j)	85.1	4.30

表 3 物体抽出の精度 (3 光源)

Table 3 The accuracy of object extraction (3 lights).

image	$R_D \times 100(\%)$	$R_I \times 100(\%)$
(k)	89.5	0.64
(l)	84.7	0.57
(m)	88.5	2.53
(n)	74.8	2.16
(o)	87.7	7.84
(p)	72.0	4.36
(q)	74.5	3.90
(r)	75.3	7.84
(s)	93.7	6.18
(t)	91.4	6.02

で述べた 10 枚の実験結果と、目視により物体抽出を行った結果との比較を行った。

いま、本手法により物体と判別された領域を A 、背景と判別された領域を B 、目視により物体と判別された領域を C 、背景と判別された領域を D としたとき、物体検出率 R_D および物体誤検出率 R_I を以下の式を用いて算出する。

$$R_D = \frac{\# \{A \cap C\}}{\# C}, \quad (23)$$

$$R_I = 1 - \frac{\# \{B \cap D\}}{\# D}. \quad (24)$$

ここで、 $\#$ は領域内の画素数を表す。なお、ここでの背景領域には影領域も含まれる。2 光源による全 10 枚の実験結果に対して評価を行った結果、表 2 のようになった。

また、3 光源を用いた実験の精度評価結果を表 3 に示す。2 光源、3 光源ともに、物体の検出漏れがやや見られるものの、影の除去はほぼできているのが確認できた。

4.3.2 しきい値変化の影響

物体抽出の性能を議論するとき、しきい値設定の変化が性能に与える影響を調べることも重要である。精度評価で用いた 10 枚の画像に対して、本手法で用いるしきい値 T_a 、 T_b 、 T_c を変化させたときの影響を調

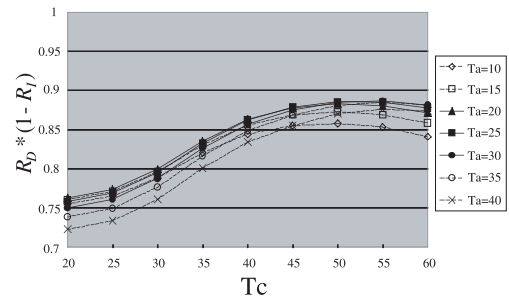
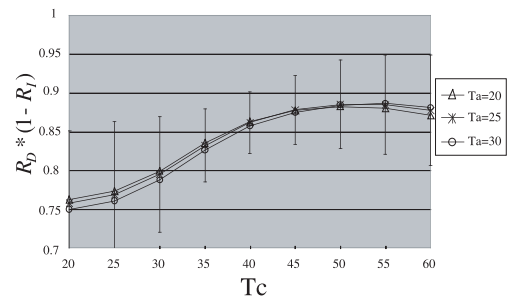
図 14 T_a 、 T_c を変動させたときの物体抽出性能の変化Fig. 14 The accuracy of object extraction under various values of T_a , T_c .

図 15 物体抽出性能の散らばり (標準偏差)

Fig. 15 Dispersion of the accuracy of object extraction (standard deviation).

べた。なお、ここでは 2 光源の場合を扱うことにする。

まず、 T_b は大津の方法で求めたしきい値とし、 T_a 、 T_c を同時に変動させたときの 10 枚の平均物体抽出性能を調べた。図 14 にその結果を示す。横軸は T_c 、縦軸は $R_D * (1 - R_I)$ である。図 14 より、 $T_a = 20 \sim 30$ 、 $T_c = 50 \sim 55$ あたりが物体抽出性能が最も良いと考えられる。しかし、標準偏差を求めてみると図 15 のようになり、 $T_c = 50, 55$ では観測シーンによって性能に散らばりがあるのが確認できる。よって、物体抽出を最も安定に行えんと考えられるのは、 $T_a = 25$ 、 $T_c = 45$ のときである。

この結果より、 T_a 、 T_c を固定し、 T_b を変動させたときの ROC 曲線を描くと図 16 のようになった。この曲線はグラフの左上に位置するほどその精度が良いといえる。10 枚の画像それぞれに対して ROC 曲線を描いているが、中には、 R_I が大きくなっても R_D が 1 に近付いていかない画像や、逆に R_D が小さくなっても R_I が 0 に近付かない画像も見られる。これは、しきい値 T_a 、 T_c を固定したことによる影響だといえる。つまり、前者は、 T_a の値が大きすぎたために、影除去をする前に物体領域を背景と誤識別したことが原因であり、後者は、 T_c の値が大きすぎたために、影領域までも物体領域として誤識別したことが原

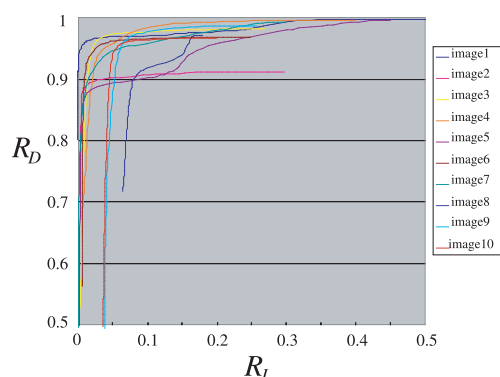


図 16 ROC 曲線

Fig. 16 ROC curve.

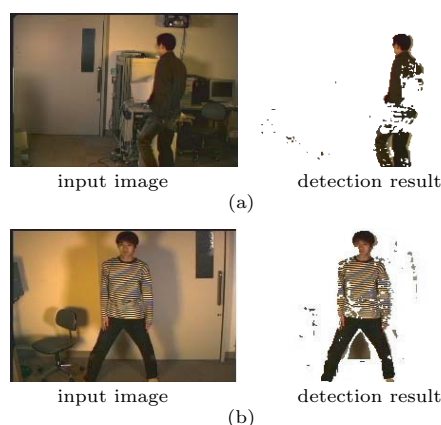


図 17 本手法による誤識別の例

Fig. 17 Examples of incorrect classification

因である．しかしながら，図 16 より，それらの画像を含めたほとんどの画像に対して，物体検出率 (R_D) 90%以上，物体誤検出率 (R_I) 10%以下を同時に実現できているのが分かる．

4.4 考 察

ここでは，本手法の問題点について考察を行う．本手法が誤識別を起こす主な要因としては，4.3.2 項で述べたしきい設定の影響のほかに，

- (1) 背景と似た色の物体の検出漏れ
 - (2) 複数の半影が重なり合った領域の誤検出
- があげられる．(1)，(2) による誤識別の例をそれぞれ図 17 (a)，(b) に示す．図 17 (a) では，人物の服の色が背景と似ているために検出漏れが起こっているが，これは本手法が色情報に基づいた手法であるため避けられないことであり，この対策としては，エッジ検出¹⁾や領域分割手法を利用することが考えられる．また，図 17 (b) では，人物の股の間の領域が，半影が重なりあった領域に相当する．これは，3.4.1 項で述べたように，半影は背景成分ベクトルで張られる空間内に分

布するため，その空間の各辺に主に分布すると仮定した本手法では誤検出が起こる．しかし，この領域は影領域全体と比べると非常に小さいといえる．このほかの問題点としては，色情報の信頼性の低下する暗い領域における誤識別，本手法では考慮していない相互反射の影響による誤識別等があげられる．

以上結論としては，若干の問題点はあるものの，本章でこれまで述べてきたように，従来の背景差分法で誤検出されていた影領域の影響を大幅に削減できていることから，本手法の有効性が確認できた．

5. ま と め

背景差分法の欠点である影の問題に対して，従来から多くのアプローチがなされてきたが，単一光源あるいは単一色光源を前提としたものが多い．本稿では，室内の複数照明環境下における照明の明るさ変化および影を考慮した背景モデルを提案し，そのモデルに基づいて照明条件の推定および影除去を行い，物体領域のみを抽出する手法を提案した．そして，実画像を用いた実験により本手法の有効性を確認した．

応用としては，本手法を人物追跡やジェスチャ認識等の前処理として用いることでそれらの精度向上が期待できる．また，照明変化が起こった場合にも物体領域と影領域を別々に抽出できることや，ある画素に対する各照明による影響の割合が計算できることから，影領域をさらに分割可能であることも本手法の特徴である．

今後の課題として，背景と似た色の物体の検出漏れの問題に対して，エッジ検出等色情報以外の情報を用いた手法と組み合わせることで解決を図ること等について検討する予定である．

参 考 文 献

- 1) 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，電子情報通信学会論文誌 D-II，Vol.J80-D-II，No.11，pp.2976-2985 (1997)．
- 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，電子情報通信学会技術研究報告，PRMU98-109，pp.23-30 (1998)．
- 3) 長谷川為春，小沢慎治：モデルに基づいた影を伴う移動物体の検出，電子情報通信学会論文誌 D-II，Vol.J83-D-II，No.2，pp.555-564 (2000)．
- 4) Koller, D., Danilidis, K. and Nagel, H.-H.: Model-Based object tracking in monocular image sequences of road traffic scenes, *Int'l Journal of Computer Vision*, Vol.10, No.3, pp.257-281 (1993)．
- 5) 渡辺真太郎，宮島耕治，武川直樹：陰影モデル

を用いた航空写真における建造物の変化抽出，MIRU'98 画像の認識・理解シンポジウム論文集 II，pp.373-378 (1998).

- 6) 緒方敏博，園田頼信，川勝正晴：監視画像における「対象物」とその「影」の分離，MIRU'96 画像の認識・理解シンポジウム講演論文集 II，pp.91-96 (1996).
- 7) Gordon, G., Darrell, T., Harville, M. and Woodfill, J.: Background estimation and removal based on range and color, *Proc. CVPR*, pp.459-464 (1999).
- 8) Stauder, J., Mech, R. and Ostermann, J.: Detection of moving cast shadows for object segmentation, *IEEE Trans. Multi Media*, Vol.1, No.1, pp.65-76 (1999).
- 9) 松山隆司，和田俊和，波部 斉：照明変化に頑健な背景差分，電子情報通信学会論文誌 D-II，Vol.J84-D-II, No.10, pp.2201-2211 (2000).
- 10) Stauffer, C. and Grimson, W.E.L.: Adaptive background mixture models for real-time tracking, *Proc. CVPR*, pp.246-252 (1999).
- 11) Ohta, N.: A statistical approach to background subtraction for surveillance systems, *Proc. ICCV*, pp.481-486 (2001).
- 12) Mikić, I., Cosman, P., Kogut, G. and Trivedi, M.: Moving shadow and object detection in traffic scenes, *Proc. ICPR*, Vol.1, pp.321-324 (2000).
- 13) Elgammal, A., Harwood, D. and Davis, L.S.: Non-parametric model for background subtraction, *Proc. IEEE ICCV'99 Frame-rate Workshop* (1999). <http://fizbin.eecs.lehigh.edu/tboulton/FAME/Horprasert/index.html>.
- 14) 奥村晃弘，岩井儀雄，谷内田正彦：屋外における移動物体の検出—照明変動と影への対策，MIRU2000，画像の認識・理解シンポジウム講演論文集 II，pp.307-312 (2000).
- 15) 佐藤いまり，佐藤洋一，池内克史：物体の陰影に基づく光源環境の推定，情報処理学会論文誌：コンピュータビジョンとイメージメディア，Vol.41, No.SIG10(CVIM1), pp.31-40 (2000).
- 16) 新宮 淳，亀田能成，角所 考，美濃導彦：背景画像合成による前景抽出とそれに基づく照明の動的制御，電子情報通信学会技術研究報告 PRMU，Vol.101, No.569, PRMU2001-218, pp.95-102 (2001).
- 17) 大津展之：判別及び最小 2 乗基準に基づく自動しきい値選定法，電子情報通信学会論文誌，Vol.J63-D, No.4, pp.349-356 (1980).

(平成 14 年 9 月 2 日受付)

(平成 15 年 3 月 28 日採録)

(担当編集委員 橋本 学)

森田 順也

平成 14 年大阪大学基礎工学部システム科学科卒業．現在同大学大学院基礎工学研究科博士前期課程在学中．



岩井 儀雄 (正会員)

平成 4 年大阪大学基礎工学部情報工学科卒業．平成 6 年同大学大学院基礎工学研究科修士課程修了．平成 9 年同大学院基礎工学研究科博士課程修了．同大学院基礎工学研究科助手を経て，現在同研究科助教授．コンピュータビジョンに関する研究に従事．IEEE，日本ロボット学会各会員．工学博士．



谷内田正彦 (正会員)



昭和 46 年大阪大学大学院工学研究科修士課程修了．同年同大学基礎工学部制御工学科助手．同助教授を経て同学部情報工学科教授，平成 6 年同学部システム工学科教授．昭和 42 年～43 年デンマーク原子力研究所留学．昭和 47 年～48 年米イリノイ大学にて Research Associate．昭和 55 年～56 年西独ハンブルグ大学 Research Fellow．昭和 57 年米ミネソタ大学 CDC Professor．ロボット学会，人工知能学会等会員．著書，ロボットビジョン (昭晃堂，大川出版賞受賞)，コンピュータビジョン (丸善，編著) 等．コンピュータ・ビジョン，画像処理，人工知能，移動ロボット等の研究を行っている．工学博士．