

## ブラインド $\gamma$ 値推定と複数明るさ可変背景を用いた屋外物体検出

阪本 光翼<sup>†</sup>      吉村 宏紀<sup>†,††</sup>      西山 正志<sup>†,††</sup>      岩井 儀雄<sup>†,††a)</sup>

Foreground Detection in Outdoor Scenes by using Blind Gamma Correction and Layered Adaptive Background Models

Kosuke SAKAMOTO<sup>†</sup>, Hiroki YOSHIMURA<sup>†,††</sup>, Masashi NISHIYAMA<sup>†,††</sup>, and Yoshio IWAI<sup>†,††a)</sup>

あらまし 本論文では、従来手法である明るさ可変背景モデル (ABM) を拡張した手法 (IGLABM) を提案する。照明変動や移動物体の影の除去に対応可能な手法として提案された明るさ可変背景モデルは、草木の揺れなどの周期的な背景変動に対応していない点、入力画像の色空間にリニア RGB 空間を仮定しており  $\gamma$  補正値が未知の公開画像データベース等を用いた場合、精度が低下してしまう点などが問題として挙げられる。これらの問題に対して、提案手法では、明るさ可変背景モデルを複数レイヤー構造にすることで、草木の揺れなどの周期的な背景変動に対応できるように拡張した。また、色空間としてリニア RGB 空間を仮定している問題に対しては、ブラインド $\gamma$ 値推定を行うことで、未知の $\gamma$ 値で撮像された画像データに対しても明るさ可変背景モデルを利用できるようにした。以上の改良により、評価用公開画像データベースを用いて提案手法の性能評価を行い、高い精度で物体検出を行うことができることを実験により検証した。

キーワード ブラインド $\gamma$ 値推定, 明るさ可変背景, 画素ベース背景差分法, 前景抽出, 物体検出

### 1. ま え が き

近年、複雑化する社会環境の大きな変化に伴い、安心・安全な社会が求められている。このような社会を実現するため、防犯システムの需要が増しており、カメラ映像から人物・車両などの移動物体を検出し、追跡や対応付けなどを行う技術が注目されている。しかし、検出対象領域を目視で検出することは負担が大きく、検出作業の自動化や省略化が望まれている。このような検出作業を自動化する手法の一つに背景差分法が存在する。

背景差分法は、その原理が簡便で高速化しやすい様々な方法が提案されてきた [1]。背景差分法の課題として一般に、照明変動への対応、物体の影の除去、草木の揺れといった周期的な背景変動への対応などが挙げられる [2]。背景差分法は、注目画素の情報のみで前景か背景かを推定する画素ベース [3]~[12]、注目画

素とその周辺領域または大域領域の情報を利用する領域ベース [13]~[19] の 2 種類に分類できる。また、近年では深層学習を用いた背景差分法 [20] なども提案されているが、事前に学習データを大量に必要とすること、オフラインでの学習が必要なことなどの欠点がある。本論文では、事前に学習データを必要とせず、オンライン学習を行う画素ベースの背景差分法に着目した。画素ベースの背景差分法の精度を高めることができれば、その精度向上の結果を利用して、領域ベースの背景差分法の精度を向上させたり、深層学習への教師信号として利用することができるからである。

画素ベース背景差分法のうち、照明変動および、物体の影の除去を同時に行う手法として、明るさ可変背景モデル (ABM) [3] が提案されている。ABM は、太陽直射光成分と環境光成分を考慮した 2 色性反射モデル [21] となっており、2 色の混合度を変化させることにより様々な明るさの背景画像を生成することができ、これにより照明変動や物体の影除去を実現している。この明るさ可変背景モデルの問題点として、草木の揺れなどの周期的な背景変動に対応していない点、入力画像の色空間にリニア RGB 空間 ( $\gamma = 1.0$ ) を仮定しており、 $\gamma$  補正値が未知の公開画像データベース等を

<sup>†</sup> 鳥取大学大学院持続性社会創生科学研究科  
Graduate School of Sustainability Science, Tottori University, Tottori 680-8550, Japan

<sup>††</sup> 鳥取大学工学部附属クロス情報科学研究センター  
Cross-informatics Research Center, Tottori University

a) E-mail: iwai@tottori-u.ac.jp

用いた場合、物体抽出精度が低下してしまう点が挙げられる。

そこで本論文では、従来手法である明るさ可変背景モデルを拡張して、問題点を解決した手法 (IGLABM, Improved Gamma corrected Layered Adaptive Background Model) を提案する。提案手法では、明るさ可変背景モデルを複数レイヤー構造にすることで、草木の揺れなどの周期的な背景変動に対応できるようにした。また、色空間としてリニア RGB 空間を仮定している問題に対しては、ブラインド  $\gamma$  値推定を行うことで、未知の  $\gamma$  値で撮像された画像に対しても明るさ可変背景モデルを利用できるようにした。以上の改良により、評価用公開画像データベースを用いて提案手法の性能評価を行い、高い精度で物体検出を行うことができることを実験により検証した。

## 2. 関連研究

### 2.1 背景差分法

前節で述べたように、本論文では画素ベース背景差分法に着目した。画素ベース背景差分法は、画像処理の分野では従来から多くの手法が提案されている [3]~[12]。最も簡便な手法として、単純にフレーム間で差分をとる手法 [4] があるが背景変動を除去することはできない。背景差分法では、背景をどのようにモデル化するかが重要であり、統計的な手法 [5]、ノンパラメトリックな手法 [6]、ファジーモデル [7]、スパース表現 [8] などが用いられている。その中でも、統計的な手法が背景のモデル化に良く用いられている。画素ベース背景差分法の代表的な統計的手法として、混合正規分布を用いた統計的な手法 [9], [10] がある。この手法は、背景画素を混合正規分布によってモデル化し、物体に相当する画素をマハラノビス距離を利用して背景差分法により検出する。また混合正規分布のため、EM 法を用いてオンラインで各画素に使われる正規分布のパラメータを推定することも可能である。また、この手法の背景モデルの更新式を改良し、移動物体の影の除去に対応した手法 MOG2 [11] が提案されている。また画素ベースのノンパラメトリックな手法として、自己組織化によって背景モデルを自動生成し物体検出を行う SOBS [12] がある。

また、領域ベースの手法としては、移動物体領域をカルマンフィルターと Gale-Shapley マッチングを使用して追跡しながら検出を行う手法 GMG [13] がある。ノンパラメトリックな手法として、画素、物体ご

とに背景の統計モデルを推定し、画素と領域の両方の情報を活用して、前景オブジェクトの検出が可能な手法 SACON [14] や、最初の数フレームの画素を初期背景モデルとして獲得し物体検出を行いつつ、更新・拡散処理により近傍へ画素情報を送る ViBe [15] などが存在する。また、SACON や ViBe の手法を改良した手法 PBAS [16] がある。PBAS は拡散処理の際に近傍画素をその位置の画素値で更新する点が、ViBe とは異なっている。同様に、注目画素の近傍パターンを辞書として更新・維持して背景か、物体かの判断を行う PAWCS [17] がある。

我々の提案手法である IGLABM は、改良元となる ABM と同様に画素ベースの背景差分法であり、ABM が対応している影の除去や照明変動に加えて、ブラインド  $\gamma$  値推定により  $\gamma$  値が未知の画像にも対応可能で、複数レイヤー構造により草木の揺れなどの周期的な背景変動にも対応可能である。

### 2.2 ブラインド $\gamma$ 値推定

一般的な画像処理では、画像処理に使用する色空間としてリニア RGB 空間が暗黙的に仮定されている。また、ABM のように画素値を物理的な観測量として使用するためには、色空間としてリニア RGB 空間が必要となる。そのため、従来手法である ABM で撮像装置や色空間、 $\gamma$  補正に関する情報がない画像を扱うのは困難となる。そのような場合においても画像処理を可能にするためにブラインド  $\gamma$  値推定と呼ばれる手法が様々に提案されている [22]~[24]。例えば、画像中の画素の部分集合の平均と分散を用いることで、 $\gamma$  値推定を行う Mean-Variance Gamma (MV-Gamma) [22] や、また、 $\gamma$  補正により周波数領域において高次相関が起こることを利用し、その相関を最小化することで  $\gamma$  値推定を行う Blind Inverse Gamma Correction (BIGC) [23] などが提案されている。その他に、リニア RGB 画像（以下、リニア画像と呼ぶ）のエッジ領域、すなわち画素の空間変動が直線上に分布するということを用いて、その直線性を評価することで 1 枚の画像から  $\gamma$  値推定を行う [24] ものもある。

我々の提案手法は、観測している背景画素の時間的変動がリニア画像では直線上に分布するということを用いてその直線性を評価することで  $\gamma$  値推定を行う<sup>(1)</sup>。

(1): 本論文では、我々の提案手法である GLABM [25] をさらに改良した IGLABM を提案する。

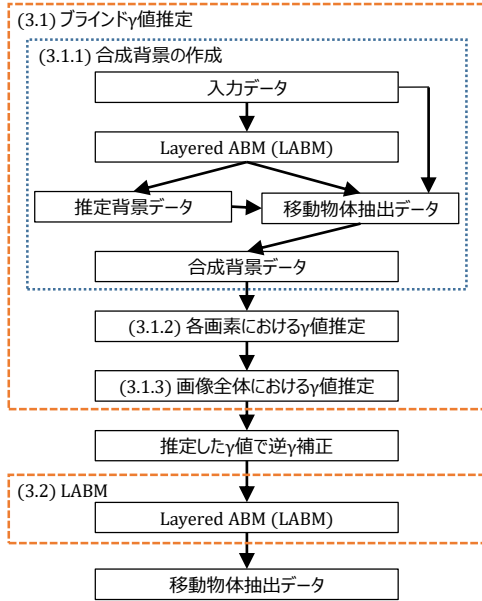


図1 IGLABMの流れ  
Fig.1 Process flow of IGLABM

### 3. 提案手法 IGLABM

ABMの問題点を改良した提案手法 (IGLABM) について説明する。IGLABM は、入力画像列の先頭部分からブラインド $\gamma$ 値推定を行う前処理部と、その後の複数の明るさ可変背景モデルを持つ背景差分法 (LABM, Layered Adaptive Background Model) で構成されている。図1に IGLABM の流れ図を示す。

#### 3.1 ブラインド $\gamma$ 値推定

入力画像中の背景画素の時間方向の短時間輝度変化を見たとき、リニア画像と $\gamma$ 補正画像とでは、RGB空間中の画素分布に違いが見られる。リニア画像の場合、図2のようにほぼ直線状に分布する。一方で、 $\gamma$ 補正画像の場合には、図3のように直線上からやや広がって分布する。この分布の性質を指標として $\gamma$ 値を推定する。この性質は観測期間中に画素に映っているものが同一対象でなければ成り立たないので、入力画像から移動物体 (前景部分) を取り除く必要がある。そこで、まず前景部分を取り除いた合成背景画像を作成し、その後 $\gamma$ 値推定を行う。

##### 3.1.1 合成背景画像の作成

入力画像  $I$  から前景部分を取り除くために、3.2で述べる LABM を適用して、物体検出画像  $O$ 、推定背

景画像  $E$  を作成する。このあと、以下の式のように、時刻  $t$  において物体として検出された入力画像  $I$  の位置  $x$  の画素  $I(x, t)$  を、推定背景画像  $E$  の対応する位置  $x$  の画素  $E(x, t)$  と置換えることで合成背景画像  $S(x, t)$  を作成する。

$$S(x, t) = \begin{cases} E(x, t) & (\text{物体領域 ; } O(x, t) = 1) \\ I(x, t) & (\text{背景領域 ; } O(x, t) = 0) \end{cases} \quad (1)$$

##### 3.1.2 各画素における $\gamma$ 値の推定

$\gamma$ 補正画像を正しい $\gamma$ 値で逆 $\gamma$ 補正した場合、背景画素は直線上に分布する。しかし、真の $\gamma$ 値は不明のため、全ての可能性のある $\gamma$ の候補値で $\gamma$ 変換を施し、その直線性を指標として用いて評価することで、より良い $\gamma$ 値を推定する。具体的には、入力画像列の先頭部分から  $N$  枚の画像を利用して、逆 $\gamma$ 補正後の各位置  $x$  の画素の共分散行列を求め、3つの固有値  $\lambda_1(x, \gamma) \geq \lambda_2(x, \gamma) \geq \lambda_3(x, \gamma) \geq 0$  を用いて、最適 $\gamma$ 値を見つけるための指標  $L(x, \gamma)$  を以下の式で計算する。図4に共分散行列より求めた3つの固有値の例を示す。

$$L(x, \gamma) = \frac{\lambda_2(x, \gamma) + \lambda_3(x, \gamma)}{\lambda_1(x, \gamma)} \quad (2)$$

この指標  $L(x, \gamma)$  は、太陽光が雲や人に遮られて照度が短時間で大きく変化する場合、共分散行列の直線性を表す  $\lambda_1$  が大きくなり、 $L(x, \gamma)$  が小さくなる。そこで、 $L(x, \gamma)$  は0に近ければ近いほど直線性があると評価できる。

次に、この指標を用いて各画素における $\gamma$ 値を以下の式で推定する。

$$\tilde{\gamma}_x = \arg \min_{\gamma \in \Gamma} L(x, \gamma),$$

ここで、 $\Gamma$  は  $\gamma$  の候補値の集合で、本論文の実験では、 $\Gamma = \{\gamma | \gamma = 0.5 + 0.1 \times k, k \in \mathbb{Z}, 0 \leq k \leq 20\}$  を使用した。

##### 3.1.3 画像全体における $\gamma$ 値推定

$\Gamma$  の様々な値ごとに、画像中の  $M$  点の位置  $x_i (i = 1, \dots, M)$  で前節で定義した指標  $L(x_i, \gamma)$  を計算し、各画素  $x_i$  における $\gamma$ 値  $\tilde{\gamma}_{x_i}$  と、そのときの指標値  $L(x_i, \tilde{\gamma}_{x_i})$  を  $M$  個取得する。例として、図5に、 $\gamma = 1.0$  の入力画像における指標値  $L$  の分布を示す。ここで、縦軸は指標値  $L$ 、横軸は対応する最大固

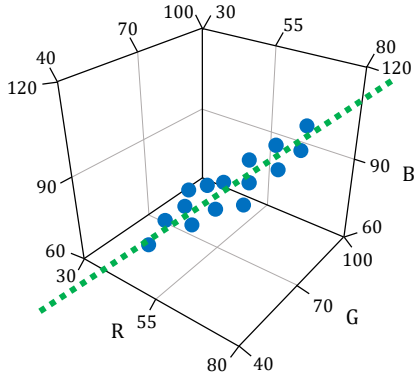


図 2 リニア画像の背景画素の RGB 分布  
Fig. 2 RGB distribution of the background pixel of linear images

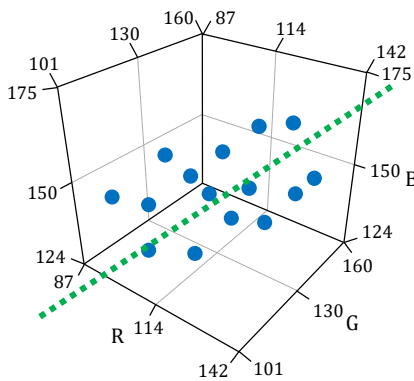


図 3  $\gamma$  補正画像の背景画素の RGB 分布  
Fig. 3 RGB distribution of the background pixel of gamma corrected images

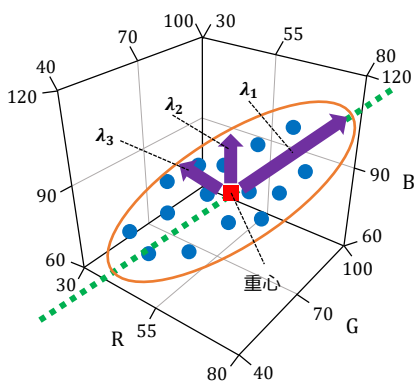


図 4 共分散行列の固有値  
Fig. 4 Eigenvalue of covariance matrix

有値  $\lambda_1$  である。  
このときに、逆  $\gamma$  補正の  $\gamma$  値が大きい場合は入力

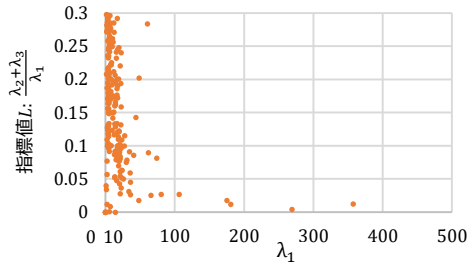


図 5 指標値と  $\lambda_1$  の分布図  
Fig. 5 distribution of index value L and eigenvalue  $\lambda_1$

画像が暗くなり、最も重要な固有値  $\lambda_1$  も小さくなる。したがって、 $\gamma$  値が大きい場合には、呼応して指標値  $L$  は小さくなり信頼性が低くなる。よって小さすぎる  $\lambda_1(x, \gamma)$  の画素は削除する必要がある。また、合成背景画像は完全に移動物体を除去しているわけではなく、背景の揺れによる他の背景色や物体色を含む場合もある。そのような画素では、共分散が大きくなり、結果として  $\lambda_1$  も大きくなってしまう。そのため、大きすぎる  $\lambda_1(x, \gamma)$  の画素も削除する必要がある。そこで、最大固有値  $\lambda_1$  に下限  $T_1$  と上限  $T_2$  を設定して、 $T_1 \leq \lambda_1(x, \gamma) \leq T_2$  を満たす画素のみを選別する。

また、背景画素が直線上に分布する場合、前節に述べたように指標値  $L$  が 0 に近くなる。そこで、大きすぎる指標値を持っている画素も削除する必要がある。そこで、上限値  $T_3$  を設定して、 $\lambda_1$  の制限に加えて、 $L(x, \gamma) \leq T_3$  も満たす画素のみを選別する。

以上のようにして外れ値を除去したあと、画像全体の  $\tilde{\gamma}$  を、候補画素の  $\tilde{\gamma}_{x_i}$  から  $\Gamma$  への投票により決定する。投票結果の上位  $K$  個から、以下の式を用いて投票数に応じた重み付き平均により推定を行う。

$$\tilde{\gamma} = \frac{\sum_{i=1}^K n_i \gamma_i}{\sum_i n_i} \quad (n_1 \geq n_2 \geq \dots \geq n_{\#\Gamma}) \quad (3)$$

ここで、 $n_i$  はビン  $i$  の投票数、 $\gamma_i$  はビン  $i$  に対応する  $\gamma$  値、 $\#\Gamma$  は、候補値の集合  $\Gamma$  の要素数を表す。本論文では、 $K = 5$  とし、 $\Gamma$  は 3.1.2 で定義したものと同一で  $\#\Gamma = 21$  である。

### 3.2 Layered Adaptive Background Model (LABM)

提案手法 IGLABM の一部である LABM (Layered ABM) について説明する。LABM は従来手法 ABM

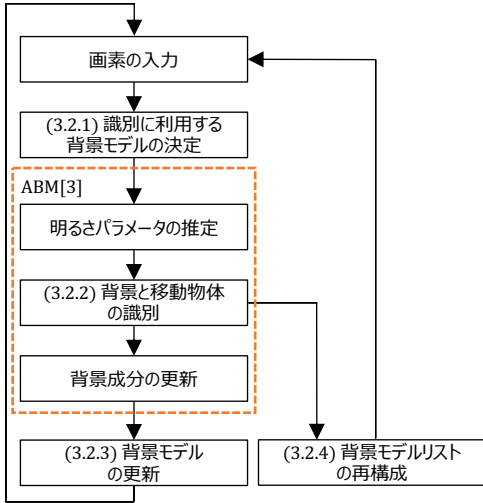


図 6 LABM の流れ図  
Fig. 6 System flow of LABM

を複数持たせることによって、草木の揺れといった背景変動に対応した手法である。基本的な考え方は、混合正規分布を利用した背景モデル [11] と同様であるが、混合正規分布の平均や分散の推定に ABM を利用する点が大きく異なる。すなわち、混合正規分布を使用する従来手法では、個々の正規分布が影の影響や背景変動、照明変動をそれぞれモデル化しているのに対し、我々の提案手法は、影の影響や照明変動を ABM で表現・吸収し、背景変動を混合正規分布で表現しようとしている点で考え方が大きく異なっている。

提案手法では、画像  $I$  の画素位置  $\mathbf{x}$  が時刻  $t$  において背景である確率密度  $p(\mathbf{x}, t|I)$  を以下のように近似して推定する。

$$\hat{p}(\mathbf{x}, t|I) = \sum_{i=1}^C w_i(\mathbf{x}, t) \mathcal{N}(I(\mathbf{x}, t); \mathbf{E}_i(\mathbf{x}, t), \hat{\sigma}_i^2(\mathbf{x}, t) \mathbf{I}_3) \quad (4)$$

ここで、 $C$  は ABM の混合数、 $w_i$  は混合重み、 $\mathbf{E}_i$  は  $i$  番目の ABM の推定背景画像、 $\sigma_i^2$  は  $i$  番目の ABM が推定している背景画素の分散で、 $\mathbf{I}_3$  は 3 次の単位行列を表す。本論文では、 $C = 6$  とした。LABM (Layered ABM) の処理の流れを図 6 に示す。

### 3.2.1 識別に使用する背景モデル数の決定

混合重み  $w_i$  は、 $i$  番目の ABM がどの程度背景であ

るかを示している尤度とも考えることができる。ABM の重み  $w_i$  を用いて背景尤度とし、背景モデルリストを背景尤度の大きい順に並び替えた ABM の中から式 (5) を満たす上位  $c$  個の ABM のみを用いる。この処理により尤度の低い ABM は使用されなくなり、背景推定の精度の向上が図れるのと、新しく追加された ABM はまだ信頼度も低いのですぐには背景として使用しないようにすることができる。

$$\frac{1}{Z} \sum_{k=1}^{c-1} w_{i_k} < T_4 \leq \frac{1}{Z} \sum_{k=1}^c w_{i_k} \quad (w_{i_1} \geq \dots \geq w_{i_c}) \quad (5)$$

ここで、 $Z$  は混合重みの正規化係数、 $T_4$  は物体抽出に使用する ABM の混合数を決定するためのしきい値である。本論文では、 $T_4 = 0.9$  とした。

### 3.2.2 背景と物体の識別

時刻  $t$  において入力画像  $I$  の位置  $\mathbf{x}$  の画素が、背景 ( $B$ ) か前景 ( $F$ ) かのベイズ決定則は、以下のようになる。

$$R = \frac{p(B|\mathbf{x}, t, I)}{p(F|\mathbf{x}, t, I)} = \frac{p(\mathbf{x}, t|I, B)p(B)}{p(\mathbf{x}, t|I, F)p(F)} \quad (6)$$

背景と前景に関して事前知識が無いので、背景と前景の事前確率密度は等しく  $p(B) = p(F)$ 、前景は一樣分布  $p(\mathbf{x}, t|I, F) = c_F$  を仮定する。以上の仮定と  $p(\mathbf{x}, t|I, B) \sim \hat{p}(\mathbf{x}, t|I)$  から、背景の決定則は以下のように変形できる。

$$O(\mathbf{x}, t) = \begin{cases} 1 \text{ (物体)} & \hat{p}(\mathbf{x}, t|I) < T_5 \\ 0 \text{ (背景)} & \hat{p}(\mathbf{x}, t|I) \geq T_5 \end{cases} \quad (7)$$

ここで、 $T_5 (= R \cdot c_F)$  は物体と背景を分離するためのしきい値である。本論文では、 $T_5 = e^{-0.9}$  を使用した。

### 3.2.3 背景モデルの更新

3.2.2 で画素が背景であると識別された場合に ABM を更新する。ここで、入力に対して全ての ABM を更新するのではなく、最も高い確率密度を出している ABM のみを更新する。その理由は、各 ABM が背景の揺れによる背景をそれぞれモデル化しており、現在観測されている画素はその背景のどれか一つだけだからである。ABM の更新は従来手法 [3] と同一である。その後、全ての ABM の重み  $w_i$  を更新する。重み  $w_i$  の更新手法は、従来手法 [11] と同様に行う。



図7 リニア画像データセット：Tree & Flag  
Fig.7 Examples of the Tree and Flag datasets

### 3.2.4 背景モデルリストの再構築

3.2.2 で対象画素が物体と識別された場合、背景尤度が低い背景モデルを棄却し、新たな背景モデルを構築する。背景モデルリストの中から、背景尤度が一番低い ABM を棄却し、入力画像  $I(x, t)$  を使用して新しい ABM を作成する。そのとき新しく作成された ABM の重み  $w_i$  の初期値は小さくしておく。これは新しい物体が観測されたときには、移動物体なのか、草木の揺れによる背景変動なのかをすぐには判別できないため、尤度の低い背景モデルとして背景モデルリストに追加するためである。もし、背景であれば定期的に観測されるので尤度が上昇し、背景モデルとして利用されるようになる。一方移動物体の場合は、尤度は上昇しないので、いずれ背景モデルリストから外れて、背景モデルとして利用されないようになる。なお、実験では、重み  $w_i$  の初期値を従来手法 [11] と同様に  $w_i = 0.002$  とした。

### 3.2.5 LABM の初期化

ここで、LABM の初期化について説明する。LABM は起動時に全てのハイパーパラメータの設定を行い、画像の重み、ABM を全て 0 で初期化する。次に、最初の入力画像が LABM に与えられたときに画像のサイズが確定するので、ここでレイヤ数  $C$  分の ABM の画像領域の確保が行われる。このときに、最初のレイヤの ABM は入力画像を背景として全画素を入れ替えて、重みとして  $w_i$  を設定する。以降は、この節で説明したアルゴリズムで背景の各画素が更新されていく。

## 4. 評価実験

### 4.1 評価用データセットの概要

提案手法の性能を評価ために、 $\gamma$  補正をオフにした工業用カメラ (POINT GREY FL2-08S2C) で撮影した2つのリニア画像データセット “Tree” と “Flag” を用意した。これらのリニア画像データセットには、雲などによる照明変動、移動物体と移動物体の影、風

表 1 真の  $\gamma$  値の推定値

Table 1 Estimation results of true  $\gamma$  values

Dataset	真の $\gamma$ 値
Tree	1.00
Flag	1.00
Dataset3	
camera1	0.80
Fountain01	0.50
Fountain02	1.80
Canoe	0.90
Boats	1.10
Overpass	1.50
Fall	2.00
Backdoor	1.00
BusStation	1.00

による草木の揺れ、旗のはためきといった周期的な背景変動が含まれている。“Tree”, “Flag” の例を図7に示す。

また  $\gamma$  値が未知の画像データセットとして、PETS2001 [26], CDW-2012 [27] から、ABM の前提条件である屋外で連続撮影された昼間の画像データセットを使用した。具体的には、PETS2001 より、“dataset3 testing camera1”, CDWS-2012 より、Dynamic Background カテゴリの “fountain01”, “fountain02”, “canoe”, “boats”, “overpass”, “fall”, Shadow カテゴリから、“backdoor”, “busStation” を使用した。なお、これらのデータセットは  $\gamma$  値が未知であるため、真の  $\gamma$  値を推定する必要がある。本論文では、ABM が  $\gamma = 1.0$  のときに最も高い性能を出すという性質を利用して、様々な  $\gamma$  値で画像データセットに逆  $\gamma$  補正を施し、その画像データセットに ABM を適用した結果、最も性能が高かった  $\gamma$  値を真の  $\gamma$  値の推定値として利用した。表 1 にその推定値を示す。

### 4.2 合成背景を作成する有効性の検証

提案手法のブラインド  $\gamma$  値推定を行う前に、合成背景を生成することの有効性を実験により検証する。

それぞれの画像データセットには真値となる物体検出の結果画像が含まれているが、各画像データセットには真値のある入力画像までに、背景のオンライン学習が可能ないようにリード部分の画像列が先頭から存在している。本実験でもこのリード部分の画像列を使用して、ブラインド  $\gamma$  推定を行う。画像のリード部分の長さは、各画像データセットによって異なるために、評価のためにはこれを揃える必要がある。そこで、各画像データより先頭フレームから人物領域の真値フレームまでの間 (リード部分) の画像列から無作為復



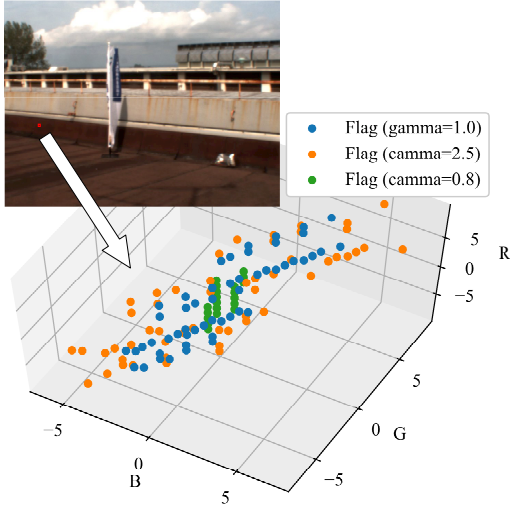


図 8 “Flag” の画素位置 (101,363) とその位置における画素の分布  
Fig. 8 Position and distribution of pixels at (101,363) in “Flag”

元抽出を行い画像 100 枚をそれぞれの画像データセットで用意した。

この画像 100 枚から、各  $\gamma \in \Gamma$  で逆  $\gamma$  補正画像を入力画像とし、入力画像から作成した合成背景画像を利用してブラインド  $\gamma$  値推定を行った。このときの“Flag”の位置 (101,363) における  $\gamma = 0.8, 1.0, 2.5$  のときの、画素の分布を図 8 の下部に示す。また、図 8 の上部に赤で囲んだエリアの中心部分が (101,363) を示している。なお、図の各画素値は平均が原点にくるように平均を引いてある。このときの指標  $L$  の値はそれぞれ 0.045, 0.038, 0.065 であった。 $\gamma = 1.0$  のときに最も指標  $L$  が小さくなっている。

さらに、 $\gamma$  値推定の比較結果を図 9 に示す。横軸は指標値  $L$  のしきい値  $T_3$ 、縦軸は  $\gamma$  値推定の平均絶対誤差である。 $\gamma$  値の推定値は他のしきい値  $T_1, T_2$  によっても変化する。この評価実験では、 $M = 400$  に固定し格子状等間隔にサンプルした。また  $T_1$  の値を、5, 10, 15, 20,  $T_2$  の値を、100, 200, 300, 400, 500, 1,000 に設定した。これによる推定誤差の結果も図 9 に示している。図 9 より、入力画像より、合成背景画像を使用したほうが  $\gamma$  値の推定精度が高いことが分かる。これより、入力画像から物体を除去した合成背景画像を作成するという提案手法の有効性が検証できた。

さらに、図 9 より  $T_3 = 0.05$  のときに最小値になっているが、このときの  $T_1, T_2$  による  $\gamma$  値推定平均誤差

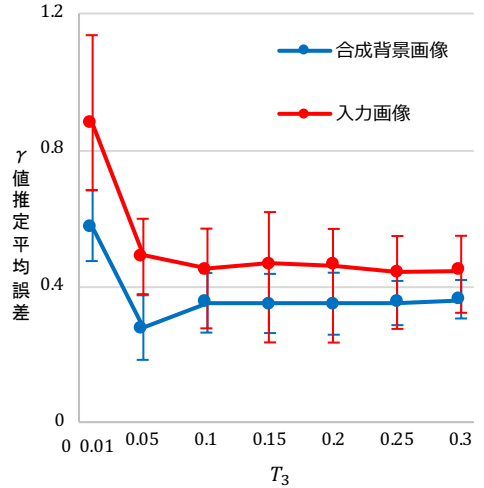


図 9 IGLABM による推定  $\gamma$  値推定平均誤差  
Fig. 9 Accuracy of gamma estimation

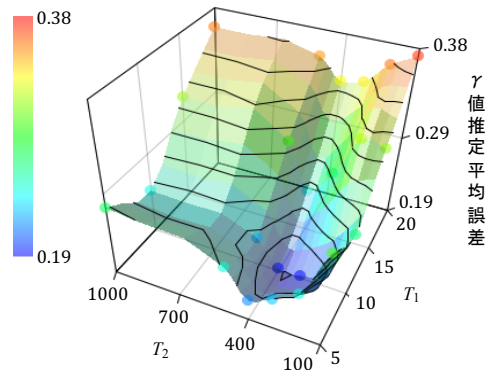


図 10  $T_3 = 0.05$  のときの  $T_1, T_2$  による  $\gamma$  値推定平均誤差

を図 10 に示す。この図より、 $T_1 = 10, T_2 = 400$  のときに  $\gamma$  値推定平均誤差が最小値をとることが分かる。以降、本論文の他の実験では、この実験で得られた最適パラメータ値  $T_1 = 10, T_2 = 400, T_3 = 0.05$  を使用した。

#### 4.3 推定点数による $\gamma$ 値推定誤差の評価

3.1.3 節で説明したように、提案手法は画像中の  $M$  点を使用して  $\gamma$  値を推定するため、選ばれた点は背景である必要がある。 $M$  が小さい場合は理想的な背景画素をサンプルする可能性が低くなり、 $M$  が大きい場合は物体領域をサンプルする可能性がある。そこで、本実験では  $M$  を  $5 \times 5$  (25 点),  $10 \times 10$  (100 点),  $20 \times 20$  (400 点),  $40 \times 40$  (1600 点),  $80 \times 80$  (6400 点) に設定

表 2  $M$  の値による  $\gamma$  値推定平均誤差

$M$ (X×Y)	5×5	10×10	20×20	40×40	80×80
$\gamma$ 値推定平均誤差	0.48	0.45	0.19	0.38	0.29

表 3 IGLABM と他手法による推定  $\gamma$  値の比較  
Table 3 Absolute estimation error of gamma values

Dataset	IGLABM	MV-Gamma	BIGC
Tree	<b>0.05</b>	1.10	2.14
Flag	0.17	<b>0.10</b>	0.36
Dataset3			
camara1	<b>0.07</b>	0.10	1.69
Fountain01	<b>0.30</b>	<b>0.30</b>	2.58
Fountain02	<b>0.38</b>	1.00	0.53
Canoe	0.25	0.20	<b>0.12</b>
Boats	0.37	<b>0.00</b>	0.30
Overpass	<b>0.06</b>	0.70	0.84
Fall	<b>0.04</b>	0.90	0.69
Backdoor	0.33	<b>0.10</b>	0.55
BusStation	<b>0.10</b>	0.20	1.07
Average	<b>0.19</b>	0.43	0.99

し、格子状等間隔にサンプルして、提案手法の  $\gamma$  値推定平均推定誤差を評価した。

推定結果を表 2 に示す。結果より、 $20 \times 20$  (400 点) が最も良い精度であった。以降、本論文の実験には  $M = 20 \times 20 = 400$  を使用した。

#### 4.4 提案手法と他手法の $\gamma$ 値推定精度の比較

提案手法における  $\gamma$  値推定の有効性を検証するため、他手法による  $\gamma$  値推定精度を比較検証した。使用した画像データセットは、前節の実験と同様である。比較手法として、提案手法と同様に環境光や撮像装置に関する情報がない場合でも、 $\gamma$  値推定が可能なブラインド  $\gamma$  推定手法を用いた。実験で比較する手法として、MV-Gamma [22], BIGC [23] を用いた。

それぞれの手法で推定した  $\gamma$  値と真の  $\gamma$  値の推定値との絶対誤差の結果を表 3 に示す。表 3 より、IGLABM の絶対誤差の平均が一番小さいことから、提案手法である IGLABM の  $\gamma$  値推定の有効性を確認した。

#### 4.5 提案手法と他手法による物体検出精度の比較

提案手法である IGLABM の有効性を検証するため、他手法との物体検出精度を比較検証した。比較には IGLABM と同じ画素ベースでの背景差分法である MOG [10], MOG2 [11], SOBS [12] と、領域ベースの背景差分法である GMG [13], ViBe [15], PBAS [16], PAWCS [17] を用いた。実験には、OpenCV または、CDW-2012 から入手したソースコードを利用し、CDW-2012 で公開されているパラメータにはその

値を、公開されていないパラメータに関してはソースコードのデフォルト値を使用した。また、手法を公平に評価するために、孤立点除去などのモルフォロジー処理を後処理として行っているソースコード<sup>(2)</sup>に対しては、その処理の直前の画像で評価を行った。また、物体検出の精度評価方法には CDW-2012 と同様の方法<sup>(3)</sup>で計算した F 値を用いた。

表 4 に人物検出結果の F 値を示す。表中の左側 5 手法が画素ベースの手法で、右側 4 手法が領域ベースの手法である。表 4 に示すように、モルフォロジー処理のような後処理を行わない場合には、IGLABM は、画素ベースの手法だけでなく領域ベースの手法を含めて、他の背景差分法と比べて高い平均性能を示しており、提案手法の有効性を確認できた。

また、我々以外の他手法の中で、人物検出結果の平均性能が高い上位 4 つの手法の実験結果の一部を図 11,12 に示す。図 11,12<sup>(4)</sup> より、IGLABM は他の背景差分法に比べ、旗のはためきと人の重なりや波の揺れなどの背景変動を物体や背景として、より正確に認識できていることがわかる。

表 5 に、参考として CDW-2012 から引用した他手法の F 値の値を示す [27]。この表の CDW の値と、SOBS や PAWCS での我々の結果 (表 4) の値とを比較すると、モルフォロジー処理は“fountain01”, “fountain02”, “canoe”, “boats”, “overpass” など小さい領域で変化の激しいときに誤識別したもの (孤立点) を除去する効果が高いが、“busStation” のような変化が少ない画像には効果がありもないことも判る。前景領域を取り囲むようにあるソフトシャドウを F 値に組み入れない CDW-2012 の評価方法は前景領域を  $3 \times 3$  のモルフォロジー処理で縮小拡大しても F 値に影響しないのでこのようなことが起きていると推定される。

## 5. まとめ

本論文では、従来手法 ABM の問題であった草木の揺れといった周期的な背景変動および、 $\gamma$  値が未知の画像に適応できない点に対応した手法 IGLABM を提案した。IGLABM と他手法を比較することで、 $\gamma$  値推定精度および、物体検出精度で高い性能を示し、

(2) : ViBe, SOBS, PAWCS

(3) : 調整されたパラメータ 1 セットのみ使用する。真値画像で背景と影になっている領域を背景とし、ソフトシャドウや評価対象外領域は評価に入れない F 値計算方法。

(4) : 下側の真値画像の灰色はソフトシャドウや境界領域を表し、CDW-2012 の評価方法では F 値の算出には入らない領域である。



表 4 IGLABM と他手法による F 値比較  
Table 4 Comparative analysis of  $F_1$ -score

Method \ Dataset	Pixel based					Region based			
	IGLABM	ABM	MOG	MOG2	SOBS	GMG	ViBe	PBAS	PAWCS
Tree	<b>0.91</b>	0.90	0.88	0.84	0.83	0.71	0.87	0.84	0.85
Flag	<b>0.72</b>	0.50	0.70	0.63	0.41	0.44	0.49	0.62	0.49
Dataset3									
camara1	0.52	0.48	0.60	0.31	0.07	0.62	0.53	<b>0.67</b>	0.50
Fountain01	0.10	0.06	0.09	0.04	0.04	0.04	0.09	<b>0.39</b>	0.18
Fountain02	0.69	0.54	0.63	0.28	0.39	0.37	0.68	<b>0.77</b>	0.48
Canoe	0.85	<b>0.88</b>	0.48	0.61	0.38	0.19	0.75	0.02	0.76
Boats	0.37	<b>0.46</b>	0.11	0.17	0.10	0.14	0.22	0.07	0.24
Overpass	0.63	0.61	0.52	0.55	0.47	0.49	0.68	0.53	<b>0.73</b>
Fall	0.53	0.23	0.55	0.34	0.18	0.16	0.43	<b>0.81</b>	0.57
Backdoor	0.81	0.76	0.84	0.72	0.76	<b>0.91</b>	0.78	0.90	0.86
BusStation	0.70	0.74	0.49	0.64	0.81	0.68	0.69	0.53	<b>0.86</b>
Average	<b>0.62</b>	0.56	0.54	0.46	0.40	0.43	0.56	0.56	0.59

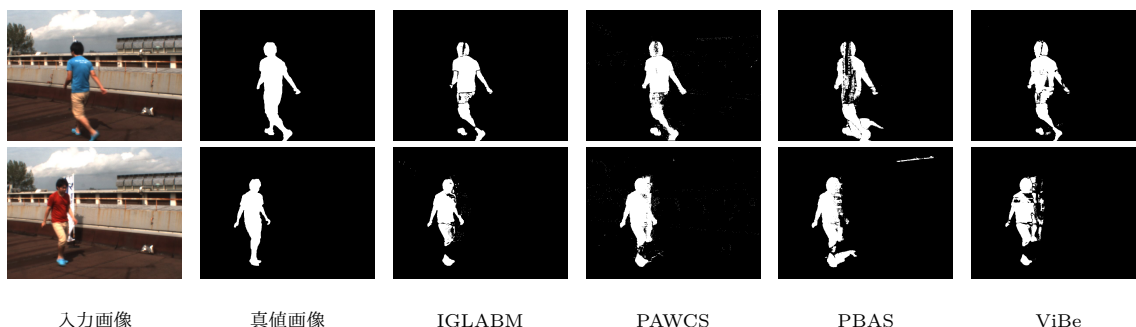


図 11 Flag (frame No. 123, 528) における抽出結果の比較  
Fig. 11 Comparison of extraction results in “Flag” (frame numbers 123, 528)

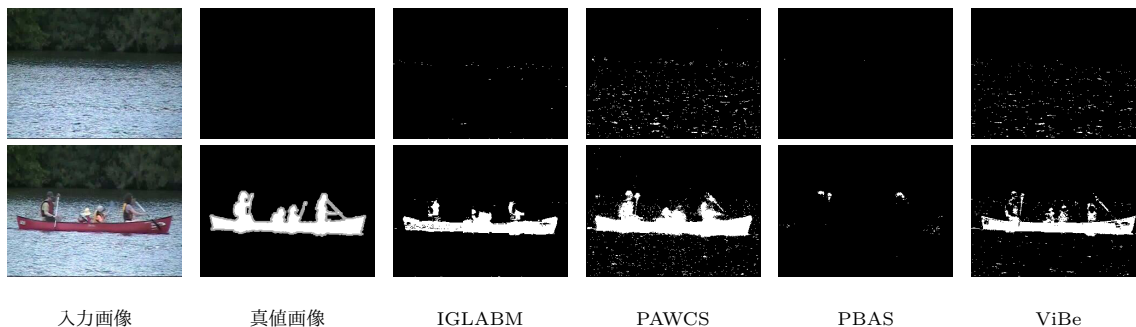


図 12 Canoe (frame No. 510, 962) における抽出結果の比較  
Fig. 12 Comparison of extraction results in “Canoe” (frame numbers 510, 962)

IGLABM の有効性を検証した。

今後の課題としては、IGLABM の各種パラメータの調整方法の自動化や IGLABM のリアルタイム推定化や領域ベースの手法への拡張などがあげられる。

謝辞 本研究の一部は、総務省 SCOPE (No. 172308003), および JSPS 科研費 JP17K00238, JP18H04114, 文部科学省による Society 5.0 実現

化研究拠点支援事業の助成を受けたものである。

文 献

[1] I. Setitra, and S. Larabi, “Background subtraction algorithms with postprocessing: A review,” Proc. Int. Conf. Pattern Recognit., pp.2436–2441, Stockholm, Sweden, Aug. 2014. DOI:10.1109/ICPR.2014.421

[2] S. K. Choudhury, P. K. Sa, S. Bakshi, and B. Majhi, “An evaluation of background subtraction for ob-

表 5 他手法の CDW-2012 での値  
Table 5  $F_1$ -scores of existing methods at CDW-2012

Method	Pixel based		Region based	
	MOG2	SOBS	PBAS	PAWCS
Fountain01	0.08	0.11	0.42	<b>0.78</b>
Fountain02	0.79	0.86	<b>0.94</b>	<b>0.94</b>
Canoe	0.89	0.94	0.36	<b>0.94</b>
Boats	0.75	0.81	0.72	<b>0.84</b>
Overpass	0.87	0.85	0.79	<b>0.96</b>
Fall	0.42	0.29	0.87	<b>0.91</b>
Backdoor	0.63	0.83	0.88	<b>0.95</b>
BusStation	0.79	0.82	0.86	<b>0.87</b>

ject detection vis-a-vis Mitigating Challenging Scenarios,” *IEEE Access*, vol.4, pp.6133–6150, 2016. DOI:10.1109/ACCESS.2016.2608847

- [3] H. Yoshimura, Y. Iwai, and M. Yachida, “Object detection with adaptive background model and margined sign cross correlation,” *Proc. Intl. Conf. Pattern Recognit.*, vol.III, pp.19–23, Hong Kong, 2006. DOI:10.1109/ICPR.2006.878
- [4] A.H.S. Lai, and N.H.C. Yung, “A fast and accurate scoreboard algorithm for estimating stationary backgrounds in an image sequence,” *Proc. Int. Symp. Circuits Syst.*, vol.4, pp.241–2444, Monterey, CA, USA, May 1998. DOI:10.1109/ISCAS.1998.698804
- [5] C.R. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, “Pfinder: realtime tracking of the human body,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.19, no.7, pp.780–785, July 1997. DOI:10.1109/34.598236
- [6] A. Elgammal, D. Harwood, and L. Davis, “Non-parametric model for background subtraction,” *Proc. Eur. Conf. Comput. Vis.*, pp. 751–767, Dublin, Ireland, June 2000.
- [7] W. Kim, and C. Kim, “Background subtraction for dynamic texture scenes using fuzzy color histograms,” *IEEE Signal Process. Lett.*, vol.19, no.3, pp.127–130, Mar. 2012. DOI:10.1109/LSP.2011.2182648
- [8] J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma, “Robust principal component analysis: exact recovery of corrupted low-rank matrices via convex optimization,” *Proc. Adv. Neural Inf. Process. Syst.*, pp.2080–2088, 2009.
- [9] C. Stauffer, and W. Grimson, “Adaptive background mixture models for real-time tracking,” *Proc. Conf. Comput. Vis. Pattern Recognit.*, vol.2, pp.246–252, Fort Collins, CO, USA, June 1999. DOI:10.1109/CVPR.1999.784637
- [10] P. KaewTraKulPong and R. Bowden, “An improved adaptive background mixture model for real-time tracking with shadow detection,” in *Video-Based Surveillance Systems*, ed. P. Remagnino, G.A. Jones, N. Paragios and C.S. Regazzoni, pp.135–144, Springer, Boston, MA, 2002. DOI:10.1007/978-1-4615-0913-4\_11
- [11] Z. Zivkovic, “Improved adaptive Gaussian mixture model for background subtraction,” *Proc. Intl. Conf. Pattern Recognit.*, vol.2, pp.28–31, Cambridge, UK, 2004. DOI:10.1109/ICPR.2004.1333992
- [12] L. Maddalena and A. Petrosino, “A self-organizing approach to background subtraction for visual surveillance applications,” *IEEE Trans. Image Process.*, vol.17, no.7, pp.1168–1177, July 2008. DOI:10.1109/TIP.2008.924285
- [13] A. B. Godbehere, A. Matsukawa and K. Goldberg, “Visual tracking of human visitors under variable-lighting conditions for a responsive audio art installation,” *Proc. 2012 Am. Control Conf.*, pp.4305–4312, Montreal, QC, Canada, 2012. DOI:10.1109/ACC.2012.6315174
- [14] H. Wang and D. Suter, “A consensus-based method for tracking: modeling background scenario and foreground appearance,” *Pattern Recognit.*, vol.40, no.3, pp.1091–1105, Mar. 2007. DOI:10.1016/j.patcog.2006.05.024
- [15] O. Barnich and M. V. Droogenbroeck, “ViBe: a universal background subtraction algorithm for video sequences,” *IEEE Trans. Image Process.*, vol.20, no.6, pp.1709–1724, June 2011. DOI:10.1109/TIP.2010.2101613
- [16] M. Hofmann, P. Tiefenbacher and G. Rigoll, “Background segmentation with feedback: the pixel-based adaptive segmenter,” *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, pp.38–43, Providence, RI, USA, 2012. DOI:10.1109/CVPRW.2012.6238925
- [17] P. St-Charles, G. Bilodeau and R. Bergevin, “A self-adjusting approach to change detection based on background word consensus,” *Proc. 2015 IEEE Winter Conf. on Appl. of Comput. Vis.*, pp.990–997, Waikoloa, HI, USA, 2015. DOI:10.1109/WACV.2015.137
- [18] J. Hao, C. Li, Z. Kim and Z. Xiong, “Spatio-temporal traffic scene modeling for object motion detection,” *IEEE Trans. Intell. Transportation Syst.*, vol.14, no.1, pp.295–302, Mar. 2013. DOI:10.1109/TITS.2012.2212432
- [19] N. Liu, H. Wu and L. Lin, “Hierarchical ensemble of background models for ptz-based video surveillance,” *IEEE Trans. Cybernetics*, vol.45, no.1, pp.89–102, Jan. 2015. DOI:10.1109/TCYB.2014.2320493
- [20] P. St-Charles, G. Bilodeau and R. Bergevin, “A self-adjusting approach to change detection based on background word consensus,” *Proc. 2015 IEEE Winter Conf. Appl. of Comput. Vis.*, pp.990–997, Waikoloa, HI, USA, Jan. 2015. DOI:10.1109/WACV.2015.137
- [21] S.A Shafer, “Using color to separate reflection components,” *Color: Research and Application*, vol.10, no.4, pp.210–218, 1985. DOI:10.1002/col.5080100409
- [22] M. Mahamdioua and B. Mohamed, “New mean-

variance gamma method for automatic gamma correction,” Intl. J. of Image, Graphics and Signal Process., vol.9, no.3, pp.41–54, Mar. 2017. DOI:10.5815/ijigsp.2017.03.05

- [23] H. Farid, “Blind inverse gamma correction,” IEEE Trans. Image Process., vol.10, no.10, pp.1428–1433, Oct. 2001. DOI:10.1109/83.951592
- [24] S. Lin, J. Gu, S. Yamazaki and H.-Y. Shum, “Radiometric calibration from a single image,” Proc. 2004 IEEE Conf. Comput. Vis. Pattern Recognit., pp.II-II., Washington, DC, USA, June 2004. DOI:10.1109/CVPR.2004.1315266
- [25] K. Sakamoto, H. Yoshimura, M. Nishiyama and Y. Iwai, “GLABM: gamma corrected layered adaptive background model for outdoor scenes,” Proc. IEEE 7th Global Conf. Consumer Electronics, pp.66–67, Nara, Japan, Oct. 2018. DOI:10.1109/GCCE.2018.8574717
- [26] D. Young and J. Ferryman. “PETS metrics: online performance evaluation service,” Proc. IEEE Int. Workshop on Vis. Surv. and Perf. Eval. of Tracking and Surv., pp.317–324, Beijing, China, Oct. 2005. DOI:10.1109/VSPETS.2005.1570931
- [27] N. Goyette, P.-M. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, “Changefetection.net: A new change detection benchmark dataset,” Proc. 2012 IEEE Conf. Comput. Vis. Pattern Recognit. Workshop, pp.1–8, June 2012. DOI:10.1109/CVPRW.2012.6238919

## 付 録

### 1. 明るさ可変背景モデル (ABM) [3]

この節では、明るさ可変背景モデルについて簡単に説明する。詳しくは参考文献を参照されたい。屋外において、雲の影響による急激な明るさ変化が起きている際に観測される物体の反射色について考える。短時間では環境光及び太陽直射光 (以下、直射光と記述する) の色度の変化は小さいことから、これを無視すると、時刻  $t$  において画素  $\mathbf{x}$  で観測される物体色  $\mathbf{Y}(\mathbf{x}, t)$  は、環境光成分  $\mathbf{X}_e(\mathbf{x}, t)$  と直射光成分  $\mathbf{X}_d(\mathbf{x}, t)$  からなると考えられる。雲は白～灰色の無彩色をしており、このことから特定の色だけを吸収することがないと思われる。そこで、光が雲を通過する前後では光のスペクトル分布に変化がないと仮定し、直射光が雲を通過するときだけに明るさが変化すると考えれば、雲の影響により明るさが変化しているときに観測される物体色は以下の2色性反射モデルで説明できる。

$$\mathbf{E}(\mathbf{x}, t) = \mathbf{X}_e(\mathbf{x}, t) + k(\mathbf{x}, t)\mathbf{X}_d(\mathbf{x}, t) \quad (\text{A}\cdot 1)$$

式 A.1 において、 $k(\mathbf{x}, t)$  ( $0 \leq k \leq 1$ ) は混合率を表す。この混合率  $k(\mathbf{x}, t)$  を明るさパラメータとして、

様々に変化させれば、急激な明るさの変化や物体の影による影響に対して頑健に対応することが可能となる。

(平成 xx 年 xx 月 xx 日受付)

### 阪本 光翼

2018 年 鳥取大学工学部知能情報工学科卒業。2019 年 鳥取大学大学院持続性社会創生科学研究科工学専攻博士前期課程に在学中。

### 吉村 宏紀 (正員)

1993 年鳥取大学工学部知能情報工学科卒業。1997 年鳥取大学大学院工学研究科博士後期課程情報生産工学専攻修了。博士(工学)。現在鳥取大学大学院工学研究科情報エレクトロニクス専攻、助教。画像、音声信号処理の研究に従事。電気学会、情報処理学会会員。

### 西山 正志 (正員)

2000 年 岡山大学工学部情報工学科卒業。2002 年 同大学院博士前期課程了。同年株式会社東芝入社。同社研究開発センターを経て、現在鳥取大学大学院工学研究科准教授。2011 年 東京大学大学院学際情報学府にて博士(学際情報学)を取得。カメラを用いた人物認識を始めとするパターン認識およびインタラクションの研究に従事。山下記念研究賞など受賞。電子情報通信学会、情報処理学会各会員。

### 岩井 儀雄 (正員)

1992 年 (平成 4 年) 大阪大学基礎工学部情報工学科卒業。1997 年 (平成 9 年) 大阪大学大学院基礎工学研究科博士課程後期修了。同年同大学院助手。2003 年 (平成 15 年) 同大学院助教授。2004 年 (平成 16 年)5 月～2005 年 (平成 17 年)3 月英国ケンブリッジ大学客員研究員。2007 年 (平成 19 年) 同大学院准教授。2011 年 (平成 23 年) 鳥取大学大学院工学研究科教授。コンピュータビジョン、パターン認識の研究に従事。博士(工学)。