

テクスチャ解析偏 GLCM を用いた猿検出実験

実験の目的

これまで、可能性尺度空間への写像方法として、CNN を用いてラスタスキャンを行う方法と、画素ごとに時系列の輝度ヒストグラムを作成し、中央値を用いて背景画像を作成する方法を実装してきた。その他の可能性尺度画像空間への写像方法として、画像処理と機械学習を組み合わせた方法などが検討できる (HOG 特徴+SVM など)。今回は、サル画像のテクスチャ解析を行う。手法としては、GLCM を利用する。

1 GLCM について

(Grey Level Co-occurrence Matrix:GLCM) は、テクスチャ画像から特徴を抽出する際に作成する行列のことをいう。日本では一般的に濃度共起行列 (Co-occurrence Matrix) もしくは空間濃度レベル依存方 (spatial gray level dependence method:SGLDM) と呼ばれている (文献 [6] 参照)。この手法は 1973 年に Haralick らによって提案された手法であり [2], テクスチャ解析の分野で広く利用されている。

1.1 GLCM の作成方法

元の画像はグレースケール変換等が行われ、濃度値で表現されているものとする。以降、GLCM を $V(s, t; d, \theta)$ で表す。ただし、元の画像において、画素 s と画素 t の対を (s, t) とする。また、基準とする画素 s に対して参照する画素 t の相対位置を (θ, d) で表す。図 1 は、 (θ, d) の位置関係を図示したものである。図 1 左下の黒点が基準点であり、青、緑、赤色の点が参照点を表す。また、距離 d の定義は式 (1) で表す。ただし、画素 s の座標を (x_s, y_s) とする。

$$d = \max(|x_s - x_t|, |y_s - y_t|) \quad (1)$$

θ には、 $0^\circ, 45^\circ, 90^\circ, 135^\circ$ を与える (理由後述)。

なお、2 点以上の n 個の画素を使用する場合もある [4]。

step1

画素の対 i , と j における濃度値を求める。そして、GLCM の対応する項に対して 1 を積算していく。手順を図 2 に示す。ただし左側が元の画像を表し、右側が濃度共起行列である。また、例では元の画像が 18 の 8 段階の離散的な値によって表現されているため、GLCM は 8 階の正方行列となっている。

step2

テクスチャ解析を行う際には、GLCM からスカラー値 (コントラスト等を表現) に変換することで、解析を行える形になる。このために作成した GLCM の転置行列を作成し、GLCM と GLCM の転置行列の各項を足し合わせる (式 (2))。

$$V(i, j) = V(i, j) + V^T(i, j) \quad (2)$$

ただし i は GLCM の行, j は GLCM の列を表す。さらに、足し合わせた転置行列の全ての項の和をとり、各項を和の値で割る (式 (3))。

$$P(i, j) = \frac{V(i, j)}{\sum_{i,j=1}^N V(i, j)} \quad (3)$$

ただし N は行もしくは列の数であり、例の場合 $N = 8$ である。また、正規化された行列を \mathbf{P} とする。

元の画像を左右、もしくは上下に反転させ、 V を作成すると、かならず $V(i, j) = V^T(i, j)$ となる。よって、逆向きの θ に対する濃度変化は同一と見なしている。このため、 θ には、 $0^\circ, 45^\circ, 90^\circ$ のみを与えている。また、画素 s の濃度 $f(s)$ を得た時に、画素 t の濃度 $f(t)$ が同時に得られる確率を \mathbf{P} は表すので、同時生起確率という。

以降の解析では、各指標の作成時にこの行列 \mathbf{P} を使用する。

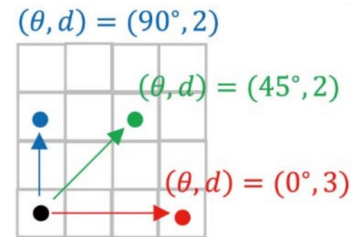


図 1: 基準点と参照点の位置関係定義 [10]

The definition of the angle (θ) and the pixel distance (d).

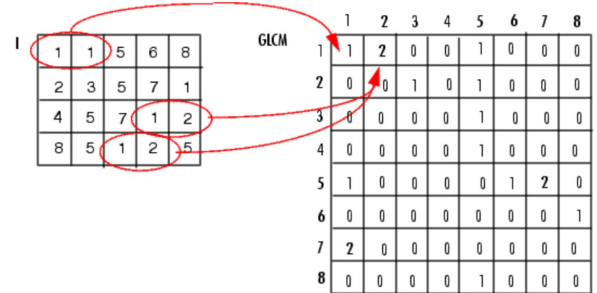


図 2: GLCM 行列の生成過程 [3]

1.2 GLCM から得られる特徴量

元論文 [2] では 14 個の特徴量が紹介されているが、一般的に使用される特徴量は数個に限られている。また python 画像処理用ライブラリの scikit-image においても全ての特徴量は実装されていないことを留意し、いくつかの特徴を紹介する。なお、特徴量を作成せずに、作成した同時共起行列 \mathbf{P} の各項自体をそのまま特徴量として NN などに入力する場合もある [7]。

1.2.1 CON (コントラストを調べる)

(Contrast : CON) (コントラスト:日訳) は,[6] においては (Inertia : 慣性) として紹介されている. 選択した2つの画素濃度の差分が, 大きいかな否かを示している. 使用例として, 画像処理によって浄水場の廃棄土を分析する際に用いた研究が報告されている [5].

CON は式 (4) で求める.

$$CON = \sum_{i=1}^N \sum_{j=1}^N (i-j)^2 P(i,j) \quad (4)$$

式 (4) は, 例えば 4×4 の正規化済み GLCM 行列 P に対して, 式 (5) のような重み行列との積をとり, 全要素の和を求めることを意味している [1].

$$\begin{bmatrix} 0 & 1 & 4 & 9 \\ 1 & 0 & 1 & 4 \\ 4 & 1 & 0 & 1 \\ 9 & 4 & 1 & 0 \end{bmatrix} \quad (5)$$

CON 値の特性として, 「細かい」画像は「粗い」画像と比較して大きい値をとる.

1.2.2 DIS (コントラストを調べる)

(Dissimilarity : 不均一性) は, 不均一性の高い画像である場合は大きな値となり, 均一な値であるほど0に近づく. IDM と逆の尺度. IDM 値の特性として, 「細かい」画像は「粗い」画像と比較して小さい値をとる. DIS は式 (6) で求める.

$$DIS = \sum_{i=1}^N \sum_{j=1}^N |i-j| P(i,j) \quad (6)$$

式 (6) は, 例えば 4×4 の正規化済み GLCM 行列 P に対して, 式 (7) のような重み行列との積をとり, 全要素の和を求めることを意味している [1].

$$\begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 1 & 3 \\ 2 & 1 & 0 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix} \quad (7)$$

1.2.3 IDM (コントラストを調べる)

(Inverse Difference Moment : IDM) は (Homogeneity : 局所一様性 [6], 均質性) と呼ばれる. 均一性の高い画像である場合は大きな値となり, 不均一な値であるほど0に近づく. Dissimilarity と逆の尺度であり, 局所的な明暗の度合いを測る. IDM 値の特性として, 「細かい」画像は「粗い」画像と比較して小さい値をとる. IDM は式 (8) で求める.

$$IDM = \sum_{i=1}^N \sum_{j=1}^N \frac{P(i,j)}{1+(i-j)^2} \quad (8)$$

式 (8) は, 例えば 4×4 の正規化済み GLCM 行列 P に対して, 式 (9) のような重み行列との積をとり, 全要素の和を求めることを意味している [1].

$$\begin{bmatrix} 1 & 1/3 & 1/5 & 1/10 \\ 1/3 & 1 & 1/3 & 1/5 \\ 1/5 & 1/3 & 1 & 1/3 \\ 1/10 & 1/5 & 1/3 & 1 \end{bmatrix} \quad (9)$$

1.2.4 ASM (規則的 or 不規則的を調べる)

(Angular second moment : ASM)(角2次モーメント:日訳) は, テクスチャ画像の一様性を表す数値である. 使用例として, テクスチャのキメを調べることで, 生ウニのテクスチャ画像から美味しさを判定する研究が報告されている [8]. ASM は式 (10) で求める.

$$ASM = \sum_{i=1}^N \sum_{j=1}^N P^2(i,j) \quad (10)$$

ASM 値の特性として, ホワイトノイズ画像に対しては数値的に大きくなり, 規則的なパターンを持つテクスチャの場合はそのパターン固有の値をとる [8]. 「細かい」画像は「粗い」画像と比較して小さい値をとる.

1.2.5 Energy (規則的 or 不規則的を調べる)

Energy は式 (11) で求められる [1]. 注意点として, [6] においては ASM が Energy として記述されている.

$$Energy = \sqrt{ASM} \quad (11)$$

1.2.6 ENT (規則的 or 不規則的を調べる)

(Entropy : ENT)(エントロピー:日訳) は, テクスチャ画像が規則的かな否かを表す数値である. (scikit-image 未実装!) ENT は式 (12) で求める.

$$ENT = \sum_{i=1}^N \sum_{j=1}^N -P(i,j) \ln P(i,j) \quad (12)$$

GLCM の各項の値は, 基本画素の濃度値 n に対して, 濃度値が m となる参照画素がどの程度共起するか (しないか) を確率によって表現した行列であると解釈できる. このため, $p(i,j)$ の値が0もしくは1の時に0となり, 0.5(意味:どっちつかずの確率) の場合大きな値を与えるエントロピー関数を使用する.

ENT 値の特性として, 大きな値ほど規則性の無いテクスチャ画像であることを示す.

1.2.7 COR (統計的評価指標を調べる)

(Correlation : COR)(相関:日訳) は, テクスチャ画像画像における基本画素に対して, 参照画素の線形依存性を測定する.

COR は式 (13) で求める.

$$COR = \sum_{i=1}^N \sum_{j=1}^N \frac{(i - \mu_x)(j - \mu_y)P(i, j)}{\sigma_x \sigma_y} \quad (13)$$

ただし $\sigma_x, \sigma_y, \mu_x, \mu_y$ は以下の式で求める.

$$\mu_x = \sum_{i=1}^N i \sum_{j=1}^N P(i, j) \quad (14)$$

$$\mu_y = \sum_{j=1}^N j \sum_{i=1}^N P(i, j) \quad (15)$$

$$\sigma_x = \sum_{i=1}^N (i - \mu_x)^2 \sum_{j=1}^N P(i, j) \quad (16)$$

$$\sigma_y = \sum_{j=1}^N (j - \mu_y)^2 \sum_{i=1}^N P(i, j) \quad (17)$$

$$(18)$$

COR 値の特性 [1]

たとえば, 地面のみを含むテクスチャ画像の相関は高くなり, 他の物体表面を含む地面テクスチャ画像の相関は低くなる傾向がある. また, テクスチャ画像のサイズを小さくすると COR 値は大きくなり, サイズを大きくすると COR 値は小さくなる傾向がある.

テクスチャ画像のサイズを決定する際に利用

地面など, 一様であると考えられる領域においてテクスチャ画像のサイズを変化させながら, 相関の値が減少しはじめるサイズが, テクスチャ画像の最大サイズであると考えられる.

2 テクスチャ画像の解析方法

テクスチャ解析の手法は大きく分けて 2 種類

2.0.1 構造レベルの解析

- Hough 変換 など

(特徴) 航空写真やリモートセンシング画像における道路や建造物など, 抽出するターゲットの構成要素が明確な人口のテクスチャには有効.

岩石, 布地, 木目など, 自然に存在するテクスチャは難しい.

2.1 統計レベルの解析

- 濃度ヒストグラム (1 次統計量)
- 自己相関関数 (2 次統計量)
- フーリエスペクトル (2 次統計量)
- 差分統計量 (2 次統計量)
- 濃度共起行列 (2 次統計量)
- ランレンジス行列 (高次統計量)

2.2 その他

(例) フラクタル情報

3 GLCM の解析方法

3.1 画素間距離 d による解析

テクスチャ画像のサイズを相関値, もしくは 2^n の値で指定したとする. このとき, 基準画素と参照画素の間の距離 d に対する各特徴量の値の変化を図 (3) のように解析することで, テクスチャ画像自身の特徴を分析することができる [10].

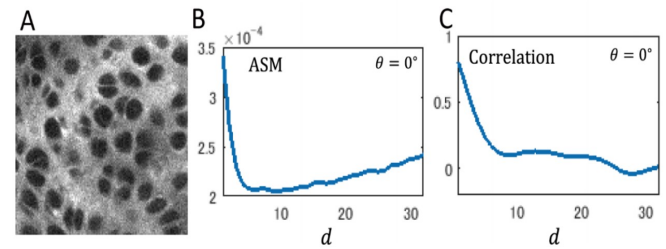


図 3: 画素間距離 d に対する ASM, COR の値 [10]

3.2 特徴量をベクトルとして扱う

得られた複数の特徴量を使用し, SVM によって判定を行う研究も行われている [9].

scikit-image 公式サイトでは, 空のテクスチャ画像と芝生のテクスチャ画像に対して, COR と DIS の値がどのように得られるかを解析している (参考図 (4)).

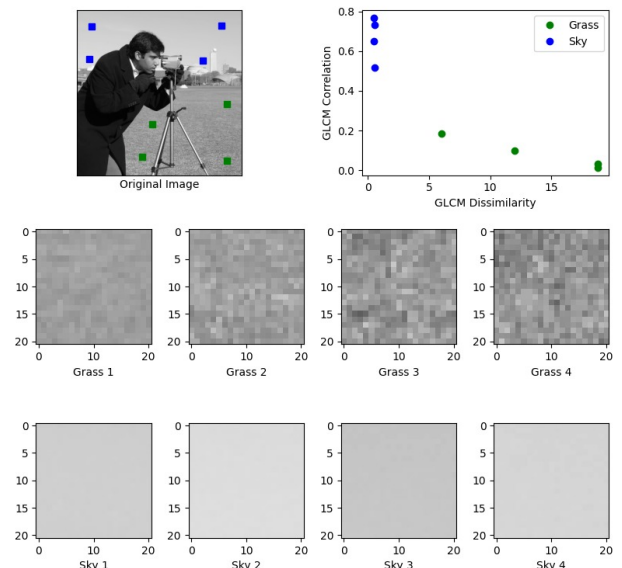


図 4: GLCM 行列による特性解析

3.3 同時生起行列をベクトルとして扱う

先述の通り、特徴量を作成せずに、作成した同時共起行列 \mathbf{P} の各項自体をそのまま特徴量として NN などに入力する場合もある [7].

実験

今回は、GLCM を使用したテクスチャ解析を行なう。具体的には、各テクスチャ尺度を特徴ベクトルの要素とし、SVM による識別を行う。

4 実験 1: texture measures の計算

実験では、事前にパラメータの調整実験を行い、GLCM を計算する際の入力画像 (window) サイズとして、基準画素を中心を含む 5×5 pixel の領域に固定した。また参照画素は $\theta = 0^\circ$, $d = 1$ の位置の画素を使用した。

なお window は図 5 のように 1pixel づつラスタスキャンを行い、各画素に対してテクスチャ尺度を与えた。また、画像の端にはミラーリング処理を施し、元の画像サイズと同じサイズの出力画像を得られる様に調整した。

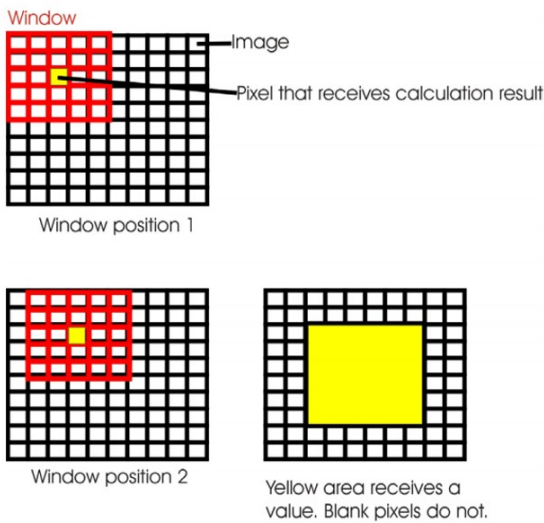


図 5: 画素毎に texture measure を与える [1]

実験 1: 結果

元の画像を図 (6) とし、gray scale 画像を図 (7) としたときの、GLCM による各テクスチャ尺度を計算した。また、各値を 0 255 の値に正規化した。これらを擬似カラー画像に変換した結果を図 (8)~図 (15) に示す。

実験 1: 考察

これらの中で correlation(図 (10)) は、(日向における) 猿の画素と地面の画素に対して安定して異なる値を与える傾向があるため、淡路島の画像に対しては比較的良好な尺度と考える。

5 実験 2: SVM による識別

テクスチャの尺度は、各画素ごとに得られる。そこで、これら尺度を特徴量とする特徴ベクトルを画素毎に作成する。さらに、SVM によって (猿領域) or (猿以外の領域) を判定する 2 クラス分類を行なった。

全体画像の位置 (x, y) における画素の、特徴ベクトル $f_{x,y}$ は 19 式で表す。

$$f_{x,y} = [f_1, f_2, \dots, f_N]^T \quad (19)$$

ただし f_1, f_2, \dots, f_N は $f_1 = COR$, $f_2 = CON$, $f_3 = DIS$, $f_4 = IDM$, $f_5 = ASM$, $f_6 = ENE$, $f_7 = ENT$, $f_8 = MEAN$, $f_9 = STD$, $f_{10} = MAX$, $f_{11} = B$, $f_{12} = G$, $f_{13} = R$, $f_{14} = L$, $f_{15} = a$, $f_{16} = b$ のテクスチャ尺度とする。

教師ラベルの付け方を述べる。猿が写っている領域の画素に対して 1, 写っていない画素に対して 0 を与えたマスク画像を作成する (図 16). (猿領域を positive 領域, 猿以外領域を negative 領域とする). positive 領域, negative 領域からそれぞれランダムに N 個の画素を選択し、訓練データとした。

5.1 実験 2: 結果

サンプリング数 $N = 2000$ のときの実験結果を図 17 に示す。また、 $N = 10000$ のときの実験結果を図 18 に示す。

訓練データ数 (サンプリング数 N) を増やしても、各画素のテクスチャ尺度ベクトルに対する SVM の予測結果に違いは見られなかった。また、図 (16) の mask 画像で表す positive, negative 領域と SVM の予測結果に類似性が見られなかった。

地面の上に入る猿については、日向に入る場合は比較的正しく猿領域を予測することができていた。しかし影の下にいる猿領域については誤った予測を与えている。また、屋根の上に入る猿は背景の山と同化して識別を上手く行うことができていない。

5.2 実験 2: 考察

サンプリング数を増やしても予測結果に大きな違いは見られなかったことから、訓練データの画素数は positive, negative それぞれ $N = 2000$ 画素程度で特徴空間を表現するのに十分なサポートベクトルを確保することができると考える。

6 実験 3: 特徴量として RGB, Lab の濃度値を追加

テクスチャ尺度のみでは背景の山や岩肌など、色情報でも分離可能と考えられる画素に対して、SVM は誤った予測をしている。そこで、特徴ベクトルに R,G,B,L*,a*,b* 各チャンネルの値を加えた。

6.1 実験3：結果

テクスチャ尺度に R,G,B,L*,a*,b* 各チャンネルの値を特徴量として加え、SVM によって予測した結果を図 (21) に示す。また比較のために、L*, a*, b*, の 3 つの濃度値のみを特徴量とする特徴ベクトルを使用した結果を図 (19) に示す。加えて、テクスチャ尺度に R,G,B の各チャンネルの値を特徴量として加えた結果を図 (20) に示す。

図 (19) を確認すると、L*a*b* の値だけを用いたにも関わらず、(正解を表す)mask 画像に近い領域を、SVM は予測している。特に背景の山と屋根の上の猿を分離することに高い精度で成功している。

図 (19) と図 (21) を比較すると、日向の領域において、猿と地面の分離精度が向上していることが分かる。

6.2 実験3：考察

図 (19) と図 (21) の結果は、閾値を用いないで緑領域 (背景の山) と猿領域を分離することが可能であることを示している。加えて、色空間において距離が近いと考えられる地面と猿の色分布の分離については、テクスチャ尺度を考慮することで精度の向上が望める。

これらの結果は、大豊の猿檻画像において、時間的に変化するため閾値による分離が難しい植物の形状変化に対して効果を期待できると考える。

(蔓植物などが柵にまわりつき、日光を受けて成長するため、画像中を動きまわる。また、風などで植物が倒れたりした際には、テクスチャ尺度によって露出した地面を検出もしくは除去できると考える。)

7 実験4：影の下にいる猿は無視する

先述の通り、図 (16) は正解ラベルを 2 値化して表現した mask 画像である。この mask 画像はあらかじめアノテーション時に図 (22) の様に影の下にいる猿の領域も positive 領域として含んでいる。しかし、日陰において猿領域と猿以外の領域を高精度に分離することができているテクスチャ尺度は、図 (8)～図 (15) の中には存在しない。そこで、影下にいる猿の検出を諦めることで、日向の猿を検出する精度を向上させることを目標とし、影下の猿領域を negative 領域としてアノテーションした。この mask 画像を図 (23) に示す。

実験では、各画素において、テクスチャ尺度と R,G,B,L*,a*,b* の各濃度値を特徴量とするベクトルを使用した。そして、それぞれの mask 画像が示す領域を訓練データとした (ランダムサンプリング画素数 5000)SVM を作成し、予測を行なった。

7.1 実験4：結果

影下の猿領域を positive として学習した結果を図 (24) に示す。また、negative として学習した結果を図 (25) に示す。図 (24) と図 (25) を比較する。日向の猿や屋根の上の猿において、日当たりの角度が悪い体の下などは、影下の猿を negative として学習した SVM の精度が低下していることが分かる。

7.2 実験4：考察

精度が低下した理由としては、日陰の猿領域と、日当たりの角度によって体の一部が暗く見える領域の色やテクスチャ尺度は、類似しているため、誤って除去してしまったと考える。

8 実験5：時間ごとの特性

実験4において作成した SVM(画像 818 の画像で作成した SVM) を、他の時間帯の画像に対しても使用し、さる領域の推定を行なった。

8.1 実験5：結果

テストした画像を図 26, 図 28, 図 30, 図 32, に示す。また、それぞれの推定結果を図 27, 図 29, 図 31, 図 33, に示す。

概ね猿領域を推定することはできていた。しかし、時間帯によっては、図 31(画像下中央左) のように、影の下を猿を検出できない場合もあった。

参考文献

- [1] HALL-BEYER, M. The GLCM Tutorial Home Page. <http://www.fp.ucalgary.ca/mhallbey/tutorial.htm> (v.3.0 March 3017 replaces).
- [2] HARALICK, R. M., SHANMUGAM, K., AND DINSTEIN, I. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics SMC-3*, 6 (1973), 610–621.
- [3] MATHWORKS. グレーレベルの同時生起行列の作成. <https://jp.mathworks.com/help/images/create-a-gray-level-co-occurrence-matrix.html>.



図 6: 元の画像 (img 818)

- [4] 橋本学, 奥田晴久, AND 鷲見和彦. 画像の濃度共起情報に基づく画像照合のための参照画素選択手法. Tech. rep., 2009.
- [5] 江口朋代, AND 五十嵐心一. 焼成した浄水場発生土の空隙構造の評価. 土木学会中部支部研究発表会 研究成果報告書概要 (和文).
- [6] 森秀二, AND 坂倉柊子. 画像処理の基礎 [2]. オーム社 (1990). ISBN:4-274-07549-4.
- [7] 梅田三千雄. ニューラルネットワークによるテクスチャ画像の識別. *Medical Imaging Technology* 9, 4 (1991), 416.
- [8] 本間朝博, AND 三上貞芳. スマートフォンによるテクスチャ解析を用いた生ウニのおいしさの簡易判別. 情報処理北海道シンポジウム (2014).
- [9] 榎本佳史, 川下郁生, 大倉保彦, 中島正光, 奥村英一郎, AND 石田隆行. 濃度共起行列を用いた胸部 x 線画像におけるじん肺の分類法の開発. *日本放射線技術学会雑誌* 67, 4 (apr 2011), 336–345.
- [10] 齋藤卓, 清松悠, 大嶋佑介, AND 今村健志. テクスチャ画像解析の医学応用～第 2 次高調波発生像による病態診断へのアプローチ～. *日本応用数理学会論文誌* 26, 2 (2016), 253–267.



図 7: 8 階調 gray scale 画像

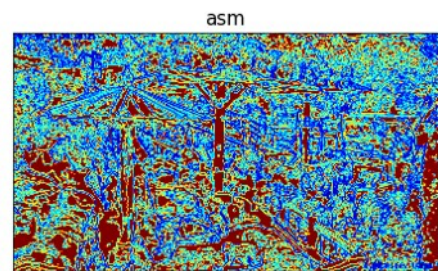


図 8: Angular second moment

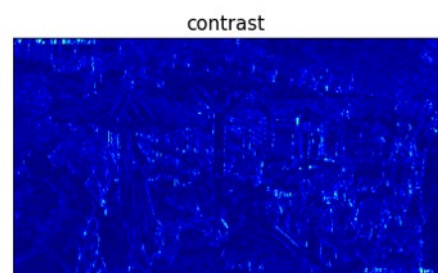


図 9: contrast

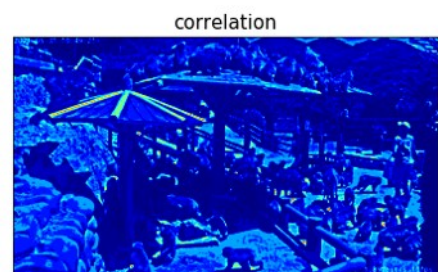


図 10: correlation

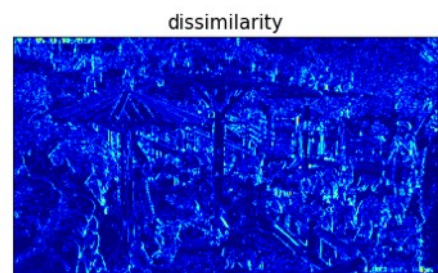


図 11: dissimilarity

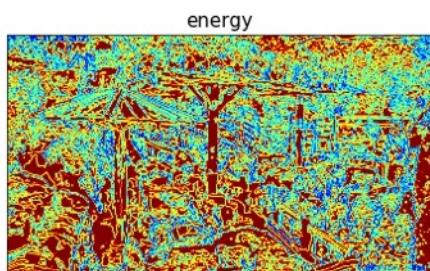


図 12: energy



図 16: mask image

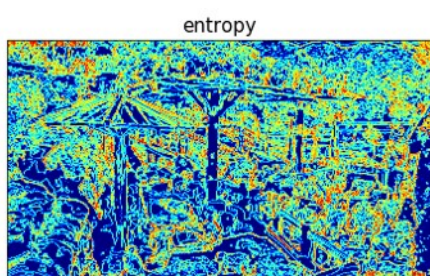


図 13: entropy



図 17: N=2000

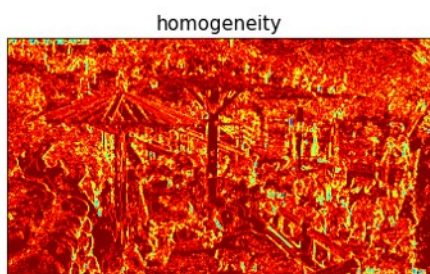


図 14: Inverse Difference Moment (homogeneity)



図 18: N=10000

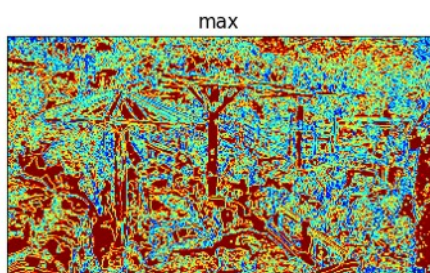


図 15: max



図 19: L*a*b*の濃度値を特徴量に使用



図 20: texture measure と RGB 濃度値を特徴量に使用



図 24: 影下の猿を positive として学習



図 21: texture measure と RGB, L*a*b* の濃度値を特徴量に使用



図 25: 影下の猿を negative として学習



図 22: 影下の猿を含む mask 画像



図 26: img 1000



図 23: 影下の猿を含まない mask 画像

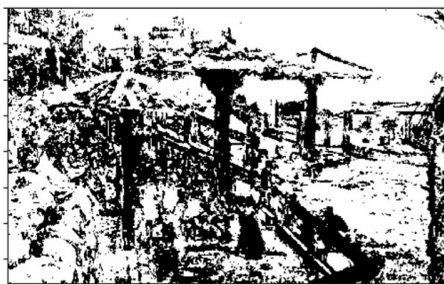


図 27: 提案手法による猿領域 img 1000



図 28: img 1074



図 29: 提案手法による猿領域 img 1074



図 30: img 1500

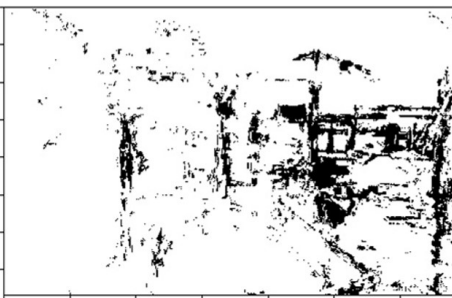


図 31: 提案手法による猿領域 img 1500



図 32: img 1779



図 33: 提案手法による猿領域 img 1779