

テンプレートマッチングにおける 類似物との識別に有効な画素選択手法

櫻本 泰憲^{1,a)} 齋藤 正孝¹ 橋本 学^{1,b)}

概要: 周辺に多数の類似物が存在する状況下でも、高信頼な照合が可能な画素を自動的に選択する手法を提案する。本研究では、あらかじめ準備した対象物と類似物のサンプル学習画像群を用いて、その2クラスの分離度を最大化する画素群を決定する問題を解く。クラス分離度の計算には判別分析法を適用する。この画素選択問題を組み合わせ最適化問題として定式化し、遺伝的アルゴリズムを用いて解を探索した。得られた画素群を用いることにより、類似物に影響されにくくなり、同時に画素数の大幅な削減により処理時間の短縮も実現した。認識性能としては、対象物の周辺に類似物を含む実画像400枚に対して画像照合したところ、テンプレート画像全画素の約0.5%の参照画素数で認識率98.5%、処理時間8msecが実現できたことを確認した。

キーワード: テンプレートマッチング, 画素選択, 識別性能, 画像照合, 組み合わせ最適化

Effective Pixels Selection method for Discriminating from Similar Objects in Template Matching

SAKURAMOTO YASUNORI^{1,a)} SAITO MASATAKA¹ HASHIMOTO MANABU^{1,b)}

Keywords: template matching, pixel selection, identification performance, image matching, combinatorial optimization

1. はじめに

生産現場では、画像センサによる物体検出が有用であるが、対象物と類似する周辺物体との識別が重要な課題になっている。

類似物体や複雑背景に対するロバスト化の研究としては、学習を用いた手法が有効とされている。Boosting Chamfer Matching[1]は、学習画像を用いて識別に有効な補助輪郭モデルを生成して相違度計算に利用することにより、複雑背景における誤検出率を低減している。また、Contour

Segment Networksを用いた物体検出手法[2]は、輪郭をセグメントに分割し、その接続関係を利用することで誤検出を低減している。しかし、これらの手法はエッジが安定して検出されることを前提としており、エッジの少ない対象物体における信頼性の低下が懸念される。画素単位の投票を用いた照合手法[3]は、M推定を用いた投票処理により部分的に類似したパターンにも頑健な照合が可能である。しかし、類似した面積の大きい類似物においては、outlierの割合が高くなりロバスト推定の精度が低下する問題がある。Multiple Instance Learning[4]を用いた手法[5]は、ポジティブサンプルに共通し、ネガティブサンプルには存在しにくい特徴をDiverse Density[6]という指標により算出し、入力クラスを識別する。しかし、Diverse Density値を求めるためには、任意のサンプルと他のサンプルの距離を

¹ 中京大学大学院 情報科学研究科
Graduate School of Information Science and Technology,
Chukyo University, Nagoya, Aichi 466-8666, Japan

a) sakuramoto@isl.sist.chukyo-u.ac.jp

b) mana@isl.sist.chukyo-u.ac.jp

全て計算する必要があるため、サンプル数が多い場合には計算コストが高くなるという問題がある。識別対象クラスごとの部分空間を利用する手法 [7] は、複数の識別対象をそれぞれ異なる部分空間として表現し、各空間に投影された未知入力パターンの距離をもとに対象を識別する。しかし、照合時に対象領域を部分空間へ投影する必要があるため、照合にかかる処理コストの高さに問題がある。対象物と周辺物の関係について人間の視覚注意モデルを定量化した Saliency Map[8] を用いた物体検出手法 [9] は、注目する物体と周辺情報の関係を考慮することで高信頼な照合が可能である。しかし、顕著性を小領域単位で評価しており、領域単位の選択は照合に不向きな画素を含む場合には精度低下が懸念される。また照合時間は画素数に比例することから高速化の効果も十分ではない。従来研究では、類似物体に対するロバスト性は考慮されているが、高速性が犠牲になっている。

本研究の目的は、実用的な処理速度を維持しつつ類似物が存在しても高信頼な物体検出を実現することである。

ところで、実用的な処理速度を目指した画像照合法としては、Canny[10] エッジや FAST[11] のコーナーなど参照データを削減するアプローチが有効とされている。近年では、対象物のエッジ方向情報を用いた照合法 [12] や、テンプレート画像中の独自性の高い画素や特徴点、周波数のみを照合に使用する手法 [13], [14], [15] が提案されている。これらの手法は、テンプレート画像におけるエッジ方向や独自性を考慮することで高信頼な照合が可能である。しかし、分析対象がテンプレート画像のみであるため、周辺状況を考慮していない。したがって、入力画像中に類似物が存在する場合には、誤認識を起し易い問題がある。テンプレートから顕著性の高い点を抽出し、入力画像の顕著性の高い点と照合する手法 [16] は、顕著性による周辺情報の考慮とデータ削減による高速な照合が可能である。しかし、この手法は追跡により探索範囲を限定しており、時系列情報を利用できない課題において、入力画像全体から顕著性の高い点を抽出すると誤対応が発生する恐れがある。また近年では、SIFT[17] や SURF[18], ORB[19] など画像中のキーポイントを用いる照合法や、Random Forests[20] や Random Ferns[21] による学習を用いたキーポイントマッチングの高速化手法 [22] も注目されている。しかし、局所情報同士の照合であることから、部分的に類似した偽物体に対しても類似性が高まり容易に誤認識が発生してしまう。

そこで本研究では、画素の削減効率の観点から文献 [13] の画素選択の枠組みをベースに、事前に与えられた類似物に関する情報を考慮することによって、識別にとって特に有効な画素群のみを抽出し、類似物の誤認識が少ない高速物体検出手法を提案する。

本研究では、対象物と類似物の識別性能を最大化する画素群を選択する。識別性能は、対象物と類似物の分離度

に基づいて評価する。すなわち、対象物画像（ポジティブサンプル）群と類似物画像（ネガティブサンプル）群を準備し、それらの分離性を最大化するような画素群を決定する問題を解く。これを組み合わせ最適化問題として定式化し、遺伝的アルゴリズム (GA) [23] を用いて解を探索する。

本手法により、ポジティブサンプル群とネガティブサンプル群の識別性能を最大化する画素群を用いることで、類似物の多い入力画像に対しても高信頼かつ高速な画像照合が実現できる。

以下、第 2 章では基本アイデアについて述べる。第 3 章では提案アルゴリズムについて説明し、第 4 章では提案手法の類似物に対する頑健性評価実験の結果を示す。

2. 基本アイデア

本研究の基本アイデアは、テンプレート画像中からマッチングに有効なごく少数の画素を自動的に選択することである。マッチングに有効な画素とは、対象物と類似物の差異のある画素である。マッチングに有効な画素のみを用いることで、類似物が多い状況下でも高信頼かつ高速なマッチングを実現できる。以降の節では、周辺物を考慮した参照画素群選択と、参照画素群の識別性能の評価について説明する。

2.1 周辺物を考慮した参照画素群選択

本節では、周辺物を考慮した参照画素群選択について説明する。対象物の周辺に物体が存在する入力画像とその周辺物を考慮した最適な参照画素群の例を図 1 に示す。

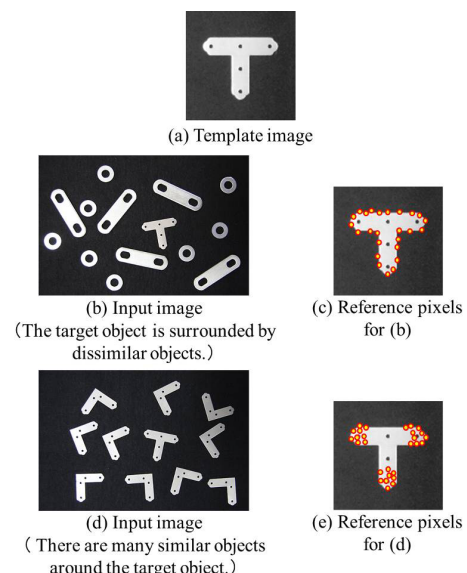


図 1 周辺物に応じて選択された参照画素群

Fig. 1 Example of reference pixels that are selected according to the surrounding objects.

(a) をテンプレート画像としたとき、入力画像が (b) のように周辺物が対象物の形状とあまり似ていない場合では、

対象物の全体的な形状を表現するために、(c)のような輪郭付近の画素をマッチングに用いるべきであると考えられる。しかし、(d)のように周辺物の形状が対象物と似ている場合は、(c)のような画素を用いたマッチングでは似た形状をもつ類似物を誤認識してしまう恐れがある。そのため、(e)のように対象物と類似物に差異のある画素を用いることで、類似物の誤認識を低減できる。このように参照画素群を周辺物に応じて最適に決定することにより、類似物の誤認識を低減できる。

2.2 2クラスの分離度に着目した識別性能の評価

本研究で扱う識別性能とは、入力画像内に類似物が存在する状況下でも、対象物を的確に識別する能力である。そのため、識別性能の評価には、対象物と類似物の分離度に着目した評価方法が望ましいと考えられる。すなわち、最適な参照画素群は、この分離度を最大化する画素群である。分離度の概念図を図2に示す。

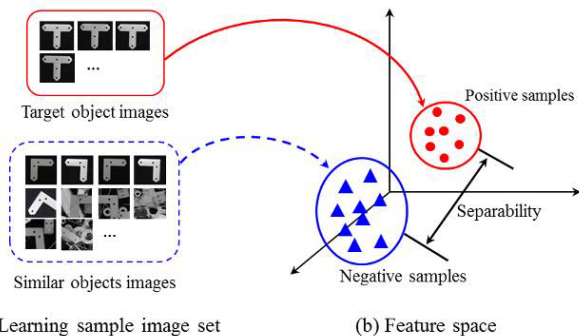


図2 分離度の概念図

Fig. 2 Schematic diagram of separability.

サンプル学習画像群 (a) から対象物画像と類似物画像を特徴空間 (b) に投影し、対象物画像を集めたポジティブサンプル群と類似物画像を集めたネガティブサンプル群の2クラスに分類する。2クラスの分離度は対象物と類似物の識別しやすさを表している。そのため、分離度を最大化することによって、類似物への誤認識を低減することができる。なお、識別性能の評価方法については3.3節で詳しく説明する。

3. 提案アルゴリズム

3.1 概要

参照画素群は、テンプレート画像の中の画素の組み合わせとして決定される。テンプレート画像には通常数万から数十万の画素が存在するため、組み合わせ数が膨大になる。そのため、総当たり法で最適解を求めることは困難である。そこで本研究では、最適な参照画素群の選択を、組み合わせ最適化問題として扱い、最適解を求めることにより照合に使用すべき画素群を決定し、その画素群を用いて照合す

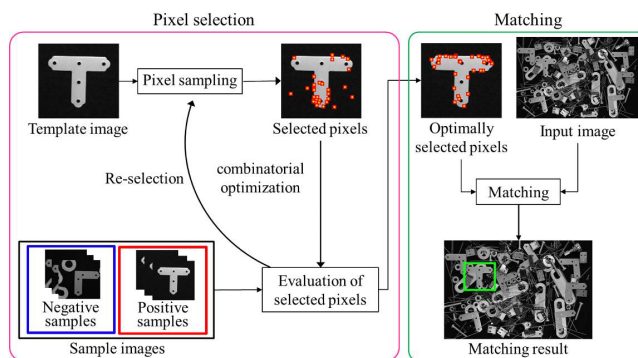


図3 提案アルゴリズムの流れ

Fig. 3 Flow of the proposed algorithm.

る。提案アルゴリズムの構成を図3に示す。

提案アルゴリズムは、画素選択と画像照合の2つの処理から構成される。まず画素選択について説明する。基本的には、与えられたテンプレート画像中から選択された少数画素群 (サブセット) の識別性能を評価し、この値が最大になる画素群を参照画素群として決定する流れとなる。まず、組み合わせ爆発を抑制するために、テンプレート画像から数割程度の候補画素群を選んでおく。これにより、元データ内における無駄なデータを排除する。これには文献 [13] の手法を用いることにより、テンプレート画像を特徴づける画素を候補画素として選択する。この画素群に対して、事前に準備したポジティブとネガティブ両サンプル群を適用して識別性能を評価する。一連の処理は組み合わせ最適化問題ととらえることができることから、本手法では遺伝的アルゴリズムを適用して効率よく実用解を求める。次に画像照合については、文献 [13] と同様に選択された画素群を用いた正規化相互相関を利用する。

3.2 候補参照画素群の事前抽出

本節では、濃度共起発生確率 [13] を用いた候補参照画素群の事前抽出について説明する。

一般的な濃度共起ヒストグラムは、図4のように画像 $f(i, j)$ における2つの画素 P と Q それぞれの濃度値 p と q のペアの発生頻度を、画像全体にわたって求め、濃度値 p, q を縦横それぞれのインデックスとする2次元マトリクス状に並べたものである。

P, Q の位置ベクトルをそれぞれ $\mathbf{v}_P = (i_P, j_P)$, $\mathbf{v}_Q = (i_Q, j_Q)$ とし、 P からみた Q の変位ベクトルを $\mathbf{d} = (k, l)$ とすると、この2画素からなる画素ペアの共起ヒストグラム $h(p, q)$ は、式 (1)(2) で定義される。

$$h(p, q) = \sum_{\mathbf{v}_P, \mathbf{v}_Q \in R} \delta(\mathbf{v}_P, \mathbf{v}_Q, p, q) \quad (1)$$

$$\begin{cases} \delta = 1 & \text{when } \{f(\mathbf{v}_P) = p\} \cap \{f(\mathbf{v}_Q) = q\} \\ \delta = 0 & \text{otherwise} \end{cases} \quad (2)$$

$$\text{where, } \mathbf{v}_Q = \mathbf{v}_P + \mathbf{d}$$

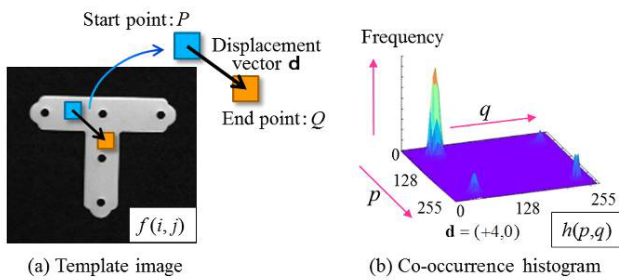


図 4 濃度共起ヒストグラム h 生成の模式図

Fig. 4 Schematic diagram of generating a co-occurrence histogram h of intensity.

領域 R は点 P と点 Q がともに画像 f に含まれるときの点 P の存在範囲である。ここで、 h の値を h の総和が 1 になるように式 (3) で正規化して濃度共起発生確率 P_h とする。 L は画像の階調数 (例えば 256 階調) であり、2 次元共起ヒストグラムの縦と横のセル数となる。 P_h はある画像 f が与えられたとき、その画像において相対変位が \mathbf{d} である画素対 P, Q がそれぞれ濃度値 p, q を同時に持つ確率を表している。

$$P_h(p, q) = \frac{h(p, q)}{\sum_{q=0}^{L-1} \sum_{p=0}^{L-1} h(p, q)} \quad (3)$$

濃度共起発生確率 P_h が低い画素は、テンプレート画像内において独自性が高いことを表している。テンプレート画像内において独自性の高い画素は、入力画像内においても独自性が高い画素であると期待できる。したがって、本研究では候補参照画素群を発生確率の低い (独自性の高い) 画素から順に抽出する。

3.3 識別性能の評価方法

本研究では、参照画素群の識別性能を図 5 に示すようなポジティブとネガティブ両クラスの間隔度により評価する。

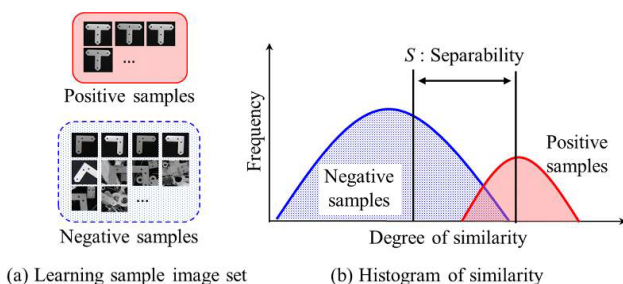


図 5 識別性能の評価方法

Fig. 5 Evaluation of the identification performance.

(a) のように対象物画像群 (ポジティブサンプル) と、類似物体を含む非対象物画像群 (ネガティブサンプル) を準備する。ポジティブサンプルは、実際の入力画像として想定される画像群から対象物領域を切り出すことで取得し、ネガティブサンプルはその画像群からランダムに領域を切

り出すことによって取得する。ある参照画素群が選ばれたとき、その画素群と両サンプル群の類似度をヒストグラム化すると、概念的には (b) のような 2 クラス分布になる。ここで、2 クラスの間隔度は、対象物と類似物の識別しやすさを表している。したがって、この間隔度 S を最大化するような参照画素群を選択することによって、類似物の誤認識を低減できる。また、識別性能を上げるためには、ポジティブサンプル群に共通する画素群を選択することも重要である。このような画素群は、さまざまな状況下に対応した汎用性に富んだ画素である。共通性の評価には、クラス内の分散値を用いる。分散値が小さいことは、クラス内に共通する画素を選択していることを表し、汎用性能が高いことを意味する。本研究では、クラス内分散とクラス間分散を用いた判別分析法 [24] により間隔度 S を求める。各サンプル画像とテンプレート画像の類似度を求め、式 (4)(5)(6) で示されるような一般的な指標により間隔度 S を計算する。全体の平均値 m_t 、ポジティブ、ネガティブサンプル群の平均値 m_p および m_n 、ポジティブ、ネガティブの分散値 σ_p, σ_n 、両サンプル数を ω_p, ω_n とする。

$$\sigma_w^2 = \frac{\omega_p \sigma_p^2 + \omega_n \sigma_n^2}{\omega_p + \omega_n} \quad (4)$$

$$\sigma_b^2 = \frac{\omega_p (m_p - m_t)^2 + \omega_n (m_n - m_t)^2}{\omega_p + \omega_n} \quad (5)$$

$$S = \frac{\sigma_w^2}{\sigma_b^2} \quad (6)$$

3.4 GA による画素パターンの最適化

提案手法では、最適な参照画素群の組み合わせは遺伝的アルゴリズム (GA) を用いて探索される。染色体は図 6 のような N ビットの 2 値の 1 次元ストリングとして定義する。

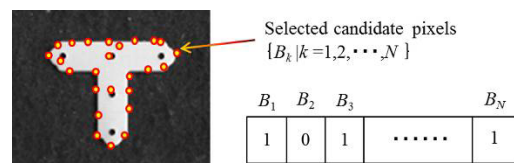


図 6 染色体の定義

Fig. 6 Definition of chromosome.

各ビット B_k は候補参照画素を表しており、値が「1」であれば参照画素として使用し、「0」であれば使用しないことを意味する。なお、候補参照画素は画素値の他にその画素の座標も保持している。最適化の目的関数は、式 (6) と定義し、評価値 S を最大とする画像群が最適参照画素群である。

集団進化のための個体の選択方法としては、ルーレット方式を用いた。子個体を生成する交叉方法には、一様交叉を遺伝子の「1」の数が一定となるように拡張した方法を

用いた。親個体 A と親個体 B の各染色体で、同一な位置にある遺伝子がともに「1」であれば、子の遺伝子も「1」とし、両親の遺伝子がともに「0」であれば、子の遺伝子も「0」とする。ここで、子の遺伝子の「1」の数が親の遺伝子の「1」の数未満であれば、両親の遺伝子が異なる遺伝子をランダムに選択し、子の遺伝子を「1」とする。これにより、遺伝子の「1」の数を一定とした交叉が可能となる。

突然変異の方法には、遺伝子の「1」の数が一定になるように拡張したビット反転方式を用いた。まず、ビット反転させる遺伝子の数 N_m 個をランダムに決定する。次に、突然変異の対象個体の「1」である遺伝子の中からランダムに反転させる遺伝子を N_m 個だけ選択し、「0」に反転させる。次に、対象の個体の「0」である遺伝子の中からランダムに N_m 個だけ選択し、「1」に反転させる。

以上の遺伝的操作を用いて集団を進化させ、過去 R 世代の間に、最大適応度が一定値を記録した時点で、進化が完了したとみなして処理を終了させる。そして、最終世代において最大の適応度をもつ個体の染色体情報を GA によって得られた解とする。なお、パラメータ R は実験的に決定する。

3.5 1次元点列テンプレートマッチング

テンプレートデータは、提案手法によって選択された離散的な画素のみが1次元的な3つの配列 $f_i(k)$, $f_j(k)$, $f_v(k)$ に格納される。 f_i は各参照画素の i 座標値、 f_j は j 座標値、 f_v は濃度値であり、 $k = 0, 1, 2, \dots, M-1$ である (M は参照画素数)。テンプレートデータを入力画像 $g(i, j)$ に対して (d_x, d_y) ずらして重ね合わせたときの正規化相互相関による類似度 $S_c(d_x, d_y)$ を、式 (7) により計算し、最大となる座標を最終的なマッチング位置とする。ただし \bar{g} , \bar{f} は領域内における参照画素の濃度値の平均値である。

$$S_c(d_x, d_y) = \frac{\sum_{k=0}^{M-1} \{g(f_i(k) + d_x, f_j(k) + d_y) \cdot f_v(k)\} - M \cdot \bar{g} \cdot \bar{f}}{\sqrt{\sum_{k=0}^{M-1} g(f_i(k) + d_x, f_j(k) + d_y) - M \cdot \bar{g}^2} \sqrt{\sum_{k=0}^{M-1} f_v(k) - M \cdot \bar{f}^2}} \quad (7)$$

4. 実験結果と考察

本章では、本研究の目的である類似物との識別性能、さらに実用的な処理速度を示すための実験結果を示す。なお、実験に用いたポジティブサンプルとネガティブサンプルの画像枚数は、それぞれ 50 枚と 1000 枚とした。また、GA の終了条件 R は、50 世代とした。

4.1 選択参照画素群の比較

提案手法により画素選択した結果について説明する。実験には、図 7 に示す対象物と類似物が交互に配

置された基板上的電子部品を用いた。比較手法としては、SIFT[17], SURF[18], ORB[19], Random, Gradient, Canny[10], CPTM[13] の計 7 手法を用いた。

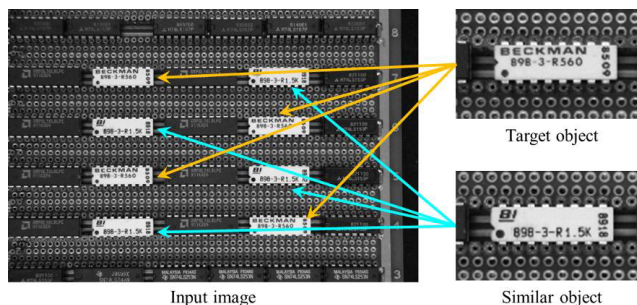


図 7 実験画像例

Fig. 7 Examples of experimental images.

テンプレート画像 (149 × 85) 全画素の約 0.24% (30 画素) の参照画素群を選択した結果を図 8 に示す。

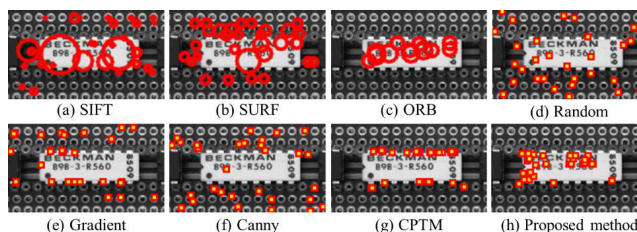


図 8 参照画素 (赤円) の選択結果

Fig. 8 Result of the selected pixels (red circles) by each method.

キーポイントマッチング手法である SIFT や SURF, ORB は、テクスチャが多い文字が印字された個所を中心にキーポイントを選択している。しかし、これらの個所は類似物にも共通して存在する個所を選択しているため、キーポイント同士の照合手法は誤照合を誘発しやすい。Random は、テンプレート画像の全体に広がって参照画素を選択している。Gradient や Canny は、勾配強度の大きい画素やエッジ画素を選択している。しかし、これらの手法は背景領域を選択した画素の影響を受けやすい問題がある。CPTM は、対象物の独自性の高い個所を中心に参照画素群を選択している。しかし、その個所は、類似物にも共通して存在している。一方、提案手法は文字が印字された個所の中でも、対象物と類似物の差異を特徴づける画素を選択している。

4.2 GA の進化過程における画素選択の変化

GA の進化過程における参照画素群の選択のされ方を評価した。図 9 は、1 世代目、100 世代目、200 世代目、400 世代目のそれぞれの時点で選択された最良の参照画素群 (30 画素) と選択された画素群を用いた類似度マップである。

世代を重ねるにつれ、対象物と類似物の差異を特徴づけ

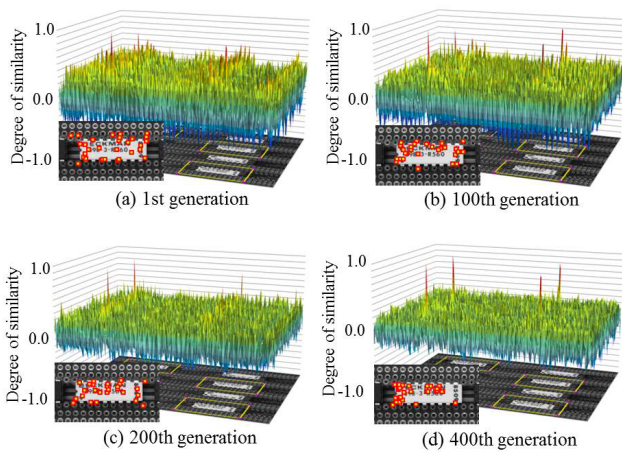


図 9 進化過程における参照画素群の変化

Fig. 9 The evolutionary behavior of selected pixels and score maps.

る対象物の印字された箇所を選択する画素が増えていることがわかる。これにより、世代を重ねるにつれ、類似度は正しい位置で急峻に高い値を示している。

4.3 識別性能の評価

提案手法と他手法の識別性能を比較した。実験には図 10 に例示する 4 種類の対象画像各 100 枚を用いて性能を評価した。比較手法としては、以下の 8 手法を用いた。

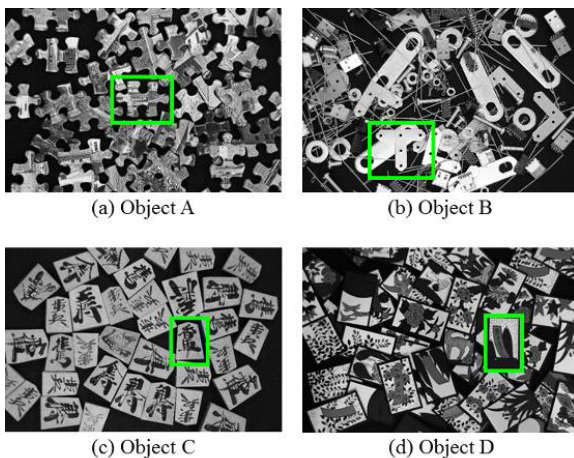


図 10 各対象物の認識結果

Fig. 10 Recognition result of each object.

- (1) NCC : Normalized Cross-Correlation を用いる手法
- (2) SIFT : SIFT キーポイント [17] からランダムにキーポイントを選択する手法
- (3) SURF : SURF キーポイント [18] からランダムにキーポイントを選択する手法
- (4) ORB : ORB キーポイント [19] からランダムにキーポイントを選択する手法
- (5) Random : テンプレート画像からランダムに選択した画素群を用いる手法

- (6) Gradient : 濃度勾配値の大きい画素から順に選択した画素群を用いる手法
- (7) Canny : Canny エッジディテクタ [10] により検出されたエッジからランダムに選択した画素群を用いる手法
- (8) CPTM : テンプレート画像の独自性が高い画素群を用いる手法 [13]

各対象物に対する参照画素数 M [pixel] と認識率 P_r [%] の関係を表 1 に示す。入力画像サイズは 600×400 であった。認識率 P_r は、 $P_r = (\text{正解座標から} \pm 2 \text{画素以内の位置で検出した対象物数}) / (\text{全対象物数}) \times 100$ と定義した。なお、SIFT 法、SURF 法および ORB 法においては、対応点が発見されたキーポイントのうち 70% が正しい対応であった場合を認識成功と定義した。

表 1 各手法の認識率

Table 1 Recognition rate of each method.

Method	Object A		Object B		Object C		Object D	
	M	P_r	M	P_r	M	P_r	M	P_r
(1)NCC	10043	93	9919	60	5525	100	8541	100
(2)SIFT[17]	30	83	60	1	30	83	30	87
(3)SURF[18]	30	55	60	0	30	20	30	62
(4)ORB[19]	30	88	27	1	30	75	30	24
(5)Random	30	32	500	39	100	62	100	71
(6)Gradient	100	58	450	62	100	59	200	73
(7)Canny[10]	30	79	450	45	100	52	100	87
(8)CPTM[13]	30	72	500	76	200	80	200	72
(9)Proposed	30	99	60	97	30	99	30	99

提案手法は、入力画像に形状的に類似する物体が多数存在する場合でも高い認識率が確認された。対象物によって NCC の認識率が大きく変動した理由は、テンプレート画像中の前景領域と背景領域の割合に起因する。また、提案手法に比べて Gradient や Canny, CPTM の認識率が低い理由は、対象物のエッジ付近の画素を中心に参照画素群として選択したため、形状的に類似する物体を誤認識したためである。一方、提案手法はサンプル画像群を用いて参照画素群を適切に選択するため、テンプレート画像に背景領域を含んでも高い認識率が期待できる。なお、Object B における SIFT や SURF, ORB の認識率が低い理由は、テンプレート画像のテクスチャが少ないため、照合に有効なキーポイントを抽出できなかったためと考えられる。また、Object D において SIFT や SURF に比べて ORB の認識率が低い理由は、キーポイントとしてコーナー以外の点が多く検出され、入力画像中の類似するキーポイントの誤対応が多いことに起因する。

次に ROC 曲線を用い、他手法と識別性能を比較した。類似度のしきい値 T_h を変数として、 T_h 以上を対象物とし、それ未満を非対象物と判断する。類似度のしきい値 T_h を変動させながら、適合率 R_p と再現率 R_r を算出する。図 11 に各手法の ROC 曲線を示す。

NCC の認識率は、表 1 より提案手法の認識率とほぼ同

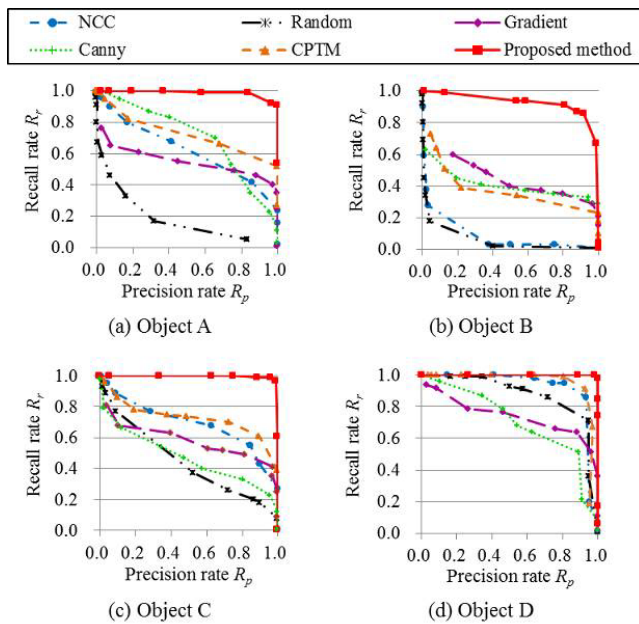


図 11 各手法の ROC 曲線

Fig. 11 ROC curve of each method.

等である。しかし、NCC の識別性能は、図 11 より CPTM や提案手法の識別性能と比べて劣っている。これは、類似パターンに共通して存在する画素を参照画素群に含んでいるからである。NCC に比べて CPTM の識別性能が高い理由は、参照画素群の中に類似パターンに共通する画素を含む割合が低かったからである。一方、提案手法は参照画素を厳選する際に、対象パターンのみが存在する画素を参照画素群として選択するため、他手法に比べ識別性能が高いことを確認した。

4.4 処理時間の評価

提案手法と他手法の照合にかかる処理時間を比較した。比較手法としては 4.3 節における実験の 8 手法を用い、4.3 節における 4 種類の対象画像を用いて性能を評価した。各対象物に対する参照画素数 M [pixel] と照合時間 T [msec] の関係を表 2 に示す。

表 2 各手法の処理時間

Table 2 Processing time of each method.

Method	Object A		Object B		Object C		Object D	
	M	T	M	T	M	T	M	T
(1)NCC	10043	1577	9919	1985	5525	1244	8541	1874
(2)SIFT[17]	30	229	60	478	30	771	30	710
(3)SURF[18]	30	125	60	253	30	181	30	176
(4)ORB[19]	30	80	27	80	30	60	30	90
(5)Random	30	8	500	109	100	24	100	24
(6)Gradient	100	23	450	95	100	24	200	42
(7)Canny[10]	30	8	450	95	100	24	100	24
(8)CPTM[13]	30	8	500	109	200	42	200	42
(9)Proposed	30	8	60	8	30	8	30	8

CPU: Intel® CORE™ i7-3.40GHz, System memory: 16GB

照合時間に関しては、提案手法は少ない参照画素数で

認識が可能であることから全画素を用いた NCC と比べて大幅に照合時間を削減できている。また、Random や Gradient, Canny, CPTM は 3.5 節で説明した照合手法を用いているため、提案手法と同等の照合時間である。なお、キーポイントマッチング手法 (SIFT, SURF, ORB) は、平行移動と回転、スケールを含めた照合時間であるが、テンプレートマッチング手法 (NCC, Random, Gradient, Canny, CPTM, 提案手法) は平行移動のみを含めた照合時間である。しかし、類似度計算回数 (探索回数) を削減する Coarse-to-Fine 探索手法と併用することにより、平行移動と回転、スケールを含めても高速に照合可能である。

提案手法における参照画素群の選択にかかる処理時間は約 16 分であった。これは他手法より長い、一般にテンプレート画像の準備はオフラインで事前に実行可能な処理である。

4.5 サンプル画像枚数と認識率の関係

提案手法におけるポジティブサンプルとネガティブサンプルの画像枚数と認識率の関係を評価した。実験画像としては図 10 の Object B における 100 枚の画像を用いた。実験結果を図 12 に示す。

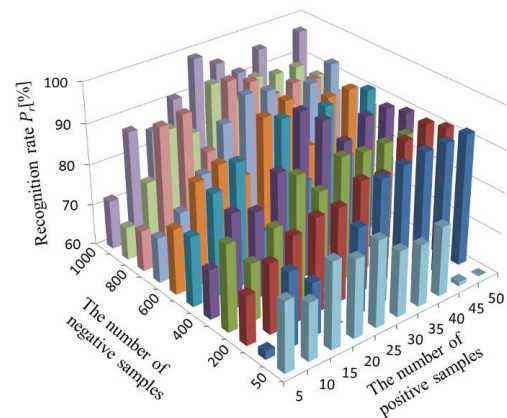


図 12 サンプル画像枚数と認識率の関係

Fig. 12 Relation between the number of learning sample images and recognition rate.

両サンプル共に画像枚数が増加するにつれ認識率が向上している。ポジティブサンプルはサンプル数が 30 枚付近から認識率が収束しているが、ネガティブサンプルはサンプル数を 1000 枚使用しても認識率が収束していない。したがって、高い認識率を得るためにはネガティブサンプルはポジティブサンプルと比べて多数必要である。これは、ネガティブサンプルの多様性を補うために多数のサンプルが必要となるからである。

4.6 参照画素数と認識率の関係

提案手法における参照画素数と認識率の関係を評価し

た. 実験画像としては図 10 の Object B における 100 枚の画像を用い, 比較手法としては 4.3 節における Gradient, Canny, CPTM の 3 手法を用いた. 実験結果を図 13 に示す.

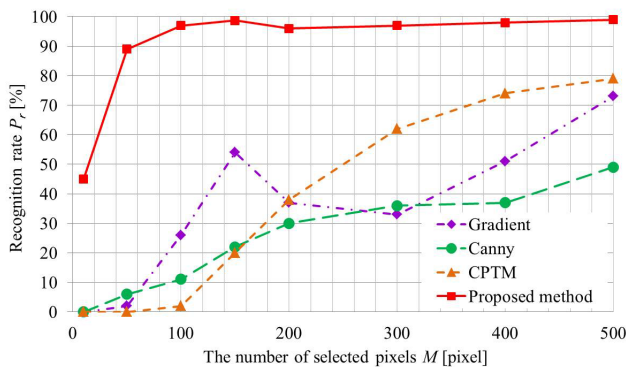


図 13 参照画素数と認識率の関係

Fig. 13 Relation between the number of selected pixels and recognition rate.

Gradient や Canny, CPTM は少ない参照画素数において認識率が低くなっている. しかし, 提案手法は少ない参照画素数においても認識率が高く, 少ない参照画素数で高速な照合が可能である.

5. おわりに

対象物と類似物の識別性能を最大化する参照画素群を自動選択する手法を提案した. その画素群をマッチングに使用することにより周辺に多数の類似物が存在する状況下でも高信頼かつ高速な物体検出を実現した. 認識性能としては, 対象物の周辺に類似物を含む実画像 400 枚に対して画像照合したところ, テンプレート画像全画素の約 0.5% の参照画素数で認識率 98.5%, 処理時間 8 msec を達成したことを確認した.

今後は, 参照画素群の選択のための処理時間の高速化と最適な参照画素数の自動決定等の各種パラメータの最適化を行う予定である.

謝辞 本研究の一部は科研費 C (23560512) の助成を受けて遂行されたものである.

参考文献

[1] Ma, T., Yang, X. and Latecki, L.: Boosting Chamfer Matching by Learning Chamfer Distance Normalization, *Proc. ECCV*, pp. 450–463 (2010).
 [2] V. Ferrari, T. T. and Gool, L.: Object Detection by Contour Segment Networks, *Proc. ECCV*, pp. 14–28 (2006).
 [3] Sibiryakov, A.: Fast and High-Performance Template Matching Method, *Proc. CVPR*, pp. 1417–1424 (2011).
 [4] Dietterich, T., Lathrop, R. and Lozano-Perez, T.: Solving the multiple-instance problem with axis-parallel rectangles, *Artificial Intelligence*, Vol. 89, No. 1–2, pp. 31–71

(1997).
 [5] 湯浅圭太, 和田俊和, 大池洋史, 坂田淳: Diverse Density に基づく画像データ検索用キーポイント抽出法について, *コンピュータビジョンとイメージメディア (CVIM)*, Vol. 2013-CVIM-185, No. 14, pp. 1–6 (2013).
 [6] Maron, O. and Lozano-Perez, T.: A framework for multiple-instance learning, *Proc. NIPS* (1998).
 [7] Fukui, K. and Yamaguchi, O.: The Kernel Orthogonal Mutual Subspace Method and its Application to 3D Object Recognition, *Proc. ACCV*, pp. 467–476 (2007).
 [8] Itti, L., Koch, C. and Niebur, E.: A Model of Saliency-Based Visual Attention for Rapid Scene Analysis, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 11, pp. 1254–1259 (1998).
 [9] Feng, J., Wei, Y., Tao, L., Zhang, C. and Sun, J.: Salient Object Detection by Composition, *Proc. ICCV*, pp. 1028–1035 (2011).
 [10] Canny, J.: A Computational Approach To Edge Detection, *Pattern Analysis and Machine Intelligence*, Vol. 8, No. 6, pp. 679–714 (1986).
 [11] Rosten, E. and Drummond, T.: Machine Learning for High-Speed Corner Detection, *Proc. ECCV*, Vol. 3951, pp. 430–443 (2006).
 [12] Liu, M., Tuzel, O., Veeraraghavan, A. and Chellappa, R.: Fast Directional Chamfer Matching, *Proc. CVPR*, Vol. 13, No. 18, pp. 1696–1703 (2010).
 [13] Hashimoto, M., Fujiwara, T., Koshimizu, H., Okuda, H. and Sumi, K.: Extraction of unique pixels based on co-occurrence probability for high-speed template matching, *Proc. ISOT* (2010).
 [14] Oerlemans, A. and Lew, M.: Interest Points based on Maximization of Distinctiveness, *Proc. MIR* (2008).
 [15] Wu, Y., Natarajan, P., Noonan, J., Prasad, R. and Natarajan, P.: Wavelet Band-pass Filters for Matching Multiple Templates in Real-time, *Proc. BMVC*, pp. 90.1–90.12 (2011).
 [16] Lee, T. and Soatto, S.: Learning and Matching Multiscale Template Descriptors for Real-Time Detection, Localization and Tracking, *Proc. CVPR*, pp. 1457–1464 (2011).
 [17] Lowe, D.: Distinctive Image Features from Scale-Invariant Keypoints, *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91–110 (2004).
 [18] Bay, H., Tuytelaars, T. and Gool, L.: SURF : Speeded Up Robust Features, *Proc. ECCV*, Vol. 3951, pp. 404–417 (2006).
 [19] Rublee, E., Rabaud, V., Konolige, K. and Bradski, G.: ORB : An Efficient Alternative to SIFT or SURF, *Proc. ICCV* (2011).
 [20] Breiman, L.: Random Forests, *Machine learning*, Vol. 45, No. 1, pp. 5–32 (2001).
 [21] Özuysal, M., Calonder, M., Lepetit, V. and Fua, P.: Fast Keypoint Recognition using Random Ferns, *Pattern Analysis and Machine Intelligence*, Vol. 32, No. 3, pp. 448–461 (2010).
 [22] Lepetit, V. and Fua, P.: Keypoint Recognition using Randomized Trees, *Pattern Analysis and Machine Intelligence*, Vol. 28, No. 9, pp. 1465–1479 (2006).
 [23] Goldberg, D.: *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison Wesley (1989).
 [24] Fukunaga, K.: *Introduction to Statistical Pattern Recognition*, Second Edition, Academic Press (1990).