

把持パターン画像の学習に基づく欠損画素復元と物体認識

笠原 啓雅[†] 松崎 淳[†] 島田 伸敬[†] 田中 弘美[†]

[†] 立命館大学大学院 理工学研究科 〒525-8577 滋賀県草津市野路東 1-1-1

E-mail: ^{††}matsuzaki@i.ci.ritsume.ac.jp, ^{†††}shimada@ci.ritsume.ac.jp

あらまし 本研究は物体の使用目的に対応した把持パターンと使用する物体形状とを同時に学習し、把持パターンとそれに対応した物体も同時に認識する手法を提案する。また認識と同時にカメラの見え方等によって物体や手領域の一部が隠蔽、欠損している場合においても、学習によって得られたデータを用いて欠損領域を復元する手法も提案する。
キーワード 物体認識、把持パターン、画像補間

Object Recognition by Observing Grasping Scene from Image Sequence

Hironori KASAHARA[†], Jun MATSUZAKI[†], Nobutaka SHIMADA[†], and Hiromi TANAKA[†]

[†] College of Information Science and Engineering, Ritsumeikan University Nojihigashi 1-1-1, Kusatsu-shi, Shiga, 525-8577 Japan

E-mail: ^{††}matsuzaki@i.ci.ritsume.ac.jp, ^{†††}shimada@ci.ritsume.ac.jp

Abstract The aim of this paper is the simultaneous restoration of incorrectly segmented the grasped object and grasping hand from the complex background and the recognition of the grasping method class.

Key words Object recognition, Grasping method, Image interpolation

1. はじめに

画像に基づく物体認識は、形状や色、模様などを手がかりにして、「同じ対象」を表している画像領域を切り出すことによって基本的に実現されてきた。「同じ対象」をどのような特徴を用いて定義・記述するかがその認識手法の適用範囲と性能を決めるが、近年ではたとえば画像のテクスチャを利用した顔検出 [1] や SIFT 特徴を利用した物体検出・照合 [2], HoG 特徴 [3] を利用した人物・車両検出など、複雑な背景の下で特定の物体を認識する優れた手法が提案されている。しかし物体の形状や見えが一見類似しているがその機能・用途が異なる物体どうし (ペンと細身のナイフなど) を識別することは、物体の見えだけに基づく手法では本質的に解決が難しい。

一方で、物体の形状や見えだけではなく、その物体に関わる人間の動作や、物体自身の持つ機能を考えあわせることで物体認識をしようとする研究も従来から行われている。Sato ら [4] は、鉗のように対象に働きかけて形状を変化させる物体について、変化の前後の見えを画像から学習することで物体の機能を推定している。北橋らは人間の動作と対象物体のもつ機能の関連性に着目し、対象認識と動作認識を同時に実現する研究を行った [5] [6]。この研究では、何か物を食べているなどの大まかな動きを認識することはできるが、握り方や何を持っているかというより詳細なレベルは取り扱っていない。

動作と物体機能の関連性については、物体を把持する際の手形状を分類する研究 [7] [8] が報告されており、人間は作業目的に応じて、物体の持ち方を適切に変えていることが指摘されている。例えばペンを使って何か物を書くときと、落ちているペンを拾うときは、異なった手の形状を使用する。このことに着目した Napier は、図 1 のように力を必要とする作業に用いる Power grasp (握力把持) と、細かな作業を行う際に用いる Precision (精密把持) の 2 つの大分類を定義した [7]。

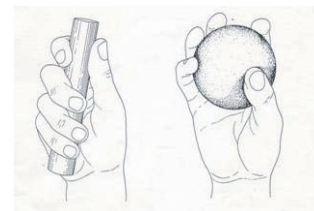


図 1 Napier の把持分類

また鎌倉らは作業療法 (リハビリテーション) による知見から、日常品 100 種余の物品を 15 個の把持パターンに分類した [8]。これらの知見に基づいて、物体の見えに加えて把持パターンを識別することによって、物体を機能的に分類・認識することが考えられる。

また画像認識の観点からは、物体と手ならびに背景を正しく分割することは正しい認識のために重要である。とくに物体把

持シーンの場合、手と物体どおしが相互に隠蔽を起こすため、その領域分割を認識前に正しく行うことは一般に困難である。今井らや浜田らは複雑背景下で運動する手の形や手話単語のための手形状を推定している [9] [10]。これらの手法では高速に運動・変形する多様な手形状を追跡・識別可能であるが、物を持っていることで手の一部が隠れる場合を扱えない。天野らや川端らは、変動する背景の前に未知物体が存在するシーンに対し、隠蔽を含んだ背景全体の推定と推定された差分に基づく未知物体領域の検出を行った (BPLP 法) [11] [12]。変動するのが前景である手と物体領域の場合にもこの手法は適用可能であるが、単一の変動背景を学習することに目的があり、複数の前景物体を識別することは行われていない。

本研究は、物体の使用目的に対応した手形状と使用する物体形状を把持パターン毎に同時に学習し、把持パターンを認識し、把持パターンから物体の使用目的を認識する手法を提案する。例えばカッターや、ペンのような細長く形状が似ている物体の把持パターンを用いて識別することを目的とする。このとき、複雑背景下の物体把持シーンでは、領域が正確に分割出来ない場合があるので、把持パターンごとの手と物体の見えの共起性を事前に主成分分析 (以下、PCA) によって学習しておき、いったん把持パターンを仮定した上で BPLP 法で欠損領域の判定と復元を試みる。複数の把持パターン候補のうち、入力画像と復元画像が十分似通っているものを識別結果として判定する。本手法では以上の方法で把持パターンを識別することによって、機能別に物体を分類する。

2. 把持パターンと物体形状に基づく物体認識

2.1 鎌倉らの把持パターン

本研究では鎌倉らの把持パターン [8] を用いて物体の用途を識別する。この把持パターンを用いる理由の一つは、我々にとって身近な日常品の分類を用いて分類を行っている点、二つめは、力を必要とする握力把握・細かな作業を行う精密把握系の他にそれらの中間に位置する中間把握系を導入しており、日常品の物体分類を詳細に行うという点において適しているためである。全 15 種類ある把持パターンから、把持形状が類似したパターンを除外した結果、6 つの把持パターンを選択し利用することにした。以下にその 6 種類の把持パターンについて簡単に整理する (図は鎌倉の論文 [8] より引用した)。

a) 握力把握-標準型 (Power Grip-Standard Type,PoS)

これは最も安定した把持の 1 つで、指で物体を上手く包みこむようにして把持する。ハンマー、ドライバー、フライパンなどを保持する場合に使用される。



図 2 握力把握-標準型 (Power Grip-Standard Type,PoS)

b) 握力把握-示指伸展型 (Power Grip-Index Extension Type,PoI)

この握力把握では、人差し指を伸ばして物体の先端に触れ、

物体の先端の 1 つの方向へ大きな力が作用し目的を達成する。フォーク、ナイフなどを使う場合に使用される。

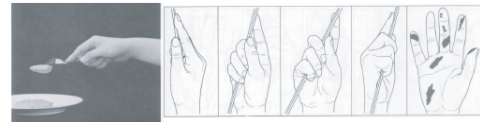


図 3 握力把握-示指伸展型 (Power Grip-Index Extension Type,PoI)

c) 精密把握-並列軽屈曲把握 (Parallel Mild Flexion Grip,PMF)

この把持は物体を拾い上げるか、それらを軽く保持する場合に使用する。全ての物体を拾うとき、カップなどの物体を持つ場合に使用される。

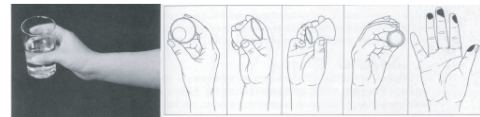


図 4 精密把握-並列軽屈曲把握 (Parallel Mild Flexion Grip,PMF)

d) 精密把握-指尖把握 (Tip Grip,Tip)

親指および人差し指の先端が物体に接している。これは小さな物体を操作するという目的がある。針、コイン、ピンなどを保持する場合、この把持が使用される。

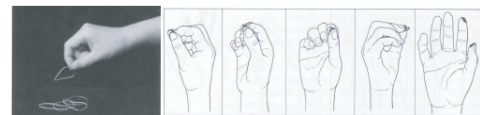


図 5 精密把握-指尖把握 (Tip Grip,Tip)

e) 中間把握-三面把握-標準型 (Tripod Grip-Standard Type,Tpd)

この把持では、親指、人差し指および中指が一般に円筒状のツールの先端で機敏な操作をするために使用される。口紅やペンなどで作業する場合、この把持が使用される。

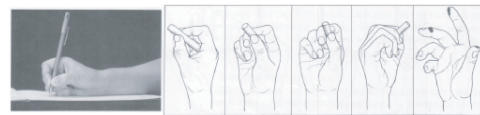


図 6 中間把握-三面把握-標準型 (Tripod Grip-Standard Type,Tpd)

f) 中間把握-側面把握 (Lateral Grip,Lat)

手と比較して、物体が小さく、完全に指によって単独で包むことができる場合に使用される。これはコーヒーカップ、爪切り、はさみなどを使う場合に使用される。

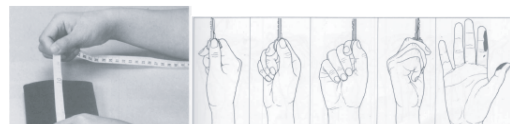


図 7 中間把握-側面把握 (Lateral Grip,Lat)

2.2 特徴画像の作成

本研究では、把持パターンごとに特徴づけられる手の見えと把持物体の見えの関連性を学習用画像を用いて学習することを試みる。その際、手と物体が映っている学習用画像から手領域や物体領域を適切に抽出し画像の正規化を行って特徴画像を生成する必要がある。特徴画像生成の手順を図 8 に示す。

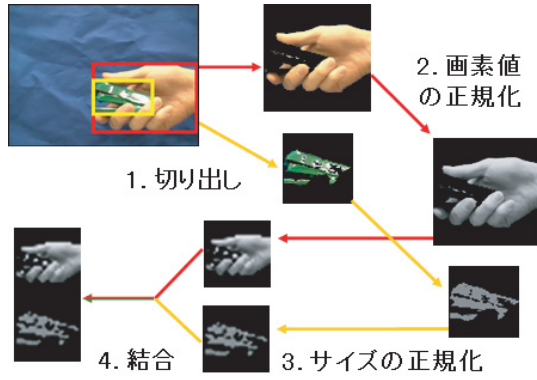


図 8 特徴画像の作成方法

1. 切り出し: 把持パターンごとに様々な姿勢の把持画像を集め、画像を手領域と物体領域に正しくラベリングする。簡単に自動ラベリングできるようにするため、単純背景とし、物体には単色のテープを巻いたものを用いる。背景差分により手と物体領域を抽出し、そのうち肌色の部分を手領域、テープ色の領域を物体領域とする。分割後微小領域を除去したのち手領域と物体領域を囲む正方形領域を求める。

2. 画素値の正規化: もとまった 2 つの領域の画素値の正規化を行う。手領域は、輝度値の平均が 100 になるようにシフトした濃淡画像を用いる。物体領域は物体ごとにテクスチャが異なるため、物体部分の輝度値を 100、それ以外を 0 とした二値画像を用いる。

3. サイズの正規化: 領域を指定するサイズ $v \times h$ に変更する。

4. 結合: 手と物体、2 つ領域を連結し、幅 v 、高さ $2h$ の特徴画像とする。

2.3 把持パターンの学習

次に手と物体形状の共起性を PCA を用いて学習する。収集した学習用サンプルから作成した特徴画像をラスタースキャンして特徴ベクトル $X = (x_1, x_2, x_3, \dots, x_p)$ を作り、特徴画像を $p = v \times h$ 次元空間上の点とみなす。各把持パターンごとに共分散行列を作成し ((1) 式)、その固有値と固有ベクトルを求める。累積寄与率が閾値 (実験では 60 % とした) に達するまでの固有ベクトルを用いて固有空間を作成する。

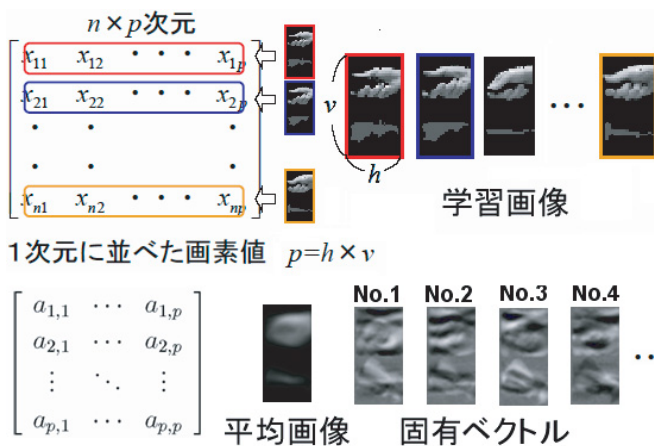


図 9 手と物体の共起性の学習

2.4 学習画像を用いた欠損領域復元

固有空間上では、把持画像から作成された特徴画像を固有ベクトルの線形結合で表現するため、入力特徴画像を各把持パターンごとの固有空間に射影し、その距離に近い把持パターンを認識結果とすることが考えられる。入力特徴画像の固有空間への単純な射影は、全ての画素の誤差の重みを一定とした場合の最小二乗法による近似に等しい。したがって、抽出された手および物体領域に欠損を含んだ画像を単純に射影すると、抽出誤りの画素値の影響が大きいため正しい近似が得られず識別を誤ることがある。

どの画素が欠損画素かを判定できれば、欠損画素以外の画素による誤差の二乗和を最小化することで、最も入力画像に近い復元画像を得ることができる (BPLP 法 [11], (2) 式)。

$$\hat{p} = \operatorname{argmin} |(E\hat{p} - \tilde{x})^T \Sigma (E\hat{p} - \tilde{x})| \quad (1)$$

- $\hat{p}$: 特徴画像ベクトルの固有空間上への射影
- $\tilde{x}$: 初期特徴画像ベクトル
- Σ : 各画素に対応する対角成分が欠損画素の場合 0、さも

なくば 1 となる対角行列

そこで、川端ら [12] の提案する反復 BPLP 法を用いて、逐次欠損画素の判定をしながら復元を行うことによって、正しい射影点ならびに復元画像を得る。処理手順を以下ならびに図 10 に示す。

欠損領域の復元手法

- (1) 2.2 節の方法より入力画像から初期特徴画像 $\tilde{x}$ を作成
- (2) 初期特徴画像 $\tilde{x}$ を把持パターン i ごとに作成した固有空間 E_i に射影し、復元画像 $E_i \hat{p}_k$ を得る
- (3) 得られた復元画像 $E_i \hat{p}_k$ と初期特徴画像 $\tilde{x}$ の残差を取る
- (4) 閾値で二値化した残差画像 (マスク画像 R_k) の 0 でない画素を欠損画素として判定し射影復元画像のうち欠損画素でない画素を初期特徴画像 $\tilde{x}$ の値に置き換えた推定画像 $\hat{x}_{k+1}$ を作る
- (5) $\hat{x}_{k+1}$ を再び固有空間 E_i に射影し、復元画像 $E_i \hat{p}_k$ を得る
- (6) (3)~(5) の処理を復元画像の変化が収束するまで繰り返す

上記復元処理を式のように (2), (3) 式のように表現できる。

$$\hat{p}_k = E^T \hat{x}_k \quad (2)$$

$$\hat{x}_{k+1} = R_k E_i \hat{p}_{k+1} + (I - R_k) \tilde{x} \quad (3)$$

- $\hat{x}_k$: k 回目の推定画像ベクトル
- $\hat{p}_k$: k 回目の固有空間上への射影画像ベクトル
- E_i : i 番目の把持パターンの固有ベクトル
- I : 単位行列
- R_k : k 回目の残差画像 (欠損画素を 1、欠損してない画素を 0 で表す対角行列)

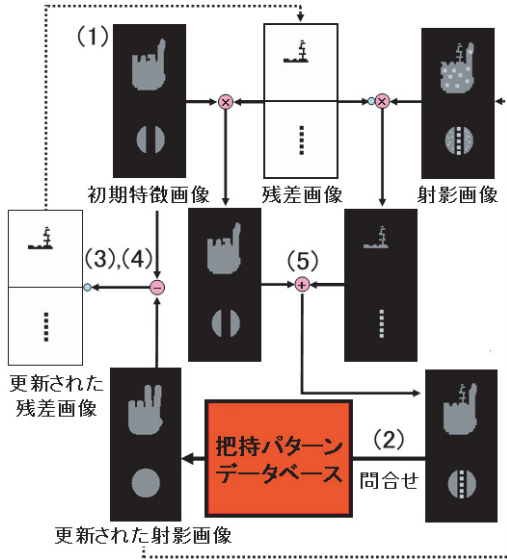


図10 欠損領域の復元プロセス

以上の計算で得られる収束先の $\hat{p}_k$ は、(1) 式において $\Sigma = I - R_k$ としたときの $\hat{p}$ に等しい。

川端らは反復 BPLP 法を変動背景の推定と未知の前景物体の検出に用いているが、本手法では、背景ではなく前景が変動する(基底により空間を張る)ことと、単なる画素ではなく、各把持パターンを特徴付ける手領域と物体領域の共起性を基底に選んでいることが異なる。そのため、手領域、物体領域のどちらかに欠損がある場合でも、他方の検出結果から欠損を復元することが期待される。

図11を用意して単純な射影と復元処理を行った。

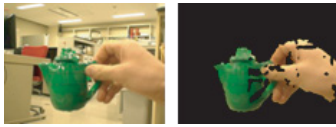


図11 右:入力画像 左:背景差分の結果

復元結果は図12のようになる。図11は、手作業で正確に背景と物体領域の分割を行った後特徴画像化したものである。そして図12と図13から差を取り図14を得た。



図12 左:図11から作成した特徴画像
中:単純射影により得た復元結果
右:反復 BPLP 法による復元

図13 図11を正確に領域分割・作成した特徴画像

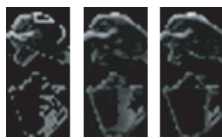


図14 図12と図13の差をとった画像

図14の画像より単純な投影を行った結果と比べ、より正確に復元ができていることが分かる。

2.5 把持パターンの認識：静止画像に対する処理

認識対象の把持シーンに対して、もし正しい把持パターンを仮定して前節の復元処理を行えば、ノイズの影響が十分小さい状況では初期特徴画像と復元画像間の残差は小さいはずである。逆に誤った把持パターンを仮定した場合は、多くの画素がノイズと判定された復元画像が得られ、結果として初期特徴画像との残差が大きくなる。すなわち、前節の復元処理を各把持パターンに対して行うことで、把持パターンを認識できる。

入力画像から作成された初期特徴画像から各クラスごとに復元処理を行い、初期特徴画像と復元画像との距離 D が閾値以下である場合に、最も D が小さい把持パターンであると判定する。

$$D(C_k) = \sqrt{\sum_{i=1}^P (\tilde{x}_i - \hat{x}_i^{C_k})^2} \quad (4)$$

- C_k : k 番目の把持パターンクラス
- $\tilde{x}_i$: 初期特徴画像の i 番目の画素値
- $\hat{x}_i^{C_k}$: k 番目の把持パターンクラスを仮定して復元された特徴画像の i 番目の画素値

画像をラスタスキャンしたベクトルを生成しているため、手の位置や物体の位置検出が平行もしくは回転移動によってずれた場合、領域の縁部分は復元の結果欠損画素と判定されやすいため、その画素の残差が大きく評価されることになる。領域の縁以外の部分でも、影がさしやすいつまみや、物体ごとの個体差が大きい部分では同様の事象が発生する。

そこで、各把持パターンごとに欠損領域の出現傾向を考慮して距離を評価する。

2.5.1 画素ごとの欠損発生率を考慮した距離評価

各把持パターンに属する全ての学習画像に対し、対応する把持パターンに基づいて2.4節の復元処理を行い、学習画像と復元画像との差分を $[0,1]$ で二値化した残差画像を作成し、さらに残差画像の平均(平均残差画像)を作成する。平均残差画像は各画素ごとの欠損発生率を表している。そこで(5)式を用いて、欠損発生率による重み付けを考慮した距離 D' を求める。

$$D'(C_k) = \sqrt{\sum_{i=1}^P (\tilde{x}_i - \hat{x}_i^{C_k})^2 / A_i(C_k)} \quad (5)$$

- $A_i(C_k)$: 把持パターンクラス C_k の平均残差画像の i 番目の画素値

2.5.2 欠損発生率の位置相関性を考慮した距離評価

2.5.1節の処理では、各画素の欠損発生率が独立であるという暗黙の仮定を置いている。実シーンにおいては欠損の発生は必ずしも独立ではなく、隣接した画素間の相関が見受けられる。そこで、各画素間の欠損発生率の相関性を考慮した距離評価を考える。

2.5.1節と同様に、各把持パターンクラスに属する学習画像から残差画像を作成し、残差画像ベクトルの2次モーメント行列を求める。これより式(8)のようにマハラノビス距離 D'' を求める。

$$D''(C_k) = \sqrt{\sum_{i=1}^P \sum_{j=1}^P (S_{ij}^{-1} (\tilde{x}_i - \hat{x}_i^{C_k}) (\tilde{x}_j - \hat{x}_j^{C_k}))} \quad (6)$$

- S_{ij}^{-1} : 残差画像ベクトルの2次モーメント行列の逆行列の (i, j) 成分

2.6 把持パターンの認識：時系列画像に対する処理

時系列画像が得られる場合には、各フレームの画像を用いて、静止画像に対する認識処理を行い、その結果を時系列にわたって積算することで把持パターンの判定を行う。各フレームごとに前節の距離 D (あるいは D' ないし D'') を求め、最小の距離に対応するクラスに票を投じる。最小の距離と次点の距離がある程度離れていれば、確実に対応する把持パターンクラスと決めることが出来るため、経験則で定めた閾値を定めて、確実に判断できる場合とあいまいな場合とで場合分けを行う。手と物体が連続して検出されているときは、同じ物体を持っていると判断できるため、手と物体の画像追跡を行いながら投票を続けていく。手と物体が画像から検出できなくなったとき (画像の外に出た場合) にはその時点で投票をリセットする。

投票条件

各把持パターンクラスの距離 $<$ 閾値 となるクラスのみ以下の投票対象とする。

if: 距離最小となる把持パターンクラスの距離 (以下最小距離) $\times 1.2 <$ 次点クラスの距離 距離最小の把持パターンクラスが、次点以下のクラスよりも十分優位と判断し、距離対応する把持パターンクラスに 1 票投票する

else if: 最小距離 $\times 1.2 >$ 他把持パターンクラスの距離、となるクラスが存在する複数の把持パターンクラス間で唯一に絞ることができないと考え、1 票を該当するクラスに均等分配して投票する

3. 実験と考察

今回は 2.1 節で記述した 6 種類の各把持パターンについて、表 2 に示す 18 種類の物体を学習した。13 人の被験者に各物体を指定したパターンで把持してもらい、持ち方を多少変化させたり手自体を微小に平行移動や回転させた様子を動画で撮影した。その動画を静止画に分解し、各把持パターンごとに 9000 枚ずつ把持画像を収集した。

把持画像から 25×50 ピクセル、1250 次元の特徴画像を作成し、作成した把持パターンごとの特徴画像から固有空間を用いて経験則より 80 次元に圧縮し把持パターンクラスを作成した。なお累積寄与率は約 65 % である。また 2.5 節より欠損発生率を表す平均残差画像を作成した。

作成した把持パターンクラスごとの平均画像と平均残差画像、固有ベクトルを図 15 に示す。図 15 より学習が行われていること、欠損が手と物体領域の輪郭に発生しやすいことが分かる。

2.2 節の方法により、各把持画像から 25×50 ピクセル、1250 次元の特徴画像を生成し、2.3 節の方法で各把持パターンクラスごとに固有空間を作成した。固有空間の次元は累積寄与率が 65 % とするところで打ち切るようにしたところ、どのクラスもほぼ 80 次元程度となったため、一律に 80 次元とした。また 2.5 節の方法により、欠損発生率を表す平均残差画像を作成した。作成した把持パターンクラスごとの平均画像と平均残差画像、固有ベクトルを図 15 に示す。図より、手の把持形状と物体の現れる位置や形状の関係性が基底ベクトルによって表現され、また領域検出の欠損が手と物体領域の輪郭に沿って発生し

やすいことが分かる。

把持パターンクラス名	学習に使用した物体
握力把握-標準型 (PoS)	ハンマー、フライパン、ホッチキス
握力把握-示指伸展型 (PoI)	カッター、ナイフ、フォーク
並列軽屈曲把握 (PMF)	グラス、コップ、小皿
指尖把握 (Tip)	画鋸、磁石、クリップ
側面把握 (Lat)	コップ (小)、コップ (中)、コップ (大)
三面把握-標準型 (Tpd)	ペン (細)、ペン (太)、スプーン

表 1 学習した把持パターンクラスと物体

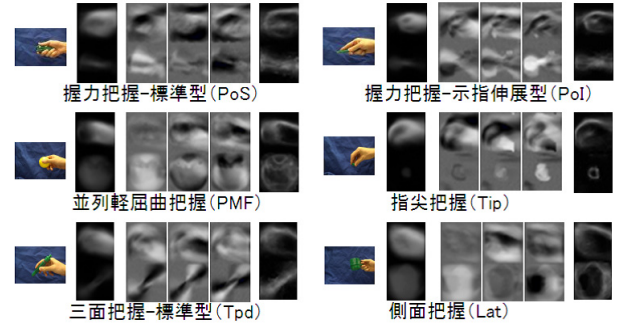


図 15 左:把持パターン毎の代表画像 中左:把持パターン毎の平均画像 中右:固有ベクトル 右:平均残差画像

評価基準の違いによる性能を比較するため、

- (1) 初期特徴と単純射影点間の距離を評価 (手法 1)
- (2) 反復 BPLP 法による復元を行い、(4) 式による単純距離を評価 (手法 2)
- (3) (5) 式による欠損発生率を考慮した距離を評価 (手法 3)
- (4) (6) 式による欠損位置の相関性を考慮した距離を評価 (手法 4)

の計 4 つの手法について物体認識実験を行い、その正答率を調べた。なお、1 つの把持パターンクラスに対しそのクラスに属する 4 種類の物体を把持した画像列を用意し、その画像列を 30 フレーム毎に分割した画像列を入力画像列とした。図 16 は入力の一例、図 17 は復元結果の抜粋である。

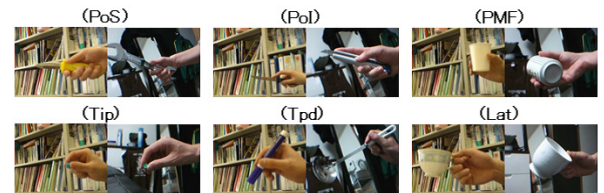


図 16 複雑背景下で撮影した動画からの抜粋

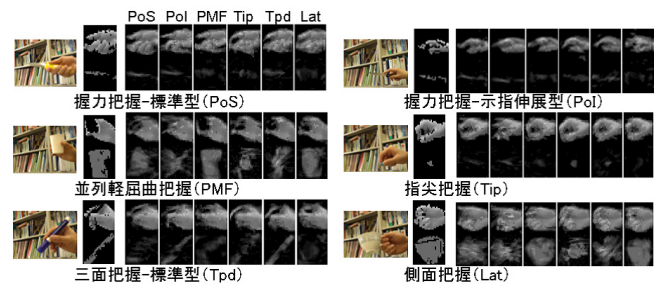


図 17 複雑背景下での入力画像、初期特徴画像、学習した各把持パターンクラスで復元した結果

表 2 は各物体の正答率である。

クラス (入力例)	手法 (1)	手法 (2)	手法 (3)	手法 (4)
PoS(ナイフ)	27(31) %	38(44) %	83(100) %	45(56) %
PoI (カッター)	67(75) %	53(75) %	65(75) %	89(100) %
PMF(紙コップ)	97(100) %	94(100) %	98(100) %	97(100) %
Tip(磁石)	84(100) %	81(94) %	9(0) %	55(69) %
Tpd(ペン)	85(94) %	85(100) %	78(94) %	38(38) %
Lat(マグカップ)	44(44) %	42(56) %	30(31) %	8(0) %

表 2 各手法で比較した結果 () 内の数字が時系列の結果

表 2 から手法 (1)(2) 間で大きな変化は見られない。これはいずれの手法においても、手や物体領域の輪郭に生じやすい欠損領域が距離に大きく影響を与えていることが原因だと考えられる。

そのことを考慮したのが手法 (3) である。手法 (3) は手法 (1)(2) に比べ、PoS, PoI, PMF, Tpd に関しては同等以上の正答率を得られた。しかし、Tip と Lat に関しては正答率が大幅に下がる結果となった。図 18 は、手法 (2)(3) を用いたときの、各把持パターンクラスに属する入力に対する各クラスへの認識率を一覧にしたものである。

(手法3)								(手法4)							
入力\認識	Pos	PoI	PMF	Tip	Tpd	Lat		入力\認識	Pos	PoI	PMF	Tip	Tpd	Lat	
PoS	39	12	2	0	1	0		PoS	45	46	6	0	1	0	
PoI	29	65	3	0	2	0		PoI	3	89	6	1	0	0	
PMF	0	1	98	0	0	0		PMF	0	1	97	0	0	2	
Tip	45	1	3	9	35	6		Tip	9	2	2	55	25	6	
Tpd	6	2	5	0	78	7		Tpd	0	13	30	1	39	16	
Lat	3	1	65	0	1	30		Lat	2	3	86	0	0	8	

■ : 各入力における最高認識率
■ : 10%以上の認識率

図 18 各クラスに属する入力に対する認識率 (%)

図 18 左の表から、手法 (3) は Tip に属する入力に対しては Pos, Tpd に誤認識、Lat に属する入力に対しては PMF に誤認識していることが分かる。これは、欠損可能性領域が中心に強く分布する PoS と Tpd が、欠損が中心に集中しやすい Tip の距離を小さくしてしまうことが原因である。Lat に関しても同様で、PMF の欠損可能性領域が Lat の中心に生じる欠損部分の距離をより小さくしてしまうことが原因である。図 19 に Lat の入力例、入力に対する PMF と Lat クラスにおける復元画像、PMF と Lat クラスの平均残差画像を示す。

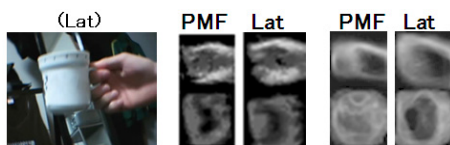


図 19 左:入力画像の抜粋 中:復元画像 右:平均残差画像 (再掲)

物体の欠損位置の相関性を考慮したのが手法 (4) である。手法 (4) では、手法 (1)(2) に比べて PoS, PoI, PMF は正答率が向上するが、Tip, Tpd, Lat の正答率は低下した。また、手法 (3) に比べて PoI, PMF, Tip は正答率は同等以上になるが、PoS, Tpd, Lat の正答率は低下した。

図 18 右の表から、手法 (4) は PoS に属する入力を PoI に誤認識しやすい。Tpd, Lat に関しても誤認識により票の分散が起こるが、これは物体の形状が類似するなどの要因で、入力に

対する欠損位置の相関が高くなるクラスが他にも存在することが原因だと考えられる。ゆえに、物体領域に対してのみ欠損位置の相関を考慮している現在の方法から、手領域を含めて欠損位置の相関を取る必要があると言える。

4. まとめと今後

本研究は、物体把持シーンにおける手形状と物体形状の共起性を学習することにより、把持パターンに由来する物体の機能に基づいた物体認識を行う手法を提案した。その際、手・物体両領域の共起性に基づいて、背景差分等による物体領域切り出しの際に生じる欠損画素の復元を同時に行うことで、復元による画像の修正が最小となる把持パターンクラスを認識結果とする判定基準を導入した。6 種類の把持パターンについて合計 24 種類の物体の識別実験を行い、本手法の有用性を確かめた。

把持パターンを決める一つの要素として手と物体の位置関係があげられる為、手と物体の位置情報を考慮することが認識率の向上に繋がるかを確認することが今後の課題である。また手や物体のモデルとして画像の見えを直接用いているので、照明や影など撮影環境の影響を強く受ける。3 次元モデルを利用した詳細な手形状の推定手法と組み合わせることで環境の変化に対する頑健性と把持パターンのより詳細な識別を行うことも将来の課題として上げられる。

文 献

- [1] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," Proc. on IEEE Conf. on CVPR Vol.1, pp.511-518, 2001.
- [2] G. Csurka, C.R. Dance, L. Fan, and C. Bray, "Visual categorization with bags of keypoints," Proc. of IEEE Conf. on ECCV, pp. 1-22, 2004.
- [3] N. Dalal, B. Triggs, "Histograms of oriented gradients for human detection," Proc. of IEEE Conf. on CVPR, pp. 886-893, 2005.
- [4] Y. Sato and T. Nagai, "Learning of Object Concept Through Function and Shape," IEEE Region 10 Conf.on TENCON 2006, pp.1-4, Nov. 2006.
- [5] 吉田成希, 北橋忠宏, "人間動作を通じての物体の機能的認識," 電子情報通信学会技術報告, PRMU2002-213, pp.13-18, Feb. 2003.
- [6] 樋口未来, 小島篤博, 福永邦雄, "物体の機能と人物の動作解析によるシーン認識," 電子情報通信学会技術報告, IE2003-36, PRMU2003-66, MVE2003-48, pp.53-58, Jul. 2003.
- [7] J. Napier, "The prehensile movements of the human hand," J. Bone and Joint Surgery, 38B,4, pp.902-913, Nov. 1956.
- [8] Noriko Kamakura, "手のかたち 手の動き", 医歯薬出版株式会社, May. 1989.
- [9] A. Imai, N. Shimada and Y. Shirai, "3-D Hand Posture Recognition by Training Contour Variation," Proc. of 6th Int. Conf. on Automatic Face and Gesture Recognition, pp.895-900, May. 2004.
- [10] Y. Hamada, N. Shimada and Y. Shirai, "Hand Shape Estimation under Complex Backgrounds for Sign Language Recognition," Proc. of 6th Int. Conf. on Automatic Face and Gesture Recognition, pp.589-594, May. 2004.
- [11] 天野敏之, 佐藤幸男, "固有空間法を用いた BPLP による画像補間," 電子情報通信学会論文誌 D-II, Vol.J85-D-II, No.3, pp.457-465 Mar. 2002.
- [12] 川端聡, 日浦慎作, 佐藤宏介, "動的背景の学習による未知物体領域の実時間抽出法," 電子情報通信学会論文誌, Vol. J89-D, No. 4 ,pp.826-835, Apr. 2006.