

食器片付けロボットのための画像処理システム

-複数手法からの選択・組合せによる食器検出と多様な動作環境対応へのアプローチ-

佐藤知正 門脇啓太 森武俊 杉正夫 福井類 (東京大学)

Image Processing System for Home Assistant Robot

-Selection and Combination of Methods for Object Detection in Various Environment-

Tomomasa SATO, *Keita KADOWAKI, Taketoshi MORI, Masao SUGI, Rui FUKUI
(The Univ. of Tokyo)

Abstract— For home assistant robot clearing a table, we aim to construct an image-processing system that obtains information the robot can use for grasping. We focus on tableware detection on a table by the system. We propose an approach that the system has several image-processing methods and the user makes a choice to have the system adjust to the environment the robot works in. For the system, we prepared these methods: (1)using a background image (2)segmentation using the model of a table represented by Ganssian Mixture Model or color histogram. By experiments, effectiveness of our proposed approach and validity of methods we choose were confirmed.

Key Words: Home Robot, Image Processing, Approach for Various Environment

1. 緒論

ロボットを活用した人の活動の円滑化を考え、本論文においてはロボットの行うタスクとして食後の食器片付けを取り上げる。このタスク実現のためには、把持対象となる食器の位置、種類、姿勢などを特定する必要がある。よって、ロボットが食器片付けを行う際に必要な食器の位置、種類といった情報を得る画像処理システム (Fig.1) の構築を目指し、本論文では特にテーブル上の食器検出を行う。ここで、ロボットとしては Fig.1 左に示すようなものを想定している。

また、人間の活動する場へのロボット導入の実現に向けて解決すべき問題のひとつが、多様な動作環境への対応である。例えば一般家庭でロボットが作業することを想定した場合、使用されている家具・家電の種類やその配置、部屋の間取りや設備、照明条件などは家ごとに異なる。また、ひとつの家庭においても、部屋の環境は時々刻々変化する。食器片付けに関しては、使用される食器やテーブル、照明条件などの家庭ごとの違いが挙げられ、テーブル上の小物やテーブルクロ

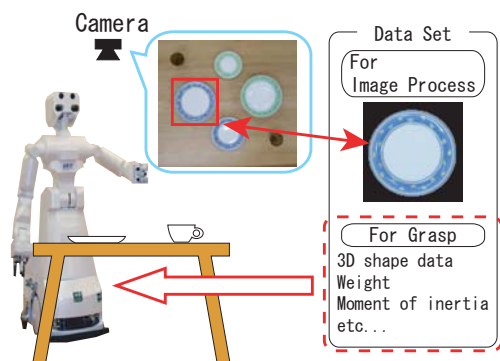


Fig.1 Conceptual sketch of image processing system

ス、ランチョンマット等の位置、種類も一回の片付け毎に変わる。このような多様な動作環境への対応として、以下のようなアプローチ (Fig.2) を提案する。

1. 特徴の異なる複数の画像処理手法を実装する
2. 処理結果をユーザに提示する
3. 提示された結果の中から、使用するものと統合方法をユーザに選択してもらう
4. ユーザの選択から、その動作環境に関する知識を得て利用する

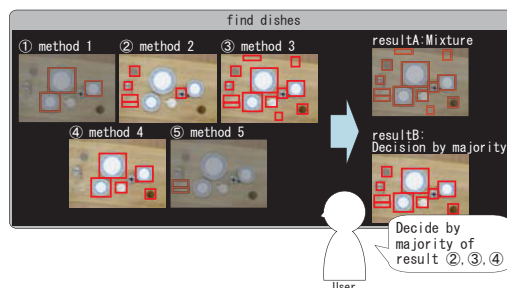


Fig.2 Conceptual sketch of Teaching interface for user's selection

このように、システム導入後にユーザの介入によって動作プログラムの調整を行うというカスタマイゼーションアプローチは、本研究における画像処理システムに限らず、人の活動する場において動作するロボットやそれに付随するシステムに有効であるといえる。上記アプローチ実現のためにシステムが持つべき画像処理手法の方向性について考察する。

2. 画像処理による情報に基づく食器の把持

画像処理により得られる情報を用いて物体を把持するシステムは、工場などの生産現場における使用を目

的とした研究が多くなされており [1, 2], 高精度かつ高速な作業が可能となっている。しかしこれはひとえに, 限られた動作環境における使用を想定し, また, 環境側がロボットののために最適化されている結果であり, これらが人の生活空間においても同様の性能を発揮するのは難しい。また, 食器の把持を目的とした研究には, 画像処理を用いるほかに, 不可視マークと RFID タグを併用したシステム [3] 等がある。画像のみによるシステムと比べ容易に食器を検出できるが, 食器に対する特殊なコーティングや RFID タグの取り付けには導入およびメンテナンスの手間とコストがかかる。この点, 画像処理によるシステムの場合, 現在使用している食器をそのままシステムが扱うことができるという利点がある。本章ではまず, 本研究において構築する画像処理システムの概要を示し, 続いてシステムの要求仕様とその実現方法を述べる。

2.1 画像処理システムの概要

本研究において構築する画像処理システム (Fig.1) は, 部屋の天井付近に設置したカメラを用いてロボットが食器の把持を行うために必要な情報を得るシステムである。

カメラを天井付近に設置する利点としては, 真上から見ることによる認識問題の簡略化, ロボットの作業エリア全体の把握といったことが挙げられる。

[認識問題の簡略化] 例えばロボットの目の位置に設けられたカメラを用いる場合, ロボットの位置や姿勢による食器の見え方の変化, 背景の多様化が生じ, また, 食器には特徴的な模様等が少ないものが多いことから, その位置や種類の特定は非常に非常に困難である。また取っ手のあるコップや楕円形の皿など, 把持する際に向きの情報が必要となる場合でも, 取っ手の向きや楕円の長軸方向を求めるのが容易となる。

[作業エリア全体の把握] 上方からの視点を有することにより, ロボットからは物体の陰に隠れて見えない場所の情報を得られ, また, テーブル上に食器が残っているかどうかの判断や, 片付ける順番の指示といったことも可能であると考えられる。

2.2 画像処理システムの要求仕様とその実現方法

本研究の最終的な目的のひとつは, ロボットによる食器片付けを実現することである。この目的を達成するための画像処理システムの要求仕様は

- 把持対象食器の位置情報を得ること
- 把持対象食器の形状, 重量等の情報を得ること

となる。食器が円形である場合を除けば, 食器の置かれている向きも必要である。これらの要求仕様を満たすため, 前提条件として以下の情報を既知であるとする。

- 食器を真上から撮影した写真のような, 画像処理に利用できる情報
- 食器の三次元形状データや重量, 慣性モーメントなどの把持に利用できる情報

これは, 家庭によって使われている食器が異なるため食器の種類の特定にはテンプレート画像が必要であり, また, お椀の内側のような詳細な形状や重量などの情報を画像処理のみで得ることは困難なためである。こ

れらの情報は, 将来的には食器メーカー等が公開しており, ユーザがダウンロードすることでシステムへ登録されるようになるものと想定している。

以上のような前提条件のもと, 画像処理システムは次のようにその要求仕様を実現する。

1. 画像処理によって食器の位置と種類を特定する
2. システムに登録済みの食器のデータと併せてロボットに渡す

こうして得られた食器の位置および形状等の情報をもとにロボットが食器を把持し, 片付けを行う。本論文では, 上記のうち食器位置の特定までを扱う。

3. 実装する複数手法の検討

本章では, 画像処理システムが本論文で提案するアプローチ実現のために用いる食器検出手法について述べる。まず, 処理手法の検討を行い, 続いて, 実装する処理手法を示す。

3.1 食器検出手法の検討

画像処理システムは, まずは食器の存在する領域の特定を行う。これは, 食器は概して模様等のテクスチャ特徴に乏しく, 鏡面反射を起こすものも多いため, テンプレート画像を用いて位置と種類を同時に特定するのが困難なためである。

食器存在領域の特定には次の 2 通りの手法を用いる。

- 背景画像を用いる手法
- テーブルのモデルを用いる手法

[背景画像を用いる手法] 食後にテーブル上にある食器を検出することは, 食事前にテーブル上に存在しなかったものを検出することと等価である。よって, ユーザの指示等により食事前の画像を取得し, 片付けを指示された際の画像との差分を取ることで食器を検出する。これは, 部屋内の物品の移動や照明の変化が小さい場合に特に有効である。

[テーブルのモデルを用いる手法] 以下のような状況に対応するには, 処理対象画像のみから操作対象を検出する手法が必要である。

- テーブル上の, 片付け対象となる食器以外の物品の移動
- テーブルを含む家具の移動

背景画像を用いない手法では, テーブル面のモデルを生成し, テーブルとそれ以外を識別することで食器の検出を行う。

3.2 背景画像を用いた食器検出手法

背景画像を用いる処理の流れは以下ようになる。

1. 食事前後の画像を用いて背景差分
2. モルフォロジー変換
3. ラベリング
4. Blob サイズによるノイズ除去

背景差分は, グレースケール画像と色相値それぞれを用いて入力画像 (Fig.3(a)) と背景画像 (Fig.3(b)) の差分を計算し, 大津の手法による 2 値化を行った後, 2 つの結果の論理和を取る (Fig.3(c))。得られた 2 値画像に縮小・拡大処理を行い, ノイズを除去する。ノイズ除去後の 2 値画像をラベリングし, Blob のサイズによる

絞り込みを行う (Fig.3(d)) . ここで, Blob のサイズとは, 画像の縦軸・横軸方向のピクセル間最大距離としている . これは, システム導入時にはカメラのキャリブレーションが必要であり, キャリブレーションで得られたデータから, 使用される食器が画像内でどの程度の大きさ以内に納まるかという事前知識が得られるため妥当である . 最後に, 得られた Blob から食器領域を特定する (Fig.3(e)) .

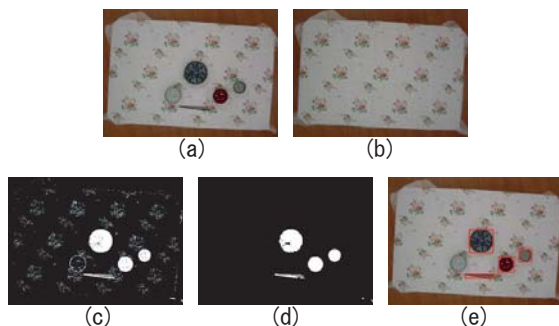


Fig.3 Tableware detection by using background image

3.3 テーブルのモデルを用いた食器検出手法

処理対象画像中のテーブル周縁部からテーブルのモデルを生成し, テーブルとそれ以外を識別する . モデル生成の際, 本論文においてはテーブル面の検出は行っておらず, 人が矩形を指定することで代用している . テーブルのモデルとして, 混合ガウス分布と色ヒストグラムの 2 種類を用いる .

3.3.1 混合ガウス分布を用いる手法

テーブルの色モデルに混合ガウス分布を用いる手法として GrabCut[4] を利用する . 指定された背景領域 (Fig.4(a) の 2 つの赤い矩形の間) と, それ以外の未知領域 (Fig.4(a) 内側の赤い矩形の内部) それぞれから RGB 色空間の混合ガウス分布を求め, 各ピクセルの色がそれぞれの混合ガウス分布に含まれる尤度を用いて前景領域と背景領域の分割を行う (Fig.4(b)) . 得られた前景領域をラベリングし, 3.2 と同様に Blob サイズによるノイズ除去を行い, 食器領域の特定を行う (Fig.4(c)) .

3.3.2 色ヒストグラムを用いる手法

色ヒストグラムは RGB 画像を用いて, R, G, B 各チャンネル毎に計 3 つ作成する . 各チャンネルは 0 から

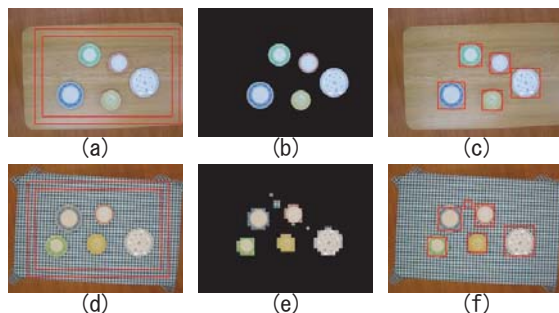


Fig.4 Tableware detection by using sole image

255 の整数値をとり, ヒストグラムのピンは 256 個である . まず, テーブル領域 (Fig.4(d) の 2 つの赤い矩形の間) の色ヒストグラムを作成する . 次に, 処理対象領域 (Fig.4(d) の内側の赤い矩形の内部) をグリッドに分割し, 各グリッドごとに色ヒストグラムを作成する . なお, 各ヒストグラムは, 全ピンの値の合計が一定の値になるように正規化する . テーブル領域とグリッドのヒストグラム間距離を計算し, 閾値を越えるグリッドを食器領域とみなす (Fig.4(e)) . なお, ヒストグラム間の距離には Bhattacharyya 距離を用いる . 食器領域とみなされた領域をラベリングし, 3.2 と同様に Blob サイズによるノイズ除去を行い, 食器領域の特定を行う (Fig.4(f)) .

4. 実験

画像処理システムの性能を確認する 2 種類の実験を行った .

4.1 画像処理手法の性能評価実験

[実験の目的] 本実験では, テーブルと食器の組合せを変えて, 実装した各画像処理手法を用いて食器の検出を行う . これにより, 各手法の性能を比較する .

[実験の設定] 本実験では, 部屋の天井にデジタルカメラを設置し, そのほぼ真下にテーブルを配置した . 使用した食器は 32 種類, テーブルはテーブルクロス等で表面のテクスチャを変更し 4 種類とした . 撮影した 1 枚の画像につき, 各画像処理手法により食器の検出を行った . 撮影した画像はテーブルクロスごとに 37 枚の計 148 枚 . 写った食器の総数はテーブルクロスごとに 67 個の計 268 個 .

[実験結果] 全画像に対して食器検出を行った結果を, 手法ごとに Table.1 に示す .

Table 1 Experiment result

method	precision(%)	recall(%)
method 3-2	62	75
method 3-3.1	78	29
method 3-3.2	21	46

Table.1 より, 単一の手法で全ての状況に対応することの困難さが確認できる . テーブルの種類ごとに実験結果を比較して明らかとなった各手法の差異を以下に述べる . 手法毎の差異が現れた典型的な例を Fig.5 に示す . Fig.5 において, 各段は同じ処理対象画像に対する結果を示しており, 左から順に, 3.2, 3.3.1, 3.3.2 で挙げた手法による処理結果となっている . 各図において, 赤い矩形が検出された食器領域を示している .

3.2 に挙げた手法は, 今回の実験では食器以外の物品の移動をしていないため, テーブルと食器の組合せが変わってもある程度食器の検出に成功しているが, それでも誤検出が見られる (Fig.5(a),(d),(g)) .

3.3.1 に挙げた手法は, テーブル面が一様なテクスチャを有している場合には非常に良い性能を発揮した (Fig.5(e)) . しかし, コントラストの大きい部分が物体の境界とされやすいという手法の特性や, テーブル周縁部が必ずしもテーブル全体の色モデルを表すのに適

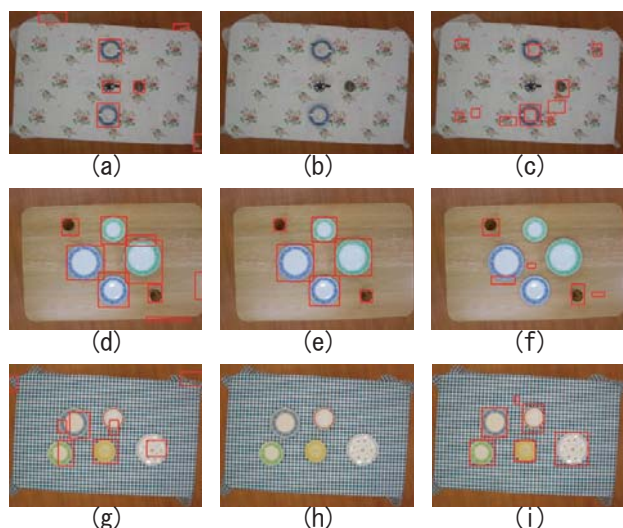


Fig.5 Typical results of each method

切でないことによる失敗も見られた (Fig.5(b),(h)) . 3.3.2 に挙げた手法は、格子柄のようなパターンが細かく繰り返すテーブル表面では他の手法より良い結果が得られた (Fig.5(i)) . しかし、テーブル上に離散的な模様がある場合 (Fig.5(c)) や、照明などによる部分的な明るさの変化がある場合 (Fig.5(f)) では非常に大きな領域を誤検出してしまふ .

以上の結果より、本論文において実装した手法にはそれぞれ得意不得意があり、それらに大きな重なりがないことが確認できた . このことから、実装した手法の方向性が妥当であること、複数の処理手法の中から選択または組み合わせることの有効性が確認された .

4.2 ロボットによる食器把持実験

[実験の目的] 本実験では、天井に設置したカメラから得られる位置情報と、そこに存在する食器の形状データを用いて食器の把持を行う . これにより、本研究で構築した画像処理システムを用いてロボットが設定タスクを実現できることを示す . なお、本実験が先行実験である

[実験の設定] 本実験は、天井に設置したカメラから得られる情報に基づくロボットによる食器の把持が可能かどうかの検証に主眼を置いている . 天井にデジタルカメラを設置し、そのほぼ真下にテーブルを配置した . 使用した食器は 25 種類であり、全て真上から見て円形である . テーブル上には把持対象となる食器をひとつだけ配置した . ロボットの手首にマーカを 2 点設置し、食器の中心位置と合わせて計 3 点を画像処理により取得する . なお、本実験においては、食器中心位置はハフ変換による円検出を用いて求めた . 得られた 3 点の座標と、事前に取得しておいたカメラキャリブレーションのデータを用いて食器とロボットの相対的な位置関係を算出する . 得られた位置情報を用いて、ロボットは食器ごとに決められた把持動作を行う .

[実験結果] 実験の様子を Fig.6 に示す . Fig.6(a),(c) がそれぞれ天井に設置したカメラからの画像であり、赤い円はハフ変換により検出されたものである . Fig.6(a),(c)

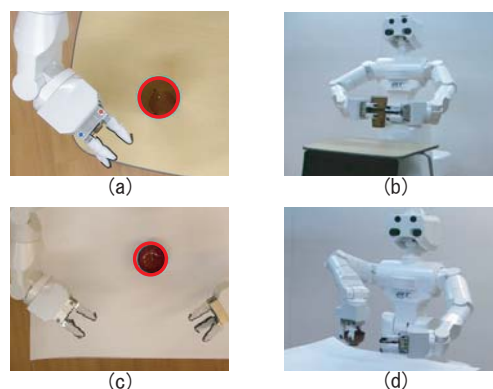


Fig.6 Robot grasps tableware using information by image processing system

各々で得られた情報に基づき食器把持を行っている様子が Fig.6(b),(d) である . 本実験では、使用した全ての食器の把持に成功し、天井に設置したカメラから得られる情報を用いて食器を把持することが不可能でないことを確認したが、把持の精度・安定性を保証するには至っていない .

5. 結論

本論文では、家事支援ロボットによる食器片付けを実現するための画像処理システムの構築を目指し、システムによるテーブル上の食器検出を行った . 人の活動する場での動作を想定しているため、画像処理システムには多様な動作環境への対応が望まれる . そこで、システムが用いる手法をユーザを介して複数から選択・組み合わせるといったアプローチを提案した . その実現のため、画像処理システムが実装すべき手法の方向性について考察し、背景画像との差分により食器を検出するものと、テーブルのモデルを生成してテーブルとそれ以外を識別するものを実装した . 実験によって実装した手法の妥当性を確認し、また、提案するアプローチの有効性を確認した .

最後に、本研究は先端融合領域イノベーション創出拠点の形成「少子高齢社会と人を支える IRT 基盤の創出」におけるトヨタ自動車株式会社との共同研究であることを記す .

参考文献

- [1] 恩田寿和, 藤原伸行, 阿部清秀, 森宣仁. 三次元円検出による部品位置決めと事前ハンド干渉チェックにより実現した視覚ベースピンピッキングシステム. 日本ロボット学会誌, Vol. 18, No. 7, pp. 995-1002, 2000.
- [2] 岡田伸廣, 長田正. 異種ビジュアルセンサを組み合わせたアクティブビジョンシステムとそのパーツピッキングへの応用. 日本ロボット学会誌, Vol. 14, No. 4, pp. 525-538, 1996.
- [3] 青柳誠司, 黒田友美, 新井泰彦, 見浪護. 不可視マークと RF タグを併用した食器の位置計測システム-計測精度の検証-. In *Proc. of JSME Conference on Robotics and Mechatronics*, pp. 10-12, 2007.
- [4] Carsten Rother, Vladimir Kolmogorov, and Andrew Blake. "grabcut": interactive foreground extraction using iterated graph cuts. *ACM Trans. Graph.*, Vol. 23, No. 3, pp. 309-314, 2004.