

画像上の輪郭特徴量を用いた食器認識システム

A camera-based dish recognition system using contour features

野村英祐, 林悠, 鏡慎吾, 橋本浩一

Eisuke Nomura, Yu Hayashi, Shingo Kagami, Koichi Hashimoto

東北大学大学院情報科学研究科システム情報科学専攻

Intelligent Control Systems Laboratory, Department of System Information Sciences,
Graduate School of Information Sciences, Tohoku University

キーワード： 食器認識(dish recognition), 特徴抽出(feature extraction), 投票処理(voting process),
ヒストグラム照合(histogram matching), テンプレートマッチング(Template matching)

連絡先： 〒980-8579 仙台市青葉区荒巻字青葉6-6-01 東北大学 工学部 機械系2号館 橋本・鏡研究室
野村英祐, TEL 022-795-7021, FAX 022-795-7019, E-mail: nomura@ic.is.tohoku.ac.jp

1. 緒言

POSシステムやセルフサービスの導入など、経営におけるあらゆる面での省力化・自動化が進む現在の外食産業において、レストランや食堂などのバックヤードで行われる食器洗浄・収納作業は、未だ自動化が実現していない分野のひとつである。ここで、食器洗浄作業とは、食事が終わり下膳された食器を食器洗浄機に並べ、洗浄する作業であり、食器収納作業とは、洗浄・乾燥後の食器を棚などに収納する作業である。これらの作業の自動化が実現すれば、作業の高効率化、作業に従事する労働力の削減などの効果が期待できる。

本研究では、食器洗浄作業の自動化を目指し、カメラから得た画像情報に基づく食器認識システムを構成する。認識には、画像から抽出した食器輪郭特徴量を累積投票処理することにより得られる、画像上での食器の位置

と回転に不変である累積ヒストグラムを用いる。この累積ヒストグラムを、食器認識システムで扱う全種類の食器について、予め作成したものを参照累積ヒストグラムとし、入力画像により与えられる一枚の食器から作成した累積ヒストグラムと各要素の投票数の二乗差分差分により比較し、最も差分の少ない参照累積ヒストグラムを持つ食器の種類を入力画像についての認識結果とするシステムである。このシステムの認識手法は輪郭の形状のみに注目するため、認識手法にテンプレートマッチングを用いるより必要な食器の情報量を減らすことができ、テンプレート画像に対し認識すべき画像が回転しているような場合についても、本研究での認識手法を用いることで認識ができる。また食器認識に適用する場合、幾何エッジヒストグラムを用いる照合^{2),3)}よりも高速に照合ができると考えられる。

本稿では、認識手法について説明し、システムが実現する食器の認識率・認識速度について考察する。

2. 認識手法

画像認識手法の手順は以下の通りである。

(1) カメラから得た8 bits グレースケール画像に対し、二値化，膨張・収縮，8連結-境界線追跡の画像処理を行い，食器の輪郭画素を取得する。(2) 食器の輪郭画素から輪郭の重心位置を求め，食器輪郭特徴量を抽出する。(3) 食器輪郭特徴量から累積ヒストグラムを作成し，食器の種類を照合を行う。

これらを順に説明していく。

2.1 食器輪郭画素の取得

一枚の食器を配置した平面に対し，鉛直下向きに設置したカメラにより撮影した8 bitsグレースケール画像がある。撮影した画像に二値化，膨張，収縮の画像処理を行う。得られた二値画像について，8連結の境界線追跡¹⁾を行い，境界画素を順次追跡する。これにより得られた境界画素は，食器の輪郭画素だけでなく，食器の模様や食器内部にできた陰影などの境界画素も含まれる。これらの境界画素の中から食器輪郭を表す輪郭画素を自動選択し，食器輪郭画素と食器輪郭の重心 G の画素を得る。

2.2 食器輪郭特徴量の抽出

輪郭上のある点 P に注目し，点 P と重心 G との距離 r ，重心 G から点 P へのベクトル p と点 P における接線方向ベクトルのなす角 θ を求める。実際のシステムにおいては，点 P は食器輪郭画素中のある一画素で与えられ，その画

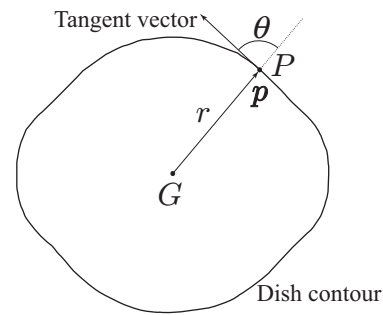


Fig. 1 食器輪郭特徴量 (r, θ)

素から食器輪郭画素の追跡順で数えて5画素分先にある画素に向かうベクトルを，接線方向ベクトルと近似し， θ を求めている。本研究ではこれらの (r, θ) を，食器の輪郭形状を記述する特徴量として，食器輪郭特徴量と呼ぶ。

2.3 累積ヒストグラムによる照合

食器輪郭特徴量 (r, θ) を，輪郭画素上の全ての画素 $P_n (n = 0, 1, \dots)$ について求め，食器輪郭特徴量の集合 $\{(r_n, \theta_n) (n = 0, 1, \dots)\}$ を抽出する。この集合に対し， (r, θ) それぞれの方向に一定区間 $(\Delta r, \Delta \theta)$ で分割し， $R_i = i\Delta r, \Theta_j = j\Delta \theta (i, j = 0, 1, \dots)$ からなる (R, Θ) 平面を作る。分割により定まる領域 $B_{ij} = \{(r, \theta) | R_i \leq r < R_{i+1}, \Theta_j \leq \theta < \Theta_{j+1}\}$ に，全ての食器輪郭特徴量 $\{(r_n, \theta_n) (n = 0, 1, \dots)\}$ を投票する。領域 B_{ij} への投票数を V_{ij} とすると，累積投票数は $C_{ij} = \sum_{k=0}^i \sum_{l=0}^j V_{kl}$ となる。この累積投票数より，図2のような累積ヒストグラム $H(R, \Theta)$ を作成する。累積投票数は代表点 (R_i, Θ_j) 上にプロットする。このように作成された累積ヒストグラムは，食器輪郭の画像上での位置と回転に不変であるため，同一の食器であれば累積ヒストグラムも同一となる。このことを照合キーとして利用し，予め十数種類の食器について作成された参照累積ヒストグラム $H_{\text{ref}}(R, \Theta)$ と，現在認識すべき食器の累積ヒストグラム $H(R, \Theta)$ を比較・照合し，食器の種類を識別す

る．しかし，実際には画像上のノイズの影響などから，同一の食器であっても累積ヒストグラムが完全に一致することは少ない．そのため本システムでは，全度数成分の差の二乗和 J が最小となる，すなわち

$$\min : J = \sum_{R, \Theta} \{H(R, \Theta) - H_{\text{ref}}(R, \Theta)\}^2 \quad (1)$$

となる累積ヒストグラムを持つ食器を識別結果とする．また，単に投票数によるヒストグラムではなく累積ヒストグラムを用いることで，抽出する特徴量に誤差が生じ，隣接する領域へ投票数が分散する場合でも，識別が可能となる．

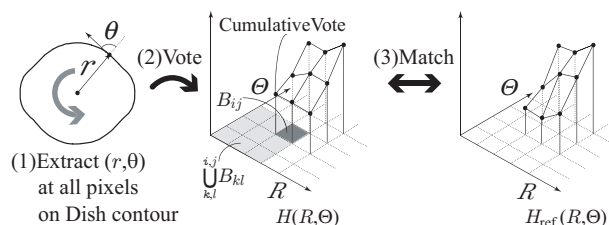


Fig. 2 累積ヒストグラムの作成と照合

3. システム構成

実験に用いたシステムを次項の図3に示す．本システムでは，計算処理にPentiumD 3.2 GHz 及びメモリ2 GByte を持つ PC を用い，カメラに Point Grey Research 社の Dragonfly Express を用いた．Dragonfly Express は最大 200 fps のフレーム速度が実現可能なグレースケールの CCD カメラで，IEEE1394 により PC に接続される．カメラは食器を配置する平面より上方850 mm の位置に鉛直下向に設置し，解像度640×480 pixels の画像を撮影した．プログラムの開発環境には Microsoft Visual Studio 2005 を用い，プログラミング言語には C++ を用いた．

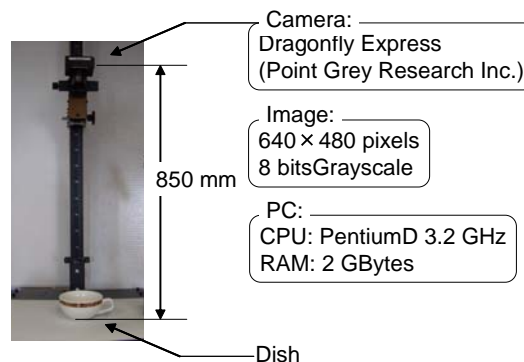


Fig. 3 実験システム

4. 実験

4.1 実験A

図4に示す9種類の異なる食器を認識対象とした．これらの食器を用いて，本研究で用いる認識手法の有効性を確認する．はじめに参照累積ヒストグラムの元になる参照モデルの食器画像を撮影した．この際に食器を画像の中心付近に配置して撮影し，レンズの歪みによる食器輪郭の歪みを軽減した．次に1種類の食器を画像の四隅に異なる姿勢で配置した画像を4枚，11種類で合計36枚撮影した．これは画像上での食器の位置・姿勢に影響を受けずに認識可能かを確認するためである．これらの画像を用いて，2で述べた認識手法により食器の認識を行った．このとき二値化の閾値，および膨張・収縮処理の回数は，食器の種類により適宜決定した．ヒストグラムの分割数については， r 方向は300 pixelsを300分割， θ 方向は3.15 radを50分割した．次に累積ヒストグラムを作成する際の (r, θ) の分割数を変化させて，同様の実験を行う．これにより，累積ヒストグラムを作成する際の分割数が，認識率と認識時間にどのような影響を与えるかを確認する．次にカメラにより撮影する画像の解像度を変化させて，同様の実験を行う．これにより，解像度の変化が認識率と認識時間にど



Fig. 4 実験に用いた9種類の食器

のような影響を与えるかを確認する。

4.2 実験B

実験Aに対し本実験では、照明条件および二値化の閾値を全ての食器について共通の値を利用して認識を行い、認識率にどのような影響を与えるかを確認する。画像を撮影する際に、撮影面の左右から蛍光灯による照明を当て、さらに実験システム全体に暗幕を被せることにより、照明条件を常に一定とした。これにより、二値化の閾値を全ての食器に共通としても、輪郭を取得することが可能となった。この条件の下で、二値化の閾値を変化させて、実験1と同様の実験を行う。また膨張および収縮の処理は、全ての食器について処理を行わないとした。

5. 実験結果と考察

5.1 実験A

2.1で述べた画像処理を行った結果を図5に示す。(A)はカメラから得た元画像，(B)は二値化，膨張，収縮後の画像，(C)は8連結の境界線追跡により得られた輪郭とその重心の画像である。(C)の輪郭から2.2で述べた方法で作成した累積ヒストグラムを図6に示す。左は図5(A)に見られるグラタン皿から作成した累

積ヒストグラム，右はコーヒーカップから作成した累積ヒストグラムである。

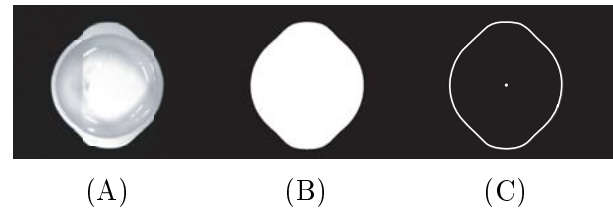


Fig. 5 実験Aにおける画像処理により得られる食器の輪郭

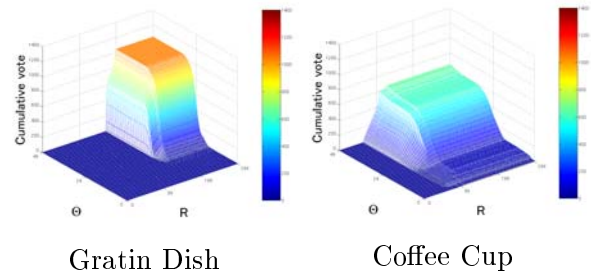


Fig. 6 実験Aにおける累積ヒストグラム

9種類合計36枚の画像についての認識率は100%であった。本研究で用いる認識手法は、食器の画像上の位置・姿勢に影響を受けずに、非常に高い認識率で認識可能であると考えられる。

9種類合計36枚の画像についての認識時間は約2.4 sであった。これより1枚の食器にかかる平均認識時間は約66 msであった。一方、1種類の食器を認識対象とし、350×320 pixelsのテンプレート画像1枚を用いたテンプレートマッチング法により、640×480 pixelsの画像について同じシステムを用いて認識を行った結果、認識時間は約120 msであった。テンプレートマッチング法によって11種類の食器を識別するには、最低でも9枚のテンプレート画像を用いる必要があるため、1枚の画像に要する認識時間も約9倍となると考えられる。食器の回転まで考慮した場合、必要となるテンプレート画像の枚数はさらに増えるため、認識の高速性という観点からも本研究で用いる認識手法は有効であると考えられる。

ヒストグラムの分割数と認識時間・認識率

の関係性について得た結果のうち、いくつかの例を表1に示す。分割数が少ない程計算処理が減るため認識時間は短くなるが、認識率は低下することが分かった。しかし分割数の減少と認識率の低下は比例関係にないため、単純に分割数を削減せず、認識速度・認識率の両面から食器認識に適切な分割数を今度検討すべきであると考え。そのため本稿では、この後の実験については特別な記述がない限り、ヒストグラムの分割数は全ての実験において、 r 方向は300 pixelsを300分割、 θ 方向は3.15 radを50分割とする。

Table 1 実験Aにおける分割数と認識時間・認識率の関係性

(r, θ) 分割数	認識時間[ms]	認識率[%]
(100,100)	2364	100
(50,50)	1146	100
(25,25)	816	100
(10,10)	701	97.2
(3,3)	677	91.6

カメラにより撮影する画像の解像度を変化させていった場合の実験結果を、表2に示す。ここで倍率とは、はじめの画像の解像度640×480 pixelsに対し、現在の縦横それぞれの解像度が何倍であるかを示している。この倍率に比例させ (r, θ) の分割数も変化させていく。この表2より、解像度が低くなる程、画像処理などの計算処理が減るため、認識時間は短くなるが、認識率は低下することが分かった。これは解像度を下げることによりカメラ画像の輪郭に関する情報の一部が失われ、正しく食器の輪郭画素を抽出できず誤認識を生じるからである。以上より、解像度を下げるとは認識時間を短くするが、大きく解像度を下げた場合は誤認識を生じる可能性があるため、安易に解像度を下げるとは避けるべきであると考え。そのため本稿では、この後の実験につ

いては特別な記述がない限り、画像の解像度は640×480 pixelsとする。

Table 2 実験Aにおける解像度と認識時間・認識率の関係性

倍率	解像度	分割数	時間[ms]	認識率[%]
1	640×480	(300,50)	2548	100
0.9	576×432	(270,45)	1966	100
0.8	512×384	(240,40)	1575	100
0.7	448×336	(210,35)	1237	100
0.6	384×288	(180,30)	986	100
0.5	320×240	(150,25)	864	100
0.4	256×192	(120,20)	670	100
0.3	192×144	(90,15)	544	97.2
0.2	128×96	(60,10)	433	97.2
0.1	64×48	(30,5)	374	83.3

5.2 実験B

実験結果の中から閾値と認識率の関係を示す例を、表3に示す。この結果によれば、今回の照明条件の下では二値化する際の輝度値の閾値は50が最適であると考えられるが、その場合においても実験Aと比較し、その認識率は低下している。また食器の種類によっては、輪郭を取得するのに適した閾値が異なるため、共通の閾値で多種類の食器を認識することは困難であり、画像に応じて動的に閾値を決定する方法を検討する必要があると考えられる。

Table 3 実験Bにおける閾値と認識率の関係性

閾値	誤認識数	認識率[%]
40	3	91.6
50	2	94.4
60	3	91.6
70	3	91.6
80	4	88.8
90	3	91.6
100	4	88.8
150	4	88.8
200	5	86.1

6. 結言

本研究では、カメラから得た画像情報から求まる累積ヒストグラムを用いた認識手法を用いることで、9種類の食器について高い認識率で認識することができることが分かった。ただし、この結果は実験条件や画像処理のパラメータを適宜決定したことによるものであり、それらを共通として決定した場合に認識率は多少低下することが分かった。今後は、実験条件や画像処理のパラメータの決定における曖昧さを改善し、より定量的に食器の認識可能性を論じることを検討する。そして、現在の食器の重心位置と種類に加え、食器の画像上の姿勢を数値として認識可能なアルゴリズムの追加を両立させることを目標に、研究を続けていく。

参考文献

- 1) 酒井幸市: デジタル画像処理の基礎と応用, 54/55, CQ出版社 (2003)
- 2) 林 宏樹, 斉藤文彦: 幾何エッジヒストグラム空間を用いた回転対応画像照合, 電気学会論文誌(C), 126-2, 203/209 (2006)
- 3) 林 宏樹, 斉藤文彦: 等距離エッジ点群の幾何特徴に基づく回転対応画像照合, 電気学会論文誌(C), 126-4, 519/525 (2006)