

# 画像処理と音声対話による冷蔵庫内の食品管理システム

松本拓也  
Takuya Matsumoto

白井良明  
Yosiaki Sirai

島田伸敬  
Nobutaka Simada

立命館大学  
Ritsumeikan University

## 1 はじめに

サービスロボットがユーザの依頼に基づいて冷蔵庫内から食品を取ってきたり、と、冷蔵庫の中身に対するユーザの質問(存在の有無、個数、場所など)に答えるサービスのために、画像処理によって冷蔵庫内の食品を認識する。飲み物、果物を対象として対象を取ってくる研究 [1] はあるが、本研究では食品全般を対象にしている。画像処理だけで認識が難しい場合には、ユーザとの音声対話による支援によって食品の認識を行う。

## 2 システムの動作

本システムの動作は、大きく 3 つの場面に分けられる。ユーザの発話によってどの場面かを判断する。

### (1) 食品を入れる場面

冷蔵庫の画像から食品の位置や形、状態(隠れているなど)を自動的に認識する。また、音声によって名前や場所をユーザから入力して貰う。その後、食品のデータ(位置、形、名前など)を登録する。

### (2) 食品を取り出す場面

取りだされた後の冷蔵庫の画像から、取り出された食品を認識する [2]。画像処理による特定が難しい場合には、ユーザとの対話によって取られた食品を特定する。その後、食品のデータの一覧から取られた食品の削除、及び食品を取った際に周りの食品が移動した場合には、食品のデータの位置の変更を行う。

### (3) ユーザが食品の確認や食品を取る依頼を行う場面

蓄積した食品のデータを基にユーザとの対話を行い、食品の場所や存在するかどうかといった質問に答える [2]。本稿では、3 章以降食品を入れた時の認識について詳しく述べる。

## 3 冷蔵庫内の食品の認識

1 つの食品の出し入れが行われるたびに冷蔵庫の画像を保存している。まず、1 つ前の時系列の画像と現在の画像との差分を取り、差分が大きい範囲を食品の認識の処理を行う範囲(以降処理範囲とする)とする。次に、処理範囲内で、冷蔵庫全体の照明の変化に対する明るさの補正、カメラのずれに対する対処を行う [2]。その後、対処した画像を用いて改めて背景差分を行う。最後に影の除去を行い、食品の領域を検出する [2]。

本研究では、食品を入れた際に既に入っていた食品を移動させた場合にも、入れた食品の認識を可能にする。既に入っていた食品が移動した場合には、食品の移動元、食品の移動先に差分が検出される。この場合には、どの領域の差分が入れた食品かが分からない。そこで、冷蔵庫内に入っていた食品が移動した後の画像を予測し、予測

した画像と現在の画像との差分を取ることで入れた食品を検出する。

### 3.1 食品のモデル(テンプレート、代表色)の登録

最初に食品を入れた時には、正確に食品を認識できる。その時に食品のモデルを登録する。その後は、食品のモデルを用いて入れた食品の認識を行い、順次食品のモデルを登録していく。

テンプレートとは食品の領域を切り取った画像である。代表色の抽出は以下の手順で行う。検出した食品に対して色のヒストグラムを用いた領域分割を行う。次に色の距離に基づいて近い色を統合する。まず、色を色ベクトル  $C = (X, Y, Z)$  で表す。ここで

$$X = \log_{10}(9S + 1) * \cos(H), \quad (1)$$

$$Y = \log_{10}(9S + 1) * \sin(H), \quad (2)$$

$$Z = \log_{10}(9V + 1) \quad (3)$$

とする。このとき  $H, S, V$  はそれぞれ色相、彩度、輝度である。  $0 \leq S, V \leq 1, 0 \leq H \leq 2\pi$  である。  $C_1(X_1, Y_1, Z_1)$  と  $C_2(X_2, Y_2, Z_2)$  の距離  $D(C_1, C_2)$  は以下の式で表す。

$$D(C_1, C_2) = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2 + w(Z_1 - Z_2)^2} \quad (4)$$

輝度の変化は、色の差に対して影響が少ないと考え、重みの値  $w$  を実験的に 0.7 とした。食品を占める面積が大きい色を代表色(複数の場合もある)として抽出する。

### 3.2 過去の食品が移動した位置の特定

既に入っていた食品が移動した場所を処理範囲内でテンプレートマッチングを行い特定する。一般的なテンプレートマッチングでは、下記の式 (5) に示す色の差の合計値  $\Delta C$  が最小となる位置を移動した位置とする。

$$\Delta C = \sum_{(i,j) \in T} |\Delta R(i,j)| + |\Delta G(i,j)| + |\Delta B(i,j)| \quad (5)$$

$$\text{ここで } \Delta R(i,j) = R_R(i,j) - R_T(i,j), \quad (6)$$

$\Delta G, \Delta B$  は  $\Delta R$  と同様にして求める。  $T$  はテンプレートを構成する画素の座標の集合であり、  $R_T$  はテンプレートの R 値、  $R_R$  は対応点領域の R 値を表す。

ただし、図 1 のように食品 B を入れた際に食品 A が回転して移動している場合には、テンプレートと見えている部分が違うためにマッチングを失敗する(図 2(a))。そこで、代表色を利用してマッチングを行い食品の移動した位置を特定する。

まず、過去の食品のデータから、処理範囲内にある食品(移動した可能性のある食品)のモデルを列挙する(図 1 の場合食品 A)。次に、列挙した食品のモデルごとに処理

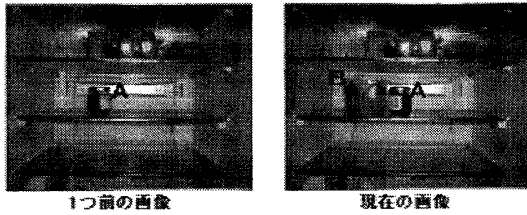


図1 食品が回転し模様が変わった例

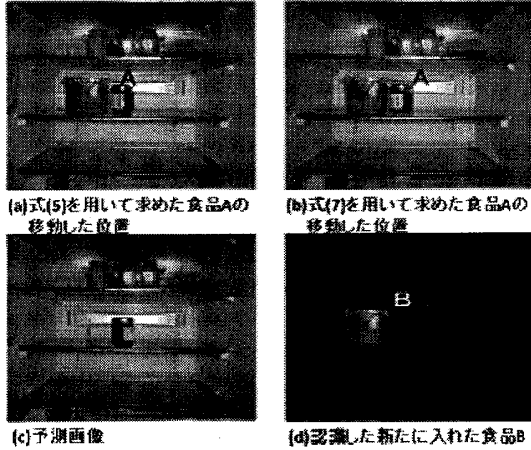


図2 図1の処理例

範囲内においてテンプレートマッチングを行い、食品の移動先を特定する。テンプレートマッチングの地点  $(x, y)$  における評価値  $\Delta S$  を下記の式で表す。

$$\Delta S(x, y; m) = \sum_{(i,j) \in T(m)} L(C_R(x+i, y+j), Q(m)) \quad (7)$$

$$L(C_R, Q) = \min_{q \in Q} D(C_R, q) \quad (8)$$

ここで、 $m$  は注目している食品のモデル、 $Q(m)$  は、モデル  $m$  における代表色の色ベクトルの集合、 $C_R(x+i, y+j)$  は  $(x+i, y+j)$  における現在の画像の色ベクトルを表す。 $\Delta S$  が最小となる位置を食品の移動先とする (図 2(b))。最後に、食品の移動元の背後の画像を以前の画像から補完し、食品の移動先をその食品で埋めた予測画像 (図 2(c)) を作成する。最後に、現在の画像と予測画像との差分を取ることで入れた食品 B を検出する (図 2(d))。

**3.3 ユーザとの対話による移動した食品の位置の修正**  
 似た色を持つ食品が入れた場合や、記録している代表の色以外の色を持つ模様が回転し現れた場合には式 (7) の方法でマッチングを行っても失敗する場合がある。そこで、移動先として可能性のある複数の位置を移動した候補の位置として、候補全てでその位置の正しさを評価する。評価によっては、ユーザとの対話を行い、移動した場所を選択して貰うことで食品が移動した場所を正確に把握する。具体的には、

- (a)  $\Delta S$  が最小の位置、(b)  $\Delta S$  が最小の値の 1.1 倍以下の値になった位置、(c) 食品が元にあった位置 を食品が移動した位置の候補とする。
- 全ての候補で、テンプレートと現在の画像の相関係数と代表色を含む割合を求める。

- 全ての候補の評価が悪い場合や、良い評価の候補が複数ある場合は、ユーザと対話を行う。良い評価の候補が単体の場合は、その位置を移動した位置とする。
- ユーザと対話を行う場合には、移動した位置の候補をユーザに示し、対話によって選択して貰い、移動した位置を特定する。

### 3.4 食品の隠べいの考慮による対話負担の軽減

食品が他の食品の裏に隠れている場合 (図 3) にユーザの手間が増える。

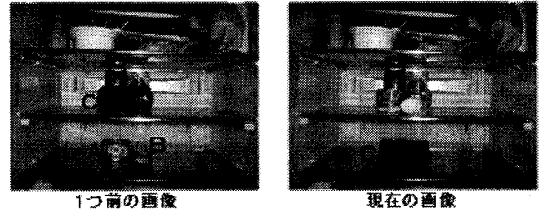


図3 他の食品によって隠された例

具体的には、図 3 の食品 A, B は、食品 C によって隠されたことで 3.3 節の評価がどの候補も低くなるため、ユーザと対話を行う。そこで、先に他の食品を隠べいしている食品の移動した食品を特定し、その食品の後ろにあると判断できる食品は、ユーザとの対話を控えることで対話負担を軽減する。

まず、3.2 節の移動した可能性のある食品を、手前にある食品、面積が大きい食品の順で移動した位置を特定する。また、位置を特定する際に以下の 2 つの条件を満たした場合に他の食品に隠れているとした。

- 3.3 節のマッチング結果の評価がどの候補も低いこと
- 食品が元にあった位置に他の食品があり、他の食品の部分の割合が高いこと

他の食品に隠れている場合にはユーザとの対話を控える。

図 3 の場合は、食品 C の移動した場所を先に特定し、その後食品 A, B の処理を行うことで食品 C の後ろに隠されていることが判断でき、ユーザに食品 A, B の移動した場所を聞かずに済んだ。

### 4 まとめ

画像認識と音声対話を用いた冷蔵庫内の食品の認識について実装した。食品を出し入れする際にすでに入っていた食品が移動した場合についても対応した。その際に、移動した食品が回転し模様が変わってしまった場合についても認識を可能にした。また、ユーザからの冷蔵庫内の食品についての問い合わせやサービスロボットによる食品を持ってこるための依頼についての対話を実装した。

### 5 今後の課題

過去に冷蔵庫に入れたことのある食品については、複数同時に入れた場合にも認識を可能にすることが今後の課題である。

### 参考文献

- [1] 榎原 靖, "対話を用いた物体認識に関する研究", 人工知能学会誌, Vol.21, No.1, p.116, 2006
- [2] 松本拓也, "冷蔵庫内の食品管理システム", 電子情報通信学会, 2008 年総合大会講演論文集, D-12-41, 2008