

インタラクティブビジョンにおいて ユーザから有用な助言を得るための手法

井本 浩靖[†] 白井 良明^{††} 島田 伸敬^{††} 三浦 純[†]

[†] 大阪大学大学院工学研究科 〒1565-0871 大阪府吹田市山田丘 2-1

^{††} 立命館大学情報理工学部知能情報学科 〒525-8577 滋賀県草津市野路東 1-1-1

E-mail: [†]{imoto,jun}@cv.mech.eng.osaka-u.ac.jp, ^{††}{shirai,shimada}@ci.ritsumeimei.ac.jp

あらまし 本論文では、指定された飲料を冷蔵庫からユーザのところに持っていくロボットのための対話を用いた物体認識システムについて述べる。システムはあらかじめ物体モデルを登録しておき、ユーザがある物体を指定すると、登録しておいた物体モデルを用いて物体認識を試みる。しかし、隠蔽や照明条件等により指定された物体を正しく認識できないことがある。そのような場合物体認識結果をユーザに伝え、指定された物体がどのような状況であるか、エッジ等の特徴がどこにあるかなどユーザから有用な助言を得ることができるような質問をする。そして、得られた助言に基づいて再度物体認識を試みる。実験は実際の冷蔵庫内のシーンを用いて行った。

キーワード 物体認識, 対話, インタラクティブビジョン, サービスロボット

A method to get helpful advice from a user in Interactive Vision

Hiroyasu IMOTO[†], Yoshiaki SHIRAI^{††}, Nobutaka SHIMADA^{††}, and Jun MIURA[†]

[†] Faculty of Engineering, Osaka University Yamadaoka 2-1, Suita-shi, 565-0871 Japan

^{††} College of Information Science and Engineering, Ritsumeikan University Noji Higashi 1-1-1, Kusatsu-shi, 525-8577 Japan

E-mail: [†]{imoto,jun}@cv.mech.eng.osaka-u.ac.jp, ^{††}{shirai,shimada}@ci.ritsumeimei.ac.jp

Abstract This paper describes an interactive vision system for a Service robot which finds an object specified by a user and brings it to the user. The system first registers object models automatically. When the user specifies an object, the system tries to recognize it automatically based on the object model. But, We have cases that the system cannot recognize an object correctly in the situations when the object is occluded by other objects. In that case, the system shows object recognition results to the user, and ask the user a question to get helpful advice what kind of situation the object is or what kind of object is occluded by. Then the system recognize the object again based on the advice. Experiments using real-world refrigerator scenes are shown.

Key words Object Recognition, Dialog, Interactive Vision, Service Robot

1. まえがき

近年、高齢化社会の到来によりサービスロボットの必要性が高まっている。そのようなロボットの機能の中で、ユーザに指定されたものを持ってくるという機能は非常に重要なものである。物体を取ってくるためには、ロボットが指定された物体の位置を知る必要があるが、その一つの方法として物体にバーコードのようなラベルを貼っておいてそれを検出するという方法が考えられる。しかし、そのような方法では、ユーザにラベルを貼る手間をかける上、ロボットからラベルが見えないような状況では位置を知ることができない。本研究では、そのようなラベルを用いずに、指定物体を画像中から認識することで

の位置を求める。

高橋ら [1] は、画像を用いて物体認識して持ってくるサービスロボットの研究を行った。しかし、部分的に隠蔽された物体の認識や照明条件が変化したときに物体モデルの色を適切な色に変換するのが困難な場合がある。

画像情報と言語情報を組み合わせた試みとしては、植物図鑑の花や果実を、図に添付された説明文を利用して認識するシステム [2] や、ユーザの発話の内容に一致する特徴を画像から抽出することで、物体を認識するシステム [3] が提案された。しかし、いずれも認識が失敗したときにそこから再度認識を試みて成功させることは行っていない。一方、高橋ら [4] は、言語と

ジェスチャによるやり取りを通して、りんごや本を認識するロボットの研究を行った。ロボットはユーザがジェスチャによって指し示した方向を見て、音声で指定された物体を画像から認識し、複数の物体が発見された場合には、ユーザとのやり取りによって物体を選択する。しかし、物体が発見されなかった場合に再認識を試みることは行っていない。

本論文では、冷蔵庫から必要なものを取ってきてもらうサービスロボットについて述べる。ユーザは手元のディスプレイを見ながら、音声でロボットに指示を出す。ここで、必要以上にユーザを煩わせないために、最初はあらかじめ登録された物体モデルを用いて可能な限り自動での物体認識を試みる。しかし、隠蔽やや照明条件等により候補領域が適当な大きさで抽出できない場合やエッジ等の形状特徴が正確に認識できない場合などには、指定された物体を認識できない。そのような場合、ユーザから物体認識を成功させるような有用な助言を得ることでユーザから指定された物体を再認識する。また、本論文のサービスロボットは、マニピュレータを用いた物体の操作も行うが、その詳細については文献 [9] を参照されたい。

最後に、本論文の構成を以下に示す。2. では物体認識の概要について述べ、3. では物体認識を行った結果とそれに基づいた候補物体のクラス分けについて述べる。4. ではユーザの助言に基づいた候補物体の限定と物体の再認識の方法について述べ、5. で結論を述べる。

2. 物体認識の概要

システムは最初にどの方向からでも認識できるように図 1(a)のような投影画像を用いて物体モデルを作成しておく [5]。ユーザから対話によって取って欲しい物体 (以下、指定物体と呼ぶ) を伝えられると、参照色 (ここでは図 1(b)の矢印で指し示した白枠で囲まれた冷蔵庫の扉の色) を用いて原画像 (図 1(b)) をモデル作成時の基準照明条件に変換した正規化画像を作成して (図 1(c))、その画像から指定された物体の代表色の範囲に入る候補領域を抽出する (図 1(d)) [6]。そして、図 2 のような物体の形状のモデルを考え、候補領域周辺からエッジやキャップなどそれぞれの形状に関する特徴 (以下、形状特徴と呼ぶ) を抽出して物体の形状を認識する (図 1(e))、黒色のグラフはエッジ強度、グラフ上の白線はその閾値、物体上の白線は候補エッジを表す)。このようにして認識した候補物体をディスプレイを通してユーザに表示する (図 1(f)) [7]。

指定物体の状況によっては、一部の形状特徴が認識できない場合や間違った場所に形状特徴が認識されてしまいすべての形状特徴が認識できてしまう場合など物体認識が成功しない場合がある。このような場合、文献 [8] のようにユーザにすぐに助言を求める方法も考えられるがユーザの負担が大きくなってしまふ。そこで本研究では、物体認識が失敗したときには候補領域の大きさと物体モデルの大きさを比較して、それをもとに候補領域の周辺をさらに詳しく調べることにより形状特徴を再検出する。ここで登録された実際の物体の大きさから画像上での物体の大きさを知る必要があるが、これはロボットが冷蔵庫に近づく際に冷蔵庫までのおおよその距離が分かるので、それを元に計算することができる。具体的には、候補領域が大きいあるいは小さく一部の形状特徴が認識できていない候補の場合は、物体モデルの許容範囲と認識できている形状特徴を考慮して、

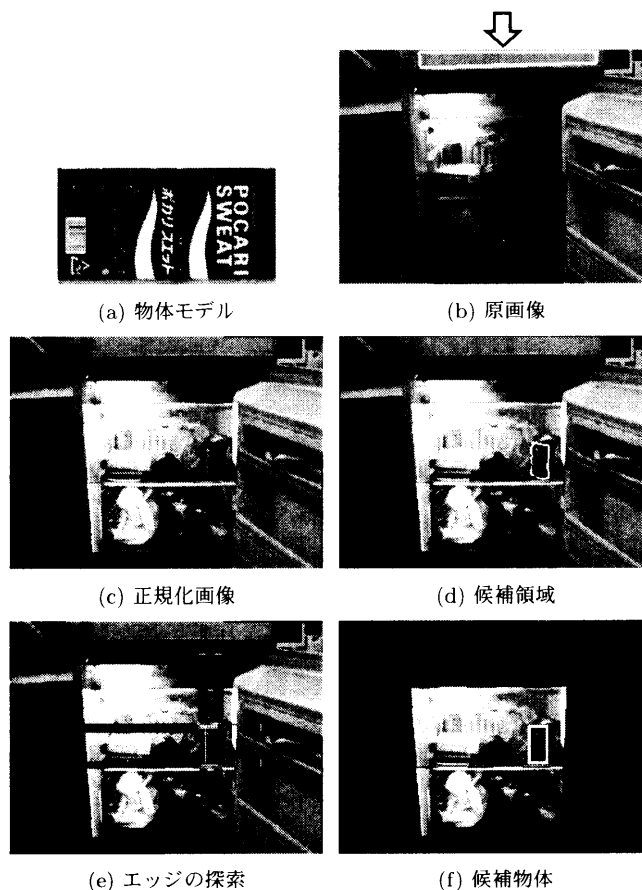


図 1 候補領域の抽出

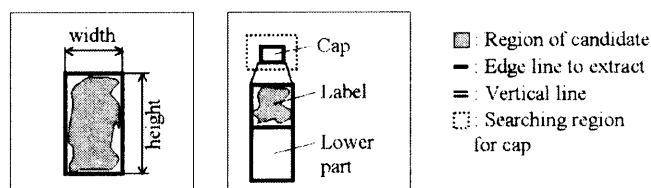


図 2 物体の形状モデル

周辺領域もしくは候補領域内で形状特徴を再認識する (図 4(a)) では、候補領域が小さく、上下のエッジが認識されている。そこで、図 4(b) で物体モデルの許容範囲まで候補領域周辺に左右のエッジを認識し、図 4(c) のような候補物体が得られる)。また、形状特徴がすべて認識されているが候補領域が大きい候補の場合は、物体モデルの許容範囲と認識された形状特徴を考慮して、候補領域内で形状特徴を再認識する (図 4(d)) では、候補領域が大きく、上下左右のエッジが認識されている。そこで、図 4(e) で物体モデルの許容範囲まで候補領域内に縦エッジを認識し、図 4(f) のような候補物体が得られる)。

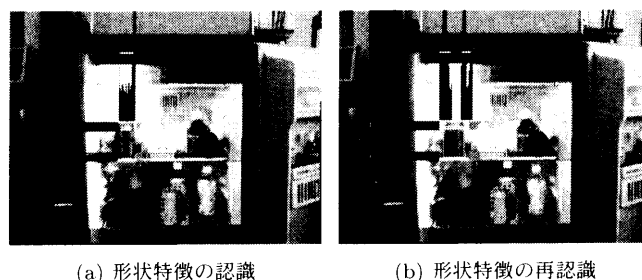


図 3 形状特徴の再認識

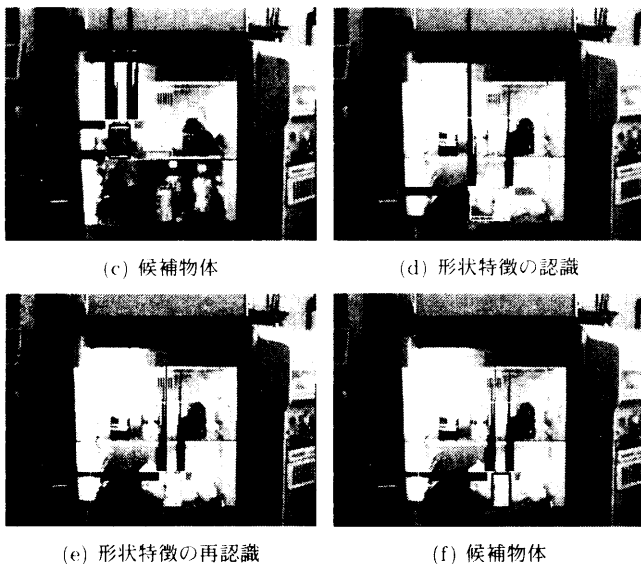


図4 形状特徴の再認識

3. 物体認識結果と候補領域のクラス分け

実際の冷蔵庫環境において、指定物体には隠蔽されているなどいくつかの状況が考えられる。そのような状況に対して2.で述べたような物体認識を行った場合、照明条件等によりエッジ等の形状特徴の一部を認識することができないなど、いくつかの物体認識結果が得られる。このとき、システムはユーザに物体としての評価の高い候補から表示するために、認識結果に基づいて候補領域のクラス分けを行う。

3.1 冷蔵庫環境に対する物体認識結果

物体認識を行う際、冷蔵庫環境内における指定物体の状況としては以下のような状況がある。

1. 他の物体に隠蔽されていない (図 5(a))
2. 違う色の物体にある程度隠蔽されている (図 5(b))
3. 違う色の物体に大きく隠蔽されている (図 5(c))
4. 似た色の物体と重なっている (図 5(d))

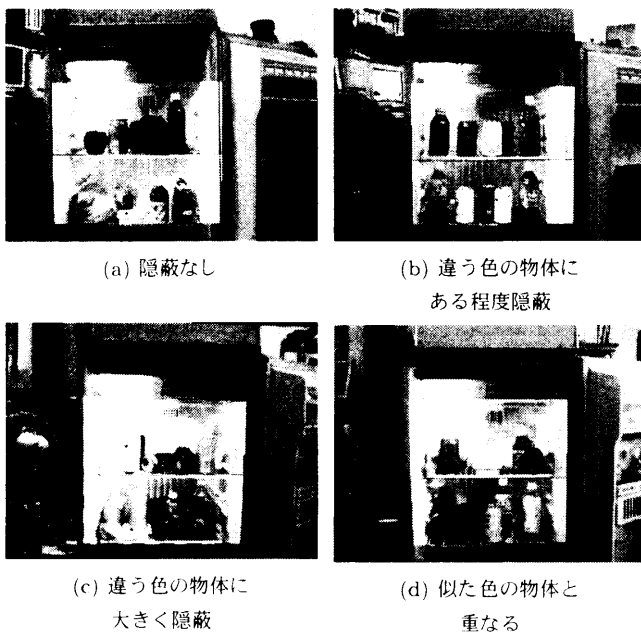


図5 冷蔵庫環境における指定物体の状況

ここで、ある程度隠蔽されている状況は候補領域は抽出される

状況で、大きく隠蔽されている状況は候補領域も抽出できないほど大きく隠蔽されている状況である。また、物体認識結果得られる認識としては以下のようなものがある。候補領域の幅、高さを物体モデルの許容範囲と比較した場合すべてで9つの場合が考えられるが、候補領域の幅が物体モデルの許容範囲であり、高さが許容範囲より大きい場合と候補領域の高さが物体モデルの許容範囲であり、幅が許容範囲より大きい場合は、似た色の物体と重なっている状況のときに得られる認識であり、ユーザへの表示方法、物体の再認識方法が同じであるので、同一の状況に分類した。状況 i-c, 状況 i-d についても同様に複数の場合を同一の状況に分類した。

- i. 候補領域が抽出できる
 - a. 候補領域の幅、高さがともに物体モデルの許容範囲
 - A. 形状特徴がすべて認識できている
 - B. 形状特徴が一部しか認識できていない
 - b. 候補領域の幅、高さのどちらか一方が物体モデルの許容範囲で、他方が許容範囲より大きい
 - A. 形状特徴がすべて認識できている
 - B. 形状特徴が一部しか認識できていない
 - c. 候補領域の幅、高さのどちらか一方が物体モデルの許容範囲で、他方が許容範囲より小さい
 - A. 形状特徴がすべて認識できている
 - B. 形状特徴が一部しか認識できていない
 - d. 候補領域の幅、高さのどちらか一方が許容範囲より大きく、他方が許容範囲より大きいか小さい
 - A. 形状特徴がすべて認識できている
 - B. 形状特徴が一部しか認識できていない
 - e. 候補領域の幅、高さがともに物体モデルの許容範囲より小さい
 - A. 形状特徴がすべて認識できている
 - B. 形状特徴が一部しか認識できていない
- ii. 候補領域が抽出できない

指定物体に対して2.で述べた物体認識を行うと、各状況においてそれぞれ以下の表1のような物体認識結果が得られる。

表1 物体の状況と得られる認識結果の関係

	状況1	状況2	状況3	状況4
認識 i-a-A	○			○
認識 i-a-B	○			○
認識 i-b-A				○
認識 i-b-B				○
認識 i-c-A	○	○		○
認識 i-c-B	○	○		○
認識 i-d-A				○
認識 i-d-B				○
認識 i-e-A	○	○		○
認識 i-e-B	○	○		○
認識 ii	○	○	○	○

3.2 候補領域のクラス分け

3.1で述べたような物体認識結果に基づいて、候補領域が得られた物体についてユーザが指定した物体らしい候補から表示するために、3つのクラスに分類する。具体的には、候補領域の幅、高さが物体モデルの許容範囲と比べてどうかと形状特徴

がどれくらい認識できたかで以下のように分類する。

クラス1：候補領域の幅、高さがともに物体モデルの許容範囲であり、すべての形状特徴が認識された場合 (認識 i-a-A)

クラス2：候補領域の幅、高さがともに物体モデルの許容範囲であるが一部の形状特徴が認識されなかった場合、または、候補物体の幅、高さのどちらか一方が物体モデルの許容範囲ではない場合 (認識 i-a-B, i-b-A, i-b-B, i-c-A, i-c-B)

クラス3：候補領域の幅、高さともに物体モデルの許容範囲ではないような場合 (認識 i-d-A, i-d-B, i-e-A, i-e-B)

4. ユーザの助言に基づいた候補物体の限定と再認識

システムは、候補領域が得られた場合は3.2で述べたクラス分けによって分類された候補領域のうち、上位クラスに含まれるものからユーザに表示し、ユーザからの助言を得る必要がある場合はユーザにそれを伝え、そして得られた助言によって候補物体の限定や物体の再認識を行う。また、候補領域が抽出できない場合は指定物体の位置を知ることができないので、その場合は指定物体の位置に関する助言が必要であることを伝え、得られた助言に基づいて候補領域の再抽出を行う。

4.1 候補領域が得られた場合

4.1.1 候補領域の表示と限定

ユーザに候補領域を表示する際には、3.2で述べたクラス分けによって分類された候補領域のうち、上位クラスに含まれるものからユーザに表示していく。クラス1を表示する際には、まず最初に候補の中に指定物体があるかどうかを尋ね、指定物体が存在する場合は、候補の中から選択してもらおう。そして、指定物体が選択されればクラス1の場合は確認を行う。もし、指定物体が存在しなければ次善クラスを表示する。

クラス2、3を表示する際には、クラス1の場合と同じくまず最初に候補の中に指定物体があるかどうかを尋ねるが、候補領域を表示する際、図6(a)のように候補領域を開いて表示するだけでは各領域をどのように表現すれば良いかわかりにくい場合や、候補領域が多く表示され位置等の特徴では選択できない場合など、ユーザが候補を選択する煩わしさがあるので、図6(b)のように各候補領域に番号をつけて、番号で候補の選択をしてもらう。そして指定物体が選択されれば、クラス2、クラス3の場合は指定物体が3.1で述べたどの状況であるかを判断する必要があるので、4.1.2で述べるようにユーザから助言をもらえるような発話を行う。もし、指定物体が表示しているクラスに存在しなければ次善クラスを表示する。次善クラスが混在しない場合は、候補領域が抽出できていないと判断できるので4.2で述べる候補領域が抽出できない場合の認識に移る。

4.1.2 ユーザへの助言の要求とその返答に基づいた物体の再認識

指定物体がクラス2もしくはクラス3であった場合、その物体がどの状況であるかと指定物体の正しい位置がどこであるかを判断するためにユーザへ助言を要求する。以下に、各クラス各認識においてユーザに要求する助言とその返答に基づいた物体の再認識の方法について述べる。



(a) 番号なし (b) 番号あり

図6 クラス2、3の候補領域の表示

●クラス2

- 候補領域の幅、高さがともに物体モデルの許容範囲であるが一部の形状特徴が認識されなかった場合 (認識 i-a-B)

認識 i-a-B が得られるのは、他の物体に隠蔽されていない状況 (状況1)、似た色の物体と重なっている状況 (状況4) の2つの状況である。そこで、物体モデルから認識されなかった特徴が存在する位置を考えて候補物体を囲み、ユーザに正しいかどうかを確認する。もし正しいのであれば状況1であると判断し、候補物体が認識できる (図7)。正しくなければ状況4と判断する。候補物体の位置を知るためには、候補領域周辺で形状特徴を探索する方法も考えられるが、どのようないくつの物体と重なっているか判断できず、どの領域を探索すれば良いかわからない。そこで状況4と判断できれば、4.2.2で述べる方法によって候補領域を拡張して再度物体認識を行う。

- 候補領域の幅、高さのどちらか一方が物体モデルの許容範囲で、他方が許容範囲より大きい場合 (認識 i-b-A, i-b-B)

認識 i-b-A, i-b-B が得られる物体の状況は、似た色の物体と重なっている状況 (状況4) である。そこで物体の正しい位置を判断するため、認識された形状特徴を含み、かつ候補領域の中で最も代表色の割合が高くなるような矩形領域を囲み、ユーザに正しいかどうかを確認する (図8)。



(a) 原画像

(b) 候補領域



(c) 形状特徴の探索

(d) 物体モデルに基づいた候補物体の予測

図7 認識 i-a-B における候補領域の認識

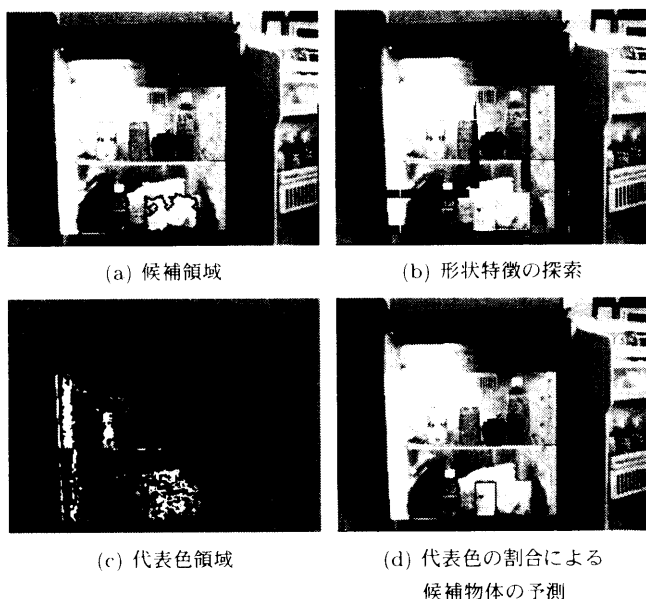


図 8 認識 i-b-A における候補領域の認識

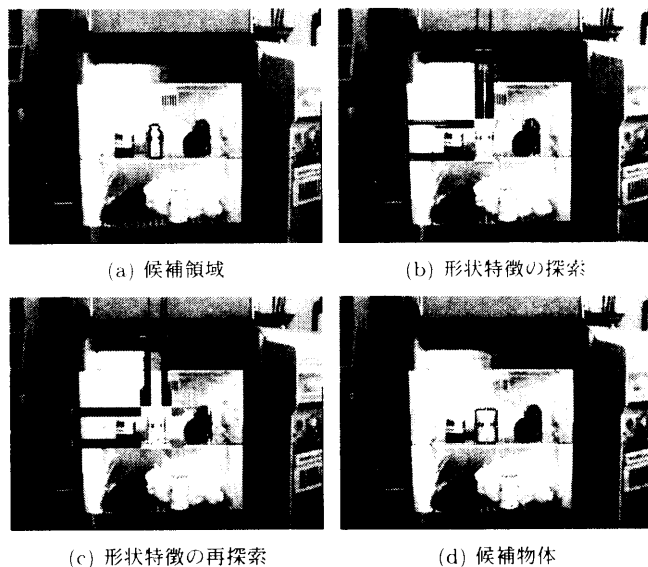


図 9 認識 i-c-A における候補領域の認識

- 候補領域の幅、高さのどちらか一方が物体モデルの許容範囲で、他方が許容範囲より小さい場合 (認識 i-c-A, i-c-B)
- 認識 i-c-A, i-c-B が得られるのは、指定物体が他の物体に隠蔽されていない状況 (状況 1)、違う色の物体にある程度隠蔽されている状況 (状況 2)、似た色の物体と重なっている状況 (状況 4) の 3つの状況である。そこでまず、指定物体が他の物体に隠蔽されていないかを尋ねる。もし隠蔽されているなら状況 2か 4であると判断でき、隠蔽物体の物体認識を試みる。認識できればマニピュレータで隠蔽物体を移動させ再度指定物体の認識を試みる。もし隠蔽物体が正確に認識できない場合は、指定物体の認識と同じようにユーザに助言を求める。また、隠蔽されていない場合は状況 1であると判断でき、物体モデルの許す幅まで形状特徴の認識を試みる。形状特徴が認識できたら候補領域を囲む矩形領域を変更し、再度ユーザに表示する (図 9)。もし、形状特徴がうまく認識できなかった場合には、4.2.2で述べる方法によって候補領域を拡張して、再度形状特徴を認識する。

● クラス 3

- 候補領域の幅、高さのどちらか一方が許容範囲より大きく他方が許容範囲より小さい場合 (認識 i-d-A, i-d-B)
- 認識 i-d-A, i-d-B が得られた場合は、クラス 2における候補領域の幅、高さのどちらか一方が物体モデルの許容範囲で、他方が許容範囲より大きい場合と同様の方法によって、どの状況であるかと物体の正しい位置を判断する。
- 候補領域の幅、高さがともに物体モデルの許容範囲より小さい場合 (認識 i-e-A, i-e-B)
- 認識 i-e-A, i-e-B が得られた場合は、クラス 2における候補領域の幅、高さのどちらか一方が物体モデルの許容範囲で、他方が許容範囲より小さい場合と同様の方法によって、どの状況であるかと物体の正しい位置を判断する。

4.2 候補領域が抽出できない場合

候補領域を抽出できない状況としては、違う色の物体に大きく隠蔽されている状況 (状況 3) が考えられる。他の状況においても画像上での色の範囲がずれている場合は候補領域を抽出できないことがある。ここでは、そのような場合において指定物体を認識する方法について述べる。

4.2.1 違う色の物体に大きく隠蔽されている物体の認識

候補領域の抽出に失敗した場合、システムはまず、違う色の物体に大きく隠蔽されている状況 (状況 3) であると判断し、ユーザに他の物体に隠蔽されていないかを尋ねる。もし隠蔽されているなら、4.1.2と同様に、隠蔽物体の物体認識を試み、認識できればマニピュレータで隠蔽物体を移動させ、再度指定物体の認識を試みる。もし隠蔽物体が正確に認識できない場合は、指定物体の認識と同じようにユーザに助言を求める。隠蔽されていない場合は 4.2.2で述べる認識を行う。

4.2.2 画像上での色の範囲がずれている物体の認識

他の物体に隠蔽されていない状況で候補領域が抽出できない場合は、候補領域を抽出する際の代表色の範囲を広げる。しかし、ただ代表色の範囲を広げただけでは、指定物体とは関係のない領域まで多く抽出してしまうことになり、物体認識が困難になったりユーザの選択が煩わしくなってしまう。そこでこのような場合は、冷蔵庫領域を大きく6つの領域に分け、その領域のうちどこにあるかをユーザに選択してもらうことにより領域が多く抽出されるのを避ける (図 10)。さらに、指定領域で代表色の信頼区間を広げて候補領域を抽出する際は、広げすぎにより領域が大きくなりすぎないように、物体モデルの幅・高さよりも大きくならない程度まで色の範囲を広げる。もし、色の範囲を広げても候補領域がある程度より大きく抽出できない場合は、学習した代表色と正規化した画像上での指定物体の色が大きく異なっていると判断できる。そこで、学習した代表色を色の分散はそのまま、色の平均値のみを抽出されている候補領域の平均値に一時的に変更することで候補領域を抽出する。このようにして候補領域が抽出できれば 2. で述べたようにその候補領域に対して形状特徴の認識を行い、ユーザに表示する (図 10)。

また、4.1.2で述べたように候補領域が物体モデルよりもかなり小さくしか抽出できない場合には、抽出されている領域の周辺で物体モデルの許容範囲を考慮して、広げすぎにより領域が大きくなりすぎてしまわないように色の範囲を広げる。そして、他の物体に隠蔽されていない状況で候補領域が抽出できない場合と同様に、色の範囲を広げても候補領域がある程度より大きく抽出できない場合は、色の平均値を変更して候補領域を抽出し、形状特徴の認識を行う。

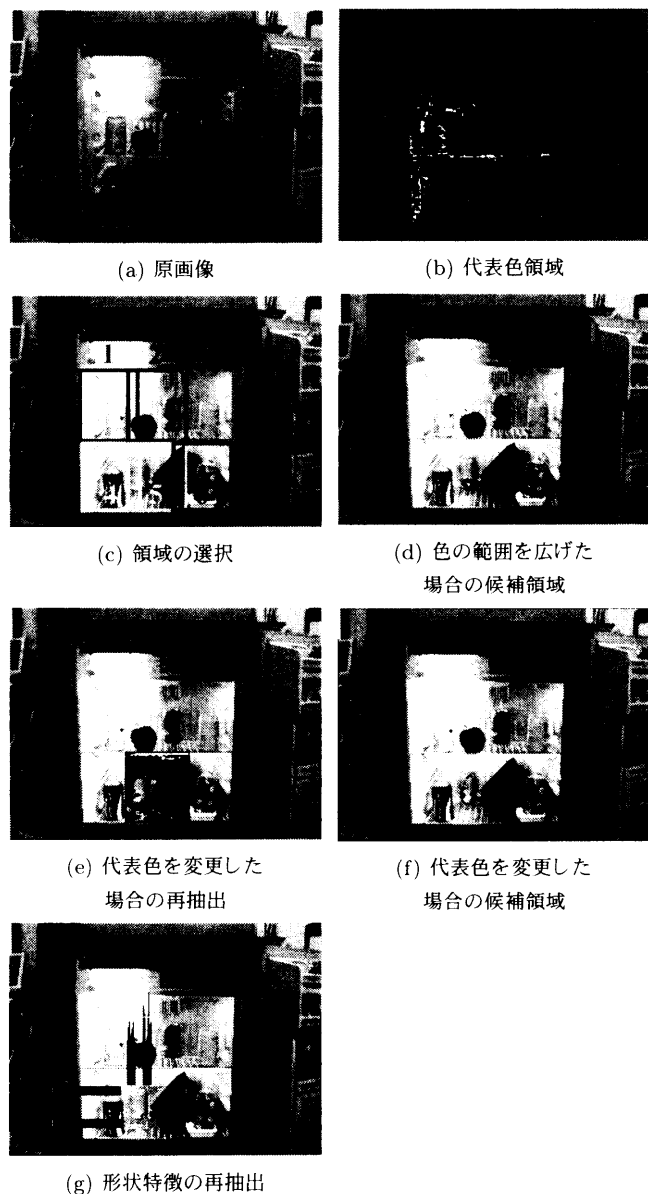


図10 候補領域の拡張

5. むすび

本論文では、対話を用いた物体認識システムにおいて、自動的な物体認識のみでは指定物体の状況や正しい位置が判断できないような場合において、ユーザから有用な助言を得て、その助言に基づいて候補物体を再認識する方法について述べた。最初に、物体認識の際に学習しておいた代表色による認識に加えて物体モデルの大きさを考慮することにより形状特徴を再認識する手法について述べた。次に、物体認識を行った結果得られる認識について考え、指定物体と得られる物体認識結果の関係

について述べた。そして、得られた結果に基づいて候補物体のクラス分けを行った。また、候補領域が得られた物体について、その物体の状況と位置を正しく知るためにユーザに助言を要求し、その助言に基づいた物体の再認識を行った。候補領域が得られない物体についても、その物体の位置に関する助言を得ることにより物体の再認識を行った。

今後の課題は以下の通りである。

- 音声認識の失敗を考慮したユーザとの音声対話
- 果物等の不定形物体の認識
- サービスロボットへの実装

文 献

- [1] Y. Takahashi, T. Komeda, T. Uchida, M. Miyagi, and H. Koyama: Development of the mobile robot system to aid the daily life for physically handicapped, Proc. of ICMA2000, pp. 549-554(2000).
- [2] 渡辺, 長尾, 岡田: 画像の内容を説明するテキストを利用した画像解析, 画像の認識・理解シンポジウム (MIRU'96) 論文集 Vol. 2, pp. 271-276(1996).
- [3] S. Wachsmuth and G. Sagarer: Connecting Concepts from Vision and Speech Processing, Workshop on Integration of Speech and Image Understanding(1999).
- [4] 高橋, 中西, 久野, 白井: 音声とジェスチャによる対話に基づくヒューマンロボットインターフェース, インタラクシオン'98 論文集, pp. 161-168(1998).
- [5] 榎原靖, 滝澤正夫, 白井良明, 三浦純, 島田伸敬, "ユーザとの対話を用いたサービスロボットのための物体登録", 第19回日本ロボット学会学術講演会 講演論文集, CD-ROM, 2001.
- [6] 榎原靖, 滝澤正夫, 白井良明, 島田伸敬, "対話を用いた物体認識のための照明変化への適応", 電子情報通信学会論文誌, Vol. J87-D-II, No. 2, pp. 629-638, Feb. 2004.
- [7] Y. Makihara, M. Takizawa, Y. Shirai, J. Miura, and N. Shimada, "Object Recognition Supported by User Interaction for Service Robots", Proc. of 16th Int. Conf. on Pattern Recognition, Vol. 3, pp. 561-564, Quebec, Canada, Aug. 2002.
- [8] Y. Makihara, M. Takizawa, Y. Shirai, J. Miura, and N. Shimada, "Object Recognition Supported by User Interaction for Service Robots", Proc. of 5th Asian Conf. on Computer Vision, Vol. 2, pp. 719-724, Melbourne, Australia, Jan. 2002.
- [9] 矢野恵生, 三浦純, 白井良明, 島田伸敬, "移動マニピュレータのための直観的作業教法", 第20回日本ロボット学会学術講演会 講演論文集, CD-ROM, 2002.