

## 画像を用いた物体の詳細情報認識手法

## Object Detail Recognition Method for Partner Robot in an Image

八木 亮†

芋野 美紗子†

土屋 誠司‡

渡部 広一‡

Ryo Yagi

Misako Imono

Seiji Tsuchiya

Hirokazu Watabe

## 1. はじめに

近年人間のパートナーとなるロボットの実現が求められている。そのようなロボットにはユーザからの命令を理解し、実行するための自律行動能力が必要である。この能力の実現には物体の詳細情報の認識が必要である。例えば、ロボットに「テレビをつける」と命令をする。テレビのリモコンは人間が扱うようにできているため、ロボットが人間と同じようにリモコンを扱おうとしてもボタン一つ一つを分けて認知することは困難である。またその認知したボタンが何を行うためのものであるかという詳細情報を認識することも難しい。この問題を解決し、ロボットが人間と同じように物体を扱うためには物体の詳細を認識することが必要であると考えられる。

本稿では物体の詳細部分の文字情報や大きさを格納している物体構成知識ベースを用いて物体の詳細な部分の認識を目指す。物体の詳細部分とは人間が直接操作を行う部分と定義する。本稿では対象物体がキーボードのときの例を挙げる。このとき、詳細部分とはキーボードのキー一つ一つのことをいう。

## 2. 研究概要

## 2.1. 前提条件

前提条件として以下の3点を設定する。

- ・入力画像中の物体は既知のもので、物体構成知識ベースが存在している。
- ・入力される画像は背景を写さず対象物体のみが写っている画像である。
- ・物体を撮影機器に対して水平、垂直方向から撮影する。

## 2.2. システムの流れ

本稿で目的としている物体の詳細認識システムの流れは図1のようになっている。

入力画像に対し、輝度差を用いて詳細部分の分割を行う。その分割部分に対し文字認識を行う。文字認識結果と物体構成知識ベースとを照合することで分割部分がどのキーで

あるかを認識する。ここで既に認識された詳細部分の位置情報と、まだ認識できていない詳細部分の位置情報を物体構成知識ベースから取得し、認識されていない詳細部分が何のキーであるかを推測する。

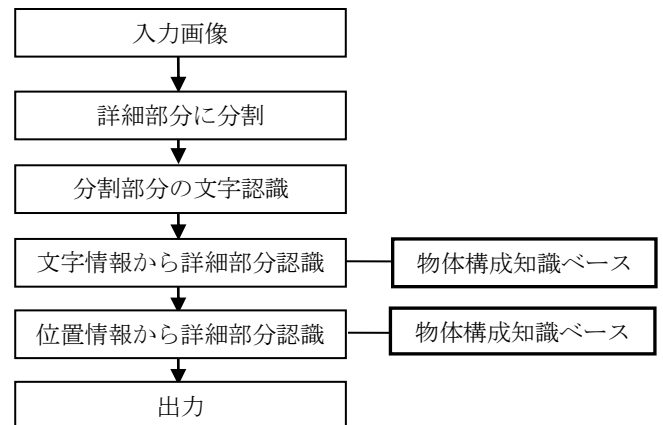


図1 システムの流れ

## 2.3. 物体構成知識ベース

物体構成知識ベースとは詳細部分の文字情報、大きさ、詳細部分同士の相対位置座標を格納した知識ベースである。物体構成知識ベースの例を表1に示す。なお大きさ、相対座標の単位はPixelである。

表1 物体構成知識ベースの例

文字情報	相対座標 (X 軸)	相対座標 (Y 軸)	大きさ (X 軸)	大きさ (Y 軸)
A/ち	50	0	150	150
S/と	225	0	150	150
D/し	400	0	150	150

## 3. 詳細部分分割

入力画像から詳細部分を分割する。詳細部分に分割することで画像内における詳細部分それぞれの領域の位置情報がわかるためである。ここで分割したときの詳細部分それぞれの位置情報と4章で説明する文字認識の結果を、物体構成知識ベースと照合することで詳細部分を認識する。

## 3.1. 輝度情報を用いた詳細部分分割

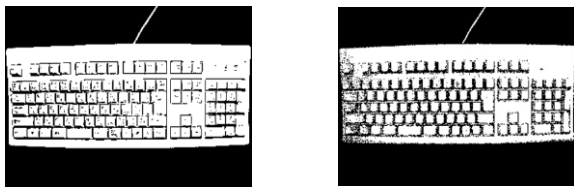
輝度差を用いて入力画像中の物体を詳細部分に分割する。詳細部分の周辺は輝度が小さくなっていることが多いため、輝度差を用いて詳細部分の分割を行う。詳細部分はボタン

‡同志社大学理工学部  
Faculty of Science and Technology, Doshisha University  
†同志社大学大学院工学研究科  
Graduate School of Engineering, Doshisha University

状になっており、詳細部分周辺が凹凸状になっているため影ができることが多いからである。輝度差を計算する際には、*YSH* 表色系を用いた。入力画像より取得した *RGB* 成分より *YSH* 表色系に変換する。輝度が閾値以下の箇所を削除することで詳細部分に分割する。図2に入力画像、図3に *YSH* 表色系を用いて詳細部分分割を行った画像、図4に *HSV* 表色系を用いて詳細部分分割を行った画像の例を示す。図4は図3と比べ雑音が多いため、今回は *YSH* 表色系を用いることで詳細部分の分割を行った。



図2. 入力画像

図3. *YSH*表色系を用いた場合 図4. *HSV*表色系を用いた場合

### 3.2. ラベリング処理を用いた詳細部分分割

物体構成知識ベースに登録されている部分以外の雑音を削除するためにラベリング処理を施し、物体構成知識ベースに登録されている詳細部分の面積の最大値を超えた部分を削除する。

### 4. 分割部分の文字認識

人間が扱う物体の詳細部分の周辺には文字が書かれていることがほとんどである。例えば、キーボードのキー上にはそのキーが何であるかの文字情報が存在している。よって詳細に分割した画像に対し、文字認識ソフトを使って文字情報を獲得する。

### 5. 文字情報から詳細部分認識

5章で獲得した文字情報と物体構成知識ベースに格納されている文字情報とを照合する。文字情報が一致していれば詳細部分が文字情報で示された部分であると認識する。一致していなければその部分は不明な詳細部分として6章の処理を行う。

### 6. 位置情報から詳細部分認識

6章で認識できていない、かつ物体構成知識ベース内に不明な詳細部分が残っていた場合、5章で認識した詳細部分と物体構成知識ベースの相対座標を用いて、不明な詳細部分を推測する。

### 7. 実験

種類の違うキーボード画像8枚、詳細部分総計208個に対して検証を行った。

評価式を式(1)に示す。 $P$ は精度、 $C\_SUM$ は画像上の位置情報と文字情報が合っていた部分の総数、 $K\_SUM$ は物体構成知識ベースに存在する詳細部分の総数を表す。

$$P = C\_SUM / K\_SUM \quad (1)$$

図5にOCR認識正答率とシステムの認識成功率を示す。OCR正答率とは、物体構成知識ベース内に存在する詳細部分に対しOCRの文字認識が正解していた割合を表す。システム全体の認識精度は68.3%、文字認識処理までの精度は27.8%だった。物体構成知識ベースの位置情報を用いて不明部分を推測した場合としなかった場合では約40%の精度の差が見られた。

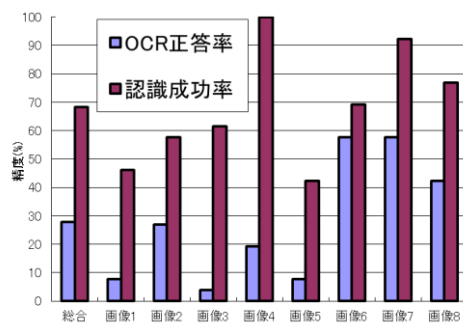


図5. 実験結果

### 8. 考察

図5からOCR正答率とシステムの認識成功率に関連性を見ることはできなかった。画像6について、OCR正答率と認識成功率が10%ほどしか変わらなかった。これは、「。」を「o」のように、OCRソフトから取得した文字情報が、物体構成知識ベースに存在していない文字情報なのに、存在している文字情報と誤認識してしまうことが原因だと考えられる。

### 9. おわりに

本稿では、物体の持つ詳細部分の認識を行った。物体構成知識ベースの位置情報を用いて不明部分を推測した場合としなかった場合では、約40%物体構成知識ベースの位置情報を用いた場合のほうが精度が高かった。

#### 謝辞

本研究の一部は、科学研究費補助金(若手研究(B)21700241)の補助を受けて行った。

#### 参考文献

- [1] 井上誠喜, 八木伸行, 林正樹, 中洲英輔, 三谷公二, 奥井誠人, “C言語で学ぶ実践画像処理”, オーム社, 1999