

# 明度情報とカラー情報を用いた情景画像中からの文字列抽出

## Character String Extraction from Scene Image using Luminance and Color information

平山 勝裕†  
Katsuhiko Hirayama

大町 真一郎†  
Shinichiro Omachi

阿曾 弘具†  
Hirotomo Aso

### 1. はじめに

屋内、屋外を移動するロボットが周囲環境の変化に適応して作業するには、自分の周囲の環境を認識することが不可欠である。近年、ロボットや画像処理・理解の研究の進展により物体の形状記述や移動経路上の障害物の判断などが可能となり、視覚による移動ロボットの実現が可能となってきている。しかし、視覚による移動ロボットが物体の形状や移動経路上の障害物の判断だけでなく、文字情報をすべて正しく抽出し、理解することができれば、より高度な動きを実現できると思われる。これを実現するにあたり、まず情景画像中に存在する文字情報を全て正しく抽出することが必要となる。

従来手法として松尾らの手法 [1] がある。これは元画像から適応閾値法を用いて文字領域と背景領域を良好に分離していると考えられる 2 値画像 (分解画像) を選択し、文字列領域を抽出する手法である。しかし画像中には文字以外の部分が多数存在するため、文字と背景を良好に分離する適切な閾値を選択することが難しい。また、一つの文字の上部と下部で照明条件が異なる場合等の特殊な照明条件下においては文字と背景とを良好に分離することも困難となる。

そこで本論文ではデジタルカメラによって撮影された種々の状況下における情景画像から文字列 (漢字・平仮名・片仮名・アルファベット) を抽出することを目的として、画像を小領域に分割し、明度情報とカラー情報から各画素に適切な閾値を与えることで、文字以外の部分や照明条件の影響を受けることなく文字と背景とを良好に分離する 2 値画像を作成し、文字列領域を抽出する手法を提案する。

### 2. 情景画像からの文字列抽出法

情景画像中に存在する文字列は、照明条件等の影響により様々な色彩や明度で存在し、その大きさや形状も画像によって異なる。これらの問題に対し良好に文字列領域を抽出するために以下の処理手順によって抽出を行う。

#### 2.1 画像の 2 値化

文字以外の部分や証明条件の影響を緩和するために画像を小領域に分割し、小領域ごとに NAT 法 [2] を用いて閾値を決定する。小領域のサイズは  $8 \times 8$  とし、4 画素ずつ移動させる。各画素は最大 4 つの小領域に属するので 1 個～4 個の閾値が求まる (図 1)。これを閾値  $th_1, th_2, th_3, th_4$  とする。

次に注目画素の周りと閾値を計算した 4 個の小領域の正規化 RGB 空間上での平均色を求め距離を比較し、各画素における 1 個～4 個の閾値に重みを付け、一つの閾値  $T$  を決定する。具体的には次のようにする。

†東北大学大学院 工学研究科

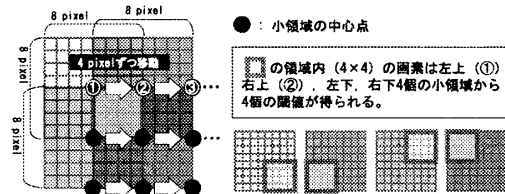


図 1: 小領域ごとの閾値計算における参照領域

RGB 空間から正規化 RGB 空間への変換は次の式で求める。

$$k = R + G + B$$

$$\text{if}(k = 0)(r, g, b) = (0, 0, 0)$$

$$\text{else}(r, g, b) = (R/k, G/k, B/k)$$

まず注目画素  $p(i, j)$  を中心とした  $5 \times 5$  領域の 25 画素の正規化 RGB 空間上での平均値 ( $avR, avG, avB$ ) を求める。そして注目画素  $p(i, j)$  の閾値を算出した 4 個の小領域 ( $8 \times 8$ ) の 64 画素の正規化 RGB 空間上での平均値を  $(R_1, G_1, B_1) \sim (R_4, G_4, B_4)$  とする (図 2)。

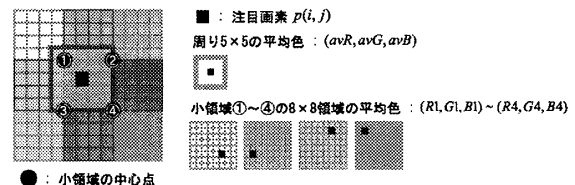


図 2: 重み付けのための平均色算出における参照領域

$(avR, avG, avB)$  との距離が最小である  $(R_n, G_n, B_n)$  ( $n = 1 \sim 4$ ) を求め、その領域から得られた閾値に大きな重みを付けて注目画素の閾値  $T$  を計算する (式 (1))。

$$\text{if}(n = 2 \text{の時距離最小}) T = \frac{th_1 + 1.2 \times th_2 + th_3 + th_4}{4.2} \quad (1)$$

全画素において閾値を計算し、分解画像を作成する。

#### 2.2 文字領域候補抽出

##### (1) 文字領域候補・部分領域候補抽出

分解画像において 4 連結で閉領域を抽出し外接方形を与え、その形状により「文字領域候補」「部分領域候補」「漢数字」「不用」の 4 種類に分類する。

##### (2) マージング処理

偏と旁など分離している文字などを一つの方形で囲むために 2 つの外接方形が交差しているか、また分離していても 5 画素以内の場合に以下の 3 種類のマージング処理を行う。

##### i) 文字領域候補方形同士のマージング処理

2 つの外接方形の面積比率が 1.5 以内かつ辺比率が 2.0 以内のときマージして 1 つの方形で囲む

- ii) 文字領域候補と部分領域候補のマーキング処理  
2つの外接方形の面積比率が 3.0 以内かつ辺比率が 2.5 以内のときマージして1つの方形で囲む
  - iii) 部分領域候補同士のマーキング処理  
2つの外接方形が 1.5 以内かつ辺比率が 2.0 以内のときマージして1つの方形で囲む
- 上記3種類のマーキング処理は以下のような場合を想定している。
- i) 偏や旁が2つの領域に分離している場合
  - ii) 偏と旁のどちらかが部分領域候補として抽出されており、文字の一部が文字領域候補として抽出されている場合
  - iii) 偏と旁それぞれが部分領域候補として抽出されている場合

### 2.3 文字列領域候補抽出

前項で抽出された文字領域候補から、文字と文字との連結性を以下の5条件で判断し文字列を構成していると思われる文字列領域候補を抽出する。

- 方形同士は交差していない
- 方形同士の面積比率が 1.5 以下
- 2つの文字領域候補間の距離が 3 文字分以下
- 2つの文字領域候補の重心のずれが垂直または水平方向に対して 3 画素以内
- 2つの方形の白画素部分か黒画素部分における正規化 RGB 空間での平均色の差が 0.025 以下

### 3. 実験

縦または横一列以上の文字列(漢字・平仮名・片仮名・アルファベット)が存在する情景画像 50 枚に対して実験を行った。50 枚の情景画像中に目視によって読みとることのできる文字の総数は 508 文字存在し、それらの文字のサイズは様々であった。撮影場所、文字列の位置、撮影時の天候などは特に限定せずに撮影した画像を使用した。

評価法として、抽出された文字列領域候補から再現率、適合率を求める。再現率は存在する文字の何割を抽出できたか、また適合率は抽出した方形の何割が正しく文字を囲っているかを示す。

表 1 に提案手法と適応閾値法を用いて分解画像を作成した場合の実験結果を示す。

表 1: 実験結果

	画像分割法	適応閾値法
再現率	75.98 % (386/508)	62.40 % (317/508)
適合率	33.68 % (386/1146)	53.27 % (317/595)

表 1 から提案手法の方が適合率は低下しているものの再現率は向上していることがわかる。すなわち、従来手法では抽出が不可能であった文字列も提案手法により抽出することが可能になったことがわかる。適合率が低下した理由は、表 1 から分かるように、抽出した方形数が非常に多かったことによるものである。これは各領域で 2 値化を行ったために分解画像に雑音が発生しているため、結果として雑音も文字列領域候補として抽出され

てしまったためである。しかし、抽出された文字列領域候補の文字としての整合性を考慮すれば、削減することのできる問題であると考えられる。

次に再現率が向上した理由は、小領域ごとに 2 値化を行ったため、文字以外の部分の影響を低減させることが可能になり小さな文字情報も抽出することができたこと(図 3)、照明条件の異なる領域の境界上に文字列が配置されている場合などの照明条件を緩和することができ、文字列領域候補として抽出することができたことなどがある(図 4)。



図 3: 文字情報が小さい画像の例(左から元画像、画像分割法による分解画像、適応閾値法による分解画像)

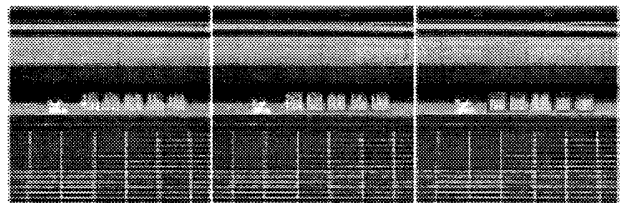


図 4: 照明条件の異なる領域の境界上に文字列が配置されている画像の例(左から元画像、画像分割法を利用した文字列領域候補抽出例、適応閾値法を利用した文字列領域候補抽出例)

小領域ごとに 2 値化したことで、図 3 から明らかなように文字以外の部分の影響を低減させることができ、小さな文字情報も抽出することができた。また図 4 から明らかなように、文字の上部と下部とで照明条件が異なっても一つの方形で抽出することができた。

### 4. まとめ

情景画像に対して、画像を分割し小領域ごとに 2 値化を行うことで分解画像を作成し、文字列を成す文字の領域を抽出する手法を提案した。画像を小領域ごとに 2 値化し分解画像を作成したため、文字以外の部分の影響を低減させることができ、種々の状況下での情景画像に対しても抽出精度を向上させることができた。

今後の課題として、小領域ごとに 2 値化することによって増加してしまった雑音領域の排除が挙げられる。また、小領域の一辺の長さや、小領域(8×8)同士の重なり具合などを最適化する必要があると考えられる。

### 参考文献

- [1] 松尾賢一、上田勝彦、梅田三千雄：“適応閾値法を用いた情景画像からの看板文字列領域の抽出” 信学論(D-II) Vol.J80-D-II No.6 pp1617-1626 June 1997.
- [2] Hon-Son Don "A noise attribute thresholding method for document image binarization" International Journal on Document Analysis and Recognition (2001) 4:131-138