

## 閉輪郭線の文字らしさに基づく画像からの文字列検出

西野友博 (東大)  
岡田慧 (東大)

山崎公俊 (東大)  
稲葉雅幸 (東大)

Tomohiro NISHINO, The University of Tokyo,  
tomohiro@jsk.t.u-tokyo.ac.jp  
Kimitoshi YAMAZAKI, Kei OKADA, Masayuki INABA

## 1 序論

環境中には文字が書かれた物品が数多くあり、人間はそれらの文字から様々な情報を得ている。人間の生活を支援する家庭用ロボットにとっても、文字情報の利用は有用であると言える。それを実現するための第一歩は、環境中の文字を検出することである。これまでの取り組みとして、閉輪郭線を利用して文字領域を抽出する手法 [1] があり、あらゆる文字は閉輪郭線を持つという特徴を利用した有効な手法であると考えられるが、抽出した輪郭線が閉じていない場合に文字を抽出できないことがある。

本論文では、文字の線幅や直線上に並んでいることなどの文字らしさから、輪郭線が文字列に属しているかを判断し、文字列を検出する手法について述べる。本手法でも文字領域検出に閉輪郭線を利用するが、その閉輪郭線から文字の線幅を推定する。文字の線幅を考慮することで、閉輪郭線として抽出できなかった輪郭線についても文字かどうかの判断を行うことができ、文字領域を補正することができる。

日常環境中では、文字が書かれた物品が奥行き方向や光軸周りに傾いていることが多いと考えられるが、本手法では輪郭線を用いているため、奥行き方向への傾きによる歪みがある文字も検出することができ、また、文字列を通る直線を検出することにより、光軸周りの傾きにも対応している。これらの処理により、日常環境中の自由な視点から文字列を検出することが可能となる。

## 2 閉輪郭線と文字線幅に基づく文字領域検出

## 2.1 閉輪郭線による文字領域の検出

まず、画像中の閉輪郭線抽出を行う。画像から Canny エッジを抽出する。このとき、長さが短いエッジは除去する。Canny エッジは角度が急な場合などに途切れてしまうことがあるため、隣接したエッジの接合を行う。その後、長さが短いエッジを1回目よりも閾値を高くして除去する。そして、エッジの両端間の距離が近いものを閉輪郭線として抽出する。この文字領域抽出処理においては、閉輪郭線のうち、他の閉輪郭線に包含されているものは除去する。ただし、除去した閉輪郭線は、それを包含していた閉輪郭線と関連付けて覚えておく。

次に、文字領域の検出を行う。文字列を構成する各文字は同一直線上に並んでいると仮定する。抽出した閉輪郭線が水平方向に連続していれば、その領域を文字領域として検出する。文字が水平方向に並んでおらず、光軸周りに傾きがある場合の対処法については第3章で述べる。

これらの処理結果を図1に示す。この手法により文字領域が検出できていることが分かる。

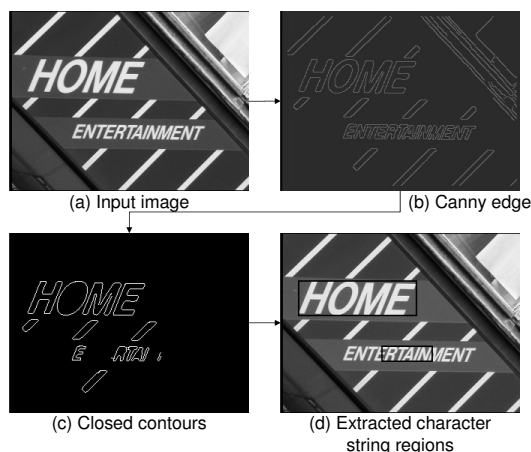


図 1: 文字領域検出処理結果

## 2.2 文字線幅を考慮した文字領域の補正

抽出した文字の閉輪郭線を利用して、その文字が属する文字列における文字線幅を推定することができる。

まず、各文字における閉輪郭線のある一定のピクセルごとにサンプリングする。そして、各サンプリング点を通り輪郭に垂直な直線を求め、その文字の閉輪郭線との交点を求める。サンプリング点とその交点の距離を求め、文字線幅候補とする。文字の閉輪郭線について、図2に示すように文字線幅候補として大きいものと小さいものが現れるが、求めるべき文字線幅は長辺におけるサンプリング点から求めることができるので、文字線幅候補を昇順に並び替えたときの中央値を採ることにより、それを文字線幅とすることができる。

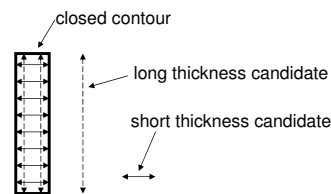


図 2: 閉輪郭における文字線幅候補

各文字領域から文字線幅を推定した後、その文字線幅に基づいて文字領域の補正を行う。輪郭線抽出で求めた輪郭線の中から、各文字領域の左右に位置する輪郭線を抽出する。それらの輪郭線の包含関係を考慮して、各輪

郭線について文字線幅を推定する．文字領域における文字線幅と輪郭線から推定した文字線幅が類似している輪郭線を抽出し，それを文字領域に加える．

図3，図4は，それぞれ図1-(d)の上の領域と下の領域の文字線幅候補を昇順に並び替えたグラフである．中央値付近でグラフの傾きが小さくなっていることから，中央値を文字線幅として選ぶことにより安定して文字線幅を推定できると考えられる．図5-(a)は図1-(d)で文字領域として抽出できなかったが文字領域中の文字の線幅と類似している輪郭線であり，図5-(b)は推定した線幅に基づいて文字領域を補正した結果である．この処理により，閉輪郭線として抽出できなかった輪郭線を文字領域中に加えることができることが分かる．

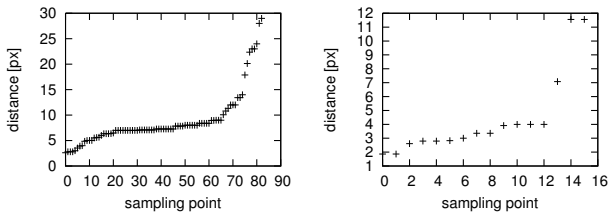


図3: 図1-(d)の上の領域における文字線幅  
図4: 図1-(d)の下の領域における文字線幅

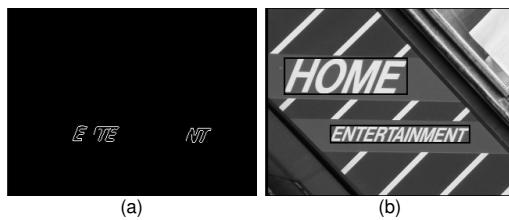


図5: 文字線幅に基づく文字領域の補正

### 3 ハフ変換を利用した文字列の傾き角度推定

この手法では輪郭線を水平方向に探索しているため，文字列が光軸周りに傾いているときに文字領域の検出に失敗するという問題がある．そこで，輪郭線の並びから文字列の光軸周りの傾き角度を推定し，その方向に輪郭線を探索する．このためには，文字列は直線上に並んだ文字から成り立つと仮定すれば，文字列を構成する各文字を通る直線を求めればよい．

この直線を求めるためにハフ変換の原理を利用する． $(r, \theta)$ 平面を格子状に区切り，各輪郭線中心を通る直線に対応する格子に投票する．このとき，投票する点数は輪郭線の端点間の距離によって変化させ，輪郭が開いている場合は点数を大きく，閉いているほど点数を小さくする．

求められた直線の傾き方向に文字を探索し，文字領域として検出した結果を図6に示す． $(r, \theta)$ 平面への投票結果を図7に示す．この処理により，文字列が光軸周りに傾いていても検出できることが分かる．

### 4 実験

これらの手法を用いて，物体上の文字列を検出する実験を行った．対象物体は，日常環境中のロボットの視点

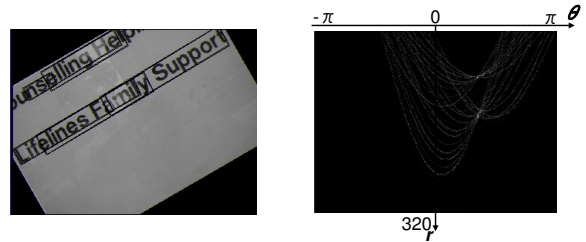


図6: 光軸周りの傾きを考慮して検出した文字領域  
図7:  $(r, \theta)$ 平面への投票結果

を想定して，奥行き方向や光軸周りに回転させている．実験結果を図8に示す．輪郭線を利用して文字検出を行っているため，奥行き方向に傾いて歪みがある文字であっても，検出することができた．また，ハフ変換を利用して光軸周りの回転を検出しているため，光軸周りに傾きのある文字列も検出することができた．

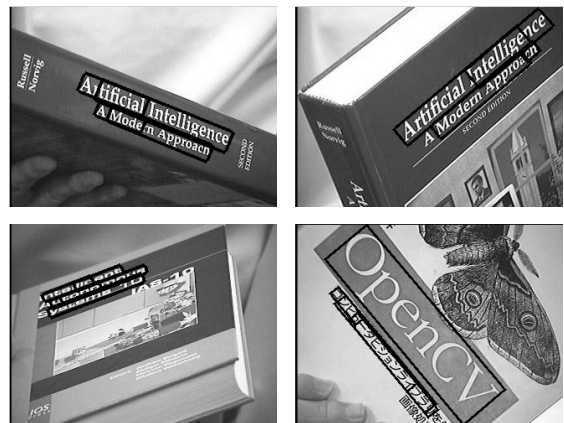


図8: 画像中の文字列検出

### 5 結論

本論文では，閉輪郭線を抽出し，直線に並んでいることを判定して文字領域を検出した後，その文字領域における文字線幅を閉輪郭線から推定することによって，文字領域を補正できることを示した．また，ハフ変換を利用して各輪郭線の中心を通る直線を求めることによって，文字列の光軸周りの傾き角度を推定することができ，光軸周りに傾いた文字列も検出できることを示した．

これらの処理によって，奥行き方向や光軸周りに文字列が傾いていることが考えられる日常環境中においても，文字領域を検出できると考えられる．

今後は，検出した文字をOCRで処理するための歪み補正や文字が小さかったり傾きが大きかったりした場合にはズームや視点の変更を行うなどの処理を実装していく予定である．

### 文献

[1] Takahashi Hiroki, Kasai Kouichi, Kim Daewoo, Nakajima Masayuki: "Extraction of text regions from scenery images using multiple features" *ITE Technical Report*, vol.25, no.11, pp.39-44, 2001.