

Toggle Mapping を用いた情景画像からの看板領域抽出

A Method for Extracting Signboard Regions from Scene Images Using Toggle Mapping

○佐々木 隆行¹, 陳 建軍¹, 高木 昇²

○Takayuki Sasaki¹, Jianju Chen¹, Noboru Takagi²

¹ 富山県立大学大学院工学研究科知能デザイン工学専攻

¹ Graduation School of Engineering, Toyama Prefectural University

² 富山県立大学工学部知能デザイン工学科

² Department of Intelligent Systems Design Engineering, Toyama Prefectural University

Abstract: In this paper, we propose an approach to extract signboard regions from scene images. We first carry out toggle mapping which is able to detect a region whose intensities are homogeneous. We therefore detect an intensity homogeneous area by using toggle mapping. A support vector machine is applied to classify a homogenous area into signboard or non-signboard classes. Finally, we conduct computer experiments to show the effectiveness for the proposed method.

1. はじめに

視覚障害者には残存視覚を有する弱視、視覚を持たない全盲の2種類に大別できる。全盲の人は外出する際に盲導犬や介護者の助けを必要とし、残存視覚を有する弱視は独力での行動を希望する反面、見知らぬ場所では目的地に辿り着けずに迷ってしまうことがある。このように、視覚障害者の自立が困難な理由の一つとして移動や歩行の困難さが挙げられる。我々は、視覚障害者が単独で街頭を移動する支援として視覚障害者歩行支援システムの開発に取り組んでいる。

現在普及している視覚障害者歩行支援には、点字ブロックや音響式信号機などがある。また、研究が進んでいる分野としては、GPS[1], RFID[2], 画像処理[3]を用いた研究が挙げられる。画像処理による歩行支援は、情景画像から情報を取得する必要がある。情景画像を入手するには様々な方法があるが我々はスマートフォンを用いた方法を考える。まず、現代はスマートフォンが普及し、誰もが外出先で画像を取得できるためインフラ整備が必要ない。また、Apple社から発売されているiPhoneであれば、視覚障害者向けにVoiceOverというスクリーンリーダーがデフォルトで搭載されているため、目印となる凹凸の少ないスマートフォンでも利用が可能である。これらのことから、スマートフォンを用いた視覚障害者歩行支援システムは、他に比べ安価に構築でき、手軽

にどこでも利用できる汎用性がある。

我々は画像処理による視覚障害者の歩行を支援するシステムの開発を目標とし、現在はシステムの一部である文字抽出を行っている。本支援システムは、情景画像から看板内の文字情報を取得し、その内容を視覚障害者に音声で提示し、視覚障害者の視認を支援するものである。情景画像からの文字抽出に関する先行研究には主に以下に示す2つの主流手法がある。

- ① エッジ情報による文字抽出
- ② 色情報による文字抽出

エッジ情報を用いた手法は古典的なエッジ抽出法を使用する研究が多い。そのため複雑な背景をもつ情景画像では文字などの輪郭をエッジとして鮮明に抽出することが困難である。この手法を使用した文献[4]では屋内での適用を前提としており、屋外の画像に適用した場合複雑なテクスチャや樹木によって誤検出が数多く見られた。また、色情報を用いた手法[5,6,7]は、文字領域が背景色と異なること、及び同じ文字列の文字は同色であること、という特徴を利用して文字を抽出している。このような色情報を用いる文字抽出では、色情報に基づくクラスタリングが良く用いられる。しかしながら、複雑な背景を持つ情景画像ではクラスター数の決定が困難な問題であり、かつ非常に多くの中間色が存在するため、適切に文字領域の画素をクラスターとしてグループ化する

ことが難しく、正確に文字領域を抽出できない。

我々が使用する情景画像は街中で撮影した情景画像を想定している。画像には建物や樹木や人物など複数のオブジェクトが入るため、背景が複雑になる。我々は、複雑な背景を持つ情景画像から文字領域を正確に抽出するためには、複雑なテクスチャや樹木のような画像領域からの誤検出を抑えることが課題であると考えた。今回の提案手法では、Toggle Mapping により輝度の均質領域である看板領域を抽出することで非文字成分を除去し、文字抽出の範囲を看板領域に限定することで文字抽出の高精度化が考えられる。処理の流れは次節で説明する。

2. 看板領域抽出処理の流れ

提案手法の概要を示す。まず、前処理として平滑化を行うことで画像内のノイズを除去する。次に均質領域抽出処理を行う。均質領域抽出処理では、toggle mapping[8,9]を用いて、輝度勾配の小さい均質領域を抽出し、抽出した領域を看板候補領域とする。看板候補領域から特徴量を抽出し、サポートベクターマシンによる看板領域判別処理を行い、看板領域を抽出する。

3. 平滑化

情景画像中の文字の画素は必ずしも同一の RGB 値を持つとは限らない。また、情景画像を取得する際何らかの原因によってノイズの混入や照明変化の影響で文字画素の RGB 値が変動する場合がある。そこで、同じオブジェクトを表す画素が似たような RGB 値を保持し、かつノイズを除去した画像を作成するために前処理として平滑化を行う。

平滑化手法としては、上下左右 4 方向からの RGB 値の変化分に対し、明度値にメディアンフィルタを掛けた値の差分で重み付けをする WMFAD(Weighted Median Filter Based Anisotropic Diffusion)を利用する。WMFAD による平滑化後の画素値は以下の式により求められる。

$$I'(x,y) = I(x,y) + \lambda \sum_D C_D \nabla_D I(x,y) \quad (1)$$

ここで $I(x,y)$ は注目画素 (x,y) の RGB 値、 $I'(x,y)$ は平滑化後の RGB 値、 D は上下左右の 4 方向、 $\nabla_D I(x,y)$ は 4 近傍の画素と注目画素 $I(x,y)$ の差分、 C_D は 4 近傍に対する重みを表し、 Y は明度値、

$$C_D = \exp(-(\nabla_D \text{median}(Y)/K)^2) \quad (2)$$

$$Y = 0.299R + 0.587G + 0.144B \quad (3)$$

で表される。 $\text{median}(Y)$ は明度値にメディアンフィルタを掛けた値、 λ 、 K は定数を表す。

式(1)において、 C_D は式(2)からも明らかなような注目画素と 4 近傍の画像との $\text{median}(Y)$ の差分に反比例する値となる。そのため、注目画素

と 4 近傍の画素との $\text{median}(Y)$ の差分が大きい画素の $\nabla_D I(x,y)$ は $I'(x,y)$ にあまり反映されず、差分が小さい画素の $\nabla_D I(x,y)$ が反映されるため、局所的なエッジ部分を残したまま平滑化することができる。また C_D は $\text{median}(Y)$ という明度値にメディアンフィルタを掛けた値で決定されるため、ノイズ画素においても 4 近傍の画素に近い $\text{median}(Y)$ 値を保持し、4 近傍の画素の $I(x,y)$ が反映されることによりノイズを除去することができる。

4. 均質領域抽出処理

看板は一般的に、背景と文字の 2 つのオブジェクトで構成され、大部分が背景を占める。さらに、背景と文字は互いに輝度勾配が大きくなるような色が使用されているという特徴がある。そこで、色の均質な領域を抽出することで看板領域抽出を行う。均質領域の抽出には toggle mapping を用いる。

4.1 Toggle Mapping

toggle mapping は、モルフォロジー演算を用いた画像内コントラスト強調手法であり、モルフォロジー演算の Dilatation, Erosion を用いてぼやけの修正などに用いられている。

toggle mapping とは、モルフォロジーの Dilatation 処理と Erosion 処理の差分のことである。均質領域では toggle mapping が小さいことを利用して以下の式で均質領域を抽出することができる。

$$s(x,y) = \begin{cases} 0 & \text{if } |d(x,y) - e(x,y)| < th \\ 1 & \text{if } |d(x,y) - e(x,y)| \geq th \end{cases} \quad (4)$$

ここで、 $d(x,y)$ は Dilatation 処理を行ったグレースケール画像 d の (x,y) 座標の画素値、 $e(x,y)$ は Erosion 処理を行ったグレースケール画像 e の (x,y) 座標の画素値、 th は閾値である。toggle mapping が閾値以下、つまり 0 の場合は、 $s(x,y)$ は均質領域の画素となる。この性質を利用して、我々は toggle mapping を画像分割に応用する。

toggle mapping による均質領域抽出の結果を図 1 にいくつか示す。



図 1：均質領域抽出結果

図1は左上が入力画像である。まず、入力画像に対してグレースケール処理を行う。次に、Dilatation 処理を行った画像 d 、Erosion 処理を行った画像 e に対して、式(4)を用いた閾値処理を行う。結果は、右上、左下、右下がそれぞれ閾値1, 5, 10 のときの結果である。いずれもラベリング処理により連結成分にはそれぞれ異なる色が割り当てられている。結果からわかるように、閾値によって均質領域が大きく異なる。そこで我々は、閾値を変化させ、得られた領域内の輝度値の標準偏差をグレースケール画像から求め、閾値による評価を行うことで、画像ごとに最適な均質領域を得られる方法を提案する。我々が行った手法を以下に示す。

4.2 提案手法

入力：濃淡画像 I_g

出力：均質領域の集合 H

Step 1. (パラメータの初期化)

$i \leftarrow 1$, $th_1 \leftarrow 1$, $th_2 \leftarrow 15$, および $H \leftarrow \emptyset$ に初期化する。(iは繰り返し回数の制御変数, th_1 は toggle mapping で得られた濃淡画像を2値化するための閾値, th_2 はある領域が均質領域であるか否か判断するための閾値, H は均質領域を保存するための集合である。)

Step 2. (初期均質領域候補の決定)

濃淡画像 I_g に toggle mapping を施し I_t を得る。閾値 th_1 を適用し I_t の2値画像 $I_b^{(1)}$ を得る。 $I_b^{(1)}$ の連結成分を $C_1^{(1)}$, $C_2^{(1)}$, ..., $C_{n(1)}^{(1)}$ とする。

Step 3. (均質領域候補の決定)

$i \leftarrow i + 1$, $th_1 \leftarrow th_1 + 1$ とする。 I_t に閾値 th_1 を適用して得られる2値画像を $I_b^{(i)}$ とする。 $I_b^{(i)}$ の連結成分をすべて求める。 H のどの均質領域にも被覆されない連結成分を $C_1^{(i)}$, $C_2^{(i)}$, ..., $C_{n(i)}^{(i)}$ とし, $C^{(i)} \leftarrow \{C_1^{(i)}, C_2^{(i)}, \dots, C_{n(i)}^{(i)}\}$ とする。

Step 4. (均質領域判別)

集合 $C^{(i)}$ の各連結成分 $C_j^{(i)}$ に対して, $C_j^{(i)}$ の画素値の標準偏差を $E(C_j^{(i)})$ とする。標準偏差 $E(C_j^{(i)})$ が閾値 th_2 より大きいとき, 連結成分 $C_j^{(i)}$ に被覆される $(i-1)$ 回目の連結成分を均質領域と判別し, 集合 H に加える。すなわち, $C_{k(1)}^{(i-1)}, C_{k(2)}^{(i-1)}, \dots, C_{k(t)}^{(i-1)}$ ($k(m) \in \{1, 2, \dots, n(i-1)\}, m = 1, 2, \dots, t, t \leq n(i-1)$)を $C_j^{(i)}$ に被覆されると $(i-1)$ 回目の連結成分とすると $H \leftarrow H \cup \{C_{k(1)}^{(i-1)}, C_{k(2)}^{(i-1)}, \dots, C_{k(t)}^{(i-1)}\}$ とする。

Step 5. 標準偏差 $E(C_j^{(i)})$ が閾値 th_2 より大きい連結成分 $C_j^{(i)}$ を集合 $C^{(i)}$ から取り除く。

Step 6. (終了条件の判定)

集合 $C^{(i)}$ の各連結成分の画素数を求め, その最大値を $s^{(i)}$ とする。さらに, 集合 $C^{(i)}$ の各連結成分の画素数を求め, その総和を $n_p^{(i)}$ とする。 $0.9 < n_p^{(i-1)}/n_p^{(i)} < 1.1$ かつ $s^{(i)}/s \leq 0.01$ ならば終了する。ここで, s は元画像の画素数である。

Step 7. Step 3 に戻る s は元画像の画素数である。

Step4 では, 均質領域の判別を行っている。この均質領域判別は, 均質領域は閾値 th_1 を増加していくことで拡大し, 他の均質領域と結合することを利用した判別法である。まず, 均質領域の標準偏差 $E(C_j^{(i)})$ を求める。 $E(C_j^{(i)})$ が閾値を超えた場合, その領域は i 回目の閾値 th_1 により他の領域同士が結合された領域であると判断できる。つまり, $(i-1)$ 回目の処理では正しく均質領域を決定できていたと考えられる。したがって, $C_j^{(i)}$ に被覆した $(i-1)$ 回目の均質領域が最適な閾値による均質領域であると判断し集合 H に加える。

Step6 の条件に当てはまるまで繰り返し処理を行い画像内すべての均質領域を抽出する。その結果, 図2の画像から図3のような均質領域が抽出される。



図2：入力画像例

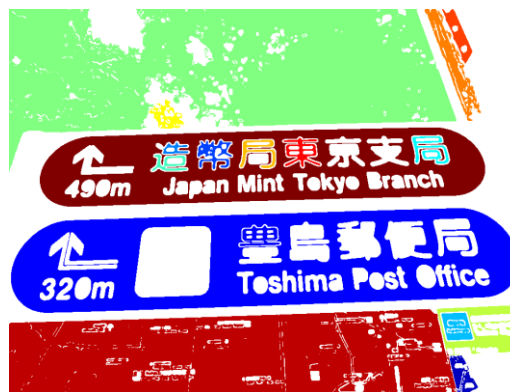


図3：図2における均質領域

4.3 均質領域抽出実験

看板を含む情景画像 154 枚に対して、提案手法による均質領域抽出を行った。154 枚の情景画像から目視で 197 個の文字認識が可能な大きさの看板領域を確認した。その結果、表 1 のような結果が得られた。

表 1：均質領域抽出実験結果

看板領域	均質領域として抽出できた看板領域	再現率
197	185	93.9%

均質領域を正しく抽出できなかった画像は、図 4 のように屋内での照明による影響や、屋外での影による影響により、同一オブジェクトであっても輝度値が異なることが原因であった。すなわち、提案手法は輝度値の大きな変化のない均質領域においては正しく処理が行われていると考える。



図 4：失敗した入力画像例



図 5：図 4 における均質領域

5. まとめ

本論文では、toggle mapping による情景画像から均質領域抽出を行った。実験から正しく提案手法が動作していることを確認した。今後は、抽出した均質領域から看板領域抽出を行う。看板領域抽出法は、均質領域から特徴量の抽出を行い、

Support Vector Machine を用いることで均質領域の看板、非看板領域の判別を行うことで看板領域を抽出する。

参考文献

- [1] 石川准, 兵藤安昭, “GPS による視覚障害者歩行支援システムの開発”, 信学技報, Vol.104, No.554, pp. 51- 56, 2005.
- [2] 国土交通省自立移動支援プロジェクト, <http://www.mlit.go.jp/seisakutokatsu/jiritsu/>
- [3] 田中奈々絵, 奥平雅士, “携帯カメラを用いた画像処理による視覚障害者行動支援に関する一検討”, 映像情報メディア学会技術報告 33(6), 173-176, 2009.
- [4] 稲見光将, 高木昇, “エッジ抽出と周波数特徴を応用した情景画像からの看板成分抽出の検討”, 日本知能情報ファジィ学会合同シンポジウム2010 (北陸信越支部・関東支部・人間共生システム研究会), 2010.
- [5] 松田友輔, 大町真一郎, 阿曾弘具, “2 値化とエッジ抽出による情景画像からの高精度文字列抽出”, 電子情報通信学会論文誌 D, Vol.J93-D, No.3, pp.336-344, 2010.
- [6] 長谷博行他, “カラー文書画像中の文字領域抽出を目的とした色分割についての検討”, 電子情報通信学会論文誌 D-II, Vol.J83-D-II, No.5, pp.1294-1304, 2000.
- [7] 芦田和毅他, “情景画像からの文字抽出”, 電子情報通信学会論文誌 D-II, Vol.J88-D-II, No.9, pp.1817-1824, 2005.
- [8] J.Fabriio, B.Marcotegui, M.Cord, “Text Segmentation in Natural Scenes Using Toggle Mapping”, 16th IEEE International Conference on Image Processing 2009, pp.2349-2352.
- [9] J.Serra, “Toggle mappings,” *From pixels to features*, pp.61-72, 1989, J.C.Simon(ed.), North-Holland, Elsevier.

連絡先

佐々木 隆行

E-mail:t354007@st.pu-toyama.ac.jp