

カラー情報を利用した情景画像中の文字列の高精度抽出

平山 勝裕[†] 大町真一郎[†] 阿曾 弘具[†]

[†] 東北大学大学院工学研究科電気・通信工学専攻 〒980-8579 宮城県仙台市青葉区荒巻字青葉 6-6-05

あらまし 情景画像からカラー情報と明度情報それぞれを利用することにより文字列領域を高精度に抽出する手法を提案する。本手法では、画像内に存在する文字とそれ以外の部分とを良好に分解する2値画像を複数枚作成し、作成された2値画像から文字列を成す文字の連結性を利用して文字列領域を抽出する。カラー情報を利用した2値画像作成では、まずカラークラスタリングを行うことで画像内に存在する各オブジェクトを代表色へと分類する。そして文字を構成している代表色とそれ以外の代表色とを分類するオブジェクトクラスタリングを行うことで複数枚の2値画像を作成する。また明度情報を利用した2値画像作成では、明度情景画像を小領域へと分割し、各小領域内に存在する各画素に周囲領域の情報のみを利用して画素閾値を決定し、各画素に異なる閾値を付加することで1枚の2値画像を作成する。このようにして得られた各2値画像において、文字列を成す文字の連結性を利用して文字列領域を決定する。
キーワード 文字列抽出, 2値画像, カラー情報, 明度情報, 情景画像

String Extraction from Scene Images Using Color Information

Katsuhiro HIRAYAMA[†], Shinichiro OMACHI[†], and Hirotomo ASO[†]

[†] Department of Electrical and Communication Engineering, Graduate School of Engineering, Tohoku University Aoba6-6-05, Aramaki, Aoba-ku, Sendai-shi, Miyagi 980-8579 Japan

Abstract In this paper, a robust character string extraction method from scene images using both color and luminance information is presented. The method makes good binary images which separate characters and other objects and extracts character strings using character connectivity. In binarization using color information, color-clustering classifies colors to some representative colors, and object-clustering classifies the representative colors. In binarization using luminance information, luminance image is divided into small regions, and a threshold to binarize is determined for each pixel in each region using only its surrounding region. Finally, in each binary image, a character string is extracted using character connectivity.

Key words String extraction, Binary image, Color information, Luminance information, Scene image

1. ま え が き

現代社会を生活していく中で文字情報はあらゆる場所に存在し、我々に様々な情報を提供してくれる。例えば、店舗の名前や住所、地名などの書かれた看板に存在する文字、また経路情報や交通状況などの道路交通情報の書かれた情報板上に存在する文字などがある。これらの文字情報を視覚機能を保持するシステムが正確に取得し理解することができれば、視覚機能を保持した自律移動ロボットの移動支援や作業支援、また自動車の自動運転支援や運転手に対する情報提供などに応用できる。

しかし、デジタルカメラ等で撮影した画像から文字情報を取得しようとする場合、画像内には文字情報以外の様々なオブジェクト(空、建物、車、人など)が存在する。また画像撮影時の天候などによって文字情報に陰影が掛かってしまう場合などがあり、文字情報のみを正確に取得することを困難にさせる要

因が多々存在する。

そのため画像内に存在する文字情報を正確に取得するためには、まず画像内のどこに文字情報が存在しているかを理解する必要がある。

そこで本研究では種々の状況下において文字情報が存在する情景をデジタルカメラで撮影した画像(情景画像)から文字がいくつか並んで文字列として存在している領域を抽出することを目的としている。

2. 情景画像中からの文字列領域の抽出

情景画像中からの文字列領域抽出の従来手法として、ここでは2つの手法を紹介する。

1つ目は、カラー情景画像から明度値を計算して作成した明度情景画像において、閾値を変化させながら得られる多数の2値画像から各画像における平均隣接数を評価して文字とそれ以

外の部分とを良好に分解している 2 値画像を複数枚選択し、各 2 値画像から文字列を成す文字の連結性を利用して文字列領域を抽出する手法である [1]。2 つ目は、カラー情景画像の正規化 R 及び正規化 G のヒストグラム上から画像内に存在するオブジェクトの分割位置を複数個選択して複数枚の 2 値画像を作成し、各 2 値画像から文字列は一直線に並んでいるという特徴を利用した AMA (Aligning-and-Merging-Analysis) 法を利用して文字列領域を抽出する手法である [2]。

しかし [1] は明度情報のみを利用して 2 値画像を作成するために文字列上に陰影が掛かっている情景画像や文字部分とその背景部分との明度値が似ている画像などにおいては文字とそれ以外の部分とを良好に分解する 2 値画像を作成することが困難となる場合がある。また [2] はカラー情報を利用して 2 値画像を作成するために、画像内の暗い部分はグレースケールに近い画像となりその部分においては文字とそれ以外の部分とを良好に分解する 2 値画像を作成することが困難となる場合がある。さらに [1]、[2] 共に画像全体から明度やカラーの特徴を利用して 2 値画像を作成するために、画像内に存在する文字列が小さい場合などに文字とそれ以外の部分とを良好に分解する 2 値画像を作成することが困難となる。良好な 2 値画像を作成できない場合、その 2 値画像から文字列領域を正確に抽出することは非常に困難となる。

そこで本研究では、カラー情報と明度情報の両者を利用することで文字とそれ以外の部分とを良好に分解した 2 値画像を複数枚作成する。これにより明度情報のみやカラー情報のみでは良好な 2 値画像を作成することが困難である画像に対しても良好な 2 値画像を作成することができる。またカラー情報を利用した 2 値画像作成においては画像内の小さなオブジェクトも選択できるカラークラスタリングを行い、明度情報を利用した 2 値画像作成においては小領域ごとに閾値を決定する。これにより画像内に存在する文字列が小さい場合でも、文字列の情報を見逃すことなく良好な 2 値画像を作成することができる。

このようにして作成された複数枚の 2 値画像から文字列を成す文字の連結性を利用して文字列領域候補を抽出し、それを利用して文字列領域を決定する。

以下、3 において前処理として行う平滑化について述べる。また 4 において 2 値画像作成法について、5 において 2 値画像からの文字列領域抽出法について述べる。そして 6 において提案手法の有効性を実験により検証する。最後に 7 で本研究の結論及び今後の課題について述べる。

3. 前処理 (平滑化)

情景画像は複数のオブジェクトで構成されている。しかし情景画像内で同じオブジェクト上の画素が必ずしも似たような RGB 値を保持しているとは限らない。また情景画像を取得する際に何らかの原因によってノイズが混入してしまう場合もある (図 1(a) 参照)。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) [3] 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。

$$I(t+1) = I(t) + \lambda \sum_D c_D \nabla_D I(t) \quad (1)$$

ここで $I(t)$ は注目画素の RGB 値、 $I(t+1)$ は平滑化後の RGB 値、 D は上下左右の 4 方向、 $\nabla_D I(t)$ は 4 近傍の画素と注目画素の $I(t)$ の差、 c_D は 4 近傍に対する重みを表し、

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

で表される。 $\text{median}(Y)$ は注目画素と 8 近傍画素の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。

c_D の定義から、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 $\nabla_D I(t)$ は $I(t+1)$ にあまり反映されず、差分が小さい画素の $\nabla_D I(t)$ 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 近傍の画素と近い値を保持し、4 近傍の画素の $I(t)$ 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。

実際に図 1(a) の画像に WMFAD を施した場合、局所的なエッジ部分を残したまま平滑化され、同じオブジェクトを示す画素が似たような RGB 値を保持することができ、かつノイズを除去できていること (図 1(b)) がわかる。

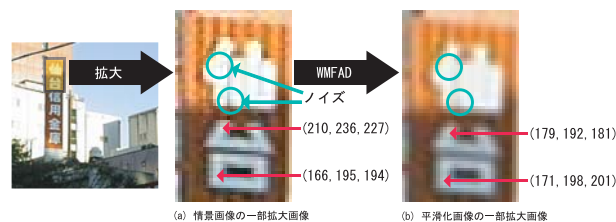


図 1 取得した情景画像の例

4. 2 値画像作成

4.1 カラー情報を利用した 2 値画像作成

平滑化画像に対し、 $L^*a^*b^*$ 空間で代表色を求めカラークラスタリングを行い、クラスタリング画像を作成する。クラスタリング画像において文字というオブジェクトを構成している代表色とそれ以外の代表色とを良好に分解するためにオブジェクトクラスタリングを行い 2 値画像を複数枚作成する。各 2 値画像で境界線の長さによる複雑度を評価基準として文字とそれ以外の部分とを良好に分解している 2 値画像を選択する。

4.1.1 カラークラスタリング

まず平滑化画像中の全画素を $L^*a^*b^*$ 空間に投影する (図 2(a))。投影された画素における $L^*a^*b^*$ の最大値、最小値を求め、それを利用して全画素を囲む直方体を作成する (図 2(b))。次に直方体の $L^*a^*b^*$ 各辺を 30 分割し、直方体の中に 27000 個の Cube を作成する (図 2(c))。各 Cube 内に存在する画素数

をカウントし、周囲 26 個の Cube よりも Cube 内に存在する画素数が多い Cube を選択し、選択された Cube の重心に代表色を配置する (図 2(d))。 (ここで周囲 26 個の Cube とは注目 Cube を中心とする $3 \times 3 \times 3$ の Cube の周囲の 26 個の Cube を指す)。各画素に対し全代表色との $L^*a^*b^*$ 空間内での距離を計算し、最も距離の小さい代表色へと変換する (図 2(e))。代表色の $L^*a^*b^*$ 値を RGB 値へ逆変換しクラスタリング画像を作成する (RGB 空間と $L^*a^*b^*$ 空間での変換式は [4] 参照)。

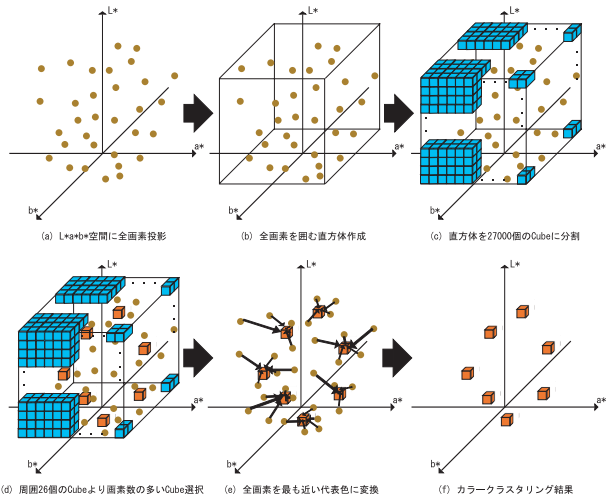


図 2 カラークラスタリング

このようにして作成されたクラスタリング画像を利用して次にオブジェクトクラスタリングを行う。

4.1.2 オブジェクトクラスタリング



図 3 平滑化画像から作成されたクラスタリング画像

照明条件下等の影響により 1 つのオブジェクトが複数の代表色で構成される場合がある (図 3 参照)。そのため文字というオブジェクトを構成している代表色とそれ以外を構成している代表色とを良好に分類するためにオブジェクトクラスタリングを行う。

まず画像の RGB 値を式 (3) によって正規化 RGB 値 (r, g, b) へと変換し、正規化 RGB 空間上でのヒストグラムを作成する。ただし $R + G + B = k$ とする。

$$(r, g, b) = \begin{cases} (0, 0, 0) & \text{if } (k = 0) \\ \left(\frac{R}{k}, \frac{G}{k}, 1 - r - g\right) & \text{otherwise} \end{cases} \quad (3)$$

正規化 RGB 空間は図 4 のように表されるので正規化 R 及び G のみを考慮すればよい。次に作成された正規化 R 及び G のヒストグラム上から高さが 100 画素以上の極大値を全て探索する (図 5(a) 参照)。ここで 100 画素という値は画像内に存在す

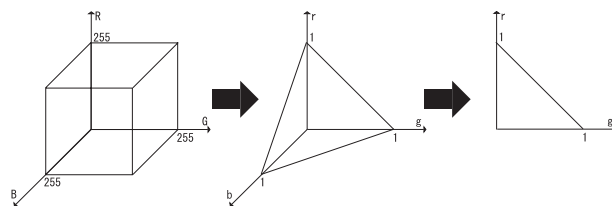


図 4 正規化 RGB 空間

る 1 文字のサイズは 25 画素以上であり、画像内に文字というオブジェクトが 100 画素以上存在すると仮定して設定した。そして正規化 R 及び G のヒストグラム上でオブジェクトの分割位置を隣り合う極大値の平均値に設定し (図 5(b) 参照)、正規化 R 及び G のヒストグラム上の分割位置の組合せ (オブジェクトクラスタ) ごとに 2 値画像を作成する。図 5 の場合、オブジェクトクラスタが 30 個作成されるため 30 枚の 2 値画像が作成される。

また入力画像がグレースケールの場合、正規化 RGB 空間上には特徴が現れないため、良好な 2 値画像を作成することができない。そこでクラスタリング画像から明度値ヒストグラムを作成して同様にオブジェクト分割位置を設定し、オブジェクト分割位置を閾値として 2 値画像を作成する。

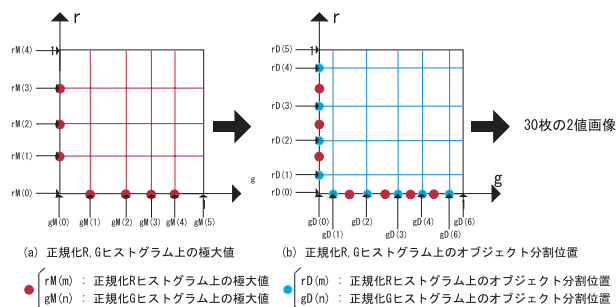


図 5 オブジェクトクラスタリング

4.1.3 複雑度による 2 値画像評価

カラークラスタリングとオブジェクトクラスタリングによって 2 値画像を作成する場合、入力画像によってはオブジェクトがほとんど存在しないような 2 値画像も多く作成される。そこで作成された 2 値画像の境界線の長さから複雑度を計算し、文字とそれ以外とを良好に分解している 2 値画像を選択する。2 値画像の複雑度は以下の式で計算する。

$$complexity = \frac{boundary}{2 \times height \times (width - 1)} \quad (4)$$

ここで $boundary$ は画像中の境界線の長さを 4 連結でカウントしたものの、また $height$ 及び $width$ は画像 1 辺の高さと幅を表す。式 (4) によって計算された複雑度が 0.01 以上である 2 値画像を文字とそれ以外の部分とを良好に分解している 2 値画像であるとして選択する。境界線の長さによる複雑度の閾値 (0.01) は予備実験により設定した。

このようにしてカラー情報を利用して、平滑化画像から文字とそれ以外の部分とを良好に分解している 2 値画像を複数枚作成する。

4.2 明度情報を利用した 2 値画像作成

平滑化画像から明度情景画像を作成する。その明度情景画像を小領域へと分割し、各小領域内の各画素において、周囲領域の情報を利用して画素閾値を決定し、各画素の異なる閾値を用いて 1 枚の 2 値画像を作成する。画素ごとの閾値による 2 値画像にはノイズが発生する場合があるため、簡単な画像補正を行う。最終的に文字とそれ以外の部分とを良好に分解した 2 値画像を 1 枚作成する。

4.2.1 画素閾値決定

(1) 画像分割及び参照領域決定

平滑化画像から明度値 $(L = (R + G + B)/3)$ を計算し、明度情景画像を作成する。それを 4×4 pixel の小領域に分割する。そして各小領域において図 6 のように参照領域を 4 つ決定する。

(2) NAT 法による領域閾値決定

各小領域に対する 4 つの参照領域それぞれにおいて NAT 法 [5] を利用して、参照領域内に存在するオブジェクトを良好に分解する領域閾値を決定する (図 6 参照)。NAT 法 [5] は、オブジェクトが似たような明度値を保持していることからオブジェクトの明度値に近い値を閾値として 2 値化を行うと 2 値画像の領域数が増加するという特徴を利用して、領域数の少ない閾値に決定する方法である。

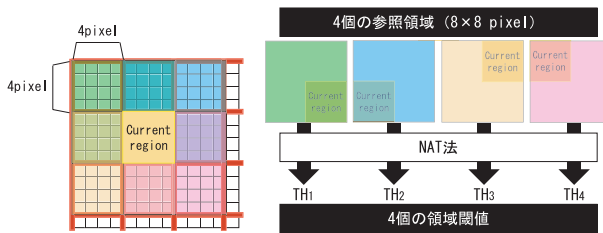


図 6 注目領域の 4 個の参照領域及び領域閾値

(3) 色比較による画素閾値決定

各小領域において定められた 4 個の領域閾値を用いて、小領域に属する各画素に対する画素閾値を決定する。

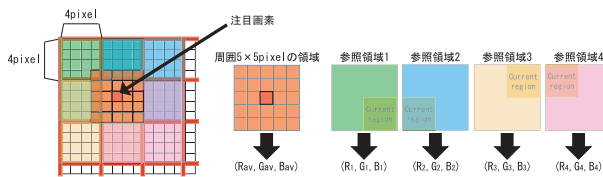


図 7 画素閾値決定法における平均色

図 7 で示されるように注目画素の周囲 5×5 pixel の正規化 RGB 空間上での平均色 (R_{av}, G_{av}, B_{av}) と 4 個の参照領域の正規化 RGB 空間上での平均色 $(R_n, G_n, B_n) (n = 1 \sim 4)$ を求める。 (R_{av}, G_{av}, B_{av}) と $(R_n, G_n, B_n) (n = 1 \sim 4)$ とのユークリッド距離 D_n を求め、 $D_n (n = 1 \sim 4)$ が最小である参照領域の領域閾値 TH_n に対し大きな重みをかけて平均をとり、注目画素の画素閾値を決定する。仮に図 7 の注目画素において、 D_3 が最小であった場合、注目画素の画素閾値 T_{ij} を以下の式

によって決める。

$$T_{ij} = \frac{TH_1 + TH_2 + 1.2 \times TH_3 + TH_4}{4.2} \quad (5)$$

全画素において画素閾値を求め、それより明度値が大きいか小さいかで 2 値化し 1 枚の 2 値画像を作成する。

4.2.2 画像補正

上記で作成された 2 値画像では図 8 のようにオブジェクトを余計に連結してしまうような画素が存在する場合があります。

そこで注目画素に対し上下及び左右の画素値がそれぞれ同じ値を保持し、かつ上下と左右の画素値が異なる画素値を保持している場合、注目画素を無視するという補正を行う。これにより 1 pixel で存在しているような線を排除することができ、オブジェクトを余計に連結してしまっているような画素が削除される (図 8 参照)。



図 8 オブジェクトを余計に連結する画素及び画像補正

このようにして明度情報を利用して、平滑化画像から文字とそれ以外の部分とを良好に分解する 2 値画像を 1 枚作成する。

5. 文字列領域の抽出

5.1 文字列領域候補抽出

カラー情報と明度情報それぞれを利用して作成された複数枚の 2 値画像に対して、文字を表している領域は黒画素または白画素の閉領域で存在しているとし、文字領域候補を抽出する。漢字や平仮名などは 2 つ以上の部分 (偏と旁) から構成される可能性があるため、同じ文字を表していると考えられる領域を統合するマージング処理を行う。以上により抽出された文字領域候補の中から文字列を成す文字の連結性を考慮し、文字列領域候補を抽出する。

5.1.1 文字領域候補・部分領域候補・漢数字候補抽出

各 2 値画像において閉領域を 4 連結で抽出し外接方形を与え、その形状により 4 種類に分類する。

- 文字領域候補：短辺 ≥ 5 かつ 長辺 / 短辺 < 1.6
- 部分領域候補：短辺 ≥ 5 かつ 長辺 / 短辺 < 4.0
- 漢数字候補：長辺 ≥ 5 かつ 方形内画素比率 ≥ 0.8
- 不用：それ以外

5.1.2 マージング処理

偏と旁など分離して存在する文字を 1 つの外接方形で囲むために文字領域候補と部分領域候補に対し以下の 3 種類の条件付きマージング処理を行う。

- 文字領域候補同士 (図 9(a) 参照)

2 つの外接方形が交差しており面積比率が 1.5 未満かつ各辺比率が 1.6 未満かつ 4 辺の内いずれか 1 辺でも 2 画素以内のずれのときマージング

- 文字領域候補と部分領域候補 (図 9(b) 参照)

2つの外接方形が交差または2画素以内に存在し面積比率が3.0かつ1辺の辺比率が1.6未満かつもう1辺の辺比率が2.5未満かつ4辺の内いずれか1辺でも2画素以内のずれのときマーキング

- 部分領域候補同士 (図 9(c) 参照)

2つの外接方形が交差または2画素以内に存在し面積比率が1.5未満かつ短辺の辺比率が1.6未満かつ長辺の辺比率が2.0未満かつ4辺の内いずれか1辺でも2画素以内のずれのときマーキング



(a) 文字領域候補同士 (b) 文字領域候補と部分領域候補 (c) 部分領域候補同士

図 9 3種類の条件付きマーキング処理の例

5.1.3 文字列領域候補抽出及び漢数字候補抽出

文字列を成す文字の連結性として次の5条件を判定し、文字列領域候補を抽出する。

- 2方形は交差していない
- 2方形の面積比率が1.5未満
- 2方形の重心のずれが5画素以内
- 2方形の重心の距離が $3 \times (\text{方形の長さ})$ 以内
- 2方形の縦、横の一辺の長さの比が共に1.2以内

このとき漢数字候補を文字列領域候補の外接方形との重心の位置と平均色の差を条件に抽出する。

このようにして各2値画像において文字列を成す文字の領域である文字列領域候補を抽出する。

5.2 文字列領域抽出

各2値画像で抽出された文字列領域候補を1つの画像に統合し、文字列統合画像を作成する。その画像の水平方向及び垂直方向における文字列領域候補の外接方形の量を示すヒストグラムを作成し、ヒストグラム上で最大値を含む山の範囲を文字列が存在する範囲とし、その範囲内に存在する文字列領域候補を全て囲む新たな方形を作成し、それを文字列領域とする。

このようにして情景画像中の文字列領域を決定し、文字列領域抽出画像を作成する。

6. 実験

種々の状況下における文字情報を含む情景画像100枚に対して実験を行った。文字情報は画像内において目視にて読み取れることを条件とし、それ以外の文字列の位置やサイズ、フォント種などは制限しない。

実験結果を以下の評価基準を利用して評価する。

- 文字列統合画像

文字列領域候補に対し、存在する文字の何割を抽出できたかを表す「再現率」と抽出した方形の何割が正しく文字を囲っているかを表す「適合率」で評価する。

- 文字列領域抽出画像

文字列が1列のみ存在する画像(63枚/100枚)に対し、文

字列領域を正確に抽出した場合は「Perfect」、余計な領域を含むが文字列領域を正確に抽出した場合は「Include」、文字列領域の一部を抽出した場合は「Part」、それ以外を「Miss」の4種類で評価する。

提案手法の有効性を検証するために従来手法との比較を行う。従来手法1として松尾らの手法[1]で2値画像を作成し、従来手法2としてWangらの手法[2]で2値画像を作成する。そして作成された2値画像から提案手法で文字列領域を抽出した結果を利用する。

6.1 実験結果

まず文字列統合画像における提案手法と2つの従来手法の再現率、適合率の結果と文字列領域抽出画像における結果を表1、表2に示す。

表 1 実験結果 (文字列統合画像)

	提案手法	従来手法1	従来手法2
再現率	69.8%	52.4%	62.1%
適合率	53.7%	34.9%	37.6%

表 2 実験結果 (文字列領域抽出画像)

	提案手法	従来手法1	従来手法2
Perfect	24列	7列	19列
Include	15列	3列	13列
Part	7列	9列	12列
Miss	17列	44列	19列

表1から、提案手法において再現率、適合率ともに向上させることができ、文字領域を最も精度良く抽出できたことがわかる。これは従来手法では抽出不可能であった文字も抽出することが可能となり、かつ少ない方形で確実に文字を抽出できたことを示している。また表2から、提案手法において「Perfect」と「Include」の枚数が最も多く、文字列領域を最も精度良く抽出できたことがわかる。これは表1からも明らかなように再現率、適合率ともに最も精度が良いために、文字列領域を最も正確に抽出できたことを示している。

図10から明らかなように、カラー情報を利用した2値画像作成により文字列が小さい画像においても文字列部分を代表色として選択することができ、文字を正確に抽出することができた。また図11から明らかなように、明度情報を利用した2値画像作成により、照明の有無で異なる閾値を与えることができ、文字列上に陰影が掛かってしまっているような画像においても文字を正確に抽出することができた。

7. むすび

カラー情報を利用して種々の状況下における文字情報を含む情景画像中から文字列を成す文字の領域を抽出する手法を提案した。

WMFADを利用した平滑化を行うことで、入力画像内に存在する文字というオブジェクトを表す画素や看板というオブジェクトを表す画素が、それぞれ似たようなRGB値を保持し、何らかの影響により混入したノイズ画素を除去した平滑化画像を作成することができる。

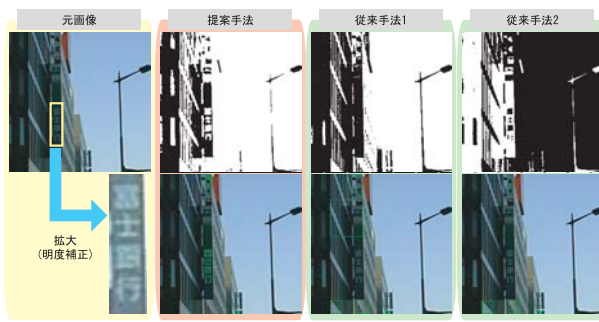


図 10 文字列が小さい画像

(作成された 2 値画像 (上) と文字列統合画像 (下))

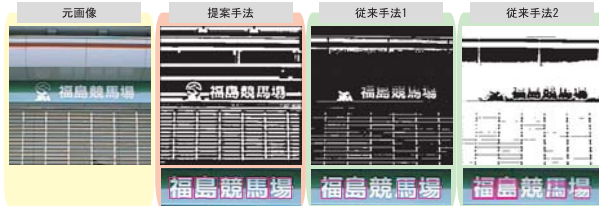


図 11 文字列上に陰影が掛かっている画像の例

(作成された 2 値画像 (上) と文字列統合画像の一部 (下))

またカラー情報を利用した 2 値画像作成では、カラークラスタリングを視覚的な色差が距離となる $L^*a^*b^*$ 空間上で行い、代表色を決定してクラスタリング画像を作成することにより、視覚的に似たような RGB 値を保持する画素が同じ代表色へと変換される。これにより平滑化画像内において、同じオブジェクトを表している似たような RGB 値を保持している画素が同じ代表色へと変換される。またカラークラスタリングにおける代表色を周囲 26 個の Cube よりも Cube 内に存在する画素数が多い Cube の重心に設定するため、画像内に存在するオブジェクトの領域が小さく画素数が少ないようなオブジェクトでも代表色として設定することができる。さらに $L^*a^*b^*$ 空間内で考慮する空間の大きさと Cube のサイズを入力画像によって可変としたために、画像内の全画素が $L^*a^*b^*$ 空間内において狭い分布しか持たない場合でも、代表色を良好に複数個設定することができる。

また明度情報を利用した 2 値画像作成では、明度情景画像を 4×4 の小領域へと分割し、各小領域に 4 個の参照領域を設定する。これにより小領域内に存在する複数のオブジェクトの明度値が似ているために、画像全体を利用して 2 値化を行うと分解することができないオブジェクトも良好に分解することができる。また各小領域に所属する各画素において平均色を利用して画素ごとに異なる画素閾値を設定する。これにより文字列上に陰影が掛かってしまい文字の上部と下部との明度値が異なってしまうような画像においても、各画素の照明の有無で異なる画素閾値を設定することができるため、文字とそれ以外の部分とを良好に分解する 2 値画像を作成することができる。

今回、デジタルカメラで撮影された種々の状況下における文字情報を含むカラー情景画像 100 枚を対象として文字列領域抽出の実験を行った。従来手法と比べて再現率、適合率共に高い評価が得られ、文字列領域の抽出に関しても最も精度よく抽出

することができ、本手法の有効性を示した。

本研究では文字列が直線的に並んでいることを仮定し文字列領域候補を抽出している。そのため曲線的に文字が配置されている文字列や斜めから文字列を撮影したために画像内において文字列が直線的に並んでいない場合、文字列を良好に抽出することができなかった。そのため曲線的や斜めに配置されている文字列をも考慮した文字列領域候補抽出法を考える必要がある。

文字情報を視覚機能を伴ったシステムにおいて利用するためには、文字認識が必要となる。また実際にこれらのシステムに組み込むことを考えると計算時間も実用的な数字を出さなくてはならない。これらの問題は本研究では考慮していないが実際に実用を考えた場合には絶対的に考慮が必要な内容である。

文 献

- [1] 松尾賢一, 上田勝彦, 梅田三千雄: “適応しきい値法を用いた情景画像からの看板文字列領域抽出,” 電子情報通信学会論文誌 D-II, Vol.J80-D-II, No.6, pp.1617-1626, June 1997.
- [2] K.Wang and Jari A.Kangas: “Character location in scene images from digital camera,” Pattern Recognition, 36, pp.2287-2299, 2003.
- [3] M.Gabbouj, P.Haavisto and Y.Neuvo: “Recent advances in median filtering,” E.Arikan(Ed.),Communication,Control, and Signal Processing, Vol.II, Elsevier Science Publishers B, Ankara, Turkey (1990), pp.1080-1094
- [4] Color Conversion Algorithms :
“http://www.cs.rit.edu/ncs/color/tc_convert.html”
- [5] Hon-Son Don: “A noise attribute thresholding method for document image binarization,” International Journal on Document Analysis and Recognition, Vol.4, pp.131-138, 2001