カラー情景画像からの文字列領域の抽出

非会員 松尾 賢一 (奈良高専)
非会員 上田 勝彦 (奈良高専)
正員 梅田 三千雄 (大阪電通大)

Extraction of Characters String Region from Color Scene Image.

Ken-ichi Matsuo, Non-member (Nara National College of Technology)

Katsuhiko Ueda, Non-member (Nara National College of Technology)

Michio Umeda, Member (Osaka Electro-Communication University)

This paper proposes an extraction method of characters string from scene image using HSL transform. In this method, an input color scene image is divided into 15 or 28 binary images by HSL transform. In case of 28 images, 13 images are obtained by dividing on H-S distribution map and 15 are obtained by dividing lightness histogram. Label allotment is carried out in each binary image and rectangle regions are extracted from label images by label allotment. These rectangles are merged by using 3 decision criteria. Candidate rectangles which surround characters are decided by the connectivity in each image and characters string area is extracted by integrating these candidates in all binary images. The experimental result shows that 92.4% of characters included in 100 types of different color scene images are correctly extracted by this method.

キーワード：情景解析，カラー画像処理，文字列領域抽出，H S L変換

1. まえがき

近年、画像処理・理解の研究によって物体の形状記述や移動経路上の障害物の判断等が可能となり、視覚による移動ロボットの実現の可能性が高まりつつある。このとき、物体そのものを認識・理解するだけでなく、その周囲に存在する文字情報を積極的に利用することが可能となれば、移動における有効な情報源となりうることが考えられる。例えば、同じようなドアが並ぶ建物内の一室を探し出す場合、ドアの形状から目的の部屋を見つけ出すことは困難である。しかし、ドア上のプレートの文字を見つけたし、文字の認識と意味の理解を行うことができれば、部屋の使用内容や、目的の部屋であるか否かを正確に知ることができる。つまり、状況によっては物体そのものを認識・理解するよりも、文字情報を利用した方が効率のよい行動を実現できる可能性がある。また、伝票や書籍等の平面上に限定されていた文字認識技術を、看板や標識などの3次元空間上の文字にまで拡張することが可能となれば、車の自動運転の支援、O AやF Aの省力化、交通監視などの幅広い応用が考えられる⁽¹⁾⁽²⁾。

これに対する要素技術として、情景画像に存在する文字領

域を抽出する研究が行われている。これらの研究は、その手法から画像中の文字領域の書式が知識として与えられているものと、そうでないものとに分けることができる。前者には、自動車のナンバープレート上の文字領域抽出⁽³⁾⁽⁴⁾、道路速度標識の文字領域抽出⁽⁵⁾⁽⁶⁾、ドアの部屋番号の文字領域抽出⁽⁷⁾などが挙げられる。また、後者の研究として、隣接画素の濃度差に基づく領域分割により、背景領域と穴領域から文字領域を抽出する方法⁽⁸⁾、文字線とその背景の濃度コントラストが高く、かつ局所的に2値であることや、多くの文字の外接矩形は正方形状であるなどの文字の普遍的な性質を利用して文字領域を抽出する方法^{(9)~(11)}、同じ色領域を定義する等色線を用いて、カラー画像を表現し、その色変化による領域分割線と同色領域の色情報とを利用して、文字領域を抽出する方法⁽¹²⁾⁽¹³⁾、文字列の単色性に着目した文字抽出の手法^{(14)~(16)}、などが報告されている。

しかし、これらの研究では、カラー文書画像のように背景領域もしくは文字領域が均一色であったり、文字認識に適した大きさの文字を含む画像や文字領域に接写された画像などを対象としていたり、あるいはテキストチャやシェーディングが存在する画像において、コントラストが十分でない場合には、必ずしも良好な結果が得られない、などの問題

があった。

これに対して、本研究では、スキャナ入力による印刷物画像や光源の安定した情景画像でなく、カメラ入力した種々の状況下での銀行看板を含むカラー情景画像を対象として、より自然な情景画像に対する文字列領域抽出法を提案し、その有効性を確認する。

従来手法では、2値化した単一の画像から文字領域を抽出するの一般的であった。この方法は、明度ヒストグラムが明確な双峰性を有する場合には有効であり、効果的な2値化が実現できる。しかし、単一の閾値による2値化の手法では、明度ヒストグラムが多峰性となりうる情景画像中の文字に対しては、文字領域と背景領域をうまく分割する閾値を決めることができず、文字抽出に適した2値画像を得ることは困難である。

提案手法では、このような情景画像に対して、単一の閾値で画像全体を2値化するのではなく、明度と色情報を等間隔に分割し、その領域に存在する画素とそれ以外の画素からなる複数枚の分解画像を用いる。これにより、閾値の選択による明度や色彩が均一でない文字の欠落を防ぎ、文字と背景領域のコントラストが十分でない画像においても、明度に依存しない色情報を用いた分解画像を基に、色差から文字領域候補を抽出することができる。テクスチャ状の領域に対しては、閉領域の面積を用いて雑音領域を排除し、文字列の幾何学的特徴と閉領域矩形内の色差を比較することで、テクスチャによって発生する偽の閉領域を排除することが可能となる。

また、単一の明度による2値画像からの文字列領域の抽出とは異なり、複数枚の分解画像ごとに文字列領域候補を抽出し、その抽出画像をすべて重ね合わせて文字列領域候補の重なる頻度を調べることによって、文字列領域の抽出位置における精度と信頼性を向上することができる。

2. 情景画像からの文字列領域の抽出

情景中の文字は、それをとりまく環境や状態によって様々な色彩や明度で存在する。また、文字自身の形状、大きさ、字形、字種なども多種多様である。これらの問題に対処するために、ここでは情景画像を明度・色彩によって複数の分解画像と呼ばれる2値画像に変換して、その画像内に存在する閉領域の外接矩形を抽出し、その形状の幾何学的特徴から、文字領域候補を抽出する。さらに、得られた文字列領域候補に文字列としての普遍的な性質を適用し、最終的に文字列領域を抽出する。つまり、以下の処理手順によって文字列領域を抽出する。

- (1) 分解画像の作成 カラー情景画像を、HSL変換⁽¹⁷⁾により明度と色彩情報に変換する。次に、明度情報から明度ヒストグラムを、色彩情報から彩度色相分布図を作成する。明度ヒストグラムに対しては、明度値を等間隔に8分割し、原画像から8枚の2値画像である明度分解画像を得る。彩度色相分布図では、彩度値を基に有彩色領域と無彩色領域に分け、さらに

有彩色領域については、色相値によって6領域に分割して計7枚の2値画像である色彩分解画像を得る。

- (2) 文字領域候補の抽出 分解画像に存在する閉領域を抽出して、その外接矩形枠を求める。この閉領域の外接矩形の形状が正方形のもの文字領域候補、長方形のもの文字部分領域候補として決定する。また、文字部分領域候補に関しては、文字領域候補の1部分とみなし、2段階のマージング処理によって、外接矩形の統合・包括を施す。
- (3) 文字列領域候補の抽出 文字領域候補に対して、それが文字列を構成している文字であるか否かを文字列の性質を用いて判断し、分解画像ごとに文字列領域候補を抽出する。各画像より得られた文字列領域候補をすべて重ね合わせ、その文字列領域候補の重なる頻度によって文字列領域を抽出する。

3. 分解画像の作成

< 3.1 > HSL変換 CCDカメラにより、情景画像をRGB信号に分解して入力する。このときのRGB信号の値は、0から255までをとり、この値を輝度 (Luminance) と呼ぶ。この輝度信号を直接色情報として扱うと、一般にR値が大きくなるとともにG、B値も相対的に大きくなる。つまり、R、G、B値の間には相関があると考えられるため、無相関に近い状態に変換する必要がある。この代表的な変換にHSL変換がある。このHSL変換により、R、G、Bの輝度信号を明るさを表す明度 (Lightness)、色の違いを表す色相 (Hue)、そして色の鮮やかさを表す彩度 (Saturation) の3つの属性に変換する。つまり、式(1)によって、RGBの輝度信号から色の明るさを表す明度値Lを求める。次に、色の彩かさを表す尺度として、式(2)より彩度値Sを求める。彩度値Sは、値が大きいと色味が増し、逆に小さくなると色味が失われる。このとき、彩度値Sが0で色味のない無彩色となる。色相は、式(3)のように、角度によって色の違いを表現する。

$$L = \frac{(R + G + B)}{3} \quad (1)$$

$$S = \frac{\text{Max}(R, G, B) - \text{Min}(R, G, B)}{\text{Max}(R, G, B)} \times 255 \quad (2)$$

$\text{Max}(R, G, B) = R$ の時

$$H = \frac{G - B}{\text{Max}(R, G, B) - \text{Min}(R, G, B)} \times 60$$

$\text{Max}(R, G, B) = G$ の時

$$H = \left(\frac{B - R}{\text{Max}(R, G, B) - \text{Min}(R, G, B)} + 2 \right) \times 60 \quad (3)$$

$\text{Max}(R, G, B) = B$ の時

$$H = \left(\frac{R - G}{\text{Max}(R, G, B) - \text{Min}(R, G, B)} + 4 \right) \times 60$$

< 3.2 > カラー情景画像の2値化 HSL変換により、情景画像を明度、彩度、色相の3つの特徴に分類する。

これらの各特徴を、画像内の各画素について求め、横軸に明度値 L 、縦軸に横軸の明度値 L に対応した明度値の累積値をとったものを明度ヒストグラムと呼ぶ。また、色相値 H を横軸に、彩度値 S を縦軸の尺度にとったものを彩度色相分布図と呼ぶ。この明度ヒストグラムや彩度色相分布図そのものから、文字領域がどこに分布しているかを判断するのは困難である。しかし、文字のもつ明度および色彩情報は、明度ヒストグラムや彩度色相分布図のある場所に必ず存在するため、明度ヒストグラムと彩度色相分布図を決められた領域ごとに分割し、分割された領域において得られる画像の閉領域のなかに、対象となる文字列領域が存在すると考えられる。

< 3・3 > 等間隔分解画像 明度ヒストグラムを等間隔に 8 分割する。明度値 L は 2 5 6 階調であるから、分割された 1 領域の明度範囲は 3 2 となる。これにより、分割された領域の範囲の明度値が閉領域となる 2 値画像が得られる。次に、色彩情報に対しては、彩度色相分布図上で図 1 に示すように彩度値 3 0 未満を無彩色領域とし、彩度値 3 0 以上では、色相値 0 度から 3 6 0 度までを、6 0 度刻みで分割する。

図 1. 等間隔分解画像における彩度色相分布図上の領域分割
Fig.1. Segmentation of H-S distribution map for generating segregated images.

例えば、入力画像の 1 画素の色相値の範囲が 3 3 0 から 3 0、彩度値の範囲が 3 0 から 2 5 5 とするとき、図 1 の彩度色相分布図上では赤色領域に属するため、赤色領域の分解画像では、入力画像の同じ座標上の画素に 1 (黒領域) が、そうでない場合は 0 (白領域) の画素値を与える。こ

れを入力画素ごとに行ない、各色に属する分解画像を作成する。これによって、各入力画像の画素は 7 領域のいずれか属し、結果的に 7 枚の 2 値画像が得られる。明度値によって作成された 8 枚の 2 値画像と、彩度・色相値によって色分割された 7 枚の 2 値画像を等間隔分解画像と定義する。

< 3・4 > 等間隔交差分解画像 等間隔分解画像では、文字の明度や色相値が図 1 の彩度色相分布図における分割境界部分にまたがると、文字に対応した閉領域内に欠落が生じ、文字領域候補としての抽出が不十分となる。これに対して、等間隔に領域を分割する際に、分割領域に交差するような分解画像を新たに作成することによって、文字領域候補の欠落の防止を図る。これを交差分解画像と定義する。明度値の分割幅は、等間隔分解画像と同じ明度幅 3 2 とし、さらに半分ずつシフトする。これにより、1 5 領域に属する分解画像が得られる。同様に色相値は図 2 で示すように 6 0 度で分割し、3 0 度ずつシフトして、中間色領域である交差分解画像を作成する。

図 2. 交差分解画像における彩度色相分布図上の領域分割
Fig.2. Segmentation of H-S distribution map for generating cross-segregated images.

4. 文字領域候補の抽出

< 4・1 > ラベルリングによる外接矩形抽出 文字や背景などの対象領域を抽出するためには、エッジ強調や、微分処理が一般的に用いられるが、ここではラベル付けされた閉領域の外接矩形を用いる。つまり、明度値と彩度・色相値によって得られた各分解画像に対して、同一の画素値をもつ閉領域にラベルを 8 連結で与えていき同じ領域として統合する。これによって、ラベルによる分解画像の領域分割が行える。この処理と並行して、ラベルによって分割された閉領域の x , y 座標の各最小値と最大値を求める。この x , y 座標の最小値を始点に、最大値を終点とする矩形から、閉領域の外接矩形が得られる。

< 4・2 > 外接矩形からの文字抽出条件 ラベル付け処理によって得られた閉領域が対象となる文字領域である

か、もしくは文字を構成している一部分であるかを外接矩形の形状特徴により判断する。ここでは、文字領域候補と文字部分領域候補の抽出条件を、次のように定義する。

- a. 外接矩形の一辺が画素数 p 以上である。
 - b. 外接矩形の縦横の辺比率が 1.5 未満の矩形は「文字領域候補」とする。
 - c. b において縦横の辺比率が 1.5 以上 10 以下の矩形は「文字部分領域候補」として扱う。
 - d. b または c 以外の矩形は「不用」とし、対象としない。
- 上記の条件により、抽出された外接矩形は「文字領域候補」、「文字部分領域候補」、「不用」に分類される。また、試行実験により画素数を $p=5$ とする。

< 4.3 > マージング処理 マージングとは、ある二つの領域の外接矩形の形状と交差状態に基づき、その領域を統合、包括する処理である。ここでは、矩形の交差状態により、2 段階のマージング処理を行う。一つは、重なった矩形同士のみをマージングする交差矩形マージング処理、もう一つは分離した矩形同士をマージングする分離矩形マージング処理である。交差矩形マージング処理では、仮に矩形 1 と矩形 2 が存在するとき、矩形 1 と矩形 2 が交差または包括状態において、この矩形 1 と矩形 2 の両矩形を囲む新たな矩形を作成する。これにより、交差矩形を一つにすることができる。

一方、分離矩形マージング処理は、偏と旁矩形が分離して存在する文字領域に対して、これを一つの矩形に統合する処理である。通常、偏や旁などの分離している矩形の多くは文字部分領域候補の状態で存在している。この文字部分領域候補を囲む正方形を考え、その正方形同士が交差していれば、文字部分領域候補を囲み得る外接矩形によって統合または包括する。これによって、文字の部分領域である偏や傍の分離を一つにする役割を果たす。

しかし、無条件に交差・分離マージング処理を続けると、文字領域候補の矩形が不用の矩形になったり、矩形が立て続けにマージされるなどの問題が起こる。これに対して、マージング後の矩形の状態と、現在の矩形の状態によりマージング処理の実行を判断する条件を付加する。具体的には、マージング前の二つの外接矩形を α 、 β として矩形 α β を辺比率により、文字を構成している 3 種類の矩形（文字領域候補、文字部分領域候補、不用）に分類する。このとき、3 種類の矩形の分類状態により、表 1 に示すようにどの候補同士が交差しているかを分類しランク分けを行う。ランクは A、B、C の 3 種類とする。

次に、矩形 α と β のマージング後の矩形状態に着目し、表 2 に示すランク分けされた 3 状態とマージング処理後の結果から、矩形 α と矩形 β のマージング処理の実行を決定する。これによって、マージングの実行が決まると、矩形 α のラベル上のある画素の RGB 値を $L_\alpha(R)$ 、 $L_\alpha(G)$ 、 $L_\alpha(B)$ とし、ラベルの面積を P_α として、式 (4) より、ラベル内の平均色 $\overline{L}_\alpha(R)$ 、 $\overline{L}_\alpha(G)$ 、 $\overline{L}_\alpha(B)$ を求める。そして、先に述べたマージング処理で、矩形 α と矩形 β 内のラベルの平均色

の色差 $D(\alpha, \beta)$ を式 (5) で表すとき、この色差 $D(\alpha, \beta) \leq e$ のものについてマージング処理を実行する。

$$\begin{aligned}\overline{L}_\alpha(R) &= \frac{\sum L_\alpha(R)}{P_\alpha} \\ \overline{L}_\alpha(G) &= \frac{\sum L_\alpha(G)}{P_\alpha} \\ \overline{L}_\alpha(B) &= \frac{\sum L_\alpha(B)}{P_\alpha}\end{aligned}\tag{4}$$

$$\begin{aligned}D(\alpha, \beta) &= |\overline{L}_\alpha(R) - \overline{L}_\beta(R)| \\ &+ |\overline{L}_\alpha(G) - \overline{L}_\beta(G)| \\ &+ |\overline{L}_\alpha(B) - \overline{L}_\beta(B)|\end{aligned}\tag{5}$$

色差の有無は、彩度色相分布図上の 2 つの領域にまたがらず、かつ人間が識別が困難である範囲として、本論文では $e=30$ と設定する。これにより、文字領域候補と他の領域矩形との併合を防止することができる。

表 1. 矩形形状によるランク分け
Table1. Ranking based on shape of rectangle.

矩形 α 矩形 β	文字領域 候補	文字部分 領域候補	不 用
文字領域候補	ランク A	ランク B	ランク C
文字部分領域候補	ランク B	ランク A	ランク C
不用	ランク C	ランク C	ランク C

表 2. ランクによるマージングの決定
Table2. Decision of marging by ranks.

統合後の矩形状態 矩形のランク	文字領域 候補	文字部分 領域候補	不 用
ランク A	可	可	不可
ランク B	可	不可	不可
ランク C	不可	不可	不可

5. 文字列候補領域の抽出

< 5.1 > 文字列候補領域の判断 情景画像中の文字部分に着目すると、文字が単独で存在することは希である。むしろ文字は、文になっていたり、文字列として存在する。このことから、文字と文字との連結性を、文字列抽出条件⁽¹⁸⁾⁽¹⁹⁾とすることができる。矩形レベルでは、文字でない雑音も文字領域候補の矩形となるため、図 3 に示す条件をもとに文字の連結性を調べ、文字列領域候補を抽出する。ここで、傾いた文字列や文字列内の文字の形状や色彩が変化するものについて考慮していく必要がある。これに対して、文字列の方向、大きさ、位置に依存しない抽出方法として、Hough 変換を用いた文字列抽出法⁽²⁰⁾や文字配置の規則性を制約条件して、この制約を満たす最小解を得る関数を用いた制約充足型文字列抽出法⁽²¹⁾などが提案されているが、本論文では銀行看板上の文字列を対象とする

ことで複雑な処理は行わず、図3に示す4つの文字の連結性の条件のみで文字列領域候補の判断を行なう手法を用いる。各分解画像で得られた全ての文字領域候補の矩形に対し、4つの文字列の条件が成り立つものを調べ、4つの条件がすべて成立するものを文字列を構成している矩形と見なし、これらの矩形を文字列領域候補と定義して抽出する。

1. 文字領域候補の矩形同士の面積比率が1.5以下である。

2. 一つの文字領域候補の矩形が水平または、垂直移動したと仮定し、その移動の軌跡上にもう一つの矩形の重心が存在する。

3. 文字領域候補の矩形間の距離が矩形3文字分以下である。

4. 文字領域候補の矩形内の対象領域におけるRGB平均値の色差 $D(\alpha, \beta) \leq e$ である。

図3. 文字の連結性に対する条件

Fig.3. Conditions for characters' connection.

<5.2> 漢数字のマーキング処理 漢数字の“一”は、それだけで文字であるが外接矩形の形状による判断では文字部分領域候補となり、文字領域候補レベルまで到達しえない。このような矩形を漢数字領域と定義する。この漢数字領域に対しては、分解画像における閉領域をラベル付け処理の過程で、各領域内の平均色が求まるため、「文字列は、同じ色によって存在する」ことを仮定した色知識による文字列内の文字候補の欠落を防ぐ処理を施す。2段階のマーキング処理が終了し、現時点における文字領域候補により文字列領域候補の範囲を決定する。次に、文字列領域候補の範囲内において、先に述べた漢数字領域の有無を調べる。漢数字領域が存在すると、文字列領域候補の平均色と漢数字領域の平均色との色差を調べる。この時の色差は、先に述べた式(5)の計算と同様であり、文字列領域候

補と漢数字領域の2つの領域内の色差 $D(\alpha, \beta) \leq e$ の時、漢数字領域を文字列領域候補として扱う。

<5.3> 文字列領域の抽出 各分解画像から得られた文字列領域候補から文字列領域を決定する。まず、分解画像から抽出された全ての文字列領域候補を文字列統合画像に書き込む処理を行なう。これにより、文字列統合画像は各分解画像から抽出された文字列領域候補を全て重ね合わせた画像となる。この画像から縦に存在する文字列領域を抽出するために、文字列領域候補の個数を1つとみなしX軸に投影する。このとき、X軸には文字列領域候補が交わった総数値を表すヒストグラムが得られる。これを文字総数ヒストグラムと定義する。この文字総数ヒストグラムは、分解画像上で同じ領域に文字列領域候補が得られるほど、文字総数ヒストグラム値が高くなる。つまり、文字列領域の抽出には、この文字総数ヒストグラムを用いて以下の手法により行なう。まず、画像サイズの一边が m のときのX座標の文字総数ヒストグラム値を H_x とし、文字総数ヒストグラム値が0である総数を H_0 とするとき、平均文字数 M_c を式(6)によって求める。

$$M_c = \frac{\sum_{x=0}^{m-1} H_x}{m - H_0} \quad (6)$$

そして、式(7)より文字総数ヒストグラム値 H_x から、平均文字数 M_c を差し引いた新たな文字総数ヒストグラム H'_x を求める。

$$H'_x = \begin{cases} H_x - M_c & H_x \geq M_c \\ 0 & H_x < M_c \end{cases} \quad (7)$$

この新たな文字総数ヒストグラム H'_x に存在する全てのヒストグラムの山に対し、山の頂点から谷に向かって下るとき、文字総数ヒストグラムのX座標上の極小値 $H_{x_{min}}$ と極大値 $H_{x_{max}}$ に囲まれたX座標の範囲内に、存在する文字列領域候補を全て囲みうる領域が最終的に文字列領域となる。

6. 実験条件

文字列を含む情景画像には、銀行看板を含むカラー情景写真を用いる。これをカラーCCDカメラを使用し、縦256×横256画素、RGB各256階調の100枚のカラー情景画像データとして入力する。抽出対象文字は、銀行看板上の縦の文字列とする。画像上の文字のサイズは未知である。撮影条件は、比較的天候の良い日に行い、撮影場所、銀行看板の位置などは特に限定していない。ただし、対象文字列には、飾り文字や手書き文字は含まれないものとする。領域抽出の判断は、目視によって、原画像の文字領域と文字領域候補の矩形が一致する、また文字列の判断は、原画像の文字列領域と抽出された文字列領域内の文字列領域候補の矩形が一致するかにより行なう。

7. 実験結果及び考察

前述した手法により、カラー情景画像からの文字列領域抽出の実験を行った。情景画像データ100枚を用いた本手

法における文字領域抽出率と文字列領域抽出率を表3に示す。また、この手法による抽出結果例を図4に示す。

表 3. 本手法による領域抽出率
Table3. Results of this method.

手法	領域抽出率	列領域抽出率
等間隔分解画像	88.2%(397 文字)	70%(70 列)
交差分解画像	92.4%(416 文字)	83%(83 列)

全画像に含まれる450文字に対して、等間隔分解画像による抽出文字数は397文字で、文字領域抽出率は88.2%の結果が得られた。また、文字列抽出数は70列で、文字列領域抽出率は70%であった。これに比べ、交差分解画像を用いた抽出結果は、同様に416文字で92.4%の文字領域抽出率、83列で83%の文字列領域抽出率となった。つまり、文字数で22文字、文字領域抽出率で4.2%、文字列領域抽出率で13%の向上がみられた。以上の結果より、等間隔分解画像を交差分解画像にする事によって文字領域抽出率、文字列領域抽出率ともに向上した。この文字列抽出実験において、以下のような結果を得た。

(1) 文字領域候補の抽出において、文字領域の外接矩形の形状の特徴のみで、文字候補を抽出することが可能である。ただし、漢数字や分離文字はこの限りでないが、矩形内の色情報や条件付きマージング処理を導入することで正形状の外接矩形を得ることができ、抽出に関してはほぼ解決できる。しかし、外接矩形の形状のみでは、文字でない領域も抽出されることがある。

(2) 文字と背景領域を確実に分割できる2値化が可能であれば、文字領域候補に4つの文字列としての定義を与えることで、文字列領域候補の抽出が可能である。特に、手書きや飾り文字を含まない限り、高い文字列領域抽出率が得られる。

(3) カラー情景画像100枚を用いた文字列領域抽出実験を行い、交差分解画像において、文字領域抽出率92.4%、文字列領域抽出率81.0%の結果を得た。

(4) 色分解画像より、明度分解画像からの文字領域抽出率が高い結果であるが、情景画像中では、明度情報と比較して色情報の安定性に欠け、入力時の環境や天候、等に大きく左右される。したがって、色情報を安定した状態に変換できる手法が必要である。

(5) 交差分解画像では、明度分解画像には有効であるが、色分解画像での有効性は確認できなかった。これは、銀行看板内の文字領域や背景領域に原色が多く用いられているためと考えられる。

8. むすび

分解画像に基づくカラー情景画像からの文字列領域抽出を行った。その方法は、まずHSL変換によって得られた明度値と彩度色相値を用いて明度ヒストグラムと彩度色相分布図を作成した。この明度ヒストグラムは8等分に領域分割し、彩度色相分布図は彩度値で無彩色と有彩色に分割

し、さらに有彩色に対しては色相値によって6領域に分割した。それぞれの分割領域に属する画素によって得られる複数枚の2値画像を分解画像と定義し、この分解画像内の閉領域を外接矩形の形状から文字領域候補、文字部分領域候補、不用の3種類に分類し、矩形間の交差、包括状態により矩形のマージング処理を施す。この処理で、得られた文字領域候補の矩形に対して、文字列の満たす条件を定義し、それを適用して文字列領域候補を抽出する。CCDカメラより入力された100枚のカラー情景画像を対象に文字列領域抽出実験によってその有効性を示した。

図 4. 分解画像からの文字列領域候補抽出例
Fig.4. Example of characters string extraction from segregated images.

文字認識過程を含まない情景画像からの文字領域抽出においては、文字と背景をうまく分離することが可能であれば、文字列のもつ普遍的特徴のみで高い文字列領域候補の抽出率が得られることを示した。また、各分解画像から抽出された文字列領域候補を重ね合わせ文字列統合画像から、文字列の出現頻度を表す文字総数ヒストグラムを用いて、文字列領域の抽出における精度と信頼性を向上させることができた。本手法では、文字と背景を分離するために明度と色彩に対して、2値画像である分解画像を作成し、その分解画像から文字列抽出を行った。本手法による分解画像の分解範囲は、試行実験によって得られた経験値を用いており、この分解画像の範囲以上に文字や背景領域の明度値や色相値が分布する場合において、文字列の欠落が生じる。この対処法として、各画像で分解範囲を変化させる処理が考えられる。残された課題として、色情報を文字領域抽出過程において、いかにして有効に使用するか、文字、文字列領域を含まない分解画像をどのように棄却するか、背景領域と文字領域の明度および色彩特徴が顕著でない領域の分離などがあげられる。今後は、明度と色彩情報による文字領域、文字列領域抽出率の向上させるための分解画像の作成手法についての検討が必要といえる。

(平成8年3月8日受付,平成8年7月19日再受付)

文 献

- (1) 美濃, "文書画像処理の現状と動向", 信学誌, Vol.76. No.5 pp.502-509 (1993)
- (2) 小川, "パターン認識・理解の新たな展開", 電子情報通信学会
- (3) 三島, 金崎 "画像処理応用車番認識装置の検討", 信学技報, PRU86-94 pp.31-36 (1986)
- (4) 安居院, 崔, 中嶋, 横前 "画像処理を用いたナンバープレート領域の抽出に関する研究", 信学論, Vol.J70-D No.3 pp.560-566 (1987)
- (5) 藪木, 三木 "道路情景画像からの最高速度標識の検出と最高速度の識別", CV76-18 pp.127-133 (1992)
- (6) 藪木, 三木 "道路情景画像中の最高速度標識の検出及び認識", 信学論, Vol.J77-D No.7 pp.1393-1394 (1994)
- (7) 恩田, 笹尾, 青木 "廊下の情景画像からのドアと部屋番号の認識システム", PRU88-145 pp.77-84 (1989)
- (8) 船越, 美濃, 坂井 "情景画像中の文字抽出の検討", 信学総全大, 6-181 (1987)
- (9) 大谷, 塩 "情景画像中からの文字抽出法の検討", PRL85-73 pp.35-46 (1985)
- (10) 大谷, 塩 "情景画像からの文字パターン抽出と認識", 信学論, Vol.J71-D No.6 pp.1037-1047 (1988)
- (11) 塩 "情景中文字の検出のための動的2値化処理法", 信学論, Vol.J71-D No.5 PP.863-873 (1988)
- (12) 上羽, 武田, 岡田 "等色線処理によるカラー画像からの文字領域の抽出", 信学技報, PRU94-28 pp.9-16 (1994)
- (13) 武田, 岡田 "等色線情報に基づくカラー画像の領域分割", MIRU'94 pp.I161-I168 (1994)

- (14) 仙田, 美濃, 池田 "色分類によるカラー文書画像からの文字パターン抽出法", MIRU'94 pp.I75-I82 (1994)
- (15) 仙田, 美濃, 池田 "階層的な色分類によるカラー画像からの文字パターン抽出法", 信学総全大, D-289 pp.297 (1994)
- (16) 仙田, 美濃, 池田 "文字列の単色性に着目したカラー画像からの文字パターン抽出法", 信学技報, PRU94-29 pp.17-24 (1994)
- (17) A.R.Smith, "Color Gamut Transformation Pairs", Computers Graphics, Vol.12, pp.12-19 (1978)
- (18) 松尾, 梅田 "濃淡及び色情報による情景画像からの文字列抽出", 信学技報, PRU92-121 pp.25-32 (1993)
- (19) 松尾, 上田, 梅田 "カラー情景画像からの文字列情報抽出", 情処学全大, 1M-6 (1994)
- (20) 池澤, 隈元 "情景画像からの文字列抽出法", 信学総全大, 7-128 (1994)
- (21) 行天, 馬場口 "制約充足型文字領域抽出の基礎検討", 信学技報, PRU92-119, pp.9-16 (1993)

松尾 賢一

(非会員) 1967年2月9日生まれ。平成5年3月大阪電通大学院修士課程了。同年4月国立奈良工業高等専門学校助手、現在に至る。画像処理、パターン認識、特に文字領域抽出に関する研究に従事。電子情報通信学会、情報処理学会会員。

上田 勝彦

(非会員) 1942年9月2日生まれ。昭和42年神戸大大学院修士課程了。同年国立奈良工業高等専門学校助手。昭和48年同助教授。現在、同情報工学科教授。工学博士。印鑑・署名の自動照合を中心とするパターン認識・画像処理の研究に従事。電子情報通信学会、情報処理学会、計測自動制御学会会員。

梅田 三千雄

(正員) 1945年5月13日生まれ。昭和43年大阪大学卒業。同年日本電信電話公社(現NTT)入社。平成元年大阪電気通信大学工学部教授。現在、同情報工学部教授。工学博士。文字認識、画像処理、認知科学等の研究に従事。電子情報通信学会、情報処理学会、テレビジョン学会、画像電子学会会員。