



Title	心理的ポテンシャルを用いたカラー情景画像からの文字列領域抽出
Author(s)	姉川, 正紀; 中村, 彰; 宮原, 未治
Citation	長崎大学工学部研究報告 Vol.31(57) p.51-54, 2001
Issue Date	2001-07
URL	<a href="http://hdl.handle.net/10069/5156">http://hdl.handle.net/10069/5156</a>
Right	

This document is downloaded at: 2020-10-20T06:52:36Z

# 心理的ポテンシャルを用いた カラー情景画像からの文字列領域抽出

姉川 正紀\*・中村 彰\*\*  
宮原 未治\*

## Extraction of Character String Areas from Color Scenery Image using Psychological Potential Field

by

Masanori ANEGAWA\*, Akira NAKAMURA\*\* and Sueharu MIYAHARA\*

As digital cameras have been used widely, a technology to extract features of character string areas from scenery (nature) image is required. However, it is very difficult to extract character string areas from a complex background of scenery image as compared with a relatively simple background of documentary image.

Human can extract of character string areas from scenery image having complex background. Therefore, a fairly high degree of accuracy in the extraction of character string areas may be attained by imitating the human visual processing system.

In this paper we proposed the method to extract character string areas from color scenery images using the psychological potential field<sup>1),2)</sup> (imitating the human visual processing system). We have also executed the experiment and obtained the favorable processing results that are close to the human sense and unlikely to be affected by amounts of color information and photographing conditions. And, this method have high degree of accuracy in the extraction of character string areas when it is applied to the color scenery image.

### 1. まえがき

近年、デジタルカメラ等の普及により、情景(自然)画像から文字領域等の特徴抽出技術が必要とされている。しかし、背景が比較的単純な文書画像からの文字抽出に比べ、背景が複雑な情景画像からの文字抽出は非常に困難である。

これに対し人間は、複雑な背景が存在する情景画像から、高精度に文字列領域を抽出する事が可能である。そこで、人間の視覚処理系を模倣し、情景画像からの文字列領域の抽出に適用する事で、高精度に文字列領域を抽出が実現出来るものと思われる。

本論文では、人間の視覚処理系を模倣し、心理的ポテンシャル<sup>1),2)</sup>を用いて情景画像から文字列等を抽出する手法を提案する。実験の結果、色情報の多少や撮

影条件に影響され難い、人間の感覚に近い良好な処理結果が得られた。また、本手法をカラー情景画像に適用した場合、より高精度な文字列領域の抽出が可能である。

### 2. 本手法の流れ

本手法ではまず、情景画像を人間の色特性に近いとされる、マンセル色票系に分解しその後の処理をおこなう(図1(a))。次に、ノイズの除去をおこなう。そして、網膜の水平細胞の機能を模倣する為に、局所的複雑度を利用して、特徴強調をおこなう(図1(b))。次に、人間の主観的(感覚的)な処理結果と比較的よく一致するとされる心理的ポテンシャルを用い、特徴抽出をおこなう(図1(c))。これらの処理の後、それ

平成13年4月20日受理

\* 情報システム工学科 (Department of Computer and Information Sciences)

\*\* 崇城大学工学部電子工学科 (Department of Electronic Engineering, Faculty of Engineering, Sojo University)

それぞれの特徴抽出結果を総合的に判断・融合し、文字領域を抽出する (図1(d)).

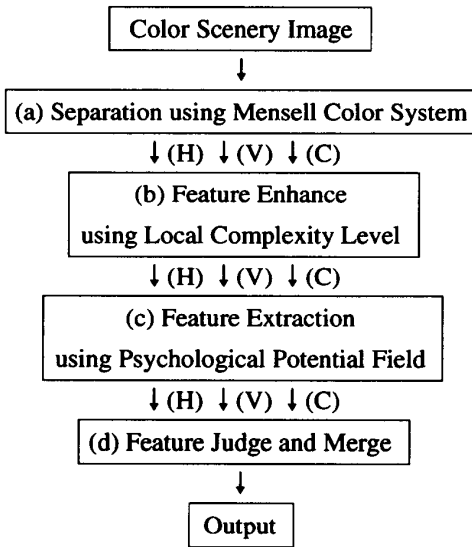


図1 本手法の流れ

### 3. マンセル色票系による分離

人間の色特性は、コンピュータ等で一般に用いられている RGB 色票系ではなく、色相(H)、明度(V)、彩度(C)で表現されるマンセル色票系に近いとされている。

そこで本手法では、入力された情景画像をマンセル色票系に分離し、その後の処理を個別の情報(H,V,C)についておこなう。コントラスト比の大きい画像では、主に明度(V)の情報を用い、色彩豊かな画像では、主に色相(H)や彩度(C)の情報を用いる。グレースケール画像の場合、明度(V)の情報のみを用いる。これによって、色情報の多少や撮影条件の影響を軽減する事が可能である。

マンセル色票系により分離した画像(H,C)を図2に示す。



図2 マンセル色票系による色分離の例

### 4. 局所的複雑度を利用した特長強調

画像の中でエッジ成分は極めて重要な特徴であると考えられる。また、人間の網膜には、エッジの抽出・強調の機能をおこなっているとされる水平細胞が知ら

れている。

まず、本手法ではそれぞれの情報(H,V,C)に対し、ノイズ除去をおこなう。そして、局所的複雑度を利用した特長強調をおこない、水平細胞の機能を模倣する。以下(a)~(c)の手順で、局所的複雑度を定義・測定する<sup>3)</sup>。

(a) 入力画像 (図3(a)) に対し、局所領域毎に画素密度  $D(x,y)$  を測定する。さらに、各画素密度  $D(x,y)$  に対し正規化をおこない、画素密度画像  $D^-$ :図3(b) を得る。

(b) 入力画像 (図3(a)) に対し、局所領域毎に高周波成分の近似として、平均周波数  $F(x,y)$  を測定する。さらに、各平均周波数  $F(x,y)$  に対し正規化をおこない、平均周波数画像  $F^-$ :図3(c) を得る。

(c) 画素密度  $D^-(x,y)$  と平均周波数  $F^-(x,y)$  より、局所的複雑度  $C(x,y)$  を、式1と定義する。さらに、各局所的複雑度  $C(x,y)$  に対して正規化をおこない、局所的複雑度画像  $C^-$ :図3(d) を得る。

それぞれ、正規化をおこなった画素密度、平均周波数と局所的複雑度の関係を図4に示す。

上記の手順で、入力画像から複雑度画像を作成した例を図5(a)に示す。図5(a)に示すように、エッジ成分の多い領域が局所的複雑度が高くなっている事が分かる。また、幾つかの情景画像を統計的に調べてみると、文字列領域は図5(a)に示すように、比較的中程度の局所的複雑度に多く分布する事が、予測実験の結果から分かった。そこで、文字領域が含まれると推測される範囲の局所的複雑度を強調する (図5(b))。この特徴強調は、色相(H)、明度(V)、彩度(C)いづれの情報に対しても特に文字領域を強調する特徴がある。

$$C(x,y) = D^-(x,y) \cdot F^-(x,y) \quad (1)$$



図3 局所的複雑度の一例

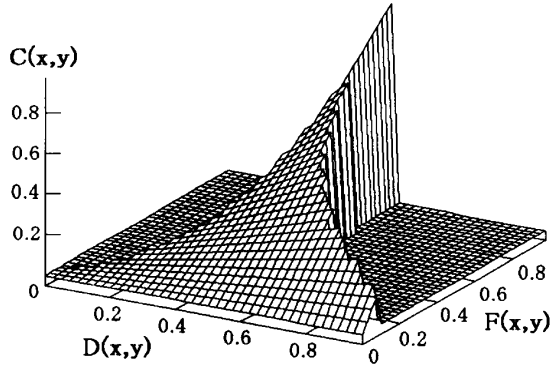


図4 画素密度・周波数と局所的複雑度の関係

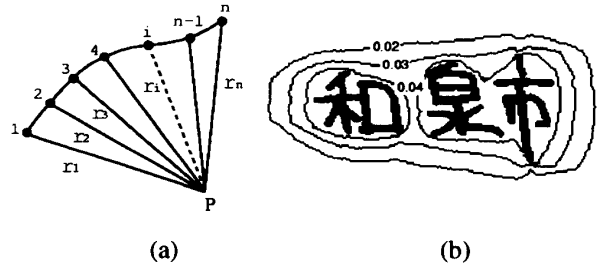
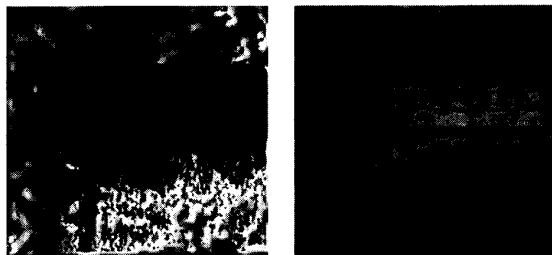


図6 心理的ポテンシャルの一例



(a)局所的画像 (b)特徴強調画像

図5 特徴強調画像の一例

5. 心理的ポテンシャルを利用した特徴抽出

DOG (Difference Of Gaussian) フィルタや LoG (Laplacian Of Gaussian) フィルタは、人間の初期視覚系の処理結果に一致する事が知られている。しかし、その処理結果は、必ずしも人間の主観的（感覚的）な処理結果と一致しない。これに対し、認知科学の分野において、人間の主観的（感覚的）な処理結果と比較的良く一致する心理的ポテンシャル（視覚の誘導場）が提案されている。

そこで、本論文では、心理的ポテンシャルを特徴抽出に用いる。心理的ポテンシャルとは、認知科学の分野において、生理学、心理学現象を基に定義されたポテンシャル場であり、式(2)で定義される。図6に示すように、視覚の誘導場のポテンシャル場は、従来の画像処理的な手法に比べ、人間の主観的（感覚的）な処理に比較的好く一致する。

心理的ポテンシャルを用いて特徴を抽出した結果を図7に示す。

$$P = \frac{1}{n} \sum_{i=1}^n \frac{1}{r_i} \quad (0 \leq P < 1) \quad (2)$$

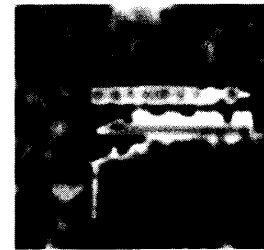


図7 特徴抽出画像の一例

6. 特徴の判断と統合

情景画像中に存在する文字領域は一般に、

- ・文字領域のポテンシャルは高い
- ・文字領域周辺のポテンシャルは平坦である（背景領域が存在する）
- ・文字領域は連続して存在する

等の特徴がある事が、予備実験の結果から分かっている。

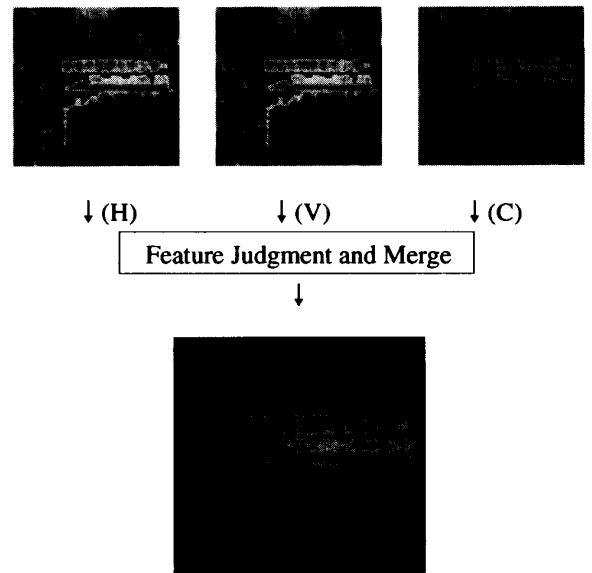


図8 特徴の判断と統合の一例

そこで、本手法では、上記の条件を満たす領域を文字領域候補として抽出する。その後、色相(H)、明度(V)、彩度(C)それぞれの情報を比較し、

- ・いずれかの情報が高い値を示している

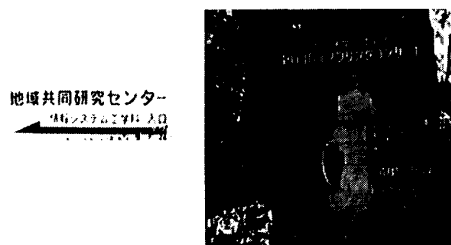
・複数の情報が同一の出力\*1

等の条件を満たした場合、文字領域として抽出をおこなう(\*1: 1つの情報が文字列領域である可能性が高く無い場合、複数の情報を比較して、同一の結果の場合、その結果を文字列領域として出力する)。

図8に特徴の判断と統合の一例を示す。

## 7. 特徴の判断と統合

本手法を情景画像に適用した結果を図9に示す。図9(a)に示すように、人間の感覚に近い文字列領域の抽出が可能である。また、図9(b)に示すように、文字列領域のコントラスト比が低く、明度情報の利用が困難な場合でも、色情報(色相(H), 彩度(C))を有効に活用する事によって、文字列領域の抽出が可能である。



(a) 高コントラスト画像 (b) 低コントラスト画像  
図9 文字列領域抽出の一例

## 8. まとめ

本手法では、人間の視覚処理系を模倣し、心理的ポテンシャルを用いて情景画像から文字領域を抽出する手法を提案した。実験の結果、色情報の多少や撮影条件に影響され難い、人間の感覚に近い良好な処理結果が得られた。

文字列領域抽出の失敗例を調査した結果、そのほとんどがあまりにも小さい文字であったり、文字列領域が小さい、文字列が途中で途切れている等であった(図10)。現在、これらの問題について解決方法を検討中である。

今回、我々は、一般的な表色系であるマンセル表色系を用いた。しかし、今回の提案手法に最適であるかどうかの検証をおこなっていない。そこで現在、我々は、L\*a\*b\*等の他の表色系との比較実験をおこなっている。

また、我々は、以前今回と同様の“心理的ポテンシャル”を用いた手書き漢字文字列からの文字抽出手法を提案している<sup>4)</sup>。今後、今回の手法と組み合わせる

事によって、カラー情景画像からの文字列領域の抽出、文字列領域からの文字抽出が、ほぼ同一のアルゴリズムによって可能であり、カラー情景画像処理システムの構築を検討している。



図10 文字列領域抽出失敗の一例

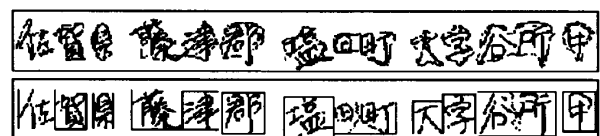


図11 心理的ポテンシャルを用いた文字抽出

## 参考文献

- 1) Zensho YOKOSE: "A Study on Character Patterns Based Upon the Theory of Psychological Potential Field", Japanese Psychological Research, Vol.12, No.1, pp.18-25, 1970
- 2) Michihiro NAGAISHI: "A Consideration on Segmentation of Handwritten Characters using Field of Induction on the Retina", Technical Report of IEICE, Japan, PRU92-120, pp.17-24, Jan.1993
- 3) Masanori ANEGAWA: "A division algorithm of a Chinese character into its component elements using local complexity of image", IEICE Transactions on Information and System, Japan, Vol.J79-D-II, No.1, pp.146-150, Jan 1996
- 4) Masanori ANEGAWA: "Handwritten Chinese Character Segmentation using Local Potential Threshold and Minimum Potential Search", Proceeding of MVA'98 IAPR Workshop on Machine Vision Applications, pp.57-60, 1998
- 5) Masanori ANEGAWA: "Extraction of Character String Areas from Color Scenery Image using Psychological Potential Field", Proceeding of MVA2000 IAPR Workshop on Machine Vision Applications, pp.197-200, 2000