

平成 28 年 5 月 19 日現在

機関番号：11301

研究種目：挑戦的萌芽研究

研究期間：2014～2015

課題番号：26540124

研究課題名(和文)文字領域推定を行わない情景画像中の文字認識アルゴリズムの構築

研究課題名(英文)Scene Text Recognition without Text Region Estimation

研究代表者

和泉 勇治 (Waizumi, Yuji)

東北大学・情報科学研究科・准教授

研究者番号：90333872

交付決定額(研究期間全体)：(直接経費) 2,800,000円

研究成果の概要(和文)：認識処理が可能な限り排除した情景画像中の文字認識システムの構築を目的とした研究課題である。従来の画像中の文字認識システムは、文字が存在する場所の推定と、推定された位置にある画像の字種を決定する二段階の認識を行うことが多く、文字が存在する場所の推定において、文字画像の検出漏れが生じた場合、その位置に存在する文字は一切認識されないという問題があった。そこで本研究課題では、画像中で同じ色合いを持つ領域と判断されたすべての箇所について、文字の一部である可能性を考慮し、すべての領域に対し認識処理を適用することで、文字領域の検出漏れの影響を受けない文字認識アルゴリズムを提案した。

研究成果の概要(英文)：The purpose of this work is building an algorithm for scene text recognition without text region estimation. Almost previous works for scene text recognition consist of two recognition stages, that is text location detection and character recognition. Undetected texts at the text location detection stage are never recognized in character recognition stage. This is a serious drawback of the two stage recognition algorithm. In this work, we propose a new scene text recognition algorithm without text location detection stage to overcome the drawback. The proposed algorithm considers an scene image as a set of closed curves and all closed curves are a part of a letter. By applying score function to all closed curves in a scene image, the proposed method can reduce un-recognition for the undetected regions.

研究分野：パターン認識

キーワード：文字認識 輪郭線 領域推定 情景画像

1. 研究開始当初の背景

デジタルカメラなどで撮影された画像中にある文字を認識する情景画像中の文字検知・認識は、著名な国際会議で毎回コンペティションが行われているように現在でも精力的に研究が推進されている。画像中の文字認識手法の多くは、画像内での文字の存在位置・範囲を推定する文字領域推定と推定領域中の字種を判断する文字認識処理を繋げる構成となっている。このような方式の場合、文字領域推定エラーと文字認識処理エラーの重畳が認識性能を低下させる原因となっている。そこで、本研究では、文字領域推定や特徴点の選択などの文字領域推定を行わない文字認識アルゴリズムの開発を行う。

2. 研究の目的

本研究では、最終的な字種決定以外での認識処理を不要とすることを目指し、画像中の全ての領域を文字の一部であると仮定し、その分布から直接認識結果を得る認識アルゴリズムの設計方式を提案する。その過程で、認識的処理を必要としない特徴量生成・抽出方式、抽出された特徴量の統合・評価による認識アルゴリズム構成方式を明らかにする。認識的処理を必要としない特徴量生成・抽出方式については、画像の領域分割などにより抽出される特徴量を利用し、文字認識の認識性能の観点で適切な手法を明らかにする。抽出された特徴量の統合・評価による認識アルゴリズムにおいては、認識対象の部分欠損などにロバストな識別関数の設計法を確立する。

3. 研究の方法

ICDAR のロバストリーディングデータセットを利用し、画像の領域分割、領域からの形状特徴量の抽出、個々の領域の形状と位置関係の評価関数の構築、その関数の評価値を用いた字種の同定を提案する。

画像の領域分割については、色空間に対する k-means, niblack や大津の方法による二値化などを利用した。色のみを用いた領域分割では、一部の画像で未分割領域、過剰な分割が行われるため、文字の線幅に基づいた領域抽出である Stroke Width Transform も利用し、色に基づいた領域分割結果と組み合わせることでより多くの領域抽出を行い、文字候補となる領域を多く抽出するようにした。抽出された形状を記述する特徴量は二次元ヒストグラムを用いた。

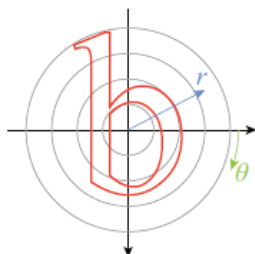


図1：二次元ヒストグラムの座標系

二次元ヒストグラムは図1にあるように、領域の輪郭線に極座標を適用し、動径方向と角度方向を任意の幅に区切り、区切られた個々の領域に存在する画素数をヒストグラムとして数値化したものである。一定の間隔で区切られた領域単位で画素数を算出するため、輪郭線の細かい変動を吸収し、ノイズのロバストとなる長所がある。図2は、図1から抽出された二次元ヒストグラムの例である。このヒストグラムを特徴量として以降の手法で利用する。

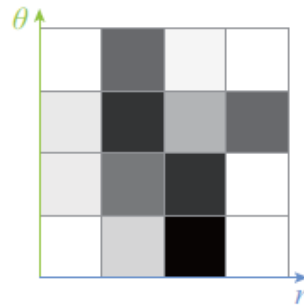


図2：16次元の二次元ヒストグラム

これまでに抽出された領域は、それ一つで文字を構成するものと、複数で一文字を構成するものがある。また、単語を考えた場合、文字の候補領域の周辺には文字としてあるべき形状や大きさの領域が存在しているはずである。この前提を定量化するための特徴量として、着目領域とその周辺に存在する領域の大きさの比と位置関係を次のように定量化する。

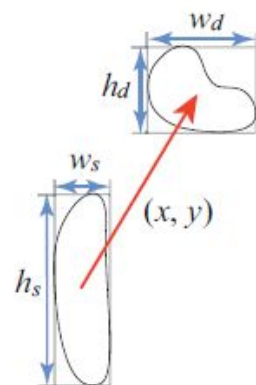


図3：領域間の関係性

図3のように、着目領域の幅、高さを、それぞれ、 w_s, h_s とし、その周辺領域の幅と高さを w_d, h_d 、着目領域の重心を原点とした周辺領域の重心の座標を (x, y) とすると、領域間の関係性を表す特徴量を

$$\left(\frac{w_d}{w_s}, \frac{h_d}{h_s}, \frac{x}{w_s}, \frac{y}{h_s} \right)$$

と定義する．領域の大きさの影響を無くするため，全ての特徴量は着目領域の大きさで規格化している．

着目領域の字種を決定するための評価関数 S を以下のように定義する．

$$S(C_i|x, X) = p(x|C_i) \sum_k \frac{1}{2} (\tanh \{ \alpha [p(y_k, r_k|C_j) - P_T] \})$$

$$= p(x|C_i) \sum_k \frac{1}{2} \left(\tanh \left\{ \alpha \left[\sum_j p(y_k|C_j) q(r_k|C_i, C_j) - P_T \right] \right\} + 1 \right)$$

ここで， x, y, r は，それぞれ，着目領域の二次元ヒストグラム，周辺領域の二次元ヒストグラム，関係性を表す特徴量である．二次元ヒストグラムに対しては，それがある字種 C である尤度を正規分布でモデル化し，関係性を評価する特徴量に関しては，着目領域とその周辺領域が，それぞれ， C_i と C_j ある尤度を利用し評価関数を定義している．この評価関数で利用している尤度は，事前に学習データを用いて学習しておくものである．画像を構成する全ての領域に対しこの評価関数の値を算出し，それに最も高い評価値を与えた字種を該当領域の認識結果とする．

4. 研究成果

ICDAR2003 dataset を利用した認識結果を表 1 に示す．評価基準は F-measure を用いる．

表 1

手法	F-measure
T.Wang et.al	0.67
提案手法(一部画像)	0.66
N.Ezaki et.al	0.62
L.Neumann et.al	0.40
提案手法(全画像)	0.33

提案手法に関しては，適切に領域抽出が行われた場合である一部画像と，適切に二次元ヒストグラムの抽出が行われなかった画像を含む全画像の 2 つの認識結果を示している．これは，本提案手法が従来の領域分割を利用しているため，提案の中心である評価関数の性能のみを適切に評価するために一部画像を利用した F-measure を示している．

提案手法においては，適切な領域分割が行われれば，従来研究と比較しても高い認識性能を示すことが明らかとなった．

本提案手法は，文字領域の推定などの認識的处理を排除し，画像を構成する全ての領域を文字の候補として扱うため，非常に多くの領域を文字として認識してしまう可能性がある．しかし，着目領域の周辺領域との関係性を適切に評価する評価関数を利用することにより，文字の一部では無い領域については，適切に文字では無いという判断が実現し，高い F-measure を獲得することが出来ている．一方，画像を構成する領域を基準に認識処理

を適用しているため，事前の画像処理により抽出されなかった領域は認識の対象とならない問題点がある．



図 4：二値化による領域抽出失敗例

図 4 は，二値化による領域抽出の失敗例である．この画像は，4 列に英単語が並んでいるのが分かるが，二値化した際，背景と間違った連結をしてしまい一文字を独立に抽出することが出来ていない．そのため，ほとんどの領域が文字ではないと判断され非常に低い認識精度となってしまった．この画像は，二値化，k-means などによる減色処理による領域分割などを利用しても，全ての文字を抽出することは出来なかった例でもある．このような場合，多数のテンプレートなどを用いたスライディングウィンドウなどでのマッチングなどにより認識する必要がある．計算量の問題から現実的では無いが，このような手法を効率的に導入した認識処理が必要であることが明らかとなり，今後の課題であると言える．

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 3 件)

Yuji Waizumi, Masako Omachi, Kazuyuki Tanaka, On Demand Calibration for Pedestrian Tracking in Non-overlapping Fields of View, IEEE Internet of Things Journal, 2016 年, 印刷中(査読有)

Yuji Waizumi, Kazuyuki Tanaka, Scene Text Recognition based on Positional Relation between Closed Curves, Proc. of International Conference on Artificial Intelligence, Modeling and Simulation, Vol.1, pp.182-186, 2016 (査読有)

Kazuyuki Tanaka, Shun Kataoka, Munekiyasuda, Yuji Waizumi, Chiou Ting Hsu, Bayesian Image Segmentation by Potts Prior and Loopy Belief Propagation, Journal of The Physical Society of Japan, Vol.83, pp.1-11, 2014(査読有)

〔学会発表〕(計 1 件)

斎藤拓馬, 和泉勇治, 田中和之, 電子情報通信学会 PRMU 研究会, 2015 年 3 月 20 日, 慶應大学理工学部矢上キャンパス

6. 研究組織

(1) 研究代表者

和泉 勇治 (WAIZUMI, YUJI)

東北大学・大学院情報科学研究科・准教授

研究者番号：90333872

(2) 研究分担者

()

研究者番号：

(3) 連携研究者

田中 和之 (TANAKA, KAZUYUKI)

東北大学・大学院情報科学研究科・教授

研究者番号：80217017