

## 画像中文字の認識に関する研究 \*

松藤果穂 (学籍番号 200621338)

研究指導教員 : 小高和己

### 1. はじめに

情景画像から文字を自動的に抽出、認識する事ができれば、看板内容の翻訳やインターネットを介した情報の取得、視覚障害者への補助など様々な分野への応用が可能になる。また、新しいメディアを用いることなく、人類が従来から持っている文字メディアをより自然に、有効に活用する事が可能となる。

情景画像から文字領域を抽出・認識する研究は、情景画像全体から文字領域を抽出・認識する手法<sup>(1)</sup>と、情景画像からユーザが要求した文字領域のみを抽出・認識する手法<sup>(2)(3)</sup>とに大別される。前者は、情景画像中に存在する全ての文字領域を完全自動で抽出・認識することを目的とする手法であり、後者は、ユーザが指定した文字領域のみを認識することを目的とする手法である。後者の手法では、情景画像中に存在する文字領域の一部だけをユーザが必要とする場合は、文字領域に似た特性を持つ非文字領域や、ユーザが必要としない文字領域を予め排除してから文字の抽出・認識を行うことができる。PDA や携帯電話などの携帯端末付属のカメラから取得した画像を用いて看板内容の翻訳などのアプリケーションに利用する場合、情景中の全ての文字領域を抽出・認識の対象とする前者の手法よりも、ユーザが必要とする文字領域だけを対象とする後者の手法の方が適していると考えられる。ここでは後者の立場から、ユーザにとって負担の少ない文字領域指示法を開発し、情景画像中からユーザが指示した文字領域を抽出・認識するシステムを構築することを目的と

する。

本論文では、文字列の始点、終点の2点のみを指示する簡便な手法を提案し、文字抽出し法、文字認識法、実現した文字認識システム、その評価結果について報告する。

### 2. 文字領域の指定

情景画像中の文字領域を指定する方法には、文字領域の正確な指定を要する方式<sup>(2)</sup>と、文字領域をラフに指定する方式<sup>(3)</sup>の二つが挙げられる。前者の方式では主に、認識したい文字列を指定枠内に収めて撮影する方式がとられている。この場合、認識対象の文字列をシステム側に明確に伝えることが可能である。しかし、同じ情景画像内にユーザが認識したい文字領域が複数存在する場合、文字領域の数だけ画像を撮りなおす必要がある。また、指定枠に合わせて画像を正確に撮影する必要があり、必ずしもユーザにとって使い易い方式であるとは言えない。後者の方式では、ユーザに認識して欲しい文字列を自由にマークしてもらう方式がとられている。この方式は、文字領域の正確な指定を要する方式に比べてユーザにかかる負担は少ない。しかし、ユーザの指定を手がかりに「看板領域」<sup>(3)</sup>を推定し、推定された看板領域の中に存在する全ての文字領域を抽出するため、ユーザが看板中の一部の文字領域だけを指定したい場合には、ユーザが必要としない文字領域まで抽出してしまう可能性がある。

ここでは、文字領域指定方式として、ユーザに認識したい文字列の始点と終点を指定してもらう方式を提案する。この方式では、ユーザが文字領域の位置を気にして撮影を行う必要はない。また、情景画像中に認識したい文字領域が複数存在する場合、同じ画像に対してユーザが改めて指定をし直せばよく、新たに画像を撮影

---

\*"Recognition of Characters from Scene Images"  
by Kaho MATSUFUJI

し直す必要も無い。一方、ユーザに認識したい文字列の始点と終点を指定させることで、認識したい文字列の位置、傾き、文字と背景の輝度などの情報をシステム側に通知することができる。

### 3. 実験

#### 3.1 文字認識システム

本研究では、情景画像中からユーザが指定した文字列を自動的に切り出し、認識を行うシステムを実装した(図1参照)。本研究では、携帯電話やPDA等の携帯可能な端末で応用可能なシステムを構築することが最終的な目標であるが、ここではパソコン(Windows XP, 1GBRAM, 2.79 GHz)上に実装した。

#### 3.2 実験方法

実装したシステムを用いて、文字の切り出し精度、文字の認識精度、文字列指定法の操作性を評価するための実験を行った。実験は、(1)ユーザ指定位置の座標データ収集と切り出し、(2)認識精度の評価実験の二つに分けて行った。

##### (1) ユーザ指定位置の座標データ収集

被験者に実際に情景画像中の文字列を指定させ、その座標位置を収集すると共に、文字列指定法に対する意見を収集する。

##### (2) 切り出し・認識精度の評価実験

収集した座標データをもとに、実際に情景画像から文字列の切り出し・認識を行い、文字の切り出し率と認識率を算出する。

また、(1)の座標データ収集を行う際、文字列の指定を「文字と文字の間」に行うよう被験者に指示した。

#### 3.3 実験条件

実験には、看板が含まれる情景画像100枚を用いた。100枚の情景画像中には総計273列の文字列が存在する。実験に参加した被験者は計10人である。画像データの一例を図2に示す。

#### 3.4 実験結果

(1) 切り出し精度 画像中の文字1775文字中(文字列数273列)、ユーザが指定した文字の総数は1044文字であった。うち、正しく切り出せた文字は755文字で、切り出し成功率は約70%であった。

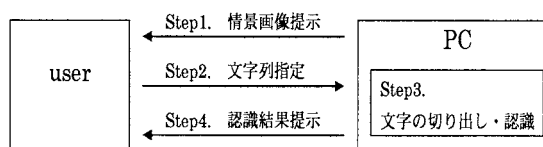


図1 インタフェースの処理手順



図2 画像データの例

(2) 認識精度 正しく切り出された文字について、累積認識率を確認したところ、第1候補で約70%、第5候補までで約80%であった。

(3) システムの操作性 収集実験終了後に被験者に文字列指定方法についての印象を質問したところ、始点と終点の2点で一つの文字列を指定する方法については、被験者10人中7人より、違和感はないとの回答を得た。

### 4. おわりに

本論文では、情景画像中の文字列を始点と終点の2点で指定する簡便な文字列抽出・認識手法を提案した。また、この手法を使用した文字認識システムを実現してその精度を評価した。さらに、被験者に情景画像を提示し、提案した文字列指定方法の操作性を評価した。

### 文献

- [1] 長井隆行, 影広達彦, 金子正秀, 樽松明: 情景画像中の文字及び看板領域の抽出, 信学技報 DSP, 100巻, 719号, p.103-108, 2001.
- [2] 仙田修司, 西山京助, 旭敏之: 携帯カメラによる日本語文字認識の手法と現実, 信学技報, 104巻, 523号, p.7-12, 2004.
- [3] 張暁暉, 長井隆行, 樽松明: ユーザの簡易指定に基づく情景中の文字抽出と認識, 情報処理学会論文誌, 46巻, 10号, p.2556-2565, 2005.