

# 情景画像中の文字の認識

電子制御工学科 志久 修

## 情景画像中の文字認識とは

情景画像とは、デジカメや携帯電話などで、周りの普通の風景をパシャッと撮影した画像のことです。また文字認識とは、キーボードからではなく、文字を撮影するだけで、コンピュータに文字入力する技術です。すなわち、情景画像中の文字認識とは、周りの風景にある文字をパチャッと撮影し、簡単にコンピュータに入力する技術です。

## どういことができるようになるか

キーボードからコンピュータに入力された文字は、いろんな便利な使い方ができます。日本語を英語に翻訳したり、インターネット検索したり、音声として読み上げたりできます。しかし、キーボードからいちいち文字を入力するのは面倒です。携帯電話の小さいキーボードからだ特にそうです。さらに、文字を打ち込むためには、文字を読める必要があります。日本語が読めない外国人が和英翻訳ソフトを持っていても、日本語の入力自体ができません。

情景画像中の文字認識技術を使うと、図1に示すように、道路の看板「佐世保高専」をカメラで撮影するだけで、「佐世保高専」の翻訳、検索などができるようになります。キーボードを使わなくていいし、文字を読めなくてもいいのです。



図1 情景画像の文字認識の応用例

## 技術的な問題点は

文字認識は昔から研究されており、特に新しい技術というわけではありません。しかし、従来の文字認識は、図2のような文書画像を対象としていたのに対し、情景画像の文字の認識では、図3のような画像を対象としなければなりません。すなわち、情景画像では、不安定な照明による文字への影の重畳、斜めから撮影することによる文字の歪み（遠近歪みなど）、ピントずれや手振れによる文字のぼけ、文字以外の背景の写りこみなどが生じています。このような文字に対して、従来の文字認識は、まったく役に立ちません。そのため、新しい文字認識技術の研究開発が、多くの研究機関で行われています。

本稿では、我々が開発してデジタルカメラや携帯電話とは、キーボードを使わずに情景画像中の文字認識では、



図2 文書画像の文字（スキャナ入力）



図3 情景画像の文字（カメラ入力）

## 開発した認識方法

図4に我々が開発している文字認識方法を示します。本方法は、文字成分抽出、文字列抽出、文字傾斜補正、文字認識の4つのステップから構成されます。まず、文字成分抽出では、「文字は細い図形である」と定義し、いろいろな図形が含まれている画像から、細い図形だけを文字成分として抽出します。次に、文字列抽出では、「文字列は、ほぼ同じ大きさの文字成分が密集している長方形領域」と定義し、前ステップで抽出した文字成分を文字列にまとめていきます。このとき、文字列が斜めに並んでいることもあるので、いろいろな方向に並びを調べます。取り出した文字列は、水平に並ぶように補正します。水平に並べた文字列では、本来は直立している文字が傾斜している場合があります。そこで、もともとは垂直（または水平）であるはずのストロークを見つけ、それらを垂直（または水平）に戻すことにより、文字をまっすぐに起こします。最後に、文字列の個々の文字ごとに、文字認識を行い、文字に対応するコードに変換していきます。



図4 開発した文字認識方法

## 応用システム

本方法をもとにいくつかの応用システムを開発しました。それらの例を紹介します。

・情報検索システム：図5は、本棚に並べてある本を撮影し、その本のタイトルを使ってインターネットを検索した例です。ここでは本を例としていますが、街角のお店の看板なら、そのお店のおすすめ情報、動物園の動物のネームプレートなら、その動物に関する生態等が、インターネットから瞬時に検索できるようになります。さらに、認識した文字を音声出力することにより、目の不自由な人に対する情報提供もできるようになります。

・携帯電話による韓日翻訳システム：図6は、携帯電話の画面で、ハングル文字の画像と、その翻訳結果「ご飯」が表示されています。これは、(株)オムロンソフトウェア社で開発されたものです。本方法の文字成分抽出、文字列抽出、文字傾斜補正(図4の、)が採用されています。

本研究の一部は、(独)科学技術振興機構(JST)の平成16年度革新技術開発研究事業「情景画像からの文字情報抽出技術の研究」、オムロンソフトウェア株式会社との共同研究によるものです。



図5 情報検索システム

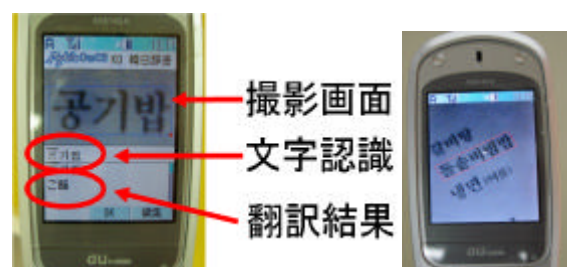


図6 携帯電話による韓日翻訳