

2016年  
6月

## てくのろじい 解体新書

暮らしを便利してくれる  
さまざまな東芝の技術。  
一体どんな技術なのか、  
ニャンダローが先生に質問します！



# カメラで撮影した文字を翻訳 カメラ画像 文字認識技術

スマホやタブレットで撮影した文字を検出・認識して  
その場で翻訳する「カメラ画像文字認識技術」。  
さまざまな場面での活用が期待されるこの技術について、  
井本先生に教えてもらいました。

### スマホやタブレットの 写真の文字を認識して翻訳

ニャンダロー：海外旅行に行ってみたく  
けど、言葉が分からないから、旅行先で  
道に迷ったり、料理の注文に困ったりし  
そうで不安だニャ。最先端の技術で何と  
かニャりませんか？

井本先生：東芝が開発した「カメラ画像文  
字認識技術」が解決してくれるよ。分から  
ない文字をスマホやタブレットで撮影すればそれを翻訳してくれるんだ（図1）。

ニャ：ニャンですか、それ！ 詳しく教  
えてください。

先生：この技術は、3つのデータ処理で  
構成されていてね。まず①文字行検出処  
理で、撮影された画像から文字列を含む  
領域を検出する。そして②文字行認識処

理で、検出された文字列領域に対して文  
字認識を行い、その結果をテキストとし  
て出力するんだ。最後に、③言語解析処  
理でそのテキストに言語解析を行い、意  
味を理解して適切な情報を表示するんだ  
（図2）。

この技術を使えば、日常生活で目に映  
る風景の画像から、文字情報を取得す  
ることができる。スマートフォンやタブ  
レットの内蔵カメラで外国語の文字を写  
すことで、それを母国語で表示できるん  
だ。つまり、外国語で書かれた案内表示  
や看板、料理のメニューなどを日本語で  
読むことができるよ。

ニャ：それは心強いです！ 「カメラで  
写す」というお手軽な操作もいいですね。

先生：東芝では、これまでもOCR（光  
学式文字認識）技術の研究開発に取り組

んでいて、郵便物の宛名を自動で読み取  
る区分機や、帳票の読み取り装置、ドキュ  
メントリーダーなど、さまざまな製品の  
実用化に成功していて、その技術を応用・  
発展させたんだ。

### 悪条件でも文字を認識する 画像特徴抽出技術

ニャ：でも、看板やメニュー表は、いろ  
いろな字体があったり、背景があったり、  
書類の文字を読み取るよりもかなり難し  
いような…。

先生：たしかに、スキャナーで読み取る  
文字画像に比べて、街中にある看板や標  
識、案内板、メニューなどは、さまざま  
な種類のフォントや背景が入り混じって  
いるし、照明環境も違うから輝度の変化  
や影の発生など、文字を認識するのに不



▶今回の先生  
研究開発センター インタラクティブ  
メディアラボラトリー  
井本和範さん  
Kazunori Imoto



● 図1 分からない文字を撮影すると…



タブレットで撮影した文字情報が翻訳される

利な条件がたくさんだね。しかも、撮影する向きによっては文字が変形してしまうこともあって、課題は多かったよ。

ニャ：やっぱりそうですね。

先生：そこで役立ったのが、顔認識や人物検出などの画像認識の研究で培った、東芝独自の画像特徴抽出技術だよ。これらを活用することで新しい文字行検出技術を開発し、複雑な背景の中に文字が書かれている場合や、影で見えにくくなっている場合など、さまざまな要因で検出が難しい文字でも検出できるようになったんだ(図3)。

この技術は、文字認識の分野で世界最高レベルの国際会議「ICDAR」の場でも、学会トップクラスの文字行検出性能と評価されるなど、世界的にもレベルの高い技術であることが証明されているよ。

日本語をはじめ、英語、中国語、韓国語に対応

ニャ：すごいですね！ 逆に日本語の文字を読んで英語に変換することもできるんですか？

先生：もちろんできるよ。現在、英語・中国語・韓国語に対応しているんだ。2020年に開催される東京オリンピック・パラリンピックを控えて、海外から日本を訪れる旅行者がどんどん増えているから、いろいろな場面で役立つと考えているよ。旅行者にとって便利だけでなく、商業施設や交通機関の職員の負担軽減にも貢献できるはずだよ。

ニャ：おもてなしにも一役買うことができそうだな。先生ありがとうございます。

● 図3 文字を検出する技術は世界トップクラス！



複雑な背景の中に文字が書かれていても大丈夫



変形して写っていても検出



影が掛かって見えにくくなっても検出

● 図2 カメラ画像を翻訳する流れ

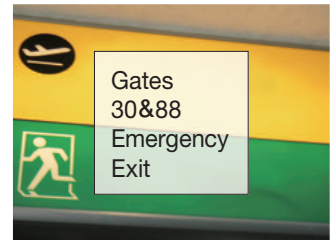
〈カメラ画像〉



画像の中から文字のある範囲を検出



文字を認識してテキストとして出力



テキストを翻訳し、適切な情報を表示

