

光学文字認識

光学文字認識(こうがくもじにんしき、[Optical character recognition](#))は、活字の文書の画像(通常イメージスキャナーで取り込まれる)をコンピュータが編集できる形式(文字コードの列)に変換するソフトウェアである。一般にOCRと略記される。OCRは、[人工知能](#)や[マシンビジョン](#)の研究分野として始まった。研究は続けられているが、OCRの中心はその実装と応用に移っている。紙に印刷された文書をデジタル化し、よりコンパクトな形で記録するのに必要とされる。さらに、文字コードに変換することで機械翻訳や音声合成の入力にも使えるようになり、[テキストマイニング](#)も可能となる。研究分野としては、[パターン認識](#)、[人工知能](#)、[コンピュータビジョン](#)が対応する。

(鏡やレンズといった光学技術を使った)光学文字認識と(スキャナーとアルゴリズムによる)デジタル文字認識は本来別の領域と考えられていた。光学技術として生き残った部分が非常に少ないため、光学文字認識という言葉は現在ではデジタル文字認識を含むものとみなされている。

初期のシステムは特定の書体を読むための「トレーニング」が必要であった(事前にその書体のサンプルを読ませることを意味する)。現在では、ほとんどの書体を高い識字率で変換することが可能である。いくつかのシステムでは読み込まれた画像からそれとほぼ同じになるようフォーマットされた出力(例えばワードプロセッサのファイルのようなもの)を生成することが可能であり、中には画像などの文書以外の部分が含まれていても正しく認識するものもある。

歴史

光学文字認識の原点は、電信技術の拡張と視覚障害者が文字を読むための機械の開発という2つの問題にまつわる活動である[1]。1914年エマニュエル・ゴールドバーグ([英語版](#))は、文字列を読み取り電信符号に変換する機械を開発した[[要出典](#)]。同じころエドモンド・フルニエ・ダルベ([英語版](#))はオプトフォン([英語版](#))という携帯型スキャナを開発した。これを印刷物のページ上ですべらせると、文字の形状によって異なる音を発する。

ゴールドバーグはデータ入力手段としてOCR技術の開発を続けた。後に彼は、画像を光電管で読み取り、必要な識別パターンを含むテンプレートと比較することでマッチングを見つけるという技法を提案した。1929年、グスタフ・タウシェク([英語版](#))も同様のアイデアを思いつき、ドイツでOCRに関する特許を取得した。アメリカでは1933年、ポール・W・ハンデルが同様のテンプレート・マッチング方式のOCR技術の特許を取得している([アメリカ合衆国特許第1,915,993号](#))。1935年、タウシェクもアメリカで特許を取得した([アメリカ合衆国特許第2,026,329号](#))。

1949年、米国復員軍人援護局から視覚障害者支援のためという要請を受け、RCAの技術者らが初期のコンピュータを使ったOCRを研究した。この研究には、単に活字をコンピュータ内の文字コードに変換するだけでなく、その文字を読み上げるという部分も含まれており、初期の音声合成の研究でもある。しかし、コストがかかりすぎることが判明し、研究は挫折した[2]。