講演番号：156

チラシ画像からの商品情報自動抽出－商品情報認識の改善－

D－11　Automatic extraction of goods information from pictures of leaflets
－Improvement of product information recognition－

石井　宏樹
Hiroki Ishii
芝浦工業大学
Shibaura Institute of Technology

高橋　正信
Masanobu Takahashi
システム理工学部
College of Systems Engineering and Science

1. はじめに

折込チラシの情報をデータベース化できれば，商品の旬や最安値などが分かり便利であるが，チラシ情報はテキストデータとして公開されておらず，また OCR ソフトでもほとんど認識できない．チラシを収集し人手でデータ化するサービス[1]もあるが，企業向けで高額である．そこで，チラシ画像から商品名と価格を自動認識してデータベース化する機能の実現を目指す．これまで，価格の自動認識機能（認識成功率99.35%），価格が税込か税抜かを識別する機能を実現した[2]．商品情報については，その候補領域を抽出し，OCR ソフトを利用して商品情報を読み取れる自動認識機能を実現したが[3]，候補領域中の商品情報領域を識別する精度や OCR による文字認識精度が低いという問題があり，精度が不十分であった．そこで商品情報の認識機能の改善を図った．なお，この機能は会社毎に実現するが，まずは埼玉県に多く店舗のあるヤオコーを対象とした．

2. Google Cloud Vision API

精度改善のため，文字認識機能を OCR ソフトから，Google Cloud Vision に変更し，それに合わせ「商品情報の認識機能」を再構築する．Google Cloud Vision API とは Google が提供する画像解析 API で，制限付きで無料で利用可能である．情景画像からの文字認識機能を備えており，従来の OCR より高精度である．取得できる情報は 1 文字毎の「テキストの認識結果」，「文字の位置情報（文字を囲う 4 隅の座標）」である．なお，価格については従来手法[2]の方が高精度なため商品情報についてのみ利用する．

3. 文字の塊をブロック化

Vision API による認識では，別の行は別の文字列として認識されるため，商品情報が複数行にわたる場合，データが断続的になる．そこで，認識結果をブロック化する．

① 文字どうしの横の間隔，文字の中心の y 座標の差の条件によって結合，行内で塊を認識．(図 1(a))．

② 行内の塊どうしの縦の間隔，先頭の文字の x 座標の差の条件によって結合しブロック化する．(図 1(b))．

(a) 行内の塊を囲んだ画像　(b) ブロックを囲んだ画像

図 1 複数行のブロック化

なお，ヤオコーのチラシでは会社名 1 つに対し商品名が箇条書きで列挙されることや同ジャンルの商品がまとめて記載されることなどがあるため，ブロック内に商品情報が複数存在する場合もある．

4. 単語の認識／種別判定および商品情報の作成

チラシ内の商品情報を構成する単語には①「会社名または産地」(以下会社名)，②「商品名」，③「内容量」などがあり，概ね上記の順序で単語が並ぶので，これを利用し以下の処理を各ブロックに対して行う．まず，典型的な場合，文字の高さは会社名と内容量が同程度で，それに比べ商品名が高いことに着目し，k-means 法を用いて高さの高いクラスと低いクラスに分ける．通常，会社名と内容量の間には商品名が存在するので，同クラスの文字が連続している場合を 1 つの単語とすることで会社名，商品名，内容量を分離する．内容量は「数字」+「単位」で構成されるので，その部分を特定し 1 つの単語となるよう修正を加える．商品名は太字で記載されており，かつ内容量より前にあるので，単語の線幅平均を算出し，その値が 1 番大きい内容量より前の単語を商品名とする．会社名は箇条書きの場合を除きほとんどの場合で商品名の直前にあるため，位置で特定する．

一つの商品情報は主に会社名，商品名，内容量から構成されるとしているが，チラシ上には会社名がない商品情報もあるため，商品名，内容量の 2 つを先に識別し，会社名があれば追加する．

5. 実験

ヤオコー公式 HP にあるチラシ画像(2018/01/10 号)を 1 面用いて商品情報の認識実験を行った．結果を表 1 に示す．Vision API により正しく認識された商品情報領域の数は 50 個であり，このうち商品情報の単語が全て正しく認識されたのは 14 個(28.0%)であった．誤認識の要因としてはブロック内の単語の誤分離が挙げられる．その主な原因は，商品名と他の単語の高さの差が僅かな場合があることや，Vision API で認識された文字の高さにばらつきがあることで，高さを用いた分離が適切に行えなかったからと考えている．しかし，ブロック化においては 86.0%という精度で正しく認識することができたため，分離精度の改善により商品情報の抽出精度の向上が期待できる．また，Vision API による認識精度自体が低かったことから，今後は価格との対応付け機能の実現と共に，Vision API の認識精度が上がるように画像を前処理するなどして商品情報の認識精度を向上したい．

表 1 認識実験の精度一覧

① 全ての文字がVision APIで正しく認識されたブロック数／全ブロック数	② ①の正解中のブロック化の正解率	③ ②の正解中の商品情報の正解率			
		会社名	商品名	内容量	商品情報
43／121 (35.5%)	37／43 (86.0%)	17／46 (37.0%)	22／50 (44.0%)	23／50 (46.0%)	14／50 (28.0%)

図 2 認識結果例(青:会社名，赤:商品名，緑:内容量)

[参考文献]
[1] 株式会社ドゥ・ハウス，全国チラシ情報サービスセンター，https://www.chirashiinfo.jp/，2017/06/30 閲覧．
[2] 染谷謙太朗，高橋正信：“チラシ画像からの商品情報自動抽出－価格認識－”，電子情報通信学会東京支部学生会研究発表会，154，2014.
[3] 亀山綾乃，高橋正信：“チラシ画像からの商品情報自動抽出－商品名認識－”，電子情報通信学会東京支部学生会研究発表会，188，2015.