

料理レシピの食材の特売情報抽出システムの考察

T120440 小泉 るう
指導教員 三好 力 教授

1. はじめに

近年、技術の進歩により電気製品が普及し、特に還暦後の女性は、料理研究を趣味とする方が多く、料理番組のレシピをよく見ている。また、高齢者社会のため、高齢者世代だけの家庭が増えています。これらの暮らしに必要な不可欠なのは、食事です。

しかし、料理好きなシニア世代や高齢者単身世帯にとって、料理番組を見ながら夕食のメニューを考える時、作りたい料理を家計的に最安金額で食材を買いたいと考え、地域周辺のスーパーの安売りチラシを自分で一枚ずつチェックし、食材の安売り金額を確認するので、手間暇かかって効率良くはありません。そこで、作りたい料理レシピに必要な食材の金額を地域周辺のスーパーの安売りチラシから自動的に抽出し、表示するシステムの実現を考えました。スーパーのPDFチラシでは、くい込み文字(文字が重なっているフォント)など特殊なフォントが利用され、また背景も複雑なので既存の文字認識では困難です。そのため、本研究では、スーパーのテキストチラシを用いて研究を行います。最終的には、レシピの食材から自動的に地域の最安スーパーを抽出し、最安買い物ルートを表示の実現を目標とするが、今回はこのうちレシピに必要な食材金額を表示する部分の実現を図ります。

2. 提案手法と実験手順

本研究では、テレビ、HP、ネット上のレシピの画像を文字認識し、食材名の認識結果と抽出したスーパーの特売金額をマッチングさせ、周辺のスーパーの特売金額を抽出するシステムを提案します。料理番組やネット配信のレシピ画面をカメラやスクリーンショットで画像に収め、文字認識機能を用いて、レシピ上の食材名を文字認識します。その結果から食材名の文字認識結果部分だけを抽出します。次に、自分が住む地域周辺のスーパーマーケットの公式サイトで、テキストチラシから食材の金額情報を抽出し、mysqlを用いて食材名と金額のデータベースを作成し、最後に、レシピから文字認識した食材名と作成したデータベースのデータと照合し、一致する食材名とその金額を出力するシステムを提案します。この実験手順を以下に示します。

(a)料理番組やネット配信のレシピ画像から食材名が取り出せることを確かめるため、実験用料理番組のレシピ画面を画像に収め、googleドライブのOCR機能を用いてレシピ画像の食材名を文字認識し、その精度を確かめます。(図1)

(b)スーパーのテキストページから取り出した食材名や金額のデータベースを作成できることを確かめるため、実験用に5枚のレシピ画像と特売食材と金額を抽出したいスーパーのテキストチラシのウェブページを解析、mysqlのテーブルにインポートしてデータベースを作成します。(図2)

(c)画面からの食材名でスーパーのウェブテキストページから抽出した食材名とその金額で作成データベースを検索し、金額を抽出できる事を確かめます。(図3)

3. 実験結果

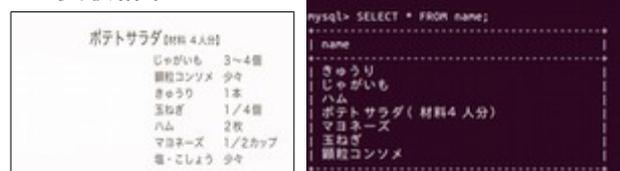


図1 画面上の食材をmysqlにインポート



図2 スーパーの食材と金額をmysqlにインポート

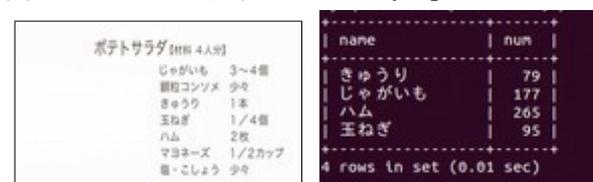


図3 レシピの食材の金額検索

(a)の実験では、正確に文字認識された部分(料理名と材料分を含め)を抽出し、mysqlにインポートすることは95.6%で成功しました。また、(b)の実験では、スーパーの特売食材とその金額を正しく抽出し、mysqlにインポートできた結果が得られました。(c)の実験では、レシピの食材金額を正しく表示できたことが分かります。結果的には、スーパーの特売情報を抽出し、レシピ画像の食材とマッチングさせてその食材と金額を表示するシステムの動作が可能であることが分かりました。

4. おわりに

本実験では、レシピ上の食材名とスーパーのウェブテキストページ上の食材名が漢字やひらかな等の違いがある場合、抽出できてもマッチングしない欠点と考えられます。また、今回は食材の金額抽出、マッチング部分で、カット売りされている食材と一個売りされている食材とレシピ上の食材を正確にマッチングできると精度が向上すると考えられます。