

料理レシピの食材の特卖情報抽出システムの提案及び検討

小泉 るう 三好 力

龍谷大学理工学部

1. はじめに

近年、技術の進歩により電気製品が普及し、人々にとって欠かせない存在になっています。特に現代社会でスマートフォン依存になっている者が多く、自身のすべき事をせず、一日の大部分をスマートフォンの使用に費やしている人や情報入手をスマートフォンのみに頼る若者のテレビ離れやコミュニケーション力の低下が近年問題化されています。その中、人々にスマートフォン以外でも人々の暮らしを快適に、楽しみながら気楽に自分の知りたい情報を入手できる事に気づいてもらうために、本研究では、スマートフォン依存を解消出来る事を目標に、身近な家電製品でも人々に家計に優しく「スマート」を感じてもらい、生活と情報社会が両立できる事を試みます。

そこで、身近にある部分のテレビに着目し、料理番組を楽しみながら、気になるレシピを紹介され、レシピ通りに料理を作りたい時、その食材が自分の住む地域周辺のスーパーにあるのか、値段はいくらか、食材の集計金額の安値段順にスーパーをランキングし、テレビに表示するシステムを提案します。

2. 提案手法

本研究は料理番組のレシピだけでなく、ネット配信中のレシピやネットサイトに載ってあるレシピも画像に収め、レシピ上の食材名を認識し、自分の住む地域周辺のスーパーでその食材の最安金額を表示するシステムの方法を提案します。

2.1 レシピ上の食材名の文字抽出

料理番組やネット配信のレシピ画面をカメラとスクリーンショットで画像に収め、既存技術である Google ドライブの光学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。

wget コマンドでソースコードをテキストファイル保存し、その結果を含めたgoogleドライブウェブページのHTMLソースを解析し、食材名の文字認識結果部分だけを抽出します。

2.2 レシピの食材名とスーパーの食材名のマッチング

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名と金額情報を抽出し、mysql を用いて食材名と金額のデータベースを作成します。

最後に、レシピから文字認識した食材名とmysqlを用いて作成した食材名と金額リストのデータベースと照合し、一致する食材名とその金額を出力します。アルゴリズムのフローチャートを図1に示す

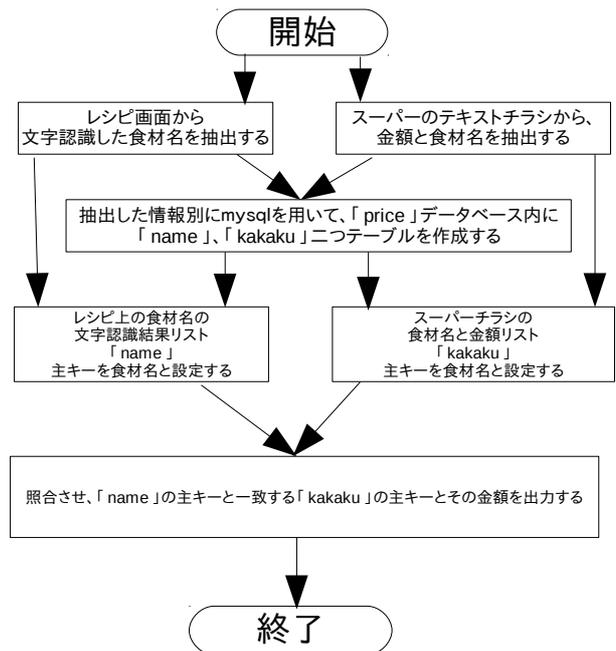


図1. レシピ画面の食材名とスーパーチラシ金額の照合アルゴリズム

3. 実験

従来のクックパッドに掲載されているレシピだけが地域周辺スーパーのお得な情報を自動的にゲットし、金額を表示してくれるシステムと比較し、それ以外な所、料理番組やレシピサイトで見たお気に入りのレシピの食材金額をいつでも表示してくれるシステムの精度を検証します。

Proposal and discussion on the system extracting sales information of the ingredients of the recipe
Ru Koizumi , Tsutomu MIYOSHI ,
Faculty of Science and Technology , Ryukoku University

3.1 実験方法

1料理番組とサイト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、その結果を抽出しテキストファイルに保存します。今回は(おしゃべりクッキングの公式サイト)と(キューピー3分間クッキング)のレシピを利用しました。地域周辺スーパーの公式サイトでテキストチラシから安売り食材の金額を抽出し、テキストファイルに保存します。今回は(イオンネットスーパー)のテキストチラシを利用しました。

食材名の文字認識結果はmysql用いて、手入力で、「price」=データベースを作り、「name」=テーブルにレシピごとの抽出したテキスト食材名を登録し、主キーは食材名とします。

スーパーの安売り食材金額はmysqlを用いて、手入力で、「price」=データベース内に「kakaku」=テーブルに抽出した地域周辺のスーパーのテキスト食材金額リスト表を予め登録し、主キーは食材名(name)とします。

そして、レシピ毎ごとに抽出した食材名テキストファイルをそのまま「name」テーブルに更新し、主キーである食材を特定情報として、「kakaku」テーブルにある食材と照合し、主キーが一致すれば食材名とその金額を抽出します。

上の手順に、レシピごとに「name」テーブルに登録した食材名の金額を「kakaku」テーブルから抽出し、コマンド画面に出力します。

その結果と元のレシピ画像を見比べ、その精度を検証します。



図2. ネットサイト掲載の元のレシピ画像



図3. 食材名文字認識の結果

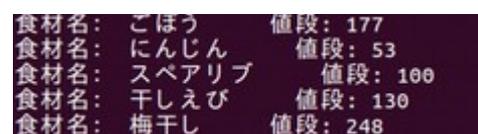


図4. 食材金額の表示

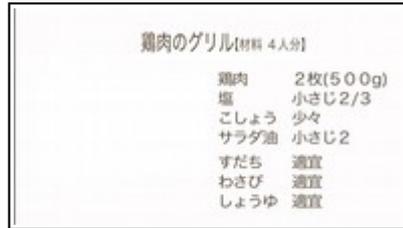


図5. 料理番組の元のレシピ画像



図6. 食材名文字認識の結果

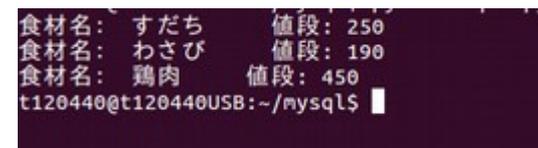


図7. 食材金額の表示

3.2 実験結果

元のレシピ画像図2と図5上の食材名を文字認識し、その結果を抽出し、地域スーパーのテキストチラシから抽出してmysqlに登録した食材名と金額リストとマッチングした結果、正確に文字認識できた食材名は全部マッチングに成功し、金額を表示することが出来ました。この実験により、クックパットに掲載されていない、それ以外な所、料理番組やレシピサイトで見たとお気に入りのレシピの食材金額をいつでも表示してくれるシステムを提案できました。

4. まとめ

本研究では、ネットサイト掲載や料理番組などを画像に取得出来るレシピ上の食材金額を自分が住む地域周辺のスーパーから抽出し、表示するシステムを提案しました。これは、レシピ画像上の食材名を既存技術 googleドライブのOCR 技術を用いて文字認識した文字を抽出し、そして、抽出したレシピの食材名とスーパーの食材名と金額をマッチングさせる手法で実験を行いました。実験の結果、食材名の金額を抽出し、表示することができました。しかし、マッチングする用に予め地域周辺のスーパーの一部の食材しか載せていないテキストチラシから食材名とその金額を抽出するため、レシピ上にマッチングできない食材があるので、レシピごとの集計金額を表示することができません。

今後の課題としては、瞬時に更新されるスーパーの食材が大量載せられてある PDF チラシから金額情報を抽出し、レシピとマッチングさせ、集計金額を表示方法の検討などがあります。