

中国のマイクロブログ (Weibo) データ解析による中国人観光客向けの注文サポートアプリ開発

情報科学科 孫 イク

指導教員：小栗 宏次

1 はじめに

訪日観光客の消費は日本に大きな経済効果を与えているが、その満足度を高めること、特に日本食についてのインバウンド対策が重要である。訪日観光客は食文化の違いによる好みの違いがあり、飲食店での注文時に客観的な情報(甘味, 辛味等)だけでなく主観的な情報(美味しい, まずい)が不足[1]しており不便を感じている。そこで本研究では、ユーザ(中国人観光客)がスマートフォンで飲食店のメニューの料理名を撮影するだけでその料理に関する主観的な情報を提示する注文サポートシステムを提案する。特に、中国のマイクロブログ(Weibo)に投稿されたつぶやきからリアルタイムにサンプリング抽出し、中国人特有の好みを反映させた主観情報を提供できるようにする。

2 注文サポートアプリと料理の主観情報収集

2.1 システム全体像

対象ユーザは日本語が全く分からない中国人観光客とする。本研究で提案するアプリの概要を図1に示す。ユーザはアプリ内のカメラ機能で料理名を撮影する程度の操作だけで、その料理を飲食した経験のある中国人の主観(感想・印象)情報を知ることができる。アプリのレスポンスを重視して、代表的な日本食に関する主観情報は事前に構築したデータベース(DB)を検索して取得する。一方、知名度の低い料理や情報鮮度の高い料理の主観情報は事前に構築することが困難なこともあるため、Weiboに投稿された中から複数件サンプリング抽出し、ポジティブ・ネガティブの主観の割合をリアルタイムに集計してユーザに提供する。

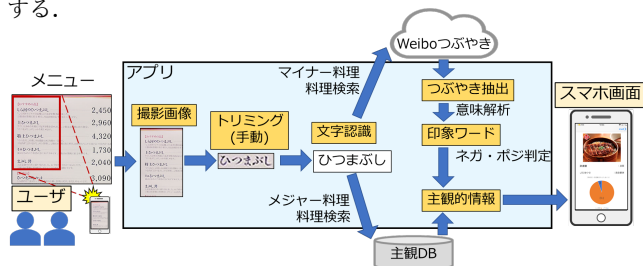


図1 アプリ概要

2.2 料理名の文字認識と代表的な日本料理の主観情報検索

ユーザはスマホで撮影した料理メニューの画像から知りたい料理名の部分をトリミングする。トリミングされた文字列画像に対してオープンソースのOCRライブラリ(Swifty Tesseract)を用いて料理名を文字情報として取得する。それが代表的な日本料理であれば事前に構築したDBを検索する。代表的な日本料理はWeibo投稿量が多いため事前にデータベースを構築しておく方がスマホアプリのレスポンス良くなる。主観情報をDBにするためには、スクレイピングによりWeibo情報を抽出し、ポジティブとネガティブの感情判定を行う。

2.3 Weiboからの主観情報の自動抽出

DBの事前構築時およびDBに登録されていない料理名が文字認識された時に対応するために、リアルタイムにWeiboから主観情報を自動抽出する方法を述べる。先行研究では「悲しい-う

れしい」及び「怒る-喜ぶ」の2つの感情尺度でWebニュース記事のテキスト感情分析をしている[2]。本研究ではそれを応用して「美味しい-まずい(中国語)」の感情尺度でそのユーザの主観をポジティブ・ネガティブに二値化し、一般的な中国人がその料理に持つ主観とする。ただし、Weiboでは全つぶやき情報から任意の文字を含むつぶやきを抽出するAPIは提供されていない。そこで、作成するアプリではWeiboのモバイルサイトにアクセスしてJsonデータを取得し、Jsonの階層構造からその料理に関するつぶやき情報を得る。そこから感情に関係する単語を抽出し、その統計量からポジティブ・ネガティブの割合を算出した。得られた主観情報の例を表1に示す。

表1 料理の主観情報

知名度	料理名	抽出件数	ポジ/ネガ割合
Major	ひつまぶし	520	143/8
Major	納豆	1033	89/23
Minor	馬刺身	673	14/123
Minor	白子	616	12/8
Latest	辛味チキン	10	4/0
Latest	カキフライ	10	1/0

3 システムの有効性検証実験

3.1 文字認識精度

文字認識対象が日本語のため、二種類の日本語学習データを用いて認識精度を比較した。100枚のメニュー写真(縦書き、手書き等除く)で検証を行った結果を表2に示す。識別精度はトリミング範囲や背景の色などに影響を受けるが、この結果からより精度の高い tessdata fast をアプリに実装することとした。

表2 料理名の文字認識精度

学習データ	特徴	正解率	平均計算時間
tessdata best	正確性重視	78 %	0.278sec
tessdata fast	速度重視	89 %	0.108sec

3.2 アプリを利用したユーザの評価

実装したアプリを中国人観光客3名、留学生3名に使ってもらい有効性を検証した。被験者は日本語習得レベルが高い人だけでなく低い人も対象にした。アンケート結果より、日本語が全く分からない・日本語習得レベルが低い中国人からアプリによるサポートについて高い評価がなされた。ただし、UIデザインの工夫や「美味しい-まずい」の二値で主観を表すことへの不満等のコメントがあった。

4 終わりに

中国人観光客向けに料理の主観的情報を提示する注文サポートアプリを提案した。今後の課題は、テキスト感情分析の際の主観情報の種類の増加や分析精度の向上、手書き等の様々なメニュー表記の文字認識精度向上が挙げられる。

参考文献

- [1] 観光庁, “訪日プロモーションの最近の動向”, 2018
- [2] 熊本他, 情処論, Vol.165, pp.15-20, 2005