

地域サイト及びブログからの観光情報の自動抽出と融合

遠藤 雅樹[†] 大野 成義[†]

[†] 職業能力開発総合大学校 〒 187-0035 東京都小平市小川西町 2-31-1

E-mail: [†]{endo,ohno}@uitec.ac.jp

あらまし Web の発展に伴い、観光に関するコンテンツも Web 上に多く情報発信され、旅行者が旅行計画時や旅先において、観光情報を Web から取得する機会も増加している。Web 上の観光情報は、検索エンジンを利用したキーワード検索結果の参照先から得られた情報を整理・分析し取得する。しかし、観光情報を収集し整理する処理は、システム化されておらず、検索を行った旅行者の情報検索能力や情報整理能力に依存している。そこで、本研究では、自治体などが提供する地域サイトと、個人が観光に関して記事を記述したブログに着目し、これらの観光情報を自動収集し、観光情報とレビューを対応付けるなど整理・分析し融合する手法を提案する。

キーワード 観光情報, 地域サイト, ブログ, 情報抽出, 情報融合

1. はじめに

近年、各地方自治体において観光は基幹産業として重要な施策と位置づけられている。観光は、観光業だけでなく、商工業・農林水産業など幅広い分野の地域経済への波及効果がある総合的な産業であり、観光振興に関する様々な取り組みが行われている [1]。観光客の目的も変化し、観るだけから参加・体験、交流、学習など目的が多様化している。また、旅行形態も団体旅行が減少し、個人・友人などの小グループ化が進んでいる。観光客の志向の変化に伴い、観光情報の発信も、都市部の旅行会社などが企画する発地型から、消費者志向の多様化に伴い現地の人しか知らない穴場や楽しみが求められるなど、より詳しい情報を持つ観光地側で情報発信をする着地型の観光情報が注目をされるようになってきている。

観光スタイルや観光情報発信の変化に伴い Web 上の観光情報が増加している。このため、観光情報収集についても、Web を利用する割合が大幅に増え一般的な手段となった。また、併せて Web における観光情報の提供や分析に関する様々な研究やシステム開発が行われている [2]。観光情報には、地方自治体・観光協会などが提供する観光ポータルサイト (以下、地域サイト) と個人が旅先での感想などを記述しているブログサイトなど (以下、ブログ) がある。観光情報提供側は、地域サイトに観光スポットやイベント情報など客観的な事実情報を記述し更新を行い情報提供している。旅行者は、旅先の感想などの情報を個人のブログなどに公開している。そして、観光計画を立てようとする人や旅先で携帯端末を用いて情報収集したい旅行者は、地域サイトとブログから自分の観光計画に必要な情報収集を行い参考となる観光情報を取得する。しかし、Web 上の膨大な情報をすべて人手によってチェックすることは現実的でない。また、地域サイトとブログを結びつけて検討する作業にも多くの時間を費やす必要があるがシステム化されていない。そのた

め観光情報収集は、旅行者の感性や情報収集力、集めた観光情報の整理能力に依存している。

そこで、本研究では、Web 上の地域サイトとブログに点在する観光情報に着目し、システムティックに観光情報を収集し分析する。そして、その結果を融合し提供する手法を提案する。本研究において観光情報は、自治体の他、公の組織である観光協会などが提供する地域サイトと個人が観光や旅行に関わらずあらゆるジャンルの情報を公開している一般ブログ内にある対象地域の内容が書かれているブログに限定する。これらの観光情報は、限定したとしても情報量が多く人手によって観光情報を収集・分析し関連情報として融合するには限界がある。本研究では、限定した地域サイトの観光スポットやイベント情報などの観光情報と対象の観光情報に対する評価や感想などをブログから得た情報を分析し融合する手法を提案する。融合した観光情報を旅行者が利用することでよりニーズに合う観光情報取得が容易になり、観光振興の一助となることを目的に研究を行うこととした。

2. システム概要

一般的に人が旅行をする際には、事前に旅行計画を立てたり、旅行先で観光地について調べたりと、いろいろな観光情報の収集を行う。Web を利用した観光情報の収集を行う場合には、観光地の地名や名称から既存の検索サイトを利用し、観光地の情報について検索する。検索の結果、自治体や観光協会の提供するポータルサイトや旅行代理店のサイトを参照して、観光地の住所や連絡先、アクセス方法などの情報収集を行う。この作業は、Web 以外の旅行雑誌やパンフレットなどを参照する場合も考えられる。Web や紙媒体を利用して情報収集を行うことで、目的の地名や名称だけでなく周辺の観光に関係する様々な地名や観光名所についての情報も入手することができる。このようにして、観光を行う候補となる名所や施設について情報収集を

行っている。更に、収集した情報について他の旅行者や地元の住民が紹介している口コミや評判などを収集し、より具体的に旅行する場所を絞り込む作業を行ったりする。この情報収集は、旅行雑誌の編集者が記述したコメントなどから得ることも可能である。しかし、より多くの情報を取得したい場合は、Web を利用することが多い。このとき、個人が旅先での感想などを記述しているブログなどを利用して情報収集を行う。このポータルサイトの情報と口コミ情報を検索する手順を、既存の検索エンジンを利用し繰り返し行うことが多い。何度も検索を行うことで、多くの観光情報が収集可能だが効率は悪い。また、収集した情報は、人が頭の中で融合させたりメモを取ってまとめることによって、実際に訪れる観光地を決める際の意思決定に役立てることができる。しかし、観光情報の収集を行うには多数のサイトを行き来し必要な情報を関連付けてまとめる作業の多くは人手によって行わなければならない、システム化されていない。

そこで、本研究では人が Web を利用して観光情報収集を行う際の観光情報収集、情報の関連付け、提示の一連の作業をシステム化することとした。システムの概要を図 1 に示す。システムは、まず観光地のある自治体や観光協会の提供する地域サイトから観光情報を抽出する。地域サイトから抽出を行う情報は、観光地の名所や施設についての概要や住所、アクセス方法などの客観的な情報である。一般的に地域サイトの情報は信頼性が高いため、地域の観光名所や施設の多くを収集できると考え収集対象とした。次に、地域サイトから抽出した観光情報に関連のある口コミや評価をブログから抽出する。ブログなどから得られる情報は、公式サイトと比較すると信頼性は劣るが、多くの人の意見が記述されているため観光に役立つ情報が多く含まれている。ブログなどの口コミや感想などを手がかりに実際に行ってみたいと考える情報を得られる可能性があり、観光を行う場所を決める際の目安となる。地域サイトとブログから情報収集後は、地域サイトとブログからそれぞれ抽出した観光情報をワードの統計情報、更にはリンク関係によるカテゴリ分類やタイムスタンプ情報により、最近の話題となる情報などを含めて関連付けを行う。地域サイトとブログの情報の関連付けを行うことによって、各ページを記述内容によって結びつけることができる。関連付けの作業は、人が行うには、頭の中で情報を融合させる作業やメモを取ってまとめる作業が必要である。情報収集に比べ煩雑な部分であるが、本研究では観光情報の融合についても自動化することとした。地域サイトとブログを関連付けてまとめた観光情報を整理し、再構成を行いユーザーに提示することで、システムを利用するユーザーに対して効果的な観光情報を提供できる可能性がある。このシステム化によって、観光情報の収集と関連付けの自動化に加え、観光情報の収集を行う個人が感覚で行っている作業をサポートする手法を提案する。

3. 関連研究

本研究では、Web 上に存在する観光情報を、地域サイト及びブログから抽出し融合させ観光情報取得の一助となる手法を提

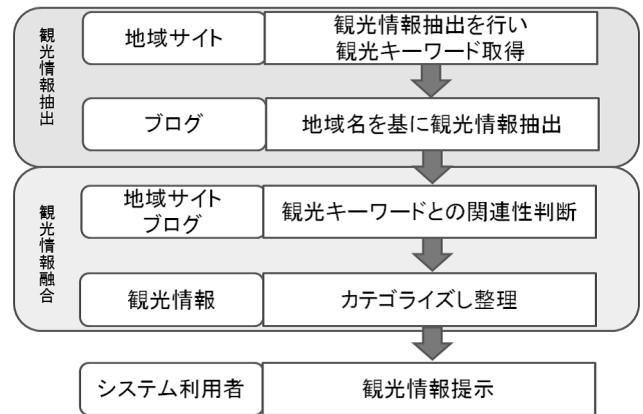


図 1 システム処理手順

Fig. 1 System processing procedure.

案している。

大槻ら [3] は、自治体が提供する地域情報サイトと、そのリンク先の地域サイトを対象としてウェブディレクトリを自動編集するシステムを提案している。地域情報サイトを対象とした情報収集を行う点は本研究と同様であるが、本研究とは地域情報サイトの情報の扱い方も異なれば、それに加えブログの情報と融合させる点でも異なる。

嶋田ら [4] は、地方都市を対象とした観光情報の抽出および分析ツールの構築に向けた手法を提案している。この提案手法は、地方都市の観光情報についての Web の意見情報と自治体などが提供するポータルサイトとの対応付けについての議論と P/N 判定を題材とした観光情報の分析を行っている。地方都市における意見情報と観光情報の対応付けを行う点においては本研究と同様であるが、嶋田らは意見情報の収集に Twitter の情報を利用しているのに対して、本研究ではブログを対象としている点と更には P/N 分析でなく意見などの情報抽出を行う点で異なる。

石野ら [5] は、観光ブログエントリから自動的に観光情報を収集する手法を提案している。この提案手法は、ブログデータベースから旅行ブログエントリを検出し、土産物情報・観光名所情報を抽出し、旅行ブログエントリからリンクを抽出することで、観光情報リンク集を構築するものである。さらに、石野ら [6] は、観光ブログエントリから収集したリンク集を広島電鉄の路線図にマッピングし旅行者に提示する観光情報提示システムを構築している。観光情報をブログエントリから収集する点においては本研究と同様であるが、本研究は一般的なリンク集でなく、自治体などが提供するポータルサイトと個人が提供するブログをマッチングさせたリンク集を構築するという点で異なる。

寺西ら [7] は、観光ガイドブックのページをカテゴリに分類することで構造化し、旅行ブログエントリと質疑応答コンテンツの対応付けを行っている。対応付けという点では本研究と同様であるが、本研究では地名と組織名だけでなく、より包括的な情報融合を行い、動的に変化する地域サイトを利用する点で異なる。

以上の研究は、観光情報の分析や情報収集を行っている。本研究は、地域サイトとブログという異質なデータから観光情報を抽出し、分析・融合してユーザに有益な情報を提供することに特徴がある。また、事前予約語などを極力利用せず低コストで特定地域の観光情報を収集する試みも特徴である。

4. 観光情報抽出

本章では、地域サイトとブログからの観光情報抽出手法と抽出結果について記述する。4.1 節に、地域サイトからの観光情報抽出手法、4.2 節にブログからの観光情報抽出手法を示す。また、4.3 節に観光情報の抽出結果を示す。

4.1 地域サイトからの観光情報抽出

地域サイトからの観光情報抽出は、Web 上に存在する収集対象の地域サイトから、観光スポットやイベントなど観光に関する情報を収集することが目的である。そこで、本研究では、地域サイトから観光情報に関する観光キーワードを収集することとした。

ここで、本研究における観光キーワードとは、地域サイト内の観光情報提供ページにおいて出現する名詞とした。地域サイト内に出現する観光キーワードは、ブログなど地域サイト以外の観光情報を記述したページでも利用される可能性がある。よって、観光キーワードは、ブログからの観光情報抽出や関連性判断及びカテゴリ化の際に利用する。

地域サイトは、地方自治体・観光協会が提供する観光情報を提供するサイトとした。本研究の実験を行う際の情報源は、本研究に関連して共同研究 [8] を行っている富山県魚津市を対象とした。

観光情報抽出対象は、魚津市の Web ページ内の観光情報^(注1)のトップページを基準にその下位ページの約 100 ページと魚津市観光協会公式サイト魚津たびナビ^(注2)内の約 300 ページを対象とした。ここで、各サイトから外部ドメインへのリンクは抽出対象外としている。抽出方法は、次のとおりである。抽出対象とした各 Web ページ上に記述された観光情報に関するテキストを取り出す。形態素解析器 MeCab^(注3)を利用し、分かち書きによって文から単語へ分割する。分割された結果から各語の品詞を判断し、名詞を観光キーワードとして収集する。ただし、名詞が連続する場合は、複合名詞を観光キーワードとして扱うこととした。収集した観光キーワードは、次節以降の処理で利用する。

4.2 ブログからの観光情報抽出

ブログからの観光情報抽出の目的は、実際に観光スポットやイベントに行った旅行者の感想や意見を抽出することが目的である。しかし、ブログのポータルサイトの数は非常に多く、旅行ブログなど特定のジャンルに特化したサイトからポータルサイトにも登録されていない個人が運用しているサイトまで幅広いサイトが存在している。より多くのブログのポータルサイト

表 1 観光情報抽出件数

Table 1 The number of sightseeing information extraction.

サイト	カテゴリ	ページ数
地域サイト	祭り・イベント	37
	史跡・歴史	19
	自然・景勝地	35
	施設・レジャー	16
	地場産品・特産品	10
	食べる・飲む	5
	その他	199
ブログ	Yahoo!ブログ	500

表 2 観光キーワード出現回数

Table 2 The sightseeing keyword appearance frequency.

単語	回数	単語	回数
魚津	509	魚津市観光協会	272
海	355	FAX	247
map	327	魚津商工会議所ビル 4F	237
TEL	326	住	228
こと	321	問	213

を抽出対象とすることで、必要とする観光情報を多く収集できるが、Web 上の全てのブログを抽出対象とすることは困難である。

よって、本研究では、特定のジャンルに特化せず多くの層の様々な話題を収集できる Yahoo Japan!が運営する Yahoo!ブログ^(注4)を対象とすることとした。

観光情報抽出対象は、Yahoo!ブログの検索においてキーワード「魚津」で検索されたブログの記事とした。観光キーワードの抽出方法は、4.1 節に示す地域サイトからの観光情報抽出と同様である。

4.3 抽出結果

地域サイトとブログからの観光情報抽出件数を表 1 に示す。観光情報抽出件数は、地域サイトとブログから観光情報を抽出するための収集対象とした Web ページ数を表す。

次に、表 2 は、表 1 に示した地域サイトの収集対象とした Web ページ (321 ページ) から抽出された観光キーワードの内、出現回数の多い上位 20 件を示す。ここで、収集対象のページが画像や PDF のみで構成されているページは除外した。このため、実際に観光キーワードを抽出したページは、321 ページ中 198 件であった。出現回数が上位の観光キーワードは、地域サイトの各ページのヘッダ部やフッタ部に利用される単語であった。よって、出現回数を利用するだけでは、観光情報を効率的に抽出できないことが確認できた。また、抽出により収集した観光キーワードは、約 1 万 1 千語であった。抽出した観光キーワードの重み付けについて次章に記述する。

5. 観光情報融合

本章では、前章で述べた地域サイト及びブログから抽出したページを基に観光情報の融合について記述する。5.1 節では、

(注1) : <http://www.city.uozu.toyama.jp/topVisit.aspx>

(注2) : <http://www.uozu-kanko.jp/>

(注3) : <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>

(注4) : <http://blogs.yahoo.co.jp/>

地域サイトから抽出した観光キーワードの重要度を得るための重み付けについて記述する。5.2 節では、ブログから抽出したページが地域サイトに関連した観光情報であるか判断する関連性判断の手法について記述する。5.3 節は、地域サイト及びブログをカテゴリに分類し整理するためのカテゴリ化の手法について記述する。そして、5.4 節に観光情報融合結果を示す。

5.1 観光キーワードの重み付け

地域サイトから抽出した観光キーワードの出現回数は、前章の処理で得ることができた。しかし、出現回数だけでは、各観光キーワードが地域サイト内においてどの程度重要なキーワードであるか判断できない。そこで、*tfidf* を用いて観光キーワードに重み付けを行うこととした。

地域サイトの各ページごとに式 (1) により文書 *d* 中における単語 *t* の出現頻度 *tf* を求める。ここで、文書は各 Web ページ、単語は観光キーワードを示す。

単語 *t* が出現する文書数と全文書数を利用し、式 (2) から各ページにおける単語 *t* の逆文書頻度 *idf* を計算する。式 (1) と式 (2) を利用し求めた *tf* と *idf* から式 (3) の *tfidf* を計算し、各観光キーワードに重み付けを行った。各観光キーワードの *tfidf* 値を次節以降の処理に利用した。

$$tf(t, d) = \frac{\text{文書 } d \text{ 中における単語 } t \text{ の出現回数}}{\text{文書 } d \text{ 中における全単語数}} \quad (1)$$

$$idf(t) = 1 + \log \frac{\text{全文書数}}{\text{単語 } t \text{ が出現する文書数}} \quad (2)$$

$$tfidf(t, d) = tf(t, d) \times idf(t) \quad (3)$$

5.2 関連性判断

関連性判断は、4.2 節の手法で抽出したブログの各ページが、抽出対象とした地域サイトの観光情報に関連があるか判断する処理である。本研究では、抽出したブログの各ページが、富山県魚津市の観光情報であるかを判断するものとする。

関連性判断には、事前に人手による学習データを利用し生成した線形判別関数を用い、機械的に地域観光情報との関連性を判断する処理である。線形判別関数は、人手によりブログの各ページの内容と地域サイトの観光情報との関連性の有無を判断した学習データを用いて生成した。今回は、ブログページ 50 件を学習データとし、地域サイトから抽出した観光キーワードと一致したブログページの単語の *tfidf* の合計と、観光キーワードに一致しなかった *tfidf* の合計を基準に関連性判断を適用した。

5.3 カテゴリ化

カテゴリ化は、抽出した地域サイトとブログの各ページを記述内容によってカテゴリに分類する処理である。カテゴリは、富山県内の 15 市町村の地域サイトのカテゴリを参考に、「見る・遊ぶ」、「イベント・祭り」、「食べる・泊まる」、「お土産・特産品」、「自然・文化」の 5 つと予め定義した。また、5 つのカテゴリには、事前に富山県内の地域サイトを基にカテゴリに含まれる語を用意した。表 3 に、事前に用意したカテゴリに含まれる語の一部を示す。5 つのカテゴリで約 60 語のカテゴリ含有語を事前に用意した。

表 3 カテゴリ含有語

Table 3 Category component word.

カテゴリ	カテゴリ含有語		
見る・遊ぶ	観光スポット	見どころ	遊ぶ
イベント・祭り	祭り	イベント	催し
食べる・泊まる	グルメ	宿泊	味覚
お土産・特産品	土産	特産	工芸
自然・文化	景勝地	歴史	史跡

この単語を基準として、地域サイトのカテゴリ化を行った。分類手法は、地域サイトの各ページにおいて各カテゴリのカテゴリ含有語の出現頻度を求め、出現頻度の多いカテゴリに分類することとした。

次に、ブログの各ページのカテゴリ化手法を示す。ブログのカテゴリ化は、カテゴリ含有語を基に分類された地域サイト群を利用した。各ブログページに対して、5 つのカテゴリに分類された地域サイトのページ群に含まれる観光キーワードの出現頻度を求め、出現頻度の多いカテゴリに分類することとした。

5.4 観光情報融合結果

前節までに記述した観光情報融合結果を記述する。

5.1 節の観光キーワードの *tfidf* による重み付けを行った結果の内、*tfidf* 合計値の大きい上位 20 件を表 4 に示す。表 2 に示した観光キーワードの出現回数の上位と比較すると、観光情報として利用可能な語が上位に現れる結果となった。尚、観光キーワードとして抽出した約 1 万 1 千語の内、人手により観光情報を表現する際に利用される可能性があるかと判断した語の数は約 30% の約 3 千語であった。

次に、5.2 節で述べたブログの線形判別関数による関連性判断結果を示す。この関連性判断結果は、表 1 に示したブログの収集対象とした 500 ページから学習データとして 50 件、実験データとして 56 件を利用し行った実験結果である。表 5 は、学習データ 50 件の判断結果である。現時点で生成した線形判別関数による正判別率は 78.0% となった。また、表 6 は、表 5 の線形判別関数を基に、実験データ 56 件について線形判別を行った結果である。正判別率は、70.0% となった。

次に、5.3 節に記述した地域サイトのカテゴリ化の結果を記述する。表 7 は、地域サイトのカテゴリ化結果である。表 7 において、地域サイトのカテゴリ化の対象としたページは、4.3 節に示した観光キーワードを取得できた地域サイトの Web ページ 198 件とした。表 7 に示す機械 (排他的) は、5.3 節に記述した手法を用いて地域サイトの各ページを、関連性が最も高いと判断したカテゴリに排他的に分類した結果、各カテゴリに分類されたページの数である。機械 (非排他的) は、同様の手法を用いて地域サイトの各ページを、関連性があると判断したカテゴリに非排他的に分類した結果、各カテゴリに分類されたページの数である。また、人手は、5.3 節に示したカテゴリ化の手法の有効性を確認するために、カテゴリ化の対象とした Web ページ 198 件を人手により分類した結果、各カテゴリに分類されたページの数である。地域サイトの各ページを人手

表 4 *tfidf* 合計値による観光キーワード上位 20 件Table 4 Sightseeing keyword high rank 20 cases by the *tfidf* total level.

観光キーワード	<i>tfidf</i> 合計値	観光キーワード	<i>tfidf</i> 合計値
海	0.704	特産品	0.566
定番商品	0.627	とき	0.557
工芸品	0.627	新鮮	0.556
イチオシ	0.627	幸	0.546
オリジナル	0.627	歴史	0.545
海産物	0.624	今	0.537
魚津漆器	0.616	魚津	0.519
伝統	0.613	map	0.507
漆器	0.612	地元	0.482

表 5 学習データ 50 件の線形判別結果

Table 5 Linear discriminant analysis result of 50 learning data.

人手\機械	関連性有	関連性無	正判別率
関連性有	25	7	78.1%
関連性無	4	14	77.8%
総合			78.0%

表 6 実験データ 56 件の線形判別結果

Table 6 Linear discriminant analysis result of 56 experimental data.

人手\機械	関連性有	関連性無	正判別率
関連性有	10	6	62.5 %
関連性無	9	31	77.5 %
総合			70.0 %

により分類した結果を正解群として、機械 (排他的) と機械 (非排他的) のそれぞれの分類結果をカテゴリ化結果群として求めた適合率と再現率も表 7 に示した。

表 8 は、5.3 節に記述したブログのカテゴリ化結果である。実験に利用したブログページ数は、表 1 に示したブログの収集対象とした 500 ページとした。表 8 に示す機械 (排他的) は、5.3 節に記述した手法を用いてブログの各ページを、関連性が最も高いと判断したカテゴリに排他的に分類した結果、各カテゴリに分類されたページの数である。機械 (非排他的) は、同様の手法を用いてブログの各ページを、関連性があると判断したカテゴリに非排他的に分類した結果、各カテゴリに分類されたページの数である。また、人手は、5.3 節に示したカテゴリ化の手法の有効性を確認するために、カテゴリ化の対象としたブログページ 500 件を人手により分類した結果、各カテゴリに分類されたページの数である。ブログの各ページを人手により分類した結果を正解群として、機械 (排他的) と機械 (非排他的) のそれぞれの分類結果をカテゴリ化結果群として求めた適合率と再現率も表 8 に示した。

6. おわりに

本研究は、地域サイトの観光情報とブログの観光情報を、Web 上から抽出し、融合する手法を提案した。低コストに観光情報を抽出・融合するために、地域サイトから観光キーワード

表 7 地域サイトのカテゴリ化結果

Table 7 Classification of web pages in the regional sites.

カテゴリ	見る 遊ぶ	イベント 祭り	食べる 泊まる	お土産 特産品	自然 文化
機械 (排他的)	9	25	20	27	66
人手	153	64	65	46	128
適合率 (排他的)	88.9%	76.0%	85.0%	51.9%	89.4%
再現率 (排他的)	5.2%	29.7%	26.2%	30.4%	46.1%

機械 (非排他的)	130	36	60	33	117
人手	153	64	65	46	128
適合率 (非排他的)	85.4%	75.0%	71.7%	57.6%	82.1%
再現率 (非排他的)	72.5%	42.2%	66.2%	41.3%	75.0%

表 8 ブログのカテゴリ化結果

Table 8 Classification of web pages in the blog sites.

カテゴリ	見る 遊ぶ	イベント 祭り	食べる 泊まる	お土産 特産品	自然 文化
機械 (排他的)	7	14	2	0	40
人手	40	27	9	0	28
適合率 (排他的)	71.4%	21.4%	0.0%	0.0%	52.5%
再現率 (排他的)	12.5%	11.1%	0.0%	0.0%	75.0%

機械 (非排他的)	34	48	31	45	33
人手	40	27	9	0	28
適合率 (非排他的)	70.6%	41.7%	12.9%	0.0%	54.5%
再現率 (非排他的)	60.0%	74.1%	44.4%	0.0%	64.3%

を抽出する手法、抽出した観光キーワードの *tfidf* 値を利用した関連性判断とカテゴリ化の手法を提案した。実験の結果、観光キーワードを利用した関連性判断は、正判別率 70.0% が得られた。また、カテゴリ化の精度は、地域サイトの適合率と再現率の平均はそれぞれ、排他的 (78.2%, 27.5%), 非排他的 (74.3%, 59.4%) が得られた。ブログではそれぞれ、排他的 (25.1%, 48.3%), 非排他的 (29.1%, 19.7%) が得られた。よって、さらに適合率や再現率の向上は必要ではあるが、低コストに特定地域の観光情報を抽出・融合する手法を構築可能であると考えられる。

今後は、今回の実験で得られた知見を発展させ、SVM などを利用したより精度の高い抽出・融合手法を検討していく。また、地域サイトから取得する観光キーワードの抽出方法についてもより効率的な手法を検討する必要がある。更に、観光情報提示の研究を進めるにあたり富山県魚津市に協力をいただきながら観光情報の効果的な提示手法を検討していく予定である。

文 献

- [1] 魚津市商工観光課：魚津市観光振興計画，魚津市，2012.
- [2] 斎藤 一：Web における観光情報の提供と分析，人工知能学会誌，26 卷 3 号，pp.234-239，2011.
- [3] 大槻 洋輔，佐藤 理史：地域情報ウェブディレクトリの自動編集，情報処理学会論文誌，Vol.42, No.9, pp.2310-2318,2001.
- [4] 嶋田 和孝，前田 裕，井上 俊右，遠藤 勉：地方都市を対象とした観光情報の分析に向けて，観光情報学会，第 3 回研究発表会，pp.8-13,2011.
- [5] 石野 亜耶，難波 英嗣，竹澤 寿幸：旅行ブログエントリからの観光情報の自動抽出，日本知能情報ファジィ学会誌，Vol.22, No.6, pp.667-679,2010.
- [6] 石野 亜耶，難波 英嗣，竹澤 寿幸：広電沿線観光情報提示システムの構築，人工知能学会 インタラクティブ，情報アクセスと可視化マイニング研究会，第 1 回，SIG-AM-01-04, pp.20-23,2012.
- [7] 寺西 拓也，野村 達二，平山 智子，石野 亜耶，難波 英嗣，竹澤 寿幸：観光ガイドブックへの旅行ブログエントリと質問応答コンテンツの対応付け，言語処理学会，第 18 回年次大会発表論文集，pp.333-336,2012.
- [8] 山中 光定，高尾 和志，遠藤 雅樹：共同研究「市民バスロケーションシステムの開発」，第 20 回職業能力開発研究発表講演会,3-15(2012).