

# 観光情報学の最前線

—観光の分散化と個人化を促進する集合知活用情報技術—

応  
専



馬 強 (京都大学大学院情報学研究科)

## 観光 2.0 へ

観光は日本経済を支えるきわめて重要な成長分野であり、雇用拡大や地域活性化に大きな効果が期待できる。2015年の訪日外国人数が2012年から2倍増で約2,000万人に達した。この好機を活かして、地域観光をさらに活発させ、「観光先進国」を実現するには、ICTの利活用が必要不可欠であり、観光業界や行政などサービス提供側の主導で、自動翻訳、外国語表示、無料Wi-Fiの整備など、既存観光資源におけるおもてなしの向上や、ICTインフラの整備が今まで重点的に行われてきた。しかしながら、現状、急増する観光客は、東京・大阪・京都などの特定の有名な観光都市に集中し、社会インフラなどの負担および住民・観光客のストレスが大きい一方、いわゆる地方への経済効果の恩恵は限定的である。持続可能な観光立国・地域社会を実現するためには、多様な観光資源を開発して細分化する観光ニーズに応え、観光の個人化や地域・時期の分散化を実現して、サービス提供側中心の観光1.0から利用者側（ユーザ）主導の観光2.0へ進化していく（図-1）必要がある。

観光客・住民の意見や観光地調査は、アンケートベースの社会統計学的手法で実施されてきたが、コストが高いため大規模な調査は難しかった。また、得られた結果の集約・分析にも多大な時間とコストを要し、タイムリーな調査分析が困難とされている。一方で、ブログやSNS(Social Network Services: ソーシャルネットワークサービス)などユーザ自らが情報発信を行って生成されるUGC (User Generated

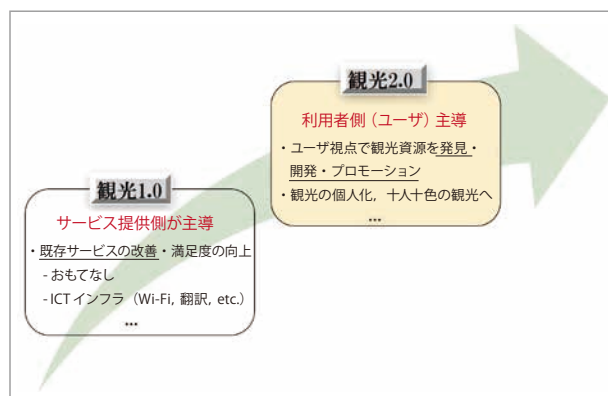


図-1 観光1.0から観光2.0へ

Contents : ユーザ生成コンテンツ) は、社会、経済、政治、医療、防災、生活など、多方面における重要な巨大情報資源となっており、地域やユーザの意見を低コスト、大規模かつタイムリーに収集・分析することが可能となりつつある。

観光情報学は、観光活動を支援するための情報技術とその利活用に関する観光学と情報学の融合分野である。観光1.0における観光情報学は、業者や行政など提供者側からの情報発信や、現地における情報提供・コミュニケーションの支援など、おもてなしの向上のための研究開発を中心に行ってきた。ユーザにとって、ガイドブック、旅行会社や観光協会のWebサイトなどサービス提供側からの情報が頼りであった。これに対して、観光2.0では、友達が旅行先で楽しんでいる写真などをSNSで見ると、その場所に行ってみたくなったり、口コミ情報を調べて旅行先を決めたりするケースが多くなってきている。フォートラベル(株)の調査では、SNSを見て旅行したくなる人が9割弱にも達している<sup>☆1</sup>。

<sup>☆1</sup> <http://internet.watch.impress.co.jp/docs/news/601717.html>

また、フジテレビ「めざましテレビ」でも紹介されたように、タイ人観光客の SNS での口コミによって、富士浅間神社が富士見の新名所となった例もあり、UGC から新しい観光スポットを発見することが可能である。このように、ユーザの観光体験に基づいて生成された UGC は、旅行プランニング、観光資源の再発見とプロモーションの重要な情報源となっている。これを背景に、観光情報学では、ユーザの視点からの“観光知”の分析と利活用、いわば観光知マイニングに関する研究が盛んに行われてきている。

## 観光知マイニングとその利活用の概要

紅葉のシーズンに京都を訪れる場合を考える。ユーザが検索エンジンなどを利用すれば、旅行ガイドサイトから紅葉の名所や人気ルートなどの情報を容易に調べられるが、多くの観光客で混雑している名所を避けて、ローカル文化を楽しんだり、のんびり紅葉狩りをしたりするスポットにはなかなか辿り着かない。このようなスポットは旅行サイトやガイドブックではあまり紹介されないが、個人のブログやツイートなどで言及されていることがある。膨大な UGC から効率よく地域観光に関する“知”をマイニングして体系化し、利活用して観光の個人化と分散化を促進することが観光情報学の一大課題である。

図-2 は観光知のマイニングとその利活用の流れを示している。まず、経度・緯度を記録したジオタグなどの位置情報を利用して、Flickr, Twitter やブログなどで発信されている UGC から地域・観光に関するデータを収集する。そして、ユーザモデリングを行い、地域を訪ねたユーザの特徴を明らかにする。次に、ユーザモデリングで明らかになったユーザ特徴を利用して、ユーザ分布を考慮した人気度(その地域・スポットが誰にとって人気なのか)、景観の美しさや地元文化の特徴を考慮した観光価値(その地域・スポットを見に行く価値があるのか)などの側面から地域・スポットの特徴を分析して、一部のユーザ(富士浅間神社の例のタイ人観光客)に

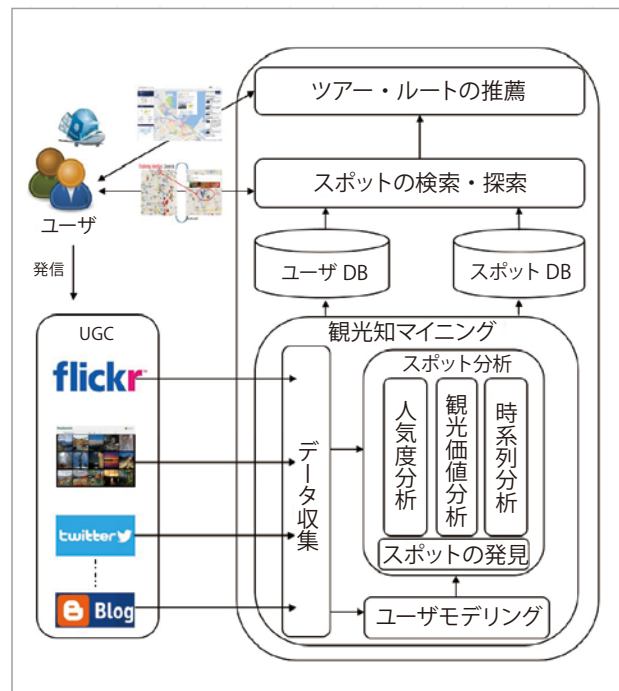


図-2 観光知のマイニングとその利活用の概要

しか知られていないが、見に行く価値の高い穴場スポットなど多様な観光資源を発見できるようにする。訪問者数が少ない新しいスポットの場合、そのほかのユーザにとって見に行く価値があるか、投資して開発する価値があるかを調査する必要があるが、観光価値の自動分析により大規模・低コストに実施できる可能性がある。

観光知マイニングによって得られたユーザや地域・スポットに関する知をデータベースに格納すれば、ユーザがスポット検索・探索システムを利用して、訪問客の分布、人気度や観光価値などの情報を閲覧しながら、自分の訪れたいスポットを選択できる。さらに、ツアー・ルートのおすすめシステムは、これらのスポットを効率よく巡回できる観光ルートを、ユーザの嗜好や時間などの制約条件を考慮し、発見してユーザに推薦する。

観光マイニングは UGC マイニングの応用であるが、多様なスポットの発見と分類、そして訪問者数やデータが少ない新しいスポットの価値分析など観光知マイニングに特有な要素技術も多くある。以下、図-2 に沿って逐次に説明していく。

## データの収集：どのようなデータが利用できるのか？

UGCにはさまざまな情報があるが、位置情報、地名やランドマークなどを手がかりとして、SNSのAPI（Application Programming Interface：アプリケーションプログラミングインタフェース）や検索エンジンを利用すれば、地域観光知のマイニングに必要なデータを収集することが可能である。

UGCは、テキスト、画像、動画などさまざまな形式で生成・公開されているが、一般に、コンテンツのほか、ユーザ属性、時間、位置などのメタデータも付与されている。画像の場合は、Exif（Exchangeable image file format）データが付与されることが多く、焦点距離、ISO感度、露出などのメタデータが利用可能である。TwitterやFacebookなどのSNSの場合は、フォロワーや友人、投稿に対する返信・反応などに関する情報も利用可能である。特に、位置情報は、膨大なUGCから観光集合知を編纂するデータを収集・整理・分析するための重要な手がかりである。

現在の観光集合知マイニングに関する研究の多くはジオタグ付きのコンテンツ（写真、ツイート、etc.）とそのメタデータ（時間、フォロワーなどのソーシャルネットワーク情報、etc.）を対象としている。特に、言語処理を必要としないジオタグ付き画像データやチェックインデータが多用されている。また、ジオタグに依存せず、コンテンツやそのタグなどのメタデータに地名やランドマークなどの手がかり表現（八坂神社、祇園、ハッ橋、舞子、etc.）が含まれていれば、そのコンテンツを、その地名やランドマークに対応する場所に関するデータとすることもできる。

## ユーザモデリング：どんなユーザが地域・スポットを訪れたか？

地域・スポットに関する情報を発信したユーザを対象に、そのプロフィールと発信履歴などを用いて

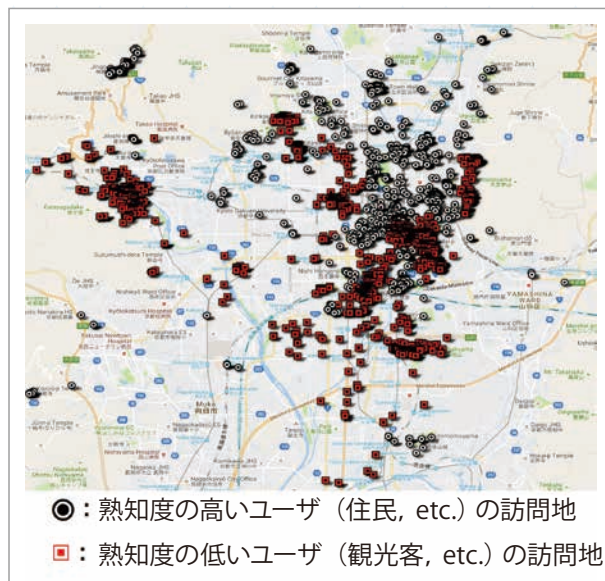


図-3 熟知度の違うユーザ群の訪問先分布（京都）

ユーザモデリングを行えば、地域・スポットを訪れたユーザの特徴が明らかになる。そのスポット・地域がどんなユーザにとって人気あるかを分析できるようになり、ひいては、個人に適したスポットの推薦や、スポットの発見・開発とプロモーションも可能となる。

画像処理、特に顔認識の技術を駆使して、観光地で撮った記念写真を分析して、訪問客の年齢・性別などの特徴を分析する手法が開発されている。また、地域についてよく知っている住民と知識のあまりない観光客の振舞いの違いに着目して、ユーザの地域熟知度の推定とそれに基づく分類手法として、フォロワーなどの友人関係に着目したソーシャルネットワークモデルと、地域への訪問回数・時間に基づく時間モデルと、訪問場所に基づく空間モデルも開発されている<sup>1)</sup>。

図-3では、京都市におけるユーザ群の訪問場所を熟知度の違いによって分類しており、観光客が観光名所に集中している様子がうかがえる。たとえば、図の左上を見れば、嵐山は観光客でにぎやかであるが、京都に詳しい人は清滝川沿いをよく訪れていることが分かる。

## スポット分析：どのようにして新たな観光地を発見するか？

スポット分析または POI (Points of Interest) マイニング<sup>☆2</sup> は世界的に活発に研究されている分野である。ユーザの訪問時間・回数に基づいて人気度を推定してホットスポットを発見する

手法が多数開発されている。これに対して、**図-4**に示されているように、観光スポットを、1) 人気度(ユーザ分布)、2) “質”(観光価値)、と3) 時間・季節<sup>☆3</sup>の3つの軸から分析するフレームワークも提案されている。これによって、ホットスポットだけではなく、まだ知られていないが潜在的に観光価値の高い穴場などの多様な観光スポットを発見可能となり、観光客の多様なニーズに応えられ、地方の観光資源の開発の支援にも効果がある。

### □ スポットの発見

スポットの発見は、地理情報システムや Wikipedia で作成・公開している既存の POI リストを利用する手法が多い。たとえば、Foursquare でのスポットのチェックイン情報を利用して、チェックインの多い場所をホットな POI とすることができる。これらの手法は、特に人気度の高い、既存のスポットの発見によく利用される。一方、UGC の位置情報を用いてスポットとその範囲を特定する手法では、既存のスポットだけではなく、今まであまり知られていなかったスポットやテーマパークなど大きな領域における特定のスポットも発見できるようになる。

位置情報を用いて写真やツイートを地図上にプ

☆2 厳密に POI は観光スポットより広い概念であるが、本稿では、観光スポットに限定して説明を行う。

☆3 UGC には時間情報が明記されている場合が多く、人気度や質の分析手法を時系列的に適用すればスポットの時間・季節の特徴とその変化を明らかにすることが可能である。本稿では、紙数の都合で時間軸の分析に関する説明を割愛する。



図-4 スポット分析の3つの軸：ユーザ分布ベースの人気度，質（観光価値）および時間

ロットして、Mean-Shift や DBSCAN などのクラスタリング手法を適用してそれらの点をグルーピングして、それぞれのグループをスポットとする手法が開発されている。

さらに、写真の Exif データを分析して、ホットスポットを景観場や視点場に分類する手法が開発されている<sup>2)</sup>。景観場とは、富士山のような、関心領域・見る対象そのものである場・スポットであり、視点場とは、関心領域・見る対象を観賞するための場・スポットである。富士見スポットの富士五湖のような記念撮影のポイントや鑑賞スポットを探することができるようになる。

### □ 人気度分析

スポットの人気度は、訪問するユーザの数、回数および滞在時間で推定することが多い。特に、滞在時間を考慮した人気度推定は個人に適した観光の実現に重要である<sup>3)</sup>。しかしながら、ユーザの特徴を考慮せず、一様に訪問回数と時間を集計するのは、新しいスポットの発見や個人に適したスポットの推薦には不十分であり、そのスポットが誰にとって人気であるかを把握することが必要である。

ユーザモデリングで紹介された技術を利用して訪問ユーザの特徴を分析すれば、ユーザ分布を考慮した人気度分析ができ、個人に適したスポットの推薦や、異なった視点からの観光資源の再発見に効果がある。たとえば、地域・スポットを訪ねたユーザの熟知度を分析すれば、観光客に知られていない地元

住民がよく行くスポット、またその逆のスポットを発見でき、ユーザの異なった観点を活かした地域の再発見につながる。

## ◻ 観光価値分析

観光価値の統一した定義はまだないが、本稿では、スポットの観光価値を、そのスポットにある景観の美

しさやローカル文化の特徴などからの魅力、つまりそのスポットを見に行く価値と定義する。今までのスポットの観光価値の分析は、現地調査に頼ってきており、コストが高い。UGCを対象とした観光価値の自動分析手法を用いれば、より多くの潜在価値の高い新しいスポットを低コストに発見できるようになる。

UGCを用いたスポットの観光価値の評価手法は、1) 評判分析、2) ユーザ行動ベースの間接評価、および、3) 景観ベースの直接評価に分けられる。

- **評判分析**：ツイートやブログなどのテキストデータを対象に、地域やスポットのツイートやブログを収集して、あらかじめ定義したパターンや極性辞書を用いてその地域・スポットに関する評判情報を抽出したり、頻度情報を用いて地域特有な表現を抽出することで、地域やスポットの特徴や評判を分析する。

- **ユーザ行動ベースの間接評価**

- UGCの作成者の行動ベースの間接評価：景観のよいスポットを訪ねたユーザは、そのスポットの全体と部分の両方の写真を撮ることが多い。この観察に基づいて、スポットの写真を遠景と近景に分類して、その割合を調べることで観光価値を評価することができる。また、ユーザのスポットでの滞在時間が長いほどスポットの質（観光価値）が高いとする手法も開発されているが、ユーザの滞在時間を容易に取得できないのが難点である。

- UGCの閲覧者の行動ベースの間接評価：よいス

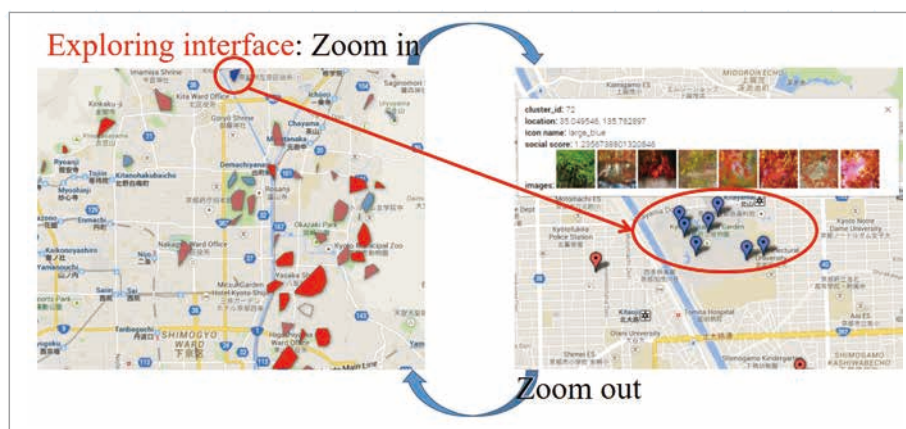


図-5 スポットの探索システムの例

ポットに関する発言やそこで撮った写真に対して、「いいね」などのポジティブなレスポンスが多くなる。これらを集計することでスポットの観光価値を推定することができるが、コンテンツの作者の友達の多さや作文・撮影のスキルを考慮して相対的に分析することが重要である。

- **景観ベースの直接評価**：スポットで撮った写真を対象に、環境心理学の知見を活用して、写真に写されている景観物の色、配置、種類などを分析して、スポットの自然や人文的景観価値を直接推定する研究も始まっている<sup>4)</sup>。景観そのものを直接分析するため、観光者数やデータの量に依存しない。穴場など訪問者数の少ない新しいスポットの発見には特に有効であるが、分析対象となるデータの量が少なく、質のバラツキも大きいため難しい。

## 観光スポット・ルートの検索・推薦

### ◻ スポットの検索・推薦

これまで紹介したユーザモデリングとスポット分析手法を用いて得られたユーザとスポット・地域に関する観光知をデータベースに格納して利用すれば、ユーザ視点の観光スポットの検索と推薦システムを構築できる。観光客は、これらのシステムを利用して、個人に適した観光スポットを選べるようになる。また、観光業界や行政などサービス提供者は、新しいスポットの発掘や、訪れた観光客の分布などの調査分析を低コスト・タイムリーに行えるようになる。

図-5は、京都市の花見スポットの探索システム<sup>5)</sup>の例である。スポットの人気度の高さを色の違いで表現し、ホットスポットと穴場スポットは異なる色（それぞれ赤と青）で示している。画面をズームインすればスポットの詳細情報や写真などが見られ、ズームアウトすれば広範囲のスポットを一覧できる。京都府立植物園は観光客にあまり知られていないが、紅葉がきれいで見に行く価値の高い穴場スポットであることが分かる。

### □ ツアー・ルートの検索・推薦

ツアー・ルートの検索・推薦システムは、複数のスポットを効率よく周遊するルートをユーザーに推薦する。観光ルートの検索や推薦には主に2つのアプローチがある。1) 頻度ベースのアプローチで人気度の高い、いわゆるゴールドルートを紹介する手法、および、2) グラフベースのアプローチで、あらかじめ選択した複数のスポットを効率よく巡る、個々のユーザーに適したルートを発見して推薦する手法である。前者のほうは、ルート・ツアーそのものを推薦するため、ユーザーによるスポット探索の必要がない。一方、後者の場合、ユーザーがあらかじめスポットの候補を決める必要があり、より細かいニーズに応えた個人に適したツアーを実現できる利点がある。

#### 人気ルートの発見と推薦

UGCを用いて過去のユーザー体験から人気ルートを抽出・発見して推薦する。これは、特に、初めて訪れる場所や旅に慣れていない初心者のユーザーなどには有用である。ルートを抽出する際、訪れたユーザーの特徴を考慮することによって、似たユーザー(群)から抽出されたルートを優先的に推薦することでユーザーの嗜好にマッチしたルートも推薦可能である。

ユーザーの訪れた場所・スポットの系列を抽出できれば、既存の頻出パターンを発見するデータマイニング手法を適用して人気ルートを発見できる。位置情報付きのUGCであれば、その位置情報を用いて訪問ルートを生成することが可能である。たとえば、Flickrに公開されているジオタグ付きの写真データとWikipediaのPOIリストを用いて、人気の高い訪問



図-6 CT-Planterの実用例<sup>6)</sup>

ルートを発見することができる。また、ジオタグを必要としない、ブログやツイートなどのテキスト情報からルート情報を抽出する手法も提案されている。

#### ルート推薦の個人化

ユーザーが選んだスポットを全部またはできるだけ多く訪問できるルートを発見して推薦する。この問題は、候補スポットを地図上にマッピングして、道路交通ネットワーク上で最適経路の探索問題としてモデリングできる。これによって、ユーザーの嗜好や時間制約などを考慮した個人に適したルートを推薦できるようになる。

従来、距離や時間の最短ルートを発見することを目的とした研究が多かったが、Yahoo! LabsのQuercia氏らの“the Shortest Path to Happiness”の推薦に関する研究をきっかけに、スポットの質(景色の美しさ、人気度、etc.)、ルートの性質(混み具合、景観、etc.)、ユーザーの嗜好、および時間(移動時間、特定のスポットの訪問開始・終了時間、etc.)などの制約も考慮するようになった。たとえば、PersTour<sup>3)</sup>など、ルートの発見・推薦をOrienteering Problemとしてモデリングして、整数計画問題で解く手法が開発されている。

さらに、インタラクティブに推薦ルート・ツアーを決めていくシステムも開発されている。たとえば、旅行プランを地図上でユーザーに提示しながら、ユーザーから指定された条件(訪問場所、時間条件など)に基づいてプランを修正していくシステムCT-Plannerがある(図-6)<sup>6)</sup>。

## 今後の展望

本稿では、UGC を用いた観光知マイニングとその利活用について紹介したが、これらはいずれもまだ改善の余地があり、今後の発展を期待したい。ユーザの生の声である UGC は、地域・観光の集合知を編纂・利活用するための重要な情報源であるが、品質のバラツキ、位置情報など必要な情報の不足、個人の主観的な情報が多く信ぴょう性が低いなど課題も山積みである。また、UGC だけではなく、政府自治体の保有する統計データ、携帯キャリアが保持するユーザ位置情報、観光施設情報、交通・気象データなど、サービス提供側のデータも観光の個人化と分散化を実現するための重要な情報源であり、これらの地域データの統合分析・利活用基盤の開発を期待したい。

円安、ビザ緩和、減税対象の拡大などの政策と東京オリンピックのシナジー効果があって、今日本の観光が熱い。この好機を活かして、世界を先導する観光知マイニングの研究開発における日本の情報学コミュニティの研究者・技術者の活躍を期待したい。観光学、環境学、都市計画、そして心理学などさまざまな分野の知見を活かして、行政・観光業・行政・

住民と協力して、観光客（外国人および日本人客）、住民、観光業界、行政などのステークホルダをはじめ、より多くの人が満足できる、観光と日常生活を調和した地域社会の実現を期待したい。

### 参考文献

- 1) Zhuang, C. Y., Ma, Q. and Yoshikawa, M. : SNS User Classification and Its Application to Obscure POI Discovery, *Multimedia Tools and Applications* (2016), pp.1-27, doi:10.1007/s11042-016-4034-6 (2016).
- 2) 白井元浩, 廣田雅春, 石川 博, 横山昌平: ジオタグ付き写真を用いた関心領域と撮影スポットの発見, *電子情報通信学会論文誌*, J97-D (4), pp.835-844 (2014).
- 3) Lim, K., Chan, J., Leckie, C. and Karunasekera, S. : Personalized Tour Recommendation based on User Interests and Points of Interest Visit Durations, *Proc. of IJCAI2015*, pp.1778-1784 (2015).
- 4) Shen, Y., Ge, M., Zhuang, C. Y. and Ma, Q. : Sightseeing Value Estimation by Analyzing Geosocial Images, *International Journal of Big Data Intelligence* (to appear) (2016).
- 5) Zhuang, C. Y., Ma, Q., Liang, X. F. and Yoshikawa, M. : Discovering Obscure Sightseeing Spots by Analysis of Geo-tagged Social Images, *Proc. of ASONAM 2015*, pp.590-595 (2015).
- 6) Kurata, Y., Shinagawa, Y. and Hara, T. : CT-Planner5 : A Computer-Aided Tour Planning Service Which Profits Both Tourists and Destinations, *Proc. of Workshop on Tourism Recommender Systems* (2015).

(2016年11月16日受付)

馬 強 (正会員) qiang@i.kyoto-u.ac.jp

京都大学大学院情報学研究所准教授。観光情報学、投資情報学、情報栄養学(情報補完)を中心にマルチメディア情報システムとWebマイニングの研究に従事。博士(情報学)。

