

ロケーションベースドサービス(LBS)利用のための 固有表現抽出を活用したコンテンツ位置情報付与支援方法の提案

A Proposal of the Mechanism for Embedding Location Metadata to Content for the Location Based Service by the Named Entity Extraction

吉田 由紀 田邊 勝義 町原 宏毅
Yuki Yoshida Katsuyoshi Tanabe Hiroki Machihara

1. はじめに

近年, Google Maps を始めとする地図 API を利用したロケーションベースドサービス (LBS) が Web 上で利用可能となっている。これまで, 著者らは, LBS に時刻バーを付加して, 位置情報と時系列情報が付与されたコンテンツの検索, および閲覧が可能な情報検索システムを提案してきた[1]。本稿では, 上記システムにコンテンツ登録機能を追加してシステムの拡張を図ると共に, 地理情報インデクシング技術を利用してコンテンツへの位置情報付与を支援する方法を提案する。

2. 情報検索システムの拡張

これまで報告してきた情報検索システムは, 予め位置情報と時系列情報を付与したコンテンツが必要なことから, コンテンツの登録・追加はシステム管理者が行っていた。このしくみでは, コンテンツの登録者と閲覧者が別々になり, ユーザが閲覧しているコンテンツに関連した情報を持っていても, 簡単にそれを載せることはできなかった。そこで, 前述のシステムに, コンテンツ登録機能を持たせ, ユーザが自由にコンテンツ登録できるように, システムの拡張を図る。具体的には, システムにコンテンツ登録機能を設け, コンテンツの登録から閲覧まで一連の処理を 1 つのシステムで実現する。

前述したように, 本システムでは, コンテンツ登録において位置情報と時系列情報を付与する必要があるが, それを自動的に行なうためには, 多くの課題がある[2]。今回, それらをすべて解決してからシステムを拡張するよりも, まずはコンテンツ登録機能をシステムに搭載することを優先させ, 属性情報付与はユーザの補完で実施することとした。時系列情報は, 数値入力, もしくはプルダウン選択で比較的容易に入力できる。それに対して, 位置情報は緯度経度の数値で管理しており, これをユーザに入力させるのは困難である。そこで, システムに搭載されている地図 API を活用し, 地図操作 (スクロールや拡大・縮小) で付与地点を決定する。このとき, 通常的位置情報付与はこの方法で十分付与できるが, 複数のコンテンツを続けて登録する際, 付与地点が前に登録したコンテンツとこれから登録するコンテンツで異なると, 地図操作が増え, 登録作業が煩雑になることが考えられる。また, 地図上の土地勘のない場所を付与地点にしたい場合も, 同様のことが起こる。この地図操作を少しでも軽減するため, 地理情報インデクシング技術[3]を利用し, 位置情報付与を支援する方法を提案する。

3. 位置情報付与支援方法

地理情報インデクシング技術は, テキストを自然言語処理で解析し, 地図上の正確な位置(省略のない完全な住所, 緯度経度)を推定する技術である。コンテンツは, 画像や映像だけでなく, 内容を示すタイトルや説明文等のテキストそのものがコンテンツになり得ることから, 1 つのコンテンツに対して, これらすべてを登録情報として入力できることとする。そして, 入力されたテキスト情報を地理情報インデクシング技術で処理し, コンテンツの内容に関連していると思われる地点を推定して, コンテンツの付与地点の候補とする。推定地点は緯度経度も算出できるので, 地図上にマークすることは地図 API で実現可能である。

一方で, 推定地点, およびその緯度経度は, あくまでも候補地点であり, 多くの場合, ユーザが想定している付与地点とは異なる。けれども, 付与地点と距離的に近ければ, 一旦, 候補地点を地図上にマーカで表示し, その後のユーザによるマーカの地図操作で微調整し, 付与地点を決定できると考えられる。ユーザは, 候補地点が地図上に表示されてから, 初めて地図操作, それも微調整を行えばよいので, 手間の軽減が期待できる。

4. システム概要

4.1 全体概要

図 1 にシステム構成図を示す。本システムは, 地図 API, ユーザインタフェース API, コンテンツアクセス API, コンテンツデータベースから構成される。ユーザは, ユーザインタフェース API を通し, コンテンツアクセス API 経由でコンテンツデータベースにアクセスする。コンテンツアクセス API には, コンテンツの検索機能と登録機能がある。前者は, ユーザに指定された位置情報と時系列情報から, 該当コンテンツを検索する機能である。後者は, 新しいコンテンツをデータベースに登録する機能で, 地理情報インデクシング技術と連携している。

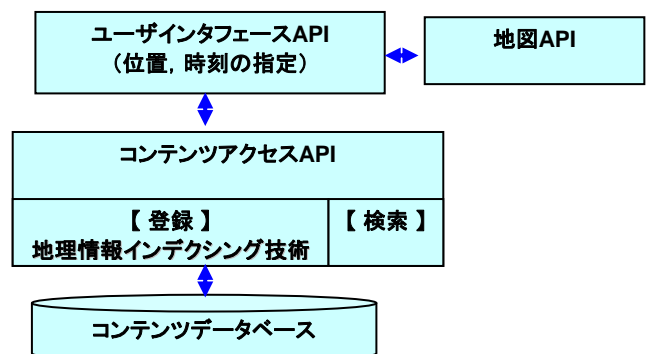


図 1 : システム構成図

4.2 登録画面

図2に登録画面を示す。画面左側には地図を、右側にはコンテンツの詳細内容(タイトル, 説明文, 時刻, 画像, カテゴリ, 緯度経度)の入力欄を表示している。ユーザに、タイトル, 説明文, カテゴリをテキストで、時刻をプルダウン選択で、画像をファイル名で入力させる。位置情報である緯度経度は、地図操作, 地理情報インデクシング技術の利用と地図操作の併用, 緯度経度の数値入力の3種類の方法で入力可能とし、地図上のマーカと緯度経度の数値は連動させる。地図の下には、地理情報インデクシング技術により推定された地名を表示する。図2では、「横浜へ行く途中、日ノ出町で下車して、桜の写真を撮った。その後、待ち合わせ場所であるみなとみらいへ行った。」という説明文から、【横浜】、【日ノ出町】、【みなとみらい】の3地名が推定され、さらに、その中から【日ノ出町】が選択されて、地図上の該当地点にマーカが表示されている状態を示している。その後、ユーザの地図操作でマーカを微調整して緯度経度を決定し、これら属性情報と共に、コンテンツをデータベースに登録する。

4.3 登録時間

本システムを用いて、コンテンツの登録時間を計測した。コンテンツ登録画面において、付与地点を地図操作のみで決定する場合と、地図操作と地理情報インデクシング技術を併用して決定する場合の2つの方法を用いて、コンテンツを登録するまでの操作時間を計測した。なお、コンテンツとしては、朝日新聞社の速報ニュースサイトにあるマイタウン内の全国アクセスランキング上位5件を利用し、見出しをタイトルに、本文の第1段落を説明文として入力した。2009年6月25日(木)と29日(月)の記事を使用し、操作の習熟度の影響をなくすため、25日分は地理情報インデクシング技術の利用なし、ありの順で、29日分はその逆の順で実施した。被験者は1名で、登録時間を表1に示す。

表1: 登録時間

地理情報 インデクシング技術	6月25日	6月29日	平均
利用なし	7分38秒	7分20秒	7分29秒
利用あり	4分6秒	4分22秒	4分14秒

コンテンツ登録の操作時間を平均すると、地理情報インデクシング技術を利用しない場合は7分29秒、利用した場合は4分14秒であり、操作時間は利用した場合の方が56.6%に減少している。また、タイトルとテキストから、地名はわかっているが場所が見つけれない場合でも、提案手法により候補地点が地図上に表示されるため、土地勘がない場所の位置情報も付与することが比較的容易であることがわかった。以上により、予備調査ではあるが、地理情報インデクシング技術と地図操作を併用して付与地点を決定する手法は、地図操作のみによる手法より効果的であることが示唆される。

5. おわりに

LBSを利用した情報検索システムにおいて、コンテンツ登録機能を搭載し、地理情報インデクシング技術と連携して、位置情報付与を支援する方法を提案し、効果を調べた。

今後は、大量データでの評価を実施し、システムの可能性を探る。また、既存のコンテンツデータベースとの連携方法を検討し、サービス実現を目指す。

参考文献

- [1] 吉田由紀, 田邊勝義, 町原宏毅: 時系列を考慮したロケーションベースドシステムの提案, 2009年電子情報通信学会総合大会, pp.121, 2009年3月
- [2] 沼見晃, 上松大輝, 大向一輝, 市瀬龍太郎, 武田英明: ActionLog:行動に着目した実世界コンテキストに基づく情報共有, 2006年度人工知能学会全国大会, 3D1-1, 2006年6月
- [3] 平野徹, 松尾義博, 菊井玄一郎: 地理的距離と有名度を用いた地名の曖昧性解消, 情報処理学会全国大会, pp. 3D-7(2008)



図2: コンテンツ登録画面