

デジタル地図と電話帳の統合による都市の詳細な時空間データの開発と既存統計情報の補完データとしての可能性検討

Development of spatio-temporal urban data by the integration of digital maps and yellow page and comparison with statistical data

秋山 祐樹 [1]; 柴崎 亮介 [2]

Yuki Akiyama[1]; Ryosuke Shibasaki[2]

[1] 東大・新領域・社会文化環境; [2] 東大・空間セ

[1] Frontier Science, The University of Tokyo; [2] CSIS,UT

<http://shiba.iis.u-tokyo.ac.jp/>

都市における事業所、店舗などの立地分布の変化は、中心市街地の空洞化に代表されるように都市の活力や魅力に大きな影響を与える。その動向を詳細かつ定量的に明らかにすることは都市空間の持続再生に不可欠である。しかしながら今日用いられている一般的な統計データは集計単位が大きく都市空間の詳細なモニタリングには不十分である。

一方、現地調査や住宅地図などの手作業による時系列判読を組み合わせれば、個別店舗などの時系列変化を追うことが出来る。しかしこれには多大な労力がかかりごく限られた広さの地域しかカバーできない。

そこで本研究では日本において全国的に整備されている住宅地図と電話帳を位置と店舗名称などにより統合し、個別店舗の時間的変遷を自動的に抽出し、詳細な時空間データを構築する手法を開発した。本データセットの開発には2つのステップで構成される。第一に同時点の住宅地図と電話帳を統合することである。そして第二に電話帳の属性情報を保有する古い年次の住宅地図と新しい年次の住宅地図を統合すること、すなわち時空間統合を行うことである。

我々の技術は建物情報の自動分析（建物名や階・部屋番号の認識）による三次元空間統合、そしてテナント名称間の表記ゆれを考慮可能な名称の同一性判定を実現した。「名称の表記ゆれ」とは住宅地図と電話帳間、あるいは異なる2時点間のテナント名称がたとえ同じテナントでも微妙に異なることを言う。本研究では自然言語処理方法の一つである「n-gram」を用いて、2つのテナント名称間の類似度を定量化することでこの問題を解決した。

さらに本研究では雑音語（ノイズワード）辞書を作成した。雑音語とはテナント名称の同一性判定結果に影響を与える頻出語、地名、駅名のことをいう。これらの辞書を用いてテナント名称から雑音語を取り除くことで、本技術は電話帳と住宅地図、そして異なる2時点のデータの統合を正確に行うことを実現した。本技術により個々の店舗・事業所の存続、転換（入替）新規出現、消滅を任意の時点間、地域でモニタリング出来る。

既に南関東全域の1995~2000年および2000~2005年の2時点間においてデータ整備が可能であることが確認出来ており、雑音語辞書は日本全土対応となっているため技術的には日本全土が整備可能な体制が整っている。

さらに本研究では既存の地域統計との突き合わせを通じて本データの利点、限界、補正方法などを明らかにする。既存の地域統計データとして、統計局発行の事業所・企業統計調査を用いる。事業所・企業統計調査は日本全国の事業所の変化の様子を業種ごとに市区町村単位で集計したデータである。本研究では東京23区において区毎に事業所・企業統計と本データを業種ごとに比較した。事業所・企業統計データと本データの業種情報は必ずしも一致しないため、業種対応表を作成し、我々のデータの業種を事業所統計の業種に置換することで、本データを事業所・企業統計調査に照合させた。図は東京23区全域における、業種大分類ごとの照合結果と事業所統計との比較である。本研究で作成したデータと事業所統計との間には1年のタイムラグ（我々のデータ：2000~2005年 事業所統計：2001~2006年）があるが、収録件数が少ない農林漁業などを除いて、多くの業種でその結果は類似したものとなった。

事業所統計は2~3年おきにしかな更新されない上に、時点間も5年に固定されている。一方多少の誤差や限界はあるものの、我々のデータセットによりこれらの制約を取り除くことが出来る。その上、テナントスケールのデータであるため、利用者は集計単位を任意に設定可能である。

これまでに本研究のような技術を開発しようとした例はほとんど見られない。また全国規模でこうした技術開発を行っている例は皆無である。我々のデータに見られるような高い柔軟性を兼ね備えた詳細な時空間都市データセットは日本ではもちろんのこと、世界でも過去に先例は見られない。

□ 存続 ■ 新設 ■ 廃業

