

口コミ分散表現ベクトルを用いた街の魅力の類型化とその空間分布

鈴木英之*

Typing of Shopping Districts Using Distributed Representations of Word-of-mouth

Hideyuki SUZUKI*

Abstract: Review sites on the Internet contain a lot of word-of-mouth data written by visitors to shopping streets and shopping malls. Many unique impressions and images of the city are described there. We obtained a distributed representation vector from those word-of-mouth data and used it to classify commercial districts. Traditionally, commercial centers have been classified in a hierarchical structure according to their size and area of attraction. It was confirmed that the same analysis is possible with the classification of commercial areas by the distributed representation vector.

Keywords: 中心市街地活性化 (city center activation), クリギング (Kriging), 分散表現 (Distributed representation), t分布型確率的近傍埋め込み法 (t-SNE), DBSCAN法 (DBSCAN)

1. 背景

衰退する多くの地域商業地区において、既存のハード（商店街）を活かしつつも、新たな社会的ニーズ（高齢世帯や子育て世帯の生活関連サービス需要）に応え、また地域文化の担い手として、これまでとは異なる新たな機能や役割を果たすことが期待される（経済産業省, 2020）ものの、それぞれの地域が今後目指すべき新たな方向性を地域の個別的な実情に応じて選択し、またそのイメージを内部で共有することは必ずしも容易でないことは想像に難くない。

かつては地域商業の振興を目的として、街の個性や魅力を明らかにする事を目的にエリアを対象にしたマーケティング調査の試みがしばしば実施された（東京商工会議所, 1984, 1987）。本研究もまた、そうした街のイメージ調査を意図するものである。

2. 研究目的

本報は、来街者によって記述された口コミ・レビューデータより得られた分散表現ベクトルを用いて商業地区を対象としたパーセプションマップを作成し、そこに満足度等の選好度の次元を加えることで選好空間モデルとし、従来の業種構成及び集客圏等による階層構造的な商業地システムとの整合を考慮しつつ「まち」の類型化を試みたうえで、その空間

分布を確認するなどし、当該分類・評価方法の有用性を検討するものである。

3. 方法

前報、鈴木（2020）を下敷きに次の手順で分析及び検討を進めた。

3.1 クチコミ・レビューの収集

本報が利用した口コミ・レビューデータはtripadvisor.comに対する2021年2月時点までの書き込みである。サイトの検索機能を用いて、全国を範囲に「商店街」、「市場」および「ショッピングモール」（近畿圏内のみ）のキーワードでリストアップされたPOIの中から、純然たる観光スポットや個別店舗等を除いたものを対象とした。こうして1,142件に及ぶ「街」に対する106,188件の口コミを収集した。

3.2 分散表現の獲得とその縮約

収集した口コミは「街」ごとに一つの文書に連結し、それぞれの「街」文書に対して分散表現ベクトルを付与する。分散表現獲得アルゴリズムにはDoc2Vecを用いて50次元のベクトルを生成した。

3.3 選考次元を加えたパーセプションマップの作成
得られた多次元データは機械学習による次元縮約手法t-SNEを利用し、二次元散布図として可視化する。

* 正会員 合同会社ファイン・アナリシス (Fine Analysis, LLC.)

〒545-0021 大阪府大阪市阿倍野区阪南町 2-24-12 Tel : 06-7503-9426 E-mail : suzuki@finea.co.jp

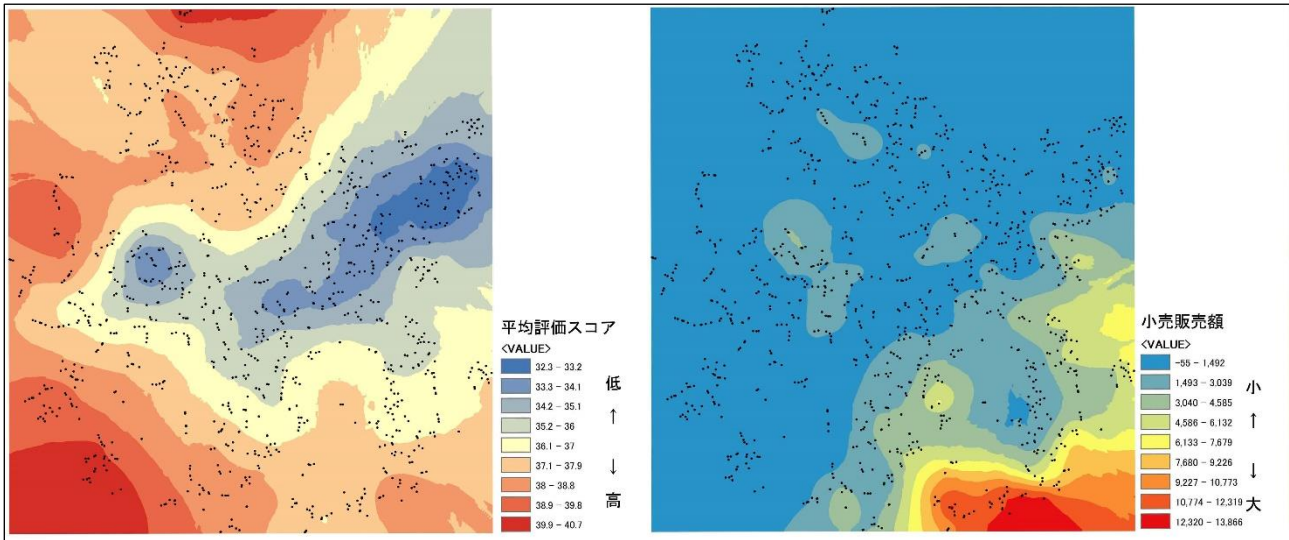


図1 分散表現ベクトルのt-SNE法による散布図可視化表現

A (左図) 平均評価スコアの推定等値線

B (右図) 小売販売額の推定等値線

この計算には Dmitry Ulyanov による Multicore-TSNE を利用した。ここで得られた散布図は高次元におけるそれぞれの街の類似度の相対的な距離を示すものであり、必ずしも縦軸および横軸に容易な解釈が成り立つものではない。そこでベクトル空間の解釈を補助するために、それぞれの街が持つ属性値(平均評価スコア, 500m メッシュの商業販売額他)にクリギング法を適用して散布図上の推計値を等値線で示した。これによってベクトル空間上のそれぞれの部分空間に布置される街の特徴が解釈される。

3.4 パーセプションマップ上の街の類型化

また散布図上の点群は、DBSCAN によってクラスタリングし、分析・解釈の便宜を図る。

4 結果と考察

4.1 パーセプションマップの概要

- ・中央部と周辺：評価スコアは、右上を除けば中心から周辺に向かうほど高くなっている(図1A)。数の上で多数を占める多くの平均的な商店街が散布図上の中央部を占め、その周辺にそれぞれに特徴的な機能や個性的なイメージを持つ街グループが位置していることが読み取れる。また中央部から見て右上部だけは評価スコアが低く、ここには書き込み数が少ないマイナーな商店街が集中していることが確認された。

- ・上部と左下部：一般的にレジャー・観光型の行動目的となる商業地の評価スコアは高い。散布図(図

1A) 上では上部のホットスポットは食をウリにした産直型市場、左下部の多くは温泉・観光地型の商店街が分布していることが確認された。

- ・右下部：百貨店が所在する都心型のターミナル型繁華街や地下街は右下部に集中する。それは小売販売額(図1C)の推計等値線からも確認される。また比較的若年者を集客していることも確認された。

4.2 DBSCAN による街のクラスタリング

パラメタの組み合わせによって、1,142 のポイント中 1,032 のポイントに対して 60 のクラスターを得ることができた。外れ値は最近隣のクラスターに所属させて分析を続けた(図2)。

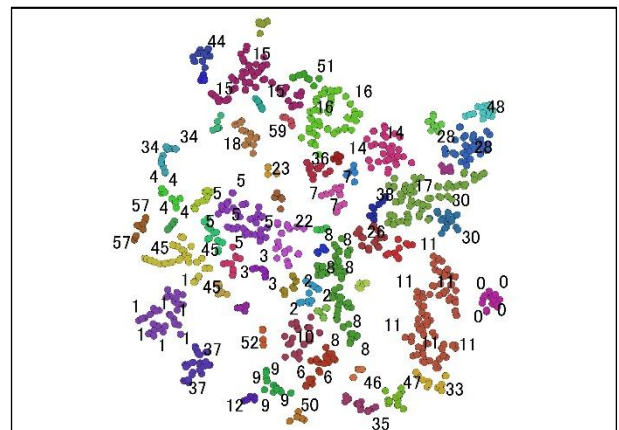


図2 散布図上点群に対するDBSCANクラスタリング結果

4.3 田村の商業地クラスタリングとの比較

商業地の階層的な中心地体系を計量的に実証した研究に田村(2008)がある。田村は商業統計等を用

いて京阪神都市圏内の基準地域メッシュにクラスター分析を適用し、商業地を5階層（大都市中核型、広域型、地域型、近隣型、街角ショップ型）及び2類型（地域型及び近隣型階層について大店型若しくは商店街型類型）に分類した（図3）。

前項で得られた60のクラスターの小売販売額中央値は4箇所を除いて200億円以上となり、広域型以上の高次の商業地がほとんどであることが判明した。これは観光・レジャー系の口コミサイトであるtripadvisorをデータ収集源としたことにより、近隣商業地について十分なデータ収集が出来なかった

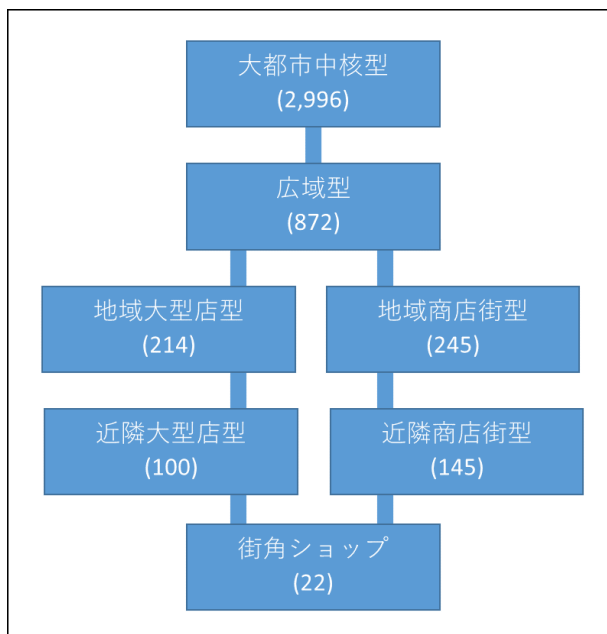


図3 田村(2008)による商業地クラスタリングと小売販売額
平均値（単位：億円） 筆者作図

事を意味する。今回の分析結果には、そうした選択バイアスがあることは留意されなくてはならない。

大都市中核型（販売額3,000億円程度以上）と思われる商業地区はパーセプションマップ右下部に集中している。中でも販売額で上位3位（クラスター35, 47, 46）はそれぞれ繁華街、地下街、専門商店街とみられるグループであり三大都市圏のターミナルや都心部でしか見られない（表2・図4下）。一方「大都市型」クラスターに隣接しているクラスター11及び33は都心部の娯楽型ショッピングセンターに加えて郊外のアウトレットモールやリージョナルショッピングセンターも含み、「大都市中核型」より一階層低い「広域型」の中心地とみられ、より遍在して

立地していることが確認できる（表1・図4右上）。

マップ右下部とは対称的に左下部には、地方の有名な観光地型商店街や温泉街とみられるクラスターが位置する。その中間に位置する一群のクラスターは興味深い。谷中ぎんざ商店街や通天閣本通り商店街等、いずれも主に東京若しくは大阪都市内部にありながら昭和レトロな町並みや情緒を魅力にもつ半ば「観光地」化した街のクラスターである。

5 まとめ

口コミ分散表現ベクトルを用いた街の類型化によって、それぞれに異なる魅力を持つ街グループが抽出できたことは収穫であった。また商業地の階層的な中心地体系に沿った街の判別が可能であることを確認した。

謝辞：本研究は、東京大学空間情報科学研究センター共同研究（No. 256）による成果である（利用データ：商業集積統計（2016年）、【空間配分版】2010年近畿都市圏 人の流れデータセット）。研究遂行にあたり関係諸氏には日頃より、有益なご助言を戴いた。ここに感謝の意を表する。

参考文献

- 経済産業省（2020）「地域の持続可能な発展に向けた政策の在り方研究会 報告書」
- 東京商工会議所（1984）『THE SAKARIBA タウンイメージ東京』
- 東京商工会議所（1987）『TOWN IDENTITY タウンイメージ東京Ⅱ』
- 田村正紀（2008）「大都市圏での中心地体系」『立地創造』p105-140 白桃書房
- 鈴木英之（2020）「観光地戦略マップ」の可視化 — 口コミ・レビューデータを用いた観光地選好空間モデル—, 地理情報システム学会学術講演論文集, 30, CD-ROM.
- 三浦展（2020）『愛される街』而立書房
- Q.Le, T.Mikolov, 2014. Distributed Representations of Sentences and Documents. International conference on machine learning. PMLR,
- VAN DER MAATEN, Laurens; HINTON, Geoffrey, 2008. Visualizing data using t-SNE. Journal of machine learning research

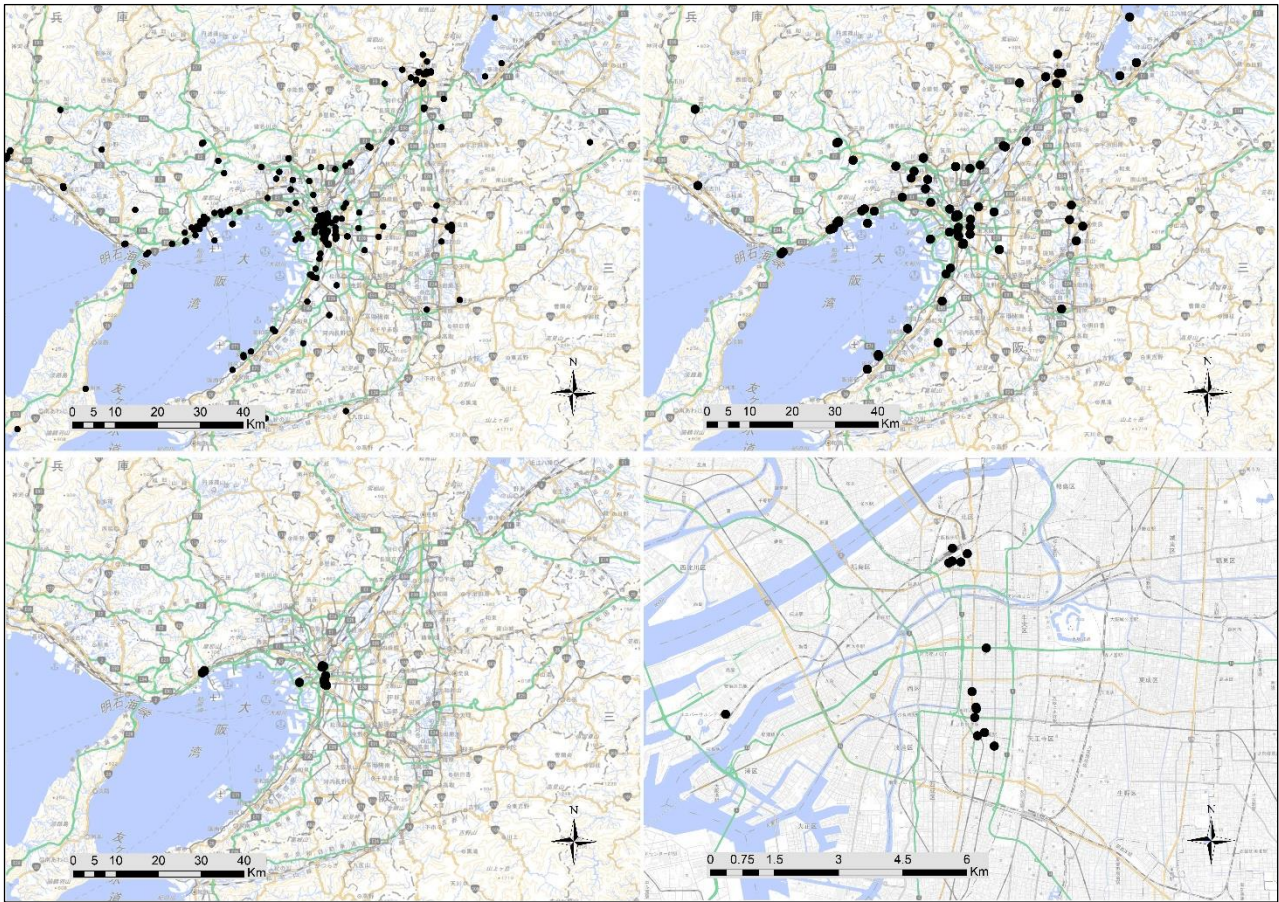


図4 街の地理的空間分布 (左上) 全ての街

(左下) クラスタ-35, 46 及び 47 (大都市中核型) の街

(右上) クラスタ-11 及び 33 (広域型) の街

(右下) クラスタ-35, 46 及び 47 (大都市中核型) の街(拡大図)

表1 クラスタ-11 及び 33 (広域型) に所属する街リスト

クラスター	名称	所在地	ロコミ数
11	阪急三番街	大阪市北区	133
11	イオンモールKyoto	京都市南区	86
11	三井アウトレットパーク 大阪鶴見	大阪市鶴見区	82
11	ミント神戸	神戸市中央区	75
11	神戸国際会館	神戸市中央区	73
11	くずはモール	大阪府枚方市	73
11	イオンモール 大阪ドームシティ	大阪市西区	70
11	六甲アイランド	神戸市灘区	67
11	りんくうタウン	大阪府泉佐野市	66
11	ヘップファイブ	大阪市北区	63
33	神戸ハーバーランド	神戸市中央区	1048
33	神戸三田プレミアム アウトレット	神戸市北区	477
33	グランフロント大阪	大阪市北区	404
33	エキスポシティ	大阪府吹田市	384
33	阪急西宮ガーデンズ	兵庫県西宮市	332
33	なんばパークス	大阪市浪速区	200
33	ららぽーと甲子園	兵庫県西宮市	180
33	大阪ステーションシティ	大阪市北区	180
33	三井アウトレットパーク マリンピア神戸	神戸市垂水区	178
33	ルクア1100	大阪市北区	147

表2 クラスタ-35, 46 及び 47 (大都市中核型) の街リスト

クラスター	名称	所在地	ロコミ数
35	心齋橋	大阪市中央区	1011
35	ミナミ (難波)	大阪市中央区	710
35	神戸元町商店街	神戸市中央区	266
35	でんでんタウン	大阪市浪速区	111
(以下は近畿圏以外の該当地区例)			
35	銀座歓楽街	東京都中央区	1158
35	横浜元町商店街	横浜市中区	634
35	秋葉原	東京都千代田区	629
35	歌舞伎町歓楽街	東京都新宿区	381
46	千日前道具屋筋商店街	大阪市中央区	231
46	船場センタービル	大阪市中央区	62
(以下は近畿圏以外の該当地区例)			
46	かっぱ橋道具街	東京都台東区	530
46	日暮里鐵道街	東京都荒川区	60
47	三宮センター街	神戸市中央区	626
47	御堂筋	大阪市中央区	354
47	ユニバーサル・シティウォーク大阪	大阪市此花区	193
47	なんばウォーク	大阪市中央区	111
47	ルクア	大阪市北区	98
47	ホワイトティウめだ	大阪市北区	94
47	ヒルトンプラザ大阪 イースト/ウエスト	大阪市北区	74
47	ハービスPLAZA/ハービスPLAZA ENT	大阪市北区	73
47	なんばCITY	大阪市中央区	67
47	ディアモール大阪	大阪市北区	49