Title:
Multi-Scale Pedestrian Walking Preference Prediction Model in Historic Area Based on Second-Order Markov Chain

Authors:
Chenlei Liu, 22013085009@stu.hqu.edu, Huaqiao University, Xiamen Key Laboratory of Integrated Application of Intelligent Technology for Architectural Heritage Protection, Xiamen University
Nan Wu, archt_wn@qq.com, Huaqiao University, Xiamen Key Laboratory of Integrated Application of Intelligent Technology for Architectural Heritage Protection, Xiamen University
Zhisheng Huang, hzs_mahakala@foxmail.com, Huaqiao University, Xiamen Key Laboratory of Integrated Application of Intelligent Technology for Architectural Heritage Protection, Xiamen University
Shaosen Wang, ymcai@xmu.edu.cn, Xiamen Key Laboratory of Integrated Application of Intelligent Technology for Architectural Heritage Protection, Xiamen University

Keywords:
Second-order Markov Chain, Multiscale, Historic Area, Walking Preference Model, Historic Area Renewal

DOI: 10.14733/cadconfP.2024.264-269

Introduction:

In recent years, within the research scope of urban planning, street vitality has often been the focus of attention concerning the spatial distribution of the macro density or its specific route choices in relation to environmental factors [1] [2]. However, these two directions of inquiry are often studied independently, with fewer studies addressing the mechanisms behind the influence of pedestrian route choices on pedestrian density distribution and their combined relationship with environmental factors. This may obscure the true pedestrian intentions behind density distributions and lack practicality. Additionally, determining effective scales and corresponding evaluation criteria in complex, geographically heterogeneous historic districts remains a challenge.

Hence, we aim to devise a method to assist in the planning decisions for historic district revitalization from a more flexible multi-scale perspective. This study introduces a model based on second-order Markov chains to predict pedestrian density and trajectory distribution influenced by the environment within the historic district, where the measurement cost of transition probabilities is significant. To address this, we propose a method based on stroke cluster to characterize the route choice features combined with deep learning to derive pedestrian route choice preference patterns, which can predict transition probabilities based on environmental factors. This approach can aid planners in guiding pedestrian trajectories by altering environmental factors, thereby achieving density distribution regulation objectives.

Main Idea:

In this section, we introduce the proposed approach, which includes four parts: data collection, pedestrian route choice preference mode, pedestrian density, and distribution prediction model, using the pedestrian walking preference model to predict.

Proceedings of CAD'24, Eger, Hungary, July 8-10, 2024, 264-269

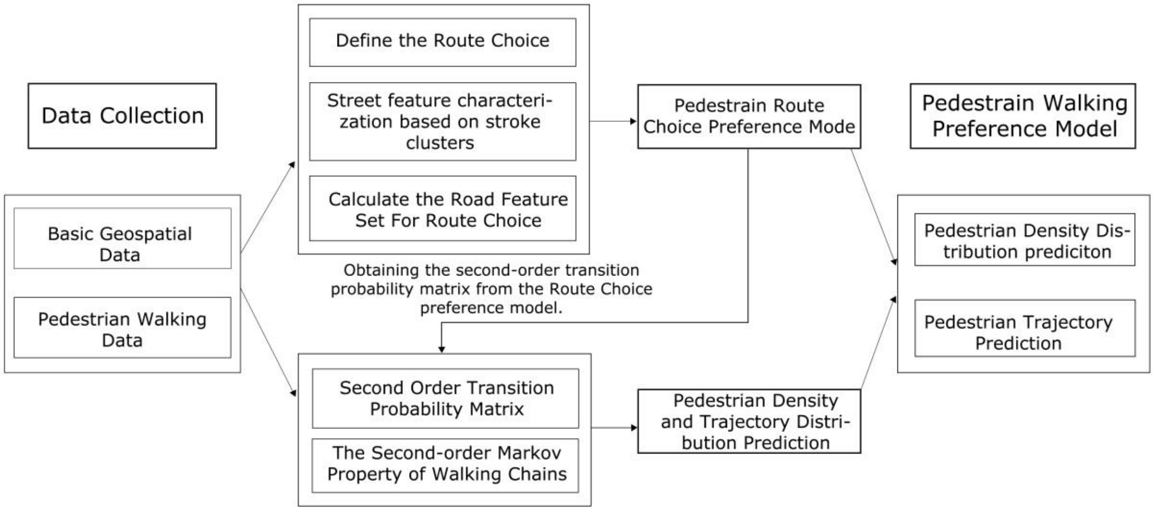

Fig. 1: Flow chart of the method.

Data Collection and Pretreatment

We selected the historic conservation district of West Street in Quanzhou, Fujian Province, China, along with its adjacent areas, as the object of our research.

Geospatial data and pedestrian walking behavior data were collected to obtain route choice features assessment indicators and probability distributions.

Route choice features were extracted from the geospatial data collected during the National Day holiday and regular holidays in October 2023. This dataset, acquired from the Baidu API, included building data, heat map data, POIs data, and road data.

Furthermore, pedestrian walking behavior data consisted of two main components: (a) obtaining tourist trajectory data for 2022-2023 from the Six Feet GPS Travel Community website (http://www.foooooot.com/), and (b) collecting pedestrian walking video data from 47 intersections within the study area between October 1st and 5th, 2023. These datasets were processed to derive the probability distribution of pedestrian route choices at intersections.

The Distribution of Pedestrian Density and Route Prediction Model Based on Second-Order Markov Chain

Pedestrians walking from point A to point B within the historic district have various possible routes, which can be described using the concept of a chain in graph theory, as illustrated in Fig. 2(a). The chain $(v_1, e_1, v_2, e_2, \cdots, v_k)$ represents a sequence of alternating vertices and edges, where $e_i = (v_i, v_{i+1})$. We divide the road network into multiple segments and abstract them as edges in the graph G, assuming that a pedestrian can transition from v_i to its adjacent vertex v_{i+1} in a unit of time. The process of a person being at v_i and selecting the next vertex v_{i+1} to move to is referred to as the route selection process. e_{i-1} is termed the previous route, and e_i is termed the next route. Influenced by the collective road environment factors, encompassing both previous routes and all next routes, the resulting probability distribution of route choices is referred to as route selection preference, as illustrated in Fig. 3(a).

The walking process exhibits memory, manifested by the combined influence of the previous route and the optional next route when choosing the next direction at the intermediate point.

Therefore, pedestrians' walking process can be described by a second-order Markov chain, as illustrated in Fig. 2(b).

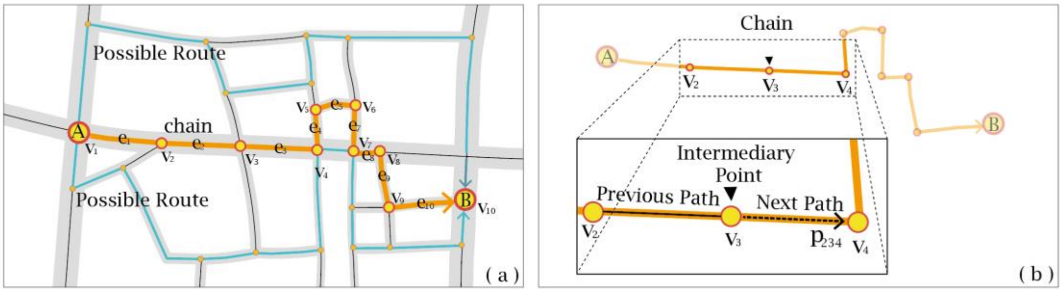

Fig. 2: (a) The chain $(v_1, e_1, v_2, e_2, \cdots, v_k)$ and possible routes, (b) The route selection process described by second-order Markov chains.

The state transition probability of a second-order Markov chain is

$$P_{kij} = P(X_{t+1} = j \mid X_t = i, X_{t-1} = k) \tag{1}$$

where $X_t = i$ indicates that a pedestrian is at vertex v_i at time t. The specific value of the transition probability can be obtained from the later-constructed pedestrian route selection preference pattern. If there is no connection between two vertices, the corresponding transition probability is 0. The second-order density $\rho_{ki}^{(t)}$ is defined as the proportion of individuals currently at v_i and were at v_k in the previous step to the total number of individuals. We can obtain

$$\rho_{ij}^{(t+1)} = \sum_{k=1}^{n} \rho_{ki}^{(t)} P_{kij} \tag{2}$$

which describes the evolution of the distribution over time with the total number of pedestrians remaining constant.

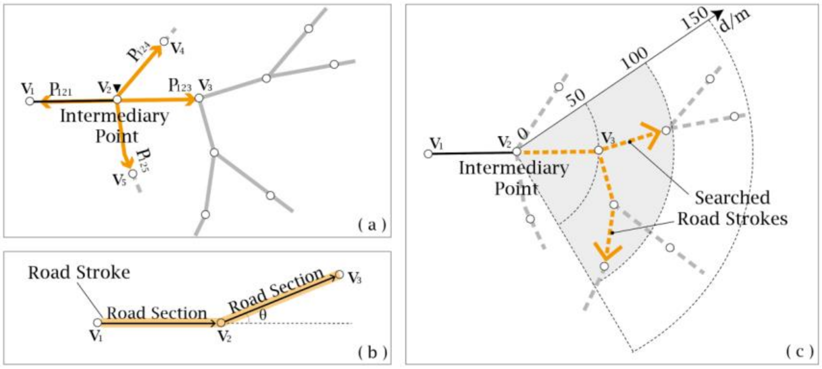

Fig. 3: (a) The route selection preferences, (b) Two smoothly connected road sections form a road stroke, (c) Searched road strokes with the same first edge (v_2, v_3) when the scale of the road network is $d = 50\text{m}$ and the cut length is $L = 100\text{m}$.

Mining Pedestrian Route Preference Patterns Influenced by Street Feature Factors

Based on former research results [3], this study employs commonly used environmental factors to evaluate pedestrian preferences for road features. In terms of the built environment, indicators for evaluating road features include road width, length, tortuosity, width-to-height ratio (W/H), plot ratio, building density, POIs (Points of Interest) diversity, as well as the density of POIs such as dining, shopping, daily necessities, public facilities, accommodation, and scenic spots [4]. Regarding the inter-In small-scale road networks, features of short roads are often characterized using stroke-based methods and shared evaluation indicators with other strokes, while in large-scale networks, long roads are often divided into smaller segments to capture road detail features. To investigate pedestrian walking preferences, it is necessary to quantify the road features that influence pedestrian route choice. However, historic district road networks, unlike regular grid systems, consist of irregular and fragmented road segments, making it challenging to find appropriate scales for computing assessment indicators. Additionally, Previous methods have not adequately captured all the road branch features that must be taken into account in route choice.

We introduce a route choice feature characterization method based on stroke cluster to compute the feature set for each route choice. Stroke [5], depicted in Fig. 3(b), is a naturally extended route created by the smooth connection of multiple road segments. Within a stroke, a smaller deflection degree between road segments corresponds to a higher overall smoothness. In small-scale road networks, the characterization capability of road stroke excels beyond that of individual road segments. In large-scale networks, potential distortion can be mitigated by restricting the length of the stroke path.

As shown in Fig. 3(c), when a cluster of strokes extends from an intermediary point, the edge from which these strokes originate is termed as the starting edge of the stroke cluster. Depending on the desired scale, one can flexibly set the truncation lengths of strokes. By employing a Depth-First Search (DFS) method, one can obtain all strokes extending outward from a central point with a length not exceeding L, which can be grouped into multiple clusters based on different starting edges. The comprehensive feature set of a stroke cluster, where each stroke's feature set is weighted by its straightness, is defined as the integrated feature set of the starting edge. During a route choice process, all available next edges are in competition. Hence, all edges connected to the intermediary point should be considered, not limited to just one pair of previous and next edges. Consequently, the comprehensive feature set that fully describes a route choice process is formed by merging the integrated feature sets of all edges connected to a central point. It is important to note that at a central point, multiple pairs of previous-next edges may undergo route choice processes, distinguished by the placement sequence of the integrated feature sets in the route choice feature set, which follows the order of the previous edge, next edge, and others.

This approach avoids redundant computations of data at different scales, thereby enhancing efficiency. Moreover, the introduction of a stroke cluster not only reflects the feature of individual roads but also captures the overlay of attributes from underlying potential roads, thereby providing a more comprehensive perspective.

Experimental Result

Multiple sets of road networks were configured with different scales paired with varying truncation lengths, utilizing a route choice feature characterization method based on stroke clusters. Multiple route choice features and their corresponding probability data were input into an NGO-BP neural network to establish mapping relationships. The mapping relationship with the best error performance was selected as the pedestrian walking preference pattern for dynamic simulation experiments to obtain predictive results. Comparing experimental predictions of density and trajectory distribution during the 2023 National Day period with actual average density and trajectory data revealed that the model accurately reflected overall density distribution differences and identified some congested and vulnerable areas. Furthermore, the Kullback-Leibler Divergence (KLD) of trajectories, used to describe the error of trajectory distribution, does not exceed 0.1 within 12 steps, indicating the model's ability to recognize route context and simulate real pedestrian walking trajectories, as shown in Fig. 4.

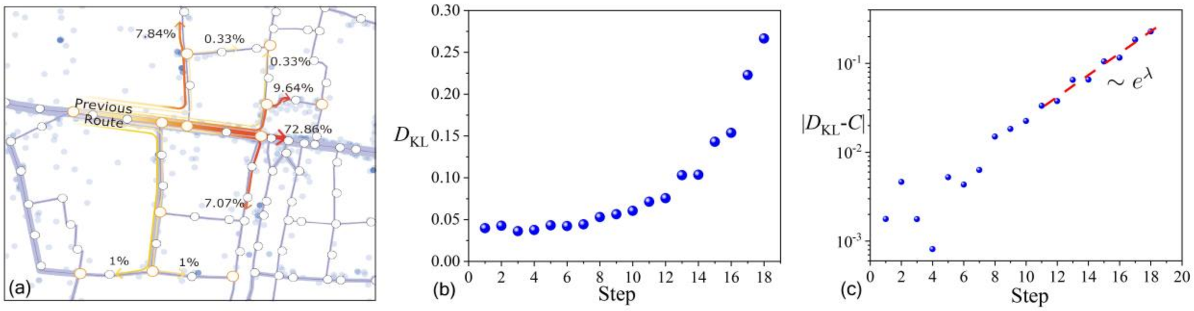

Fig. 4: (a) In a 50-meter-scale road network, with a truncation length of L=200 meters, and given the previous direction, the probability distribution of pedestrian trajectories for eight steps, (b) The relationship between KLD and the number of steps between predicted probability and actual probability under the setting conditions of Figure 4(a), (c) The KLD in Figure 4(b) is approximately an exponential function of the number of steps, where $C = 0.0038$ and $\lambda = 1.22$. Please note that selecting a logarithmic vertical axis will amplify the fluctuations in the image within fewer steps.

Conclusions:

Considering both trajectory and density distributions can better reveal the microscopic essence of pedestrian movements within historic district and obtain more comprehensive pedestrian walking information. This facilitates direct guidance on altering environmental factors to influence pedestrian trajectories, thereby regulating density distributions of historic district.

We propose a multi-scale pedestrian preference model based on a second-order Markov chain to predict pedestrian density and trajectory distributions. By considering the influence of stroke clusters, we mitigate model distortion when the road network scale changes. Incorporating memory-based route preference patterns enables a more accurate simulation of pedestrian route choices influenced by the historic districts. Predicting and comparing density and trajectories during the National Day holiday in 2023 with regular holidays demonstrate the model's ability to capture variations in environmental factors within historic districts, reflected in density and trajectory distributions. This approach assists planners in regulating the pedestrian distribution and walking trajectory by revitalizing environments in historic districts, facilitating more nuanced preservation and revitalization efforts.

Chenlei Liu, https://orcid.org/0009-0007-5726-0259.
Nan Wu, https://orcid.org/0009-0000-6349-6347.
Zhisheng Huang, https://orcid.org/0009-0006-5688-9023

References:

[1] Zhang, L.; Zhang, R.; Yin, B.: The impact of the built-up environment of streets on pedestrian activities in the historical area, Alexandria Engineering Journal, 60(1), 2020, 285-300. https://doi.org/10.1016/j.aej.2020.08.008
[2] Ben-Akiva, M.; Bierlaire, M.: Discrete choice models with applications to departure time and route choice, 2003, 7-37. https://doi.org/10.1007/0-306-48058-1_2
[3] Wu, J.; Zhao, C.; Li, C.; Wang, T.; Wang, L.; Zhang, Y.: Non-linear Relationships Between the Built Environment and Walking Frequency Among Older Adults in Zhongshan, China, Frontiers in Public Health, 9, 2021. https://doi.org/10.3389/fpubh.2021.686144
[4] Zou, H.; Liu, R.; Cheng, W.; Lei, J.; Ge, J.: The Association between Street Built Environment and Street Vitality Based on Quantitative Analysis in Historic Areas: A Case Study of Wuhan, China, Sustainability, 15(2), 2023, 1732-1754. https://doi.org/10.3390/su15021732

[5] Thomson, R. C.: The 'stroke' Concept in Geographic Network Generalization and Analysis, Progress in Spatial Data Handling: 12th International Symposium on Spatial Data Handling, Springer Berlin Heidelberg, 2006, 681-697. https://doi.org/10.1007/3-540-35589-8_43

Proceedings of CAD’24, Eger, Hungary, July 8-10, 2024, 264-269