CityScouter: Exploring the Atmosphere of Urban Landscapes and Visitor Demands with Multimodal Data

Yuki Kubota
Tokyo Institute of Technology
Tokyo, Japan
kubota@miubiq.cs.titech.ac.jp

Soto Anno
Tokyo Institute of Technology
Tokyo, Japan
anno@miubiq.cs.titech.ac.jp

Tomomi Taniguchi
TANIGUCHI TOMOMI DESIGN
Tokyo, Japan
i.talk.him@gmail.com

Kosei Miyazaki
Hyogo University
Hyogo, Japan
kosei@hyogo-dai.ac.jp

Akira Tsujimoto
Nikken Sekkei Ltd
Tokyo, Japan
akira.tsujimoto@nikken.jp

Hiraki Yasuda
Nikken Sekkei Ltd
Tokyo, Japan
hiraki.yasuda@nikken.jp

Takayuki Sakamoto
Nikken Sekkei Ltd
Tokyo, Japan
sakamotot@nikken.jp

Takaaki Ishikawa
Yahoo Japan Corporation
Tokyo, Japan
tishikaw@yahoo-corp.jp

Kota Tsubouchi
Yahoo Japan Corporation
Tokyo, Japan
ktsubouc@yahoo-corp.jp

Masamichi Shimosaka
Tokyo Institute of Technology
Tokyo, Japan
simosaka@miubiq.cs.titech.ac.jp

ABSTRACT

This paper proposes a novel demo application named CityScouter that utilizes multimodal data to analyze various aspects of urban characteristics quantitatively. Existing studies have proposed systems to examine either the physical characteristics of cities or the nature of people residing there. However, there is a lack of systems that analyze the characteristics of cities from both the physical and the residents' aspects. CityScouter addresses this challenge by leveraging computer vision technologies to quantify the quality of the urban landscape atmosphere and combining it with location information and user search history to reveal the desires of people visiting the area. The application is user-friendly and compatible with mobile devices, enabling users to conveniently enhance their understanding of cities while exploring them. Additionally, we provide reviews from urban development experts, offering insights into the applicability of our application. Furthermore, we showcase the usefulness and user experience of CityScouter through live demonstrations at the conference venue.

CCS CONCEPTS

• **Human-centered computing → Ubiquitous and mobile computing systems and tools**.

KEYWORDS

Urban ladscape analysis, search history, urban planning

ACM Reference Format:
Yuki Kubota, Soto Anno, Tomomi Taniguchi, Kosei Miyazaki, Akira Tsujimoto, Hiraki Yasuda, Takayuki Sakamoto, Takaaki Ishikawa, Kota Tsubouchi, and Masamichi Shimosaka. 2023. CityScouter: Exploring the Atmosphere of Urban Landscapes and Visitor Demands with Multimodal Data. In *Proceedings of (UbiComp/ISWC '23 Adjunct).* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

UbiComp/ISWC '23 Adjunct, October 8–12, 2023, Cancún, Mexico

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In recent years, there has been an increasing interest in quantitatively analyzing urban characteristics based on the accumulated large-scale data in cities. Quantifying urban characteristics offers various benefits, such as their application in urban development and urban planning projects. Numerous studies have developed systems for analyzing the physical characteristics of cities by leveraging diverse data such as mobility logs [8] and POI (Points of Interest) information [3]. These systems enable us to reveal cities' characteristics, which highly contribute to urban development projects.

On the other hand, some studies shift the focus from the characteristics of the cities to the features of the people staying in them [4, 6]. For example, Sakamoto et al. [7] proposed a framework and VR system to understand the latent desires of visitors in a specific region by utilizing users' GPS logs and search histories. This approach allows us to understand what people seek in a city, providing insights that cannot be obtained solely by analyzing the city's physical characteristics.

However, while there have been systems analyzing the characteristics of the city itself or focusing on city visitors, there has been no attempt to analyze both aspects simultaneously. It has been suggested that the physical characteristics of a city and the characteristics of its visitors or residents mutually influence each other, contributing to the formation of the city's features [5]. That is, by analyzing these two aspects simultaneously, it becomes possible to gain a deeper understanding of the nature of a city.

To address this challenge, we propose a novel mobile application called **CityScouter** that leverages the diverse multimodal data available in urban environments to uncover both the physical characteristics of cities and the characteristics of visitors. Specifically, our proposed application aims to quantitatively assess the goodness of the atmosphere of urban physical appearance (i.e., urban landscapes) and the latent desires of people visiting the city. We utilize computer vision technologies to quantitatively assess the quality of urban landscapes based on image-based features. Furthermore, our application incorporates the method proposed in previous study [7], which analyzes the desires of users based on their GPS and search history to reveal the demands of people staying in a city.

CityScouter runs on mobile devices so that users can better understand the cities while walking through them on-site. We invite urban development experts to use the application and share their reviews, enabling us to discuss the potential contribution of CityScouter to urban development projects. Additionally, we provide a live demonstration of our application at the conference venue, allowing users to experience it firsthand.

Figure 1: The overview workflow of CityScouter. Users can launch the app and capture a landscape image to check the atmosphere score of the captured landscape and the desires of nearby visitors.

2 DEMO APPLICATION

In this chapter, we provide detailed information about our proposed demo application. CityScouter is a mobile application that runs on iPad and iPhone devices. Figure 1 illustrates the workflow of our application. To begin using CityScouter, users simply need to launch it and capture a landscape image. The application automatically calculates the atmosphere score of the captured image, which can be viewed on the screen. Moreover, based on the user's location when capturing the photo, the application provides information on frequently searched queries and visitors' desires in the nearby area. These analysis results are logged within the application, enabling users to review them at any time and make comparisons between different places. Additionally, users receive notifications when entering areas with exceptionally high desire scores. This feature allows users to passively receive information, enriching their city exploration experience and making it more engaging.

In Figure 2, we present a detailed view of the screen that displays the analysis results of CityScouter. The analysis screen provides users with comprehensive information about the captured landscape image, the location's regional statistics, and visitors' characteristics based on nearby search queries. The screen's top section demonstrates a radar chart showing the atmosphere score of the taken picture. It also includes statistical values related to the landscape image, such as the average color distribution and ImageNet [1] class labels. The middle section provides the user's location information and regional statistical data of the surrounding area, including crime rate, percentage of the elderly population, and land prices. The bottom section presents frequently searched queries within the vicinity, highlighting the topics or POIs that attract visitors. Additionally, it displays scores representing the desires or interests of the visitors in that specific location. By presenting this comprehensive information set, CityScouter enables users to gain a deeper understanding of a city's characteristics from multiple perspectives.

3 METHODS

In this chapter, we provide detailed explanations of the two statistical calculation techniques utilized in CityScouter: (1) the method for quantifying the atmosphere score for landscape images and (2) the method for calculating potential desires based on user GPS and search histories.

3.1 Landscape Image Atmosphere Analysis

The "atmosphere" of a landscape is an important indicator of the impression people have of a place, but at the same time, it is a challenging metric to measure due to its subjective and abstract nature. Therefore, we attempt to measure the atmosphere of a landscape comprehensively by considering four aspects: atmosphere, consistency, weirdness, and sense of luxury. By designing multiple attributes, we can reveal a landscape's detailed characteristics. Furthermore, instead of providing a scalar value as the ground truth for the atmosphere of images, we assign it based on the voting rates from users across multiple evaluation grades. This allows us to capture the variability and uncertainty associated with users' perceptions of the atmosphere.

To quantify these attribute score distributions for urban landscapes, we develop a model that leverages computer vision techniques. Specifically, we utilize a pre-trained MobileNet [2], which has been trained on the ImageNet [1] dataset, as an image feature extractor. The extracted features are then fed into two-layer fully connected layers, which predict the scores for each attribute. To estimate the scores in the distribution form, we set the number of neurons in the final layer as the same number as in the ground truth evaluation grade, and a softmax function is applied to normalize the output.

Figure 2: The details of the screen display the CityScouter's analysis results. The screen provides a comprehensive overview of the analysis results, combining information about the captured landscape, the user's location, and the search history of visitors in the vicinity.

3.2 Search Query-based Users' Desire Analysis

CityScouter utilizes large-scale user GPS and search history data to gain a quantitative understanding of the preferences and desires of visitors in a city [7]. First, for each location, users with a history of visiting the area will be specified from the GPS records. Next, the search histories of these users are extracted to identify search queries associated with the visited regions.

For each query, we perform weighting adjustment using TF-IDF based on the total number of searches in the entire region and the total number of searches within the specific target region. This adjustment is made to give significant weight to queries that are prominently searched within the region. In addition, we conduct an online survey using crowdsourcing to assess the strength of the relationship between search query terms and 12 desire categories designed in advance. This allows us to transform each search query into a 12-dimensional vector. Finally, the desire vector associated with a specific region is calculated by taking the weighted sum of the vectors corresponding to each search query in the analysis area, where the weights are determined based on TF-IDF.

4 EXPERIMENTS

4.1 Datasets

In this chapter, we provide details about the datasets used to enable the key functionalities of CityScouter.

4.1.1 Image Atmosphere Score Data. We first gather a dataset of 2,305 landscape images from various locations. Then, the atmosphere scores for these images are assigned using the Yahoo! CrowdSourcing platform. This platform allows us to outsource simple tasks to an unspecified number of users. Each participant is assigned a masked ID to ensure privacy, protecting their personally identifiable information.

Through the crowdsourcing platform, we collect responses from multiple users using a five-scale rating system to evaluate the atmosphere quality. Users are presented with five options: -2, -1, 0, +1, and +2, along with the landscape image. They choose the option that best represents their perception of the atmosphere in a single-choice format. The user responses are then used to calculate the distribution of scores for each evaluation grade. As a result, we obtain a five-dimensional ground truth score distribution vector that represents the atmosphere quality for each of the 2,305 landscape images in our dataset.

4.1.2 User Mobility Log and Search Query Data. In this study, we utilize users' GPS records and search histories from Yahoo! JAPAN's mobile application as our source of location information and search query data. To ensure the privacy and confidentiality of users, we obtain location information only from users who have provided their consent explicitly. Furthermore, we apply masking techniques to anonymize each user's ID, ensuring that no personally identifiable information is disclosed or associated with the data.

4.2 Model Performance Evaluation

Since CityScouter provides the atmosphere scores predicted by the model, quantitatively evaluating the model inference performance is a critical perspective to understand the reliability of our application. In this section, we describe the result of the performance evaluation of our atmosphere scoring model. To evaluate the performance of the scoring model, we use two standard evaluation metrics, mean absolute error (MAE) and Pearson linear correlation coefficient (PLCC). Since both the ground truth and the estimated value are given as a distribution over a five-point scale, we first convert this distribution to a scalar value by taking the weighted average of each evaluation grade score (i.e., $-2 \sim +2$) multiplied by the voting rate from users.

Table 1: Model performance of for each attribute.

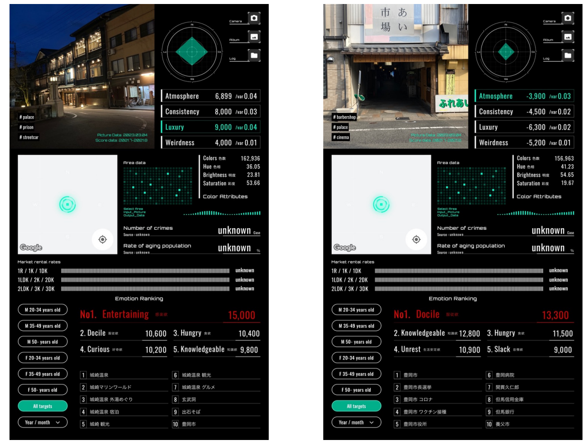
Attribute	MAE	PLCC
Atmosphere	0.255 ± 0.013	0.755 ± 0.016
Consistency	0.289 ± 0.030	0.728 ± 0.051
Weirdness	0.319 ± 0.008	0.694 ± 0.029
Luxury	0.302 ± 0.016	0.748 ± 0.035

Table 1 shows the model's performance for each attribute score based on the dataset divided using a 5-fold cross-validation approach. The results demonstrate that the models can accurately infer scores for all attributes, indicating the effectiveness of applying CNN models pre-trained on a large-scale image dataset.

4.3 User Study

In this section, we present reviews from actual urban policy and urban development experts who used the application. These reviews discuss how our proposed application can be effectively utilized for urban projects.

- **Reviewer 1 (Assistant Chief Data Officer of governments)**
  In Japan, numerous local governments aim to attract tourists worldwide. Enhancing the scenery to appeal to tourists is vital in this endeavor. However, the concept of an attractive landscape is subjective, making it challenging for local government officials to judge. Additionally, even if tourism consultants or architectural experts suggest improvements, these efforts may be incomplete, temporary, or unsustainable if the residents do not comprehend what constitutes a good landscape. With CityScouter, both local government officials and ordinary citizens can identify attractive landscapes and areas needing improvement. This empowers them to continuously enhance their surroundings by themselves. Moreover, enabling residents to sustainably improve the landscape encourages a consistent influx of tourists and contributes to economic development.
  Figure 3 shows the analysis results of the landscapes of Kinosaki Onsen, a renowned tourist destination in Japan. Based on the analysis results, the locations where the landscape has been improved exhibit higher atmosphere scores (Figure 3(a)), while the areas that require improvement show lower scores (Figure 3(a)). This demonstrates the potential of the application to achieve objective evaluation of landscapes.



(a) (b)

Figure 3: Examples of applying CityScouter to a recently renovated tourist site: (a) Analysis results at a location where landscape improvement has been implemented. (b) Analysis results at a location where landscape improvement is not sufficient.

- **Reviewer 2 (City Planner)**
  In this application, the atmosphere of places previously considered subjective and difficult to measure can now be obtained as scores of both urban visitors' desires and landscapes' appearance perspectives. This allows for sharing and discussing the vague indicator of urban atmosphere in a comparable manner with others. Furthermore, comparing places with similar desire scores or landscape atmosphere scores makes it possible to discuss the underlying factors and other related aspects in detail. Moreover, CityScouter includes several convenient features to enhance the user experience. For instance, it provides notifications when entering areas with prominent desire scores, which enable users to grasp new insights and discoveries. The application also allows users to sort the log of analyzed results based on desire scores or image atmosphere scores, facilitating easier access to and review of the analysis outcomes. These functionalities can allow us to discover the region characteristics that have not been fully explored yet and to understand the relevance among different cities.

5 CONCLUSION

We proposed CityScouter, a new mobile application that integrates landscape image data with city visitors' GPS history and search history to analyze both the physical characteristics of urban environments and the latent desire of visitors. By leveraging multimodal datasets, CityScouter offers unique insights that were previously inaccessible. The application is designed for mobile devices and features a user-friendly interface, making it convenient for on-site usage in cities. We demonstrate the effectiveness of CityScouter through numerical experiments and feedback from urban development experts.

REFERENCES

[1] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. ImageNet: A large-scale hierarchical image database. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

[2] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. 2017. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv preprint arXiv:1704.04861* (2017).

[3] Antti Jylhä, Yi-Ta Hsieh, Valeria Orso, Salvatore Andolina, Luciano Gamberini, and Giulio Jacucci. 2015. A Wearable Multimodal Interface for Exploring Urban Points of Interest. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*. 175–182.

[4] Fabio Miranda, Harish Doraiswamy, Marcos Lage, Kai Zhao, Bruno Gonçalves, Luc Wilson, Mondrian Hsieh, and Cláudio T. Silva. 2017. Urban Pulse: Capturing the Rhythm of Cities. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 791–800.

[5] Nikhil Naik, Scott Duke Kominers, Ramesh Raskar, Edward L Glaeser, and César A Hidalgo. 2015. *Do People Shape Cities, or Do Cities Shape People? The Co-evolution of Physical, Social, and Economic Change in Five Major U.S. Cities*. Working Paper. National Bureau of Economic Research.

[6] Muhammad Rizwan, Wanggen Wan, Ofelia Cervantes, and Luc Gwiazdzinski. 2018. Using Location-Based Social Media Data to Observe Check-In Behavior and Gender Difference: Bringing Weibo Data into Play. *ISPRS International Journal of Geo-Information* 7, 5 (2018).

[7] Takayuki Sakamoto, Hiraki Yasuda, Akira Tsujimoto, Mutsuhiro Iwamoto, Yuta Hayakawa, Takuto Usami, Kota Tsubouchi, and Masamichi Shimosaka. 2019. City-Atmosphere: VR Image to Glimpse Wishes in the Air. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*. 342–345.

[8] Arianna Salazar-Miranda, Fan Zhang, Maoran Sun, Pietro Leoni, Fábio Duarte, and Carlo Ratti. 2023. Smart curbs: Measuring street activities in real-time using computer vision. *Landscape and Urban Planning* 234 (2023), 104715.