Building and Environment 225 (2022) 109563

Contents lists available at ScienceDirect

Building and Environment

journal homepage: www.elsevier.com/locate/buildenv

Small public space vitality analysis and evaluation based on human trajectory modeling using video data

Tong Niu [a], Linbo Qing [a,*], Longmei Han [b], Ying Long [c,d], Jingxuan Hou [e], Lindong Li [a], Wang Tang [a], Qizhi Teng [a]

[a] College of Electronics and Information Engineering, Sichuan University, Chengdu, Sichuan, 610064, China
[b] Chengdu Institute of Planning & Design, Chengdu, Sichuan, 610072, China
[c] School of Architecture and Hang Lung Center for Real Estate, Tsinghua University, Beijing, 100084, China
[d] Key Laboratory of Eco Planning & Green Building, Ministry of Education, Tsinghua University, Beijing, 100084, China
[e] School of Architecture, Tsinghua University, Beijing, 100084, China

ARTICLE INFO

Keywords:
Small public space
Vitality
Computer vision

ABSTRACT

Small public spaces are important for citizens to live and socialize with a high utilization rate. The vitality of small public space plays an important role in evaluating space quality and attraction and provides reference for urban governance issues such as vitality evaluation of public space, quality optimization, and site micro-renewal. Previous studies of vitality based on low-throughput surveys or big data with low positioning accuracy are not suitable for the high-efficiency study of small public space. In this study, a systematic framework of vitality quantification in small public spaces is built on fine-grained human trajectories extracted from videos for more efficient and refined human-oriented vitality evaluation. A multi-indicator vitality quantification method is first proposed to comprehensively represent human vitality, including number of people, duration of stay, motion speed, trajectory diversity and trajectory complexity. Furthermore, a video dataset of small public space along with our sub-index-assisted expert assessing scheme is proposed to evaluate our vitality quantification framework. Finally, we analyze the correlation between quantitative vitality indicators and the expert-assessing vitality through multiple linear regression and obtain the optimal vitality quantification model. The experimental results indicate that our dataset is reliable and the vitality quantification model constructed with our quantitative indicators can better characterize urban vitality than the previous model based on number of people and staying time.

1. Introduction

In urban areas, small public spaces exist in the gap of urban space and can be in the form of courtyard, patio, atrium, building gap, roof garden, etc. These spaces constitute a significant part of the public open space and are seen as the important symbols of the public realm because of their high utilization rate [1–3]. How to effectively evaluate the concepts and connotations of small public space vitality from a user's perspective and improve the happiness of citizens' lives has also been a significant concern of urbanologists, sociologists and urban planners [3–5].

It is still a challenge that how to collect information of urban space that represents vitality and evaluate space vitality in a cost-effective way [6]. In micro-scale public space, previous studies have been conducted using low-throughput surveys such as on-site observation [7–9], questionnaire survey [5,10,11], and so on. Although the traditional approaches can provide detailed and high-quality data, it is high-cost, time-consuming and hard to be implemented on a large scale. With the development of information technology, mobile signaling data [6,12,13], GPS data [14,15], LBS data [6,16–19] and social media data [20–22] become readily accessible data to facilitate vitality study in a large scale. Compared with traditional methods, they have the advantages of extensive coverage, long observation period, and low acquisition cost. However, it is difficult to transform these methods for large-scale space to small public space research because of low data accuracy [3]. Therefore, it is necessary to propose a more systematic and complete vitality representation system based on high precision data for small public space.

* Corresponding author.
E-mail address: qing_lb@scu.edu.cn (L. Qing).

https://doi.org/10.1016/j.buildenv.2022.109563
Received 9 May 2022; Received in revised form 10 August 2022; Accepted 29 August 2022
Available online 9 September 2022

As video data intuitively records fine-grained spatio-temporal information about humans, some researchers utilized video surveillance to depict the spatio-temporal activity information of pedestrians [3,23–26] and facial information [27]. These methods provided a novel approach to explore vitality in small public spaces. Based on video datasets and computer vision methods, Hou et al. [3] recorded the location of people to quantify the usage of small public spaces. Williams et al. [24] identified flows of people and generated social cohesion images with tracking algorithms. Yan et al. [25] extracted trajectories of humans from videos and analyzed the spatial and temporal distribution of stay. Likewise, Liang et al. [26] analyzed the relationship between the walking speed and climate based on trajectories. Wan et al. [27] recognized the faces and the emotional intention of every face using deep learning algorithms and identified that pedestrians' emotion is also an important vitality representation factor. Yunqin et al. [23] utilized multi-object tracking algorithm to quantify vitality based on the pedestrian number and proposed classification algorithm to obtain activity-based vitality. These methods dealt with the data acquisition problems of traditional methods, avoided the shortcomings of non-visual big data, and enhanced the granularity and precision of crowd information. However, some of these methods [3,24,25] mainly focused on quantitative statistics of location information in spatial or temporal dimensions and qualitative discussion of human activity. Thus, there are still some unresolved problems. First of all, it is not comprehensive to quantify pedestrian activities using location information. The location reflects the number and distribution of people, but does not reflect the dynamic activity information of people. Secondly, compared with quantifying space vitality, the qualitative discussion is difficult to effectively carry out spatial comparison of multiple small public spaces, which is not conducive to researchers' overall grasp of space. At last, it has not been verified the accuracy of the large-scale or micro-scale methods used to characterize spatial vitality. Although the information of activity type was utilized in Ref. [23], this method also ignored the details of individual activity and did not verify the accuracy of vitality quantification values.

Gehl emphasized that vitality is not only the product of the number and staying duration of people, but the diverse activities combine to make communal spaces in cities and residential areas meaningful and attractive [1], which is mostly ignored by previous studies. On the other hand, a large number of studies show that the location distribution of people, state of a crowd in public space and rich dynamic activity information are contained in people's trajectories [28–30]. Furthermore, the trajectory structure can reflect behavioral characteristics and behavioral types intuitionally [31,32]. Therefore, the analysis of trajectories is of great interest to explore activity types and activity diversity so as to implement the comprehensive measurement of space vitality.

In this study, a new vitality quantification framework is developed to understand the vitality representation of small public spaces from multiple perspectives based on human trajectories. According to the vitality connotations proposed by Gehl [1], we introduce the analysis methods of human trajectory to construct five quantitative vitality indicators which are *the number of people*, *the duration of staying*, *motion speed*, *trajectory diversity* and *trajectory complexity* to represent vitality. Among them, through analyzing human trajectories to mine the details of activity information comprehensively, *trajectory diversity* and *trajectory complexity* are proposed to fill the gaps of vitality quantification in small public space. We also collect a video dataset of small public space and propose a vitality expert rating system to verify the reliability of each quantitative indicator. In order to make expert panelists pay attention to people-oriented multifaceted vitality representation factors during overall vitality assessment [1,27,33–35], we formulate four subjective sub-indexes: *activity intensity*, *activity diversity*, *activity sustainability*, *pedestrian emotion* and an overall vitality evaluation index: *overall vitality*. Correspondingly, we propose an expert rating process in which panelists assess the overall vitality referring to their expertise and their understanding of vitality. The overall vitality will be used as the ground truth of vitality in experimental analysis. Finally, we obtain the optimal vitality quantification model by analyzing correlation between the quantitative indicators and the overall vitality.

The main contributions of this paper are as follows:

1. A novel vitality quantification framework is proposed. This framework includes the definitions and calculation methods of five quantitative indicators of vitality which are number of people (*num*), duration of stay (*dur*), motion speed (*sp*), trajectory diversity (*TD*) and trajectory complexity (*TC*). Specifically, sophisticated *TD* and *TC* are proposed to improve the comprehensiveness and accuracy of vitality representation.
2. A dataset along with expert evaluated scores is proposed specifically for small public space research to evaluate our vitality quantification framework. It will also contribute to the study of vitality representation of small public space from a user's perspective. Correspondingly, a people-oriented expert rating method for vitality is proposed to guide the panelists to perceive various vitality representation factors and assess overall vitality comprehensively.
3. We analyze the relationship between our quantitative indicators of vitality and overall vitality. Meanwhile, we obtain the optimal vitality quantification model. And it is verified that the trajectory diversity and trajectory complexity have great contributions to vitality representation and our vitality quantification model yields higher goodness of fit than the model constructed with shallow attributes which are number of people and staying time.

The remainder of this paper is organized as follows: Section 2 gives details of our study framework; Section 3 presents the experiment and the results of the proposed framework; Section 4 summarizes the experiment results and discusses the significance and limitations of this study; Section 5 concludes remarks.

2. Methodology

According to the vitality connotations proposed by Gehl [1], how many people use the public spaces, how long their activities last, which activity types can develop and the diversity of these activities are of great significance to the creation of vitality. In addition, some classical researches [33,36] also indicated that it is more meaningful and reasonable to quantify vitality as a multifaceted indicator. Generally speaking, as a multifaceted indicator, vitality should be composed of the following elements: number of people, activity duration, activity diversity and activity type as shown in Fig. 1. Thus we must improve the usual method to explore more fine and dynamic activity information in small public spaces so as to represent vitality completely. Owing to video data which contains rich human information and their activity details, we would analyze video data to measure multifaceted vitality in a human-oriented perspective with computer vision algorithms. Furthermore, at the micro spatial scale, the fine-grained trajectory contains details of human activity including location distribution, grouping state, activity duration, activity diversity and behavior types. Therefore, through analyzing human trajectories extracted from videos, we propose a vitality quantification framework.

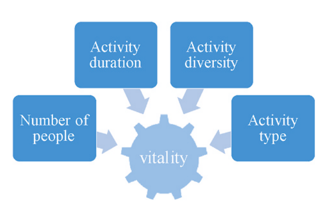


Fig. 1. Representation elements of vitality.

2.1. The overall methodological framework

As shown in Fig. 2, firstly, based on the representation elements of vitality proposed by classical researches [1,33,36] as shown in Fig. 1, we construct a set of vitality representation factors which include number of people, duration of staying people, speed of passing people, activity diversity and activity type. Among them, duration of staying and speed of passing people are separately utilized to represent the activity duration of staying people and passing people. And then, we utilize computer vision algorithms to process videos and extract human trajectories. After that, corresponding to the set of vitality factors, we explore human activities contained in trajectories and propose five quantitative indicators which are the number of people (*num*), the duration of stay (*dur*), motion speed (*sp*), trajectory diversity (*TD*) and trajectory complexity (*TC*). Their specific calculation methods are given in Section 2.2. The correspondence between vitality elements, the set of representation factors and quantitative indicators is shown in Table 1. In Section 2.3, we collect a video dataset of small public spaces and propose a subjective expert rating method to evaluate the overall vitality (*vitality*). Afterward, through the correlation analysis of quantitative indicators and vitality scores, we measure the contributions of our quantitative indicators to the overall vitality and the optimal vitality quantification model is obtained in Section 2.4.

2.2. Quantitative vitality indicators of small public space

A quantitative study of space vitality brings new insights for evaluating space quality and attraction. However, because shallow attributes such as number of people and staying time are often used as the proxy for vitality, it still has limited knowledge of the relations between people's diverse social activities and urban vitality in small public spaces. Therefore, based on video data, we propose trajectory diversity and trajectory complexity to characterize people's activities besides the number of people, staying time and speed of movement. In the following part, we will explain the meaning and calculation method of each indicator.

2.2.1. The number of people (num)

The number of people that can be observed outdoors in a small public space is an indication of the attractiveness of the space [1]. We use multi-object detecting and tracking algorithms to extract pedestrian trajectories in videos and the number of trajectories represents the number of people in the shooting space during the shooting time. Because there are tracking errors, we use the trajectories after filtering which is introduced in Section 3.1. That is to say, the results don't include pedestrians that are heavily occluded or too small to unclear in the videos.

2.2.2. Trajectory diversity (TD)

Previous classic researches [1,6,33,36] stressed that a variety of activities make a significant contribution to the creation of vitality. For the activity diversity, Gehl proposed that multiple groups and a variety of social activities correspond to activity richness [1]. Thus the number of activities, the number of people participating in each activity and the difference between human activities contain the important information of activity diversity. On the one hand, the structure difference of trajectories can reflect the heterogeneity of activity intuitively and the trajectory clustering can recognize the activity groups effectively. On the other hand, since Shannon entropy [37] is determined by "uncertainly" and "specificity" in Information Theory, it is well-suited to measuring the specificity and diversity of activities in public space. Therefore, we propose a new diversity proxy that is trajectory diversity based on Shannon entropy and human trajectory. We recognize the trajectory clusters of different activities at first, and meanwhile, we calculate the structure difference between trajectories participated in diverse activities. Based on these differences and clustering results, we utilize the Shannon entropy to measure the specificity of each activity which adds up to trajectory diversity.

Table 1
The correspondence between vitality elements, factors and quantitative indicators.

Vitality representation elements	Vitality representation factors	Quantitative vitality indicators
Number of people	Number of people	Number of people (*num*)
Activity duration	Duration of staying people	Duration of stay (*dur*)
	Speed of passing people	Motion speed (*sp*)
Activity diversity	Activity diversity	Trajectory diversity (*TD*)
Activity type	Activity type	Trajectory complexity (*TC*)

2.2.2.1. The recognition of trajectory clusters of different activities. In order to distinguish the clusters participating in different activities and calculate the structure difference of trajectories, we design a two-step clustering method using spatio-temporal distance and structure distance between trajectories respectively.

As shown in Fig. 3(a), at first, we extract and preprocess trajectories with detecting and tracking algorithms. And then, in our first step of the clustering method, we detect each group in which people move together because the same group of people walking together have similar activity types. We use Fréchet distance [38] to measure the distance between each spatio-temporal trajectory pair and obtain the spatiotemporal distance matrix. After that, hierarchical clustering [39] is adopted to cluster those people who move together, that is, keep consistent in their behaviors. For instance, as shown in Fig. 3(b), these people with id 1, 4, 8 and 10 sitting together and eating hot pot are detected to be group1, the people with id 13 and 14 walking together are grouped, and the people with id 5 and 3 walking together are grouped. At the same time, the others are divided into solo groups. We can also observe the result in 3D system of coordinates and these 3D trajectories with the same color are clustered into the same group.

Though the first step achieves group clustering, it can only guarantee

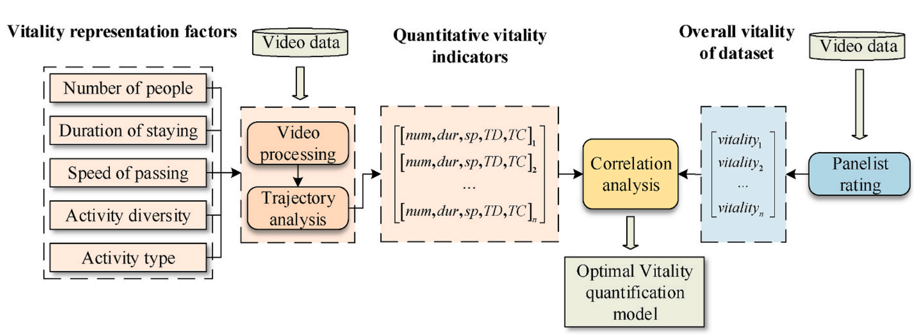
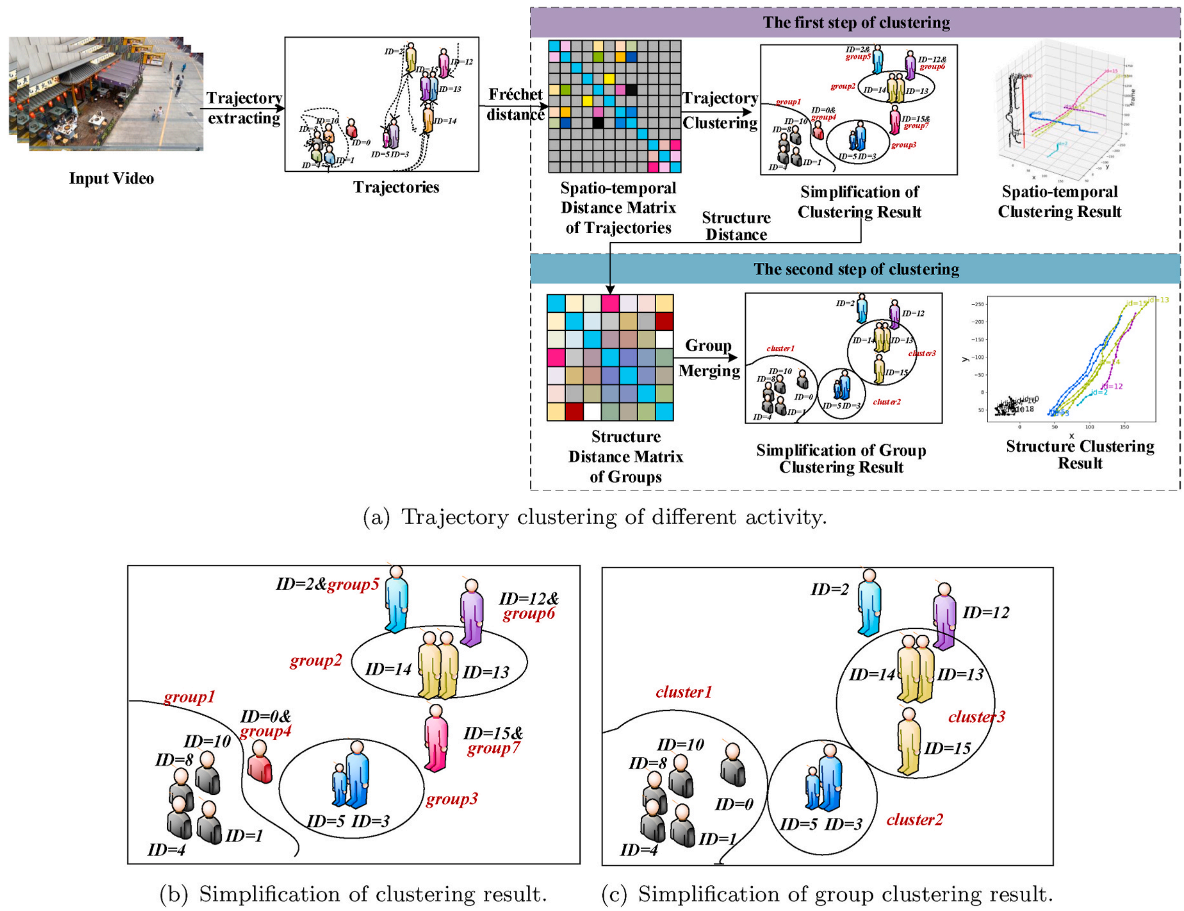


Fig. 2. Research framework.



(a) Trajectory clustering of different activity.

(b) Simplification of clustering result.

(c) Simplification of group clustering result.

Fig. 3. Distinguishing the groups participating in different activities.

that people in the same group have similar behavior patterns, but cannot guarantee that each group includes all of the people with a unique activity type. Thus in our second step of the clustering method, in order to achieve more accurate clustering, we detect each cluster by merging groups in which people participate in similar activity even though these groups are far away from each other in spatio-temporal dimension. In this step, we design the calculation method of structure distance matrix which is given in Section 2.2.2.2. On the basis of the matrix, groups are merged into clusters by hierarchical clustering methods. For example, as shown in Fig. 3(c), those people in group2 and group7 are merged to cluster3 because of their same direction and the similar moving path. We also can observe the final result in plane-coordinate system and these trajectories with the same color are merged into the same cluster.

2.2.2.2. The calculation method of structure difference. As shown in Fig. 4, each trajectory consists of time-discrete sampling points. The vectors composed by these discrete sample points contain the key information of trajectory structure. Among them, the red global vector represents moving direction and speed. In addition, these black local vectors embody trajectory shape and structure. Thus, through measuring angle difference and length difference at both local and global scales, we propose trajectory structure difference (*SD*). In order to fuse as many groups as possible that participate in the same behavior to reduce the error, the minimum value in structural distances of all of the permutations and combinations of trajectories from two groups is designed to the distance between these two groups. Finally, we obtained the structure distance matrix of groups.

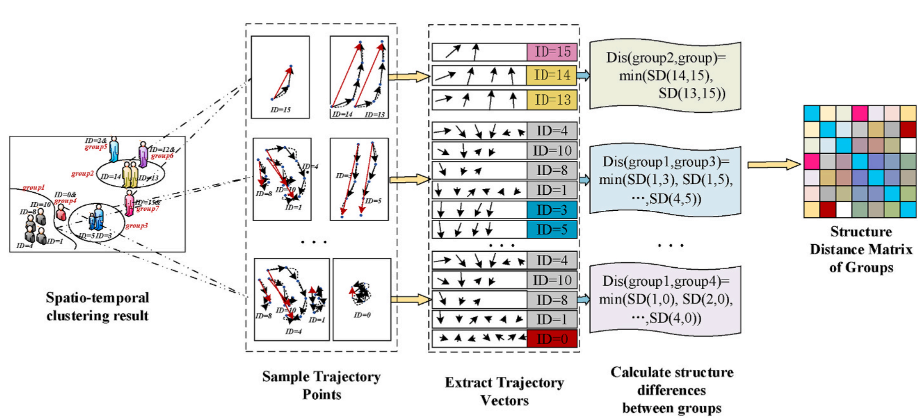


Fig. 4. Calculating structure difference matrix of groups.

We use angle difference and length difference at both local and global scales to calculate the morphological structure difference between trajectory P and Q. Trajectory P is represented by local vector $\{\overrightarrow{p}_i\}_{i=1}^{n}$ and global vector $\overrightarrow{p}_{global}$, and trajectory Q is represented by local vector $\{\overrightarrow{q}_i\}_{i=1}^{m}$ and global vector $\overrightarrow{q}_{global}$. Let θ_{global} be the angle difference between global vector $\overrightarrow{p}_{global}$ and $\overrightarrow{p}_{global}$, and l_{global} be the length difference between global vectors of two trajectories. Then, θ_i is the cosine difference between $\overrightarrow{p}_i$ and $\overrightarrow{q}_i$, and l_i is the length difference between $\overrightarrow{p}_i$ and $\overrightarrow{q}_i$. The morphological structure difference between trajectory P and Q are defined as Eq. (1). The larger the distance is, the greater the morphological difference is.

$$SD(P,Q)=\theta_{global}l_{global}\cdot\sum_{i=1}^{n}\theta_i l_i \tag{1}$$

where the cosine difference is calculated as Eq. (2) and the length difference is calculated as Eq. (3).

$$\theta_i\left(\overrightarrow{p}_i,\overrightarrow{q}_i\right)=arc\cos\left(\frac{\overrightarrow{p}_i\cdot\overrightarrow{q}_i}{\|\overrightarrow{p}_i\|_2\|\overrightarrow{q}_i\|_2}\right) \tag{2}$$

$$l_i\left(\overrightarrow{p}_i,\overrightarrow{q}_i\right)=\frac{\max\left(\|\overrightarrow{p}_i\|_2,\|\overrightarrow{q}_i\|_2\right)}{\min\left(\|\overrightarrow{p}_i\|_2,\|\overrightarrow{q}_i\|_2\right)} \tag{3}$$

2.2.2.3. The calculation of trajectory diversity. For a particular trajectory, the lower its similarity with all types of trajectory in space, the larger its specificity, and the greater its contribution to diversity is. In addition, the Shannon entropy can be utilized to measure the specificity of components of a system to some extent [40]. Therefore, based on clustering results and trajectory structure differences, we calculate trajectory diversity in terms of entropy.

Firstly, we define the calculation formula of spatial trajectory entropy as Eq. (6) to represent the specificity of each trajectory. In the Eq. (4), $p_i(t_i,t_c)$ is the proportion of the structural difference $SD(t_i,t_c)$ between trajectory t_i and center trajectory t_c of each clustering center, and *clus* is the number of trajectory clusters. In the Eq. (5), $H(t_i)$ is the trajectory entropy of the single trajectory t_i which represents the heterogeneity between t_i and other trajectories of different types. And finally, as shown in Eq. (6), we defined the trajectory diversity as the average of each spatial trajectory entropy, where $H(S)$ is the final trajectory diversity and n is the number of trajectories. The larger $H(S)$ is, the more uniform the distribution of different trajectory types in space is, and the richer the behavior types are.

$$p_i(t_i,t_c)=\frac{SD(t_i,t_c)}{\sum_{c=1}^{clus}SD(t_i,t_c)} \tag{4}$$

$$H(t_i)=-\sum_{c=1}^{clus}p_i(t_i,t_c)\log(p_i(t_i,t_c)) \tag{5}$$

$$H(S)=\frac{\sum_{i=1}^{n}H(t_i)}{n} \tag{6}$$

2.2.3. Trajectory complexity (TC)

Gehl [1] has proposed that in a good environment, a completely different, broad spectrum of human activities is possible and when outdoor areas are of poor quality, only strictly necessary activities occur. Concretely, Gehl divided outdoor behaviors in public spaces into three types: necessary behaviors, optional behaviors and social behaviors [1]. Necessary behavior is an inevitable life affair and daily work, so it has little relationship with the external environment and has a low leisure degree. Optional behavior can only occur when the time and place are suitable and people have the willingness to participate. Such activities depend on external material conditions and can only occur in a large number of high-quality outdoor activity spaces. Social behavior depends on the participation of others and the external material environment to a certain extent. That is to say, different types of behavior represent different degrees of vibrancy. In the small public space, passing is usually a necessary behavior, while other activities such as resting, sitting, stopping to communicate, sightseeing and taking photos belong to optional and social behaviors. The passing pedestrians have a strong moving purpose and low liveliness, and the trajectory shape is a mostly simple straight line. While the optional and social activities have strong randomness and great difference in the trajectory. Therefore, based on trajectories we propose a new indicator which is the trajectory complexity to represent the activity liveliness.

As shown in Fig. 5, a sample scene in our dataset is selected to display the characteristics and differences of trajectory structures. The Fig. 5(a) shows the trajectories in the sample scene. We select four trajectories to show the details. Among them, two trajectories of passing people and two trajectories of leisure people are shown in Fig. 5(b). The trajectory structures of passing people approximate straight lines, however, the trajectories of leisure people are complicated and confused. In addition, we utilize the minimum description length(MDL) principle [41] to extract inflection points which are shown in Fig. 5(c).There are few inflection points in the trajectory of passing people and the vectors extracted by connecting adjacent inflection points are similar. And conversely, the trajectories of leisure people are composed of multiple inflection points and the vectors are great variables. That is, differences between behaviors are mainly reflected in the inflection points and the variations of these vectors. Thus, based on these inflection points and vectors, we propose the measurement method of trajectory complexity.

Firstly, we extract the inflection points of each trajectory using MDL. And then through connecting adjacent inflection points, we extract their vectors. Finally, in order to explore the complexity contained in the variations of vectors, we design the calculation method as Eq. (7):

$$C_j=\sum_{i=1}^{m-1}\theta_{ij}l_{ij} \tag{7}$$

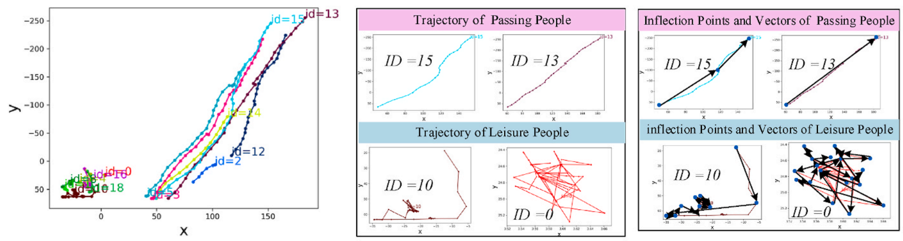
$$TC=\frac{\sum_{j=1}^{n}C_j}{n} \tag{8}$$

where C_j is the complexity of the *j-th* trajectory and m is the number of the trajectory segment vectors of the *j-th* trajectory. θ_{ij} and l_{ij} is the angle difference and length difference between the *i-th* trajectory segment vector and i+1-th trajectory segment vector of the *j-th* trajectory respectively. They are calculated according to Eq. (2) and Eq. (3). The greater the difference is, the greater the change in motion direction and velocity will be, and the more complex the trajectory will be. And finally, as shown in Eq. (8), we calculate the average of C_j to measure our trajectory complexity of the scene, where n is the number of trajectories.

2.2.4. The duration of stay (dur)

The overall social activity or liveliness of an environment is a product of the number of people and the duration of their stay [1]. In this paper, the duration of stay is the total length of stay including each person engaged in fixed or continuous activities. As same as Section 2.2.3, the trajectory of staying people is composed of more inflection points or staying points. Thus, we proposed a method to detect staying people based on extraction of inflection and staying points. As shown in Fig. 6, we utilize a stop point identification method based on density [42] to detect the trajectory stop points of each pedestrian. And we extract the inflection point of each trajectory using the minimum description length (MDL) principle [41]. Only those who have a stop point or multiple inflection points in their trajectories participate in fixed or continuous activities. Therefore, if the number of stop points or the number of inflection points exceeds the threshold, these people are identified as staying people.

Finally, we calculate the duration of stay by summing each person's activity duration. These staying person's activity duration is the



(a) Trajectories. (b) Trajectory local vectors. (c) Inflection points and vectors of trajectories

Fig. 5. The structure difference between trajectories.

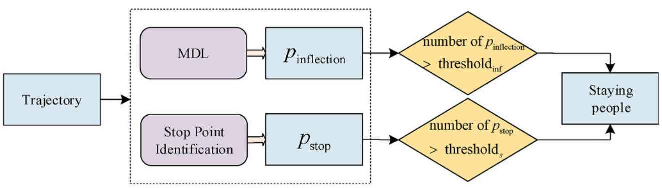


Fig. 6. Staying people detection.

difference between the end frame and the start frame of his/her trajectory and passing people have no staying time. The calculation of staying time is shown in Eq. (9) and Eq. (10), where T_{stay} is the trajectory set of staying people, t_i is the *i-th* trajectory in the set, e_i and s_i are the end frame and start frame of t_i, d_i is the staying duration of t_i, and *dur* is the duration of stay. The larger duration of stay indicates that there are more people participated in fixed or continuous activities, these activities last longer. And this space tends to grow in vitality. In our study, the longest d_i of each trajectory is 1 min which is limited by the length of our video data.

$$d_i = \begin{cases} e_i - s_i & t_i \in T_{stay} \\ 0 & t_i \notin T_{stay} \end{cases} \tag{9}$$

$$dur = \sum_{i=1}^{len(T_{stay})} d_i \tag{10}$$

2.2.5. Motion speed (sp)

The speed of people's walking and the duration they stay reflect the quality of physical places in public space to some extent. The spaces that make people slow down and want to stay always provide high quality environment and fun [34]. For the space without staying people, pedestrian speeds also reflect the space vitality to a certain extent. Thus, motion speed is a very significant representation factor of vitality and we propose '*sp*' to quantify motion speed. The recognition method of passing people is as same as Section 2.2.4, and these people not belonging to the set of staying people are passing people. Motion speed is the average motion speed of passing people. As shown in Eq. (11), we calculate the movement speed of each pedestrian based on the trajectory points at first, where $t_j = \{p_1, p_2, \ldots, p_{len(t)}\}_j$ represents the *j-th* trajectory consisting of $len(t)$ trajectory points, and $|p_i - p_{i-1}|$ is the euclidean distance between two consecutive points. And then, we calculate the average of these movement speed as shown in Eq. (12), where n is the number of trajectories.

$$speed_j = \frac{\sum_{i=1}^{len(t)} |p_i - p_{i-1}|}{len(t)} \tag{11}$$

$$sp = \frac{\sum_{j=1}^{n} speed_j}{n} \tag{12}$$

2.3. Video dataset construction based on people-oriented vitality assessment

In order to verify the reliability of our vitality quantitation method, we collect 48 videos of small public spaces at first. Before collecting videos, we formulated a data collection and research proposal referring to the previous researches [3,24] so as to conduct the study following research ethics. The specific measures of ethics maintenance in data collection are as follows: on the one hand, we obtained the permission from each sub-district office for conducting social observation experiments. We also signed agreements with each sub-district office, stating that collected videos would only be used for research, the data will not be shared with the third party, and the personal attributes needed to be blurred after collecting data. On the other hand, we posted notice to explain to people who entered the shooting space that they were being recorded, the videos are only used for observation research, the research is related to crowd behavior, not to individuals and it is promised that the facial and vocal information will be blurred. In addition, we shot in outdoor public spaces, thus, people could still see that they were being observed, even if they did not read the notices. Ethical approval for this study was granted by Ethics Committee on Biomedical Research, West China Hospital of Sichuan University (Reference No.: 2021-1072).

2.3.1. Dataset collection

In Wuhou District, Jinjiang District and Shuangliu District in Chengdu City, we randomly selected 48 small-scale public spaces as sample spaces including commercial areas, leisure areas, education areas and residential areas with different intensity of vitality. Sampling sites are small public spaces, including gaps between buildings, street corners, extension plazas in front of buildings, and pocket parks. Among them, street-corner small-scale public space is located at the intersection of several roads (two or more) instead of the two sides of roads. The size of each research area is equivalent to the size of scene that the camera can capture from a height of 4–5 m with a pitching angle of 45° down.

From April to June 2021, between 2 and 6 p.m. on workdays, we shoot more than a minute video of each selected site within normal weather and clipped a minute clip from each video in our experiment. In order to reduce the error caused by the occlusion between pedestrians and prevent the shaking of the shooting, we set up a 1.7-m tripod on the second floor of the buildings or the overpasses and set the shooting angle to 45° down which is similar to the surveillance camera. And we use the same mobile phone to film each video with the resolution of 1920 × 1080 to ensure consistent video quality. Besides these, in order to explore the crowd gathering situation and the true structure of the trajectory, we record a rectangle on each site by placing landmarks and measured its size, which was used to convert pixel position coordinates into world coordinates [43,44] in our experiment.

2.3.2. People-oriented vitality assessment indexes

At present, multiple perceptual qualities have been developed to guide experts to perceive the space design quality and measure urban design [45]. These measurement indexes of space design quality are mainly used to evaluate the quality of built environment elements and they are not direct representation indicators of vitality. According to vitality representation factors proposed by Gehl [1], people and their activity are the main and intuitive representations of spatial vitality. Therefore, we propose a people-centered expert assessing the method of vitality. Our expert scoring method is composed of sub-evaluation indexes and a process in which the sub-indexes guide panelists mainly focus on Gehl's theories [1,34]. These sub-indexes are defined with conceptual description of vitality and proposed to unify assessment criteria so as to reduce the impact of other vitality theories on expert assessments. Based on previous researches about vitality representation, the flow of people, duration of staying and various activities have great contributions to vitality representation [1]. Moreover, the dispersion or aggregation of crowds [35], passing speed [34] and people emotion [27] also have significance for vitality representation. Thus, we summarize the previous researches and propose our expert assessment system which is composed of the overall vitality and four subjective sub-indexes which are activity intensity (*AI*), activity diversity (*AD*), activity sustainability (*AS*) and pedestrian emotion (*PE*). In our scoring process, these four sub-indexes are evaluated by panelists at first, so as to make panelists understand and pay attention to people-oriented vitality representations comprehensively. And then, based on their perception and expertise, it is more possible for them to assign more reasonable and reliable overall vitality scores. However, the scores of sub-indexes have no direct effect on the final vitality values and panelists are completely unaware of the quantitative indicators. In general, the panelists' understanding of vitality and their own professional experience is the basis of assessing vitality. The rating standards and samples of sub-indexes are shown in Fig. 7.

2.3.2.1. Activity intensity (AI). *AI* is defined as the degree of the liveliness of the people in a small public space. A place would appear lively if large numbers of people stay for short durations or fewer people stay for longer durations [1]. Therefore, *AI* is composed of two conditions: the number of people passing by that is staying for a short duration and the number of people staying for a longer duration. We divide the *AI* into five grades and the specific rating standard and samples are shown in Fig. 7(a).

2.3.2.2. Activity diversity (AD). *AD* is defined as the richness of activities in a small public space. Since characteristic of best used small public space is a higher proportion of groups rather than solitary individuals with a human-oriented perspective [35]. If there are people divided into multiple groups to carry out a variety of social activities, activities are

(a) Rating standard and samples of activity intensity. (b) Rating standard and samples of activity diversity.

(c) Rating standard and samples of activity sustainability. (d) Rating standard and samples of pedestrian emotion.

Fig. 7. Rating standards and samples of sub-indexes.

richness [1]. In addition, a small public space with scattered walking directions is more abundant than scenes with a single walking direction if there are only passerby. Therefore, *AD* is composed of three conditions: activity category, walking direction and grouping situation. We also divide the *AD* into five grades and the specific rating standard and samples are shown in Fig. 7(b).

2.3.2.3. Activity sustainability (AS). *AS* is defined as the degree of people staying in spaces, the activity sustainability of staying people is related to their staying time, and passing people is related to their moving speed [34]. Therefore, *AS* is composed of two conditions: staying time and moving speed. We also divide the *AS* into five grades and the specific rating standard and samples are shown in Fig. 7(c).

2.3.2.4. Pedestrian emotion (PE). Considering that pedestrian emotion is also an important vitality representation factor [27], we invite *PE* and define it as the overall pedestrian emotion and the atmosphere of the scene. We also divide the *PE* into five grades relying on whether the atmosphere is relaxed or hurried and nervous. The specific rating standard and samples are shown in Fig. 7(d).

2.3.2.5. Overall vitality. Inspired by Jacobs [33] and Gehl [1], the overall vitality is defined as the people and their activities that can be observed in space, and it is the product of the quantity, types and duration of various activities. Experts need to combine the understanding of the sub-indexes with knowledge of the urban planning field to assign the spatial vitality ratings subjectively which is divided into five grades.

2.3.3. Vitality assessment process

We propose the expert scoring process that the sub-indexes guide the overall vitality scoring with a multiple-cycle rating and score checking, so as to reduce individual subjective deviations. Especially, we design rating one video after another and rating one index after another which are all the sub-index scoring. Of these, during rating one video after another, panelists need to consider the overall information of each video, so as to strengthen their familiarity with the videos. During rating one indicator after another, it is easier for experts to measure each video using uniform criteria and familiarize themselves with each sub-index. After sub-index scoring, panelists would have a more comprehensive view of vitality and would assign a fairer score. The rating process is as show in Fig. 8 and the details of process are as follows:

1. **Familiar with vitality indicators**: Present the evaluation indicators with panelists in advance, and explain their doubts. At the same time, show them sample videos in order to let the experts better understand the score range.
2. **Familiar with videos**: The panelists browse each video at least twice to get the overall information of the video, such as the flow of people, behavior, passing speed, activity continuity, atmosphere, etc.
3. **Rating one video after another**: The panelists assign ratings to the four sub-indexes of a video on a 1–5 Likert scale before moving on to the next video. Panelists would watch the videos one time in a loop.
4. **Rating one indicator after another**: The panelists assign rating to one sub-indexes of each videos on a 1-5 Likert scale and then moving on to the next sub-index. Panelists would watch the videos four times in a loop.
5. **Score checking**: The panelists compare two scores of each video to check the sub-index scores and watch this video again to determine values if the scores are different.
6. **Rating overall vitality**: The panelists assign quantitative ratings to the overall vitality of a video on a 1–5 Likert scale based on professional knowledge in the field of urban planning and referring to the four sub-indexes.

2.4. Construction of vitality quantification model

In order to explore the relationship between our quantitative vitality indicators with overall vitality scores and obtain the vitality quantification model, we conduct correlation analysis based on stepwise multiple linear regression using SPSS software [6,46].

For our analysis, let num_i, dur_i, sp_i, TD_i and TC_i, $i = \{1, 2, \cdots, 48\}$ represent our quantitative indicators: the number of people, duration of stay, motion speed, trajectory diversity and trajectory complexity of the *i-th* scene in dataset respectively. Consequently, for a typical scene *i*, the vitality descriptor can be represented by the vector matrix:

$$\mathbf{Q} = [num_i, dur_i, sp_i, TD_i, TC_i]^T \tag{13}$$

Finally, taking these vitality descriptors as independent variables and the vitality scores that panelists assigned to our dataset as dependent variables, we can proceed for stepwise multiple linear regression analysis. Using this, the multiple linear regression analysis model of the typical scene *i* is as follows:

$$Y_i = B_1 num_i + B_2 dur_i + B_3 sp_i + B_4 TD_i + B_5 TC_i + c = \mathbf{BQ} + c \tag{14}$$

$$\mathbf{B} = [B_1, B_2, B_3, B_4, B_5] \tag{15}$$

where *c* is the constant of the model which represents unobserved error term, the **B** is the regression coefficient matrix, which indicates the quantitative relationship between our quantitative indicators and the result of overall vitality scoring, but not the contribution of each indicator in the multi-linear regression analysis model. To compare the contribution of each independent variable to the vitality representation, the influence of the unit must be eliminated [47]. That is to say, it is necessary to standardize the variable values after regression and the standardized value is calculated by the usual method that the variable minus its mean and divided by its standard deviation estimate. The standardized regression model:

$$Y_i^* = Beta_1 num_i^* + Beta_2 dur_i^* + Beta_3 sp_i^* + Beta_4 TD_i^* + Beta_5 TC_i^* + c \tag{16}$$

$$Y_i^* = (Y_i - \overline{Y}) / \sigma_Y \tag{17}$$

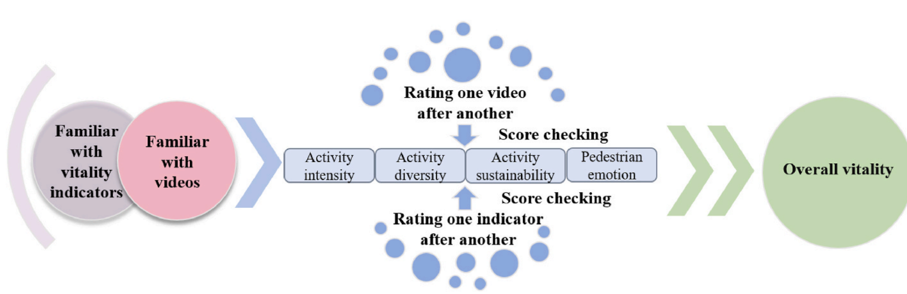


Fig. 8. The vitality assessment process.

$$num_i^* = (num_i - \overline{num}) / \sigma_{num} \tag{18}$$

$$dur_i^* = (dur_i - \overline{dur}) / \sigma_{dur} \tag{19}$$

$$sp_i^* = (sp_i - \overline{sp}) / \sigma_{sp} \tag{20}$$

$$TD_i^* = (TD_i - \overline{TD}) / \sigma_{TD} \tag{21}$$

$$TC_i^* = (TC_i - \overline{TC}) / \sigma_{TC} \tag{22}$$

where the $Beta_i$ denotes the standardized coefficient. Finally, the effect of each quantitative indicator on the model of vitality quantification could be measured by the standardized coefficient of each independent variable. The larger the coefficient value, the greater the influence on the dependent variable.

3. Experiment and analysis

In order to verify our vitality quantification method, we have conducted comprehensive experiments. Before calculating our quantitative vitality indicators, we preprocess dataset and extract human trajectories at first. And then, we first demonstrate the effectiveness of our dataset and the reliability of the expert rating system. On the other hand, we explore the relationship between four subjective sub-indexes with overall vitality scores. And finally, in order to verify the reliability of vitality characterization, we explore the correlation between overall vitality scores and five quantitative vitality indicators that we proposed. Furthermore, we obtain the optimal vitality quantification model.

3.1. Data processing

Data preprocessing method includes four aspects: a) video clip, b) pedestrian trajectory extraction, c) position coordinate transformation, d) trajectory denoising. The specific process for each video is as shown in Fig. 9: first, we clip a minute clip from each video, where the frame is smooth and stable. Second, in order to yield trajectories accurately and fast, we select the latest YOLOv5 detector [48] to locate multiple pedestrians in every image of the video and use a classic tracking method DeepSort tracker [49] to associate the same pedestrians between two images. Third, because the shape and distribution of trajectories are influenced by the cameras' angle, it is necessary to transform the original perspective video into the bird's eye view to study the behavior in the world coordinates, and we use the perspective transformation method [43,44] to accomplish the coordinate transformation. Forth, we propose a trajectory denoising method that the incomplete short trajectories are deleted by filtering the trajectories with less than 90 frames to reduce occluded and blur objects and the trajectories are subsampled in 30:1 to reduce trajectory point drift. Although we cannot tackle tracking error to extract perfect trajectory, from the human-oriented perspective of vitality station, the close and clear individual motion trajectory has a greater contribution to vitality, while the tracking error of the severely obscured and too fuzzy individuals have little impact on the quantitative analysis of vitality. With the development of multi-target tracking technology, higher performance algorithms can be directly introduced into the method of this paper to improve the accuracy of vitality quantization.

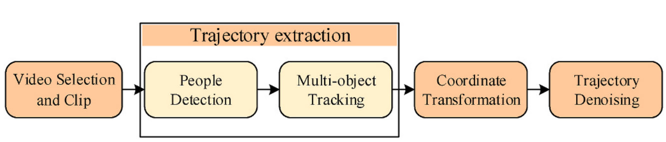


Fig. 9. Data preprocessing.

3.2. Experiment design and setting

The expert assessment is the basis for the construction of dataset and evaluation of proposed vitality quantification method. In order to verify the reliability of assessment results and the efficiency of assessment system, we utilize intra-class correlation coefficients (ICCs) to evaluate score consistency of each sub-index and overall vitality. On the other hand, we explore the relationship between four sub-indexes and the overall vitality using the stepwise multiple linear regression analysis method to explore the panelists' focus so as to verify the efficiency of proposed expert assessment system.

Through our vitality quantification system, we calculate the quantitative vitality indicators. And among them, in our measurement of duration of stay, we set $threshold_{\text{inf}}$ as 3 and $threshold_{\text{s}}$ as 1. Based on verified dataset, we implement the stepwise multiple linear regression analysis to explore the correlation between overall vitality scores and our five quantitative vitality indicators, so as to verify the reliability of proposed quantitative vitality indicators and explore the contribution of each indicator to the overall vitality.

3.3. Evaluation methods

3.3.1. Reliability evaluation of expert assessment

In order to evaluate and verify the reliability of expert scoring results, we choose ICCs and Cronbach's alpha as the evaluation metrics:

1) The ICC is the preferred measure of inter-rater reliability when cases are rated in terms of some interval variable or interval-like variable, such as the Likert scales our expert panel used to rate scenes [50]. And it is sensitive to both random error and systematic error. According to the criteria [51] used in Measuring Urban Design [45], ICCs > 0.8 are perfect, 0.8 > ICCs > 0.6 are substantial, 0.6 > ICCs > 0.4 are moderate, 0.4 > ICCs > 0.2 are fair and ICCs < 0.2 are slight.
2) Cronbach's alpha is also a method of examining inter-rater reliability and is the most commonly used reliability analysis method in social science research. 0.8 > Cronbach's alpha > 0.7 are regarded as satisfactory [52].

We carry out these two metrics utilizing SPSS software [53]. In order to evaluate the absolute consistency and repeatability of our expert rating result, we choose two-way random model, single rater type and absolute agreement definition in SPSS software [53].

3.3.2. Correlation analysis

Our correlation analysis is based on the stepwise multiple linear regression. The basic idea of the stepwise multiple linear regression is to automatically select the most important variables from a large number of available variables to establish a predictive or explanatory model for regression analysis. SPSS software is used to help identify the most relevant variables of the regression model through a step-by-step introduction, evaluation and deletion of the variables [54]. Specifically, the model would include the variable most highly correlated to the dependent variable at first. Then it will add the next most correlated variable until no further variables are significant. In this process, it is possible to delete a variable that has been included at an earlier step but is no longer significant. In the process of stepwise multiple linear regression, we employ commonly adopted evaluation indexes of regression algorithms as the evaluation metrics for the regression results [54]. The details of these indexes are as shown in Table 2:

3.4. The analysis of expert assessment results

3.4.1. Inter-rater reliability of subjective ratings

Reliability measure results are summarized in Table 3. From their ICC values, most indicators including activity intensity (*AI*), activity diversity (*AD*), activity sustainability (*AS*) and overall vitality

Table 2
The evaluation metrics for the regression results.

Evaluation Metric	Definition
B	The estimate of the regression coefficients.
Std.Error	The standard error of the regression coefficients.
Beta	The normalized regression coefficients. It represents the degree of influence of the independent variable on the dependent variable. The larger Beta, the greater influence of independent variable.
t	The intermediate statistic in the T-test. T-test is the significance test of the coefficient of each independent variable (Beta) and is utilized to determine whether each independent variable is indeed linearly related to the dependent variable. The larger the test statistic, the less likely it is that the results occurred by chance.
Sig. in T-test	The final significance in T-test. It is generally used to measure the significance of Beta. Sig. $\geq$ 0.05 indicates that the coefficient of the independent variable has no statistical significance in this model, and the corresponding variable should be deleted from the regression model. Sig. $<$ 0.05 indicates that it was statistically significant in the model and the variable should be retained. The lower Sig., the more significant the Beta of the independent variable.
tolerate	The measure of the severity of multicollinearity in multi-linear regression models. Multicollinearity means that one independent variable can be a linear combination of one or several other independent variables. If there is multicollinearity, the correlated variables should be deleted. tolerance $<$ 0.1 indicates that there is a severe collinearity problem between independent variables [55].
VIF (variance inflation factor)	The measure of the severity of multicollinearity in multi-linear regression models. VIF $>$ 10 indicates that there is a severe collinearity problem between independent variables [55].
R^2	The goodness of fit of the model to the truth value. R^2 ranges from 0 to 1. The closer R^2 is to 1, the higher the goodness of fit is.
Adjusted R^2	The adjusted goodness of fit. It is formulated by introducing the number of the independent variable to R^2, so as to eliminate the influence of the number of independent variables on the goodness of fit. Since R^2 will approach 1 as the number of independent variables increases, Adjusted R^2 is commonly utilized to evaluate the goodness of fit. The closer Adjusted R^2 is to 1, the higher the goodness of fit is.
F	The intermediate statistic in the F-test. F-test is the significance test of the linear regression result and is utilized to determine whether there are independent variables indeed linearly related to the dependent variable.
Sig. in F-test	The final significance in F-test. It is generally used to measure the significance of the model. Sig. can be obtained by querying F distribution table. Sig. $\geq$ 0.05 indicates that the multi-linear regression result isn't significant and the dependent variable has no linear correlation with any of the independent variables. Sig. $<$ 0.05 indicates that the multi-linear regression result is significant, however, this does not mean that all independent variables are significant, so the T-test is essential.

demonstrate substantial inter-rater reliability among panelists (0.8 $>$ ICCs $>$ 0.6); the pedestrian emotion (*PE*) shows moderate inter-rater reliability (0.6 $>$ ICCs $>$ 0.2). For purpose of comparison, Cronbach's alpha is also reported for these ratings. All of our indicators demonstrate high reliability (Cronbach's alpha $>$ 0.8) [52].

Of these, the reason why '*PE*' shows lower ICCs is that the facial expressions of the people in clips are fuzzy, and panelists only perceive the atmosphere of the scene to assess emotion scores, which is highly subjective. There are two sample shown in Fig. 10: almost all wear masks and their facial expressions are invisible in Sample 1; in Sample 2, the expressions of those people walking in the opposite direction or sitting far away are also fuzzy. All of these would impact the judgment of experts.

Most of all, the highest ICCs of the overall vitality value indicate that overall vitality score is more reliable than sub-indexes, which verifies the effectiveness of the overall rating method guided by sub-indexes.

3.4.2. The relationship between the overall vitality and sub-indexes

In order to further verify the efficiency of sub-indexes guidance for overall rating, we utilize stepwise multiple linear regression to explore the relationship between overall vitality and four subjective sub-indexes. In particular, independent variables are '*AI*', '*AD*', '*AS*' and '*PE*', whereas the vitality score is the dependent variable. Table 4 displays models produced in the process of stepwise multiple linear regression. Among them, Model 5 displays the optimal regression model results between sub-indexes and overall vitality. In addition, the adjusted R^2 of the Model 5 reaches 0.869 and its goodness of fit is higher than all the other models. Its independent variables are '*AD*', '*AI*' and '*AS*'. All the tolerances of these independent variables are larger than 0.1 and all the VIFs are less than 10, which means that there is no multi-collinearity between independent variables in Model 5 [55].

Table 3
Inter-rater reliability for ratings of sub-indexes and overall vitality.

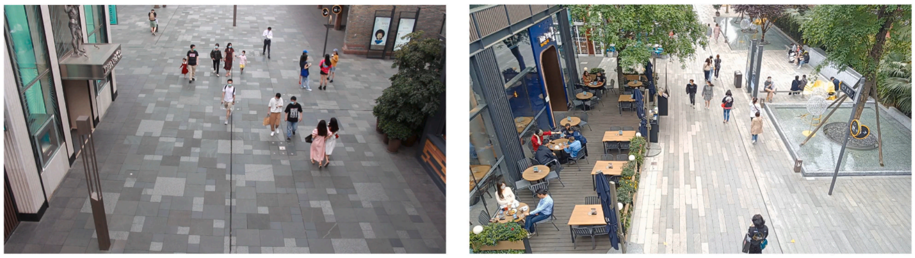
	Intra-class Correlation↑	95% Confidence Interval	Cronbach's Alpha↑
Activity Intensity	.662	.544–.769	.915
Activity Diversity	.603	.481–.722	.883
Activity Sustainability	.631	.460–.764	.926
Pedestrian Emotion	.551	.381–.699	.898
Overall Vitality	.701	.595–.798	.921
N	48		

Through analyzing the Sig. of each independent variable, the regression coefficients of '*AI*' and '*AS*' are the most significant, followed by '*AD*'. Though the Sig. value of '*AD*' is higher than 0.05, the model including this index has highest Adjusted R^2 which also indicates the significance of '*AD*'. On the contrary, '*PE*' has no significant association with vitality assessed by panelists and it is deleted. That is, panelists tend to pay little attention to pedestrian emotion when perceiving vitality. Therefore, apart from '*PE*', the other three indicators appear to play a guiding role in assessing the overall vitality from different dimensions. The reason why '*PE*' is ignored by panelists may be that they were affected by the vitality factors previously proposed, such as the number of people and staying time, and almost hardly took into account pedestrian emotion, and the facial blur led to the difficulty in identifying emotion in the clips.

On the other hand, the values of standardized regression coefficients Beta of Model 5 show that '*AI*', '*AS*' and '*AD*' have great contributions to the evaluation of vitality from panelists and their contributing degree minishes in order. That is, when the panelists evaluated the overall vitality of each scene, they were likely to focus on vitality factors in Gehl's theories [1,34]. The signs of the coefficients Beta are also as expected, panelists percept the larger activity intensity, activity sustainability and activity diversity, meanwhile, panelists tend to percept the larger vitality and assign the larger overall vitality. In general, the vitality assessing results obtained through our vitality assessment method are consistent with the vitality connotation proposed by Gehl [1,34].

3.5. The analysis of vitality quantification indicators and model

In order to verify the reliability of our quantitative vitality indicators,



(a) Sample 1. (b) Sample 2.

Fig. 10. Samples of dataset.

Table 4
Regression model results between vitality and sub-indexes.

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics		Adjusted R^2↑
		B	Std. Error	Beta			Tolerance	VIF	
1	(Constant)	3.106	1.176		2.641	.011			0.717
	AD	.865	.079	.850	10.949	.000	1.000	1.000	
2	(Constant)	−.354	1.289		−.275	.785			0.794
	AD	.679	.080	.668	8.480	.000	.707	1.415	
	PE	.358	.084	.337	4.274	.000	.707	1.415	
3	(Constant)	−2.769	1.193		−2.320	.025			0.858
	AD	.251	.115	.246	2.186	.034	.238	4.193	
	PE	.377	.070	.355	5.411	.000	.704	1.420	
	AI	.509	.111	.483	4.594	.000	.274	3.653	
4	(Constant)	−1.535	1.299		−1.182	.244			0.867
	AD	.210	.112	.207	1.871	.068	.231	4.325	
	PE	.087	.157	.082	.553	.583	.130	7.697	
	AI	.548	.109	.521	5.047	.000	.265	3.770	
	AS	.247	.120	.300	2.054	.046	.132	7.555	
5	(Constant)	−1.074	.988		−1.087	.283			**0.869**
	AD	.210	.111	.206	1.884	.066	.231	4.325	
	AI	.556	.107	.528	5.208	.000*	.270	3.702	
	AS	.307	.051	.373	5.991	.000*	.717	1.394	

a.Dependent Variable: OverallVitality.
b.Model 5:N = 48, R = .937, R^2 = .878, F = 105.301, Sig. = .000.
c.*Sig. < 0.05.

stepwise multiple linear regression is utilized to explore the relationship between the overall vitality and five quantitative vitality indicators that we propose. Furthermore, through the regression analysis, we establish the optimal vitality quantification model. In our stepwise multiple linear regression, independent variables are the number of people (*num*), the duration of stay (*dur*), motion speed (*sp*), trajectory diversity (*TD*) and trajectory complexity (*TC*) extracted by our vitality quantification system, whereas the overall vitality is the dependent variable. Table 5 displays models produced in the process of stepwise multiple linear regression. Among them, Model 4 displays the optimal regression mode. In addition, the adjusted R^2 of the Model 4 reaches 0.761 and its goodness of fit is higher than all the other models. Its independent variables include '*num*', '*dur*', '*TD*' and '*TC*' except '*sp*'. At the significance level of 0.05, all variables of Model 4 are significant. All the

Table 5
Regression model results between overall vitality and quantitative indicators.

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics		Adjusted R^2↑
		B	Std. Error	Beta			Tolerance	VIF	
1	(Constant)	9.825	.746		13.166	.000			0.618
	num	.291	.033	.791	8.775	.000	1.000	1.000	
2	(Constant)	6.790	1.075		6.314	.000			0.696
	num	.301	.030	.819	10.147	.000	.991	1.010	
	dur	.005	.001	.290	3.592	.001	.991	1.010	
3	(Constant)	5.967	1.035		5.767	.000			0.740
	num	.221	.039	.602	5.708	.000	.497	2.011	
	dur	.004	.001	.236	3.069	.004	.934	1.071	
	TD	1.742	.597	.308	2.918	.006	.496	2.014	
4	(Constant)	5.513	1.014		5.436	.000*			**0.761**
	num	.214	.037	.582	5.729	.000*	.493	2.027	
	dur	.004	.001	.254	3.420	.001*	.923	1.084	
	TD	1.734	.573	.307	3.027	.004*	.496	2.015	
	TC	.015	.007	.159	2.189	.034*	.967	1.034	

a.Dependent Variable: OverallVitality.
b.Model 4:N = 48, R = .884, R^2 = .781, F = 38.314, Sig. = .000.
c.*Sig. < 0.05.

tolerances of these independent variables are larger than 0.1 and all the VIFs are less than 10. Therefore, there is no multicollinearity between independent variables in Model 4 [55]. In general, as important supplements of vitality representation, '*TD*' and '*TC*' that we propose, are added into the vitality representation system to make the model fit better and the representation system more perfect. Particularly, compared with model 2, it can be inferred that '*num*' and '*dur*' cannot fully reflect small public space vitality and combining four quantitative indicators in the model can better characterize urban vitality.

Through analyzing the Sig. of each independent variable, the regression coefficient of '*num*', '*dur*', '*TD*' and '*TC*' have statistical significance in vitality prediction model. However, the regression coefficient of '*sp*' isn't significant and it is deleted. Through analyzing '*sp*' and '*dur*', we find that '*sp*' is the moving speed of pedestrian passing by and passing is usually a necessary behavior which makes a smaller contribution to vitality creation, but '*dur*' is the staying time of who participated in fixed or continuous behaviors which have great potential for vitality creation. Thus the regression result is in line with Gehl's theory [34]. Further discussions are reported in Section 4.

In addition, based on the regression analysis, we also obtain the optimal vitality quantification model as shown in Eq. (23) and the standardized model is as shown in Eq. (24). The signs of the coefficients B in the quantification model are as expected, a place would appear lively if there are many people or fewer people staying for longer durations. Diverse trajectory, that is, rich behavior categories add to space vitality. Moreover, the higher the trajectory complexity is, the more optional and social activities, and space vitality is very possible to be greater due to the considerable potential of optional and social activities for vitality creation. On the other hand, in Eq. (24), the values of Beta show that '*num*', '*dur*', '*TD*' and '*TC*' have great contributions to the representation and prediction of vitality, and their contributing degree minishes in order. Therefore, according to the analysis of our regression results and the optimal vitality quantification model, a place would appear vibrant, if large numbers of people participate in diverse activities and these activities are mainly optional activities and social activities with long duration. In general, the regression results are consistent with the vitality theories of Gehl [1,34].

$$vitality = 0.214num + 0.004dur + 1.734TD + 0.015TC + 5.513 \tag{23}$$

$$vitality^* = 0.582num^* + 0.254dur^* + 0.307TD^* + 0.159TC^* \tag{24}$$

4. Discussion

In this paper, we construct a video dataset of small public spaces based on people-oriented vitality assessment method of experts in which sub-evaluation indexes guide overall vitality scoring. Furthermore, through extraction and exploration of human trajectories in video data, we propose a systematic vitality quantification system that is composed of number of people, motion speed, duration of stay, trajectory diversity and trajectory complexity.

Since the expert assessment of overall vitality is the basis of our experimental analysis, we utilize the intra-class correlation coefficient to verify the reliability of expert assessment results at first. And the experimental results demonstrate that assessment results have substantial reliability and the guidance of sub-evaluation indexes and proposed expert scoring process effectively improve scoring consistency and reliability. In addition, through analyzing the correlation between sub-evaluation indexes and vitality values, we verify that the panelists comprehensively pay attention to activity intensity, activity sustainability and activity diversity during overall vitality assessment. These results also indicate that our vitality assessment system makes panelist perception focus coherence with the vitality connotation [1] proposed by Gehl. However, it can also be noticed that panelists pay little attention on pedestrian emotion during assessing vitality. Therefore, how to improve the inter-rater reliability and how to make panelists take account of pedestrian emotion still need to explore. In addition, although our expert scoring method can guide experts to be more familiar with videos and indicators, and better grasp the overall distribution, it takes time and effort to circulate repeatedly. We also need to study to reduce the assessing complexity.

Meanwhile, through analyzing the correlation between proposed quantitative vitality indicators and overall vitality, we obtain the optimal vitality quantification model and it is certified that number of people, duration of stay, trajectory diversity and trajectory complexity make significant contributions to the representation of vitality. However, motion speed(*sp*) is insignificant to the overall vitality in our dataset including various behaviors. Nevertheless, '*sp*' is also an effective indicator to quantify the vitality relevant to the passing pedestrian and it is likely to represent the vitality of other public spaces whose main crowd are pedestrians. Thus, '*sp*' is still significant for vitality researches to quantify vitality comprehensively in the future. And our optimal model can better represent vitality than the model composed with shallow attributes (i.e., number of people and duration of stay), which further demonstrates the effectiveness of sophisticated trajectory diversity and trajectory complexity. The optimal model also indicates that the contribution of quantitative vitality indicators to vitality creation minishes in order of number of people, activity diversity, duration of stay and trajectory complexity and these four quantitative indicators have positive influences on vitality creation. That is, a vibrant public space should firstly have the ability to attract a large number of people. Secondly, it should provide opportunities for people to participate in various activities. And then, it should attract people to participate in continuous activity and stay for a long time. Lastly, it should provide opportunities for optional behaviors and social behaviors which have greater contributions for vitality creation than necessary behavior. These results are consistent with the vitality connotation [1] proposed by Gehl. Since both of expert scoring sub-indexes and quantitative indicators are mainly based on Gehl's researches [1,34], there are implicit correlations between them. However, on the one hand, sub-indexes are the conceptual description of vitality. And panelist assessment is a traditional vitality evaluation from the perspective of urban planner. Quantitative indicators are numerical and objective statistics of human activities representing vitality. And the vitality quantification is a vitality evaluation method based on computer vision. On the other hand, panelists do not know the quantitative indicators and the scores of sub-indexes have no direct effect on the final vitality values. That is, the two approaches evaluate vitality from different perspectives in parallel. Therefore, these implicit correlation does not affect the reasonability of regression analysis and the validity of the vitality quantitative indicator. Therefore, our method is effective and it is convenient for urban designers and planners to grasp the representation of spatial vitality more accurately and efficiently. However, we should always be aware of the biases of these data sources, such as sampling bias. Due to the limitation of data acquisition, we only studied a 1-min video of each space. That is, we only analyze the spatial vitality of a current observation point in time. Future work includes exploring the urban vibrancy over a longer period and how the vitality changed over time.

5. Conclusion

Small public spaces are important places for citizens to live and socialize and with a high utilization rate. The vitality of small public space not only reflects the happiness of citizens but also determines whether the physical environment of space is lively or attractive on the human scale. Meanwhile, convenient video data collection and high-performance computer vision methods provide opportunities to quantify small public space vitality from the individual perspective. Therefore, we propose five quantitative vitality indicators of small public spaces based on individual trajectory data which is extracted from videos with computer vision algorithms. The five indicators are the number of people, motion speed, the duration of stay, trajectory diversity and trajectory complexity. Meanwhile, we collect a video dataset of small public spaces and propose a people-oriented expert assessing system to assess the overall vitality and construct a vitality dataset. And through analyzing the correlation between the overall vitality and these five indicators, we obtain the optimal vitality quantification model. The experimental results indicate that the guidance of our expert assessment method efficiently improves the reliability of overall vitality assessment. And our vitality quantification model yields higher goodness of fit to the overall vitality than the model constructed with shallow attributes which are the number of people and staying, which also verify the great contribution of trajectory diversity and trajectory complexity.

Based on our extensive experiments, we can conclude that our vitality quantification method is feasible for fine-scale and comprehensive spatial-temporal vitality research, and it is convenient to objectively measure vitality as a uniform and independent variable, which is urgently needed. Moreover, due to the accessibility of data collection, it can be used to analyze the small public spaces in different areas and time. The subjective expert assessing system proposed in this article presents a method to assign vitality labels for small public spaces according to multidimensional representation factors. Researchers can measure the effectiveness of their proposed vitality evaluation method using vitality labels to improve their methods.

CRediT authorship contribution statement

Tong Niu: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation. **Linbo Qing:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Formal analysis, Conceptualization. **Longmei Han:** Writing – review & editing, Project administration, Formal analysis, Conceptualization. **Ying Long:** Writing – review & editing, Formal analysis, Conceptualization. **Jingxuan Hou:** Writing – review & editing, Formal analysis. **Lindong Li:** Writing – review & editing, Investigation, Data curation. **Wang Tang:** Writing – review & editing, Data curation. **Qizhi Teng:** Supervision.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Linbo Qing reports financial support was provided by National Natural Science Foundation of China.

Data availability

The authors do not have permission to share data.

Acknowledgment

This work was supported by the National Nature Science Foundation of China under Grant 61871278.

References

[1] J. Gehl, Life between Buildings, Van Nostrand Reinhold, New York, 1987. https://www.academia.edu/download/49875924/Jan Gehl-Life Between Buildings Using Public Space-Island Press 2011.pdf.

[2] J.A. Sanchez-Vaquerizo, D.C. Llach, The social life of small urban spaces 2.0, in: International Conference on Computer-Aided Architectural Design Futures, 2019, pp. 295–310, https://doi.org/10.1007/978-981-13-8410-3 21.

[3] J. Hou, L. Chen, E. Zhang, H. Jia, Y. Long, Quantifying the usage of small public spaces using deep convolutional neural network, PLoS One 15 (10) (2020), e0239390, https://doi.org/10.1371/journal.pone.0239390.

[4] H. Ji, L. Qing, L. Han, Z. Wang, Y. Cheng, Y. Peng, A new data-enabled intelligence framework for evaluating urban space perception, ISPRS Int. J. Geo-Inf. 10 (6) (2021) 400, https://doi.org/10.3390/ijgi10060400.

[5] M.Z. Pakoz, M. Isık, Rethinking urban density, vitality and healthy environment in the post-pandemic city: the case of istanbul, Cities 124 (2022), 103598, https://doi.org/10.1016/j.cities.2022.103598.

[6] C. Jia, Y. Liu, Y. Du, J. Huang, T. Fei, Evaluation of urban vibrancy and its relationship with the economic landscape: a case study of beijing, ISPRS Int. J. Geo-Inf. 10 (2) (2021) 72, https://doi.org/10.3390/ij gi10020072.

[7] V. Mehta, Lively streets: determining environmental characteristics to support social behavior, J. Plann. Educ. Res. 27 (2) (2007) 165–187, https://doi.org/10.1177/0739456X07307947.

[8] M.G.H. Lunecke, R. Mora, The layered city: pedestrian networks in downtown santiago and their impact on urban vitality, J. Urban Des. 23 (3) (2018) 336–353, https://doi.org/10.1080/13574809.2017.1369869.

[9] H.G. Sung, D.H. Go, C.G. Choi, Evidence of jacobs's street life in the great seoul city: identifying the association of physical environment with walking activity on streets, Cities 35 (2013) 164–173, https://doi.org/10.1016/j.cities.2013.07.010.

[10] D.O.S. Jalaladdini, Urban public spaces and vitality: a socio-spatial analysis in the streets of cypriot towns, Procedia Soc. Behav. Sci. 35 (2012) 664–674, https://doi.org/10.1016/j.sbspro.2012.02.135.

[11] O. Marquet, C. Miralles-Guasch, The walkable city and the importance of the proximity environments for barcelona's everyday mobility, Cities 42 (2015) 258–266, https://doi.org/10.1016/j.cities.2014.10.012.

[12] W. Wu, X. Niu, Influence of built environment on urban vitality: case study of shanghai using mobile phone location data, J. Urban Plann. Dev. 145 (3) (2019), 04019007, https://doi.org/10.1061/(ASCE)UP.1943-5444.0000513.

[13] S. Liu, L. Zhang, Y. Long, Y. Long, M. Xu, A new urban vitality analysis and evaluation framework based on human activity modeling using multi-source big data, ISPRS Int. J. Geo-Inf. 9 (11) (2020) 617, https://doi.org/10.3390/ijgi9110617.

[14] X. Delclo's-Alio, A. Gutierrez, C. Miralles-Guasch, The urban vitality conditions of jane jacobs in barcelona: residential and smartphone-based tracking measurements of the built environment in a mediterranean metropolis, Cities 86 (2019) 220–228, https://doi.org/10.1016/j.cities.2018.09.021.

[15] J. Wu, N. Ta, Y. Song, J. Lin, Y. Chai, Urban form breeds neighborhood vibrancy: a case study using a gps-based activity survey in suburban beijing, Cities 74 (2018) 100–108, https://doi.org/10.1016/j.cities.2017.11.008.

[16] C. Ratti, D. Frenchman, R.M. Pulselli, S. Williams, Mobile landscapes: using location data from cell phones for urban analysis, Environ. Plann. Plann. Des. 33 (5) (2006) 727–748, https://doi.org/10.1068/b32047.

[17] X. Zhang, Y. Sun, T.O. Chan, Y. Huang, Z.L.A. Zheng, Exploring impact of surrounding service facilities on urban vibrancy using tencent location-aware data: a case of guangzhou, Sustainability 13 (2) (2021) 444, https://doi.org/10.3390/su13020444.

[18] W. Yue, Y. Chen, P.T.M. Thy, P. Fan, Y. Liu, W. Zhang, Identifying urban vitality in metropolitan areas of developing countries from a comparative perspective: Ho chi minh city versus shanghai, Sustain. Cities Soc. 65 (2021), 102609, https://doi.org/10.1016/j.scs.2020.102609.

[19] C. Zeng, Y. Song, Q. He, F. Shen, Spatially explicit assessment on urban vitality: case studies in chicago and wuhan, Sustain. Cities Soc. 40 (2018) 296–306, https://doi.org/10.1016/j.scs.2018.04.021.

[20] Y. Meng, H. Xing, Exploring the relationship between landscape characteristics and urban vibrancy: a case study using morphology and review data, Cities 95 (2019), 102389, https://doi.org/10.1016/j.cities.2019.102389.

[21] C. Wu, X. Ye, F. Ren, Q. Du, Check-in behaviour and spatio-temporal vibrancy: an exploratory analysis in shenzhen, China, Cities 77 (2018) 104–116, https://doi.org/10.1016/j.cities.2018.01.017.

[22] S. Lu, C. Shi, X. Yang, Impacts of built environment on urban vitality: regression analyses of beijing and chengdu, China, Int. J. Environ. Res. Publ. Health 16 (23) (2019) 4592, https://doi.org/10.3390/ijer ph16234592.

[23] Y. Li, N. Yabuki, T. Fukuda, Exploring the association between street built environment and street vitality using deep learning methods, Sustain. Cities Soc. 79 (2021), 103656, https://doi.org/10.1016/j.scs.2021.103656.

[24] S. Williams, C. Ahn, H. Gunc, E. Ozgirin, M. Pearce, Z. Xiong, Evaluating sensors for the measurement of public life: a future in image processing, Environ. Plann. B: Urban Analytics and City Science 46 (8) (2019) 1534–1548, https://doi.org/10.1177/2399808319852636.

[25] W. Yan, D.A. Forsyth, Learning the behavior of users in a public space through video tracking, in: 2005 Seventh IEEE Workshops on Applications of Computer Vision (WACV/MOTION'05), 1, 2005, pp. 370–377, https://doi.org/10.1109/ACVMOT.2005.67.

[26] S. Liang, H. Leng, Q. Yuan, B. Wang, C. Yuan, How does weather and climate affect pedestrian walking speed during cool and cold seasons in severely cold areas? Build. Environ. 175 (2020), 106811 https://doi.org/10.1016/j.buildenv.2020.106811.

[27] N.B.W.T. Xu, H. Ma, The influence of spatial vitality around subway stations in beijing on pedestrians' emotion, in: 13th International Symposium for Environment-Behavior Studies (EBRA), 2018, pp. 268–274. https://www.researchgate.net/publication/328759238.

[28] W. Ge, R.T. Collins, R.B. Ruback, Vision-based analysis of small groups in pedestrian crowds, IEEE Trans. Pattern Anal. Mach. Intell. 34 (5) (2012) 1003–1016, https://doi.org/10.1109/TPAMI.2011.176.

[29] K. Yamaguchi, A.C. Berg, L.E. Ortiz, T.L. Berg, Who are you with and where are you going?, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition CVPR), 2011, pp. 1345–1352, https://doi.org/10.1109/CVPR.2011.5995468.

[30] J. Sun, Q. Jiang, C. Lu, Recursive social behavior graph for trajectory prediction, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern

Recognition, CVPR), 2020, pp. 660–669, https://doi.org/10.1109/CVPR42600.2020.00074.
[31] B. Solmaz, B.E. Moore, M. Shah, Identifying behaviors in crowd scenes using stability analysis for dynamical systems, IEEE Trans. Pattern Anal. Mach. Intell. 34 (10) (2012) 2064–2070, https://doi.org/10.1109/TPAMI.2012.123.
[32] D. Mařcetic, L. Malës, S. Ribaric, Crowd motion pattern detection at the microscopic level, in: 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2019, pp. 1093–1098, https://doi.org/10.23919/MIPRO.2019.8756660.
[33] J. Jacobs, The Death and Life of Great American Cities, Vintage, 2016, https://doi.org/10.2307/794509.
[34] J. Gehl, B. Svarre, How to Study Public Life, Springer, 2013. https://link.springer.com/content/pdf/10.5822/978-1-61091-525-0.pdf.
[35] W.H. Whyte, The Social Life of Smal Urban Spaces, Project for Public Spaces, New York, NY, USA, 1980. https://www.academia.edu/38809877/The social life of the_small urban_space?from=cover page.
[36] J. Montgomery, Editorial urban vitality and the culture of cities, Plann. Pract. Res. 10 (2) (1995) 101–110, https://doi.org/10.1080/02697459550036649.
[37] C.E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (3) (1948) 379–423, https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.
[38] T. Eiter, H. Mannila, Computing discrete fréchet distance, Technical Report, http://citeseerx.ist.psu.edu/viewdoc/do wnload?doi=10.1.1.90.937&rep=rep1&type=pdf, 1994.
[39] S.C. Johnson, Hierarchical clustering schemes, Psychometrika 32 (3) (1967) 241–254, https://doi.org/10.1007/BF02289588.
[40] R.R. Yager, On the specificity of a possibility distribution, Fuzzy Set Syst. 50 (3) (1992) 279–292, https://doi.org/10.1016/0165-0114(92)90226-T.
[41] J.G. Lee, J. Han, K.Y. Whang, Trajectory clustering: a partition-and-group framework, in: Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data, 2007, pp. 593–604, https://doi.org/10.1145/1247480.1247546.
[42] Y. Li, H. Ren, Visual analysis of vessel behaviour based on trajectory data: a case study of the yangtze river estuary, ISPRS Int. J. Geo-Inf. 11 (4) (2022) 244, https://doi.org/10.3390/ijgi11040244.
[43] T.I. Ohta, K. Maenobu, T. Sakai, Obtaining surface orientation from texels under perspective projection, Int. Joint Conf. Artif. Intell. 81 (1981) 746–751. http://www.researchgate.net/publicati on/220816317 Obtaining Surface Orientation from Texels under Perspective _Projection.
[44] J. Su, X. He, L. Qing, T. Niu, Y. Cheng, Y. Peng, A novel social distancing analysis in urban public space: a new online spatio-temporal trajectory approach, Sustain. Cities Soc. 68 (2021), 102765, https://doi.org/10.1016/j.scs.2021.102765.
[45] R. Ewing, O. Clemente, K.M. Neckerman, M. Purciel-Hill, J.W. Quinn, A. Rundle, Measuring Urban Design: Metrics for Livable Places, Springer, 2013. https://link.springer.com/content/pdf/10.5822/978-1-61091-209-9.pdf.
[46] F. Aletta, J. Kang, Towards an urban vibrancy model: a soundscape approach, Int. J. Environ. Res. Publ. Health 15 (8) (2018) 1712, https://doi.org/10.3390/ijerph15081712.
[47] H. Wang, F. Yang, L. Liu, Comparison and application of standardized regressive coefficient and partial correlation coefficient, J. Q. Tech. Econ. 9 (2006) 150–155. http://en.cnki.com.cn/Article_en/CJFDTOTAL-SLJY200609016.htm.
[48] G. Jocher, A. Stoken, J. Borovec, C. Liu, Laughing, A. Hogan, Lorenzomammana Tkianai, Yxnong AlexWang1900, L. Diaconu, Marc Wanghaoyang0106, Ml5ah, Doug Hatovix, J. Poznanski, L. Yu, Changyu98, P. Rai, R. Ferriday, T. Sullivan, X. Wang, YuriRibeiro, E.R. Claramunt, Hopesala, P. Dave, Yzchen, ultralytics/yolov5: v3.0. https://zenodo.org/record/3983579#.YUWeQSviuM8, 2020. NanoCode012, ChristopherSTAN.
[49] N. Wojke, A. Bewley, D. Paulus, Simple online and realtime tracking with a deep association metric, in: 2017 IEEE International Conference on Image Processing, ICIP, 2017, pp. 3645–3649, https://doi.org/10.1109/ICIP.2017.8296962.
[50] J.L. Fleiss, B. Levin, M.C. Paik, Statistical Methods for Rates and Proportions, John Wiley & Sons, 2013. http://www.emis.de/MATH-item?1034.62113.
[51] J.R. Landis, G.G. Koch, The measurement of observer agreement for categorical data, Biometrics 33 (1) (1977) 159–174, https://doi.org/10.2307/2529310.
[52] J.M. Bland, D.G. Altman, Statistics notes: Cronbach's alpha, BMJ 314 (7080) (1997) 572, https://doi.org/10.1136/bm j.314.7080.572.
[53] T.K. Koo, M.Y. Li, A guideline of selecting and reporting intraclass correlation coefficients for reliability research, J. Chiropractic Med. 15 (2) (2016) 155–163, https://doi.org/10.1016/j.jcm.2016.02.012.
[54] M. Tranmer, M. Elliot, Multiple linear regression, Cathie Marsh Centre Census Surv. Res. (CCSR) 5 (5) (2008) 1–5. http://hummedia.manchester.ac.uk/institutes/cmist/archive-publications/working-papers/2020/multiple-linear-regression.pdf.
[55] J.I. Daoud, Multicollinearity and regression analysis, J. Phys. Conf. 949 (2017), 012009, https://doi.org/10.1088/1742-6596/949/1/012009.