Green View Index Analysis and Optimal Green View Index Path Based on Street View and Deep Learning

Anqi Hu
Osaka Metropolitan University
anqihu1028@gmail.com

Jiahao Zhang
Osaka University
zhangjiahao@hp-info.med.osaka-u.ac.jp

Hiroyuki Kaga
Osaka Metropolitan University
kaga@envi.osakafu-u.ac.jp

Abstract

Streetscapes are an important part of the urban landscape, analysing and studying them can increase the understanding of the cities’ infrastructure, which can lead to better planning and design of the urban living environment. In this paper, we used Google API to obtain street view images of Osaka City. The semantic segmentation model PSPNet is used to segment the Osaka City street view images and analyse the Green View Index (GVI) data of Osaka area. Based on the GVI data, three methods, namely corridor analysis, geometric network and a combination of them, were then used to calculate the optimal GVI paths in Osaka City. The corridor analysis and geometric network methods allow for a more detailed delineation of the optimal GVI path from general areas to specific routes. Our analysis not only allows for the calculation of specific routes for the optimal GVI paths, but also allows for the visualisation and integration of neighbourhood landscape data. By summarising all the data, a more specific and objective analysis of the landscape in the study area can be carried out and based on this, the available natural resources can be maximised for a better life.

1. Introduction

The urban streetscape is an important part of the urban landscape. The environmental resources that people live in possibly affect their lifestyle. Several studies have shown a positive relationship between the availability of the urban landscape and the health of inhabitants, which may provide opportunities for health improvement [1, 2]. It can also make some extent contributes to noise mitigation [3, 4]. In addition, streetscapes have a relationship to heat regulation in cities, as street vegetation can reduce the heat island effect through shading and transpiration [5, 6]. Therefore, street greening is not only a significant presence in people in terms of aesthetics, but also a convenient strategy for adaptive environmental design in urban life to create thermally comfortable and more attractive living environments.

There are many evaluation indexes for the evaluation of street greening, among which the current one is more widely focused on the Green View Index (GVI). The GVI can quantify the amount of greenery within the pedestrians’ field of view and is more suitable for describing the environment observed by the human eye, which can partially compensate for the shortcomings of traditional assessment indexes [7]. With the development of science and technology, there has been a great progress from the collection of street view image to image processing [8, 9]. However, most studies stay at the level of collecting basic GVI data [10], focusing on how to use new technical tools to turn urban streetscapes into visual data charts to analyze the distribution characteristics of urban visual greenery, such as GVI maps, heat maps, etc. Therefore, there have some problems proposed: how to take advantage of the GVI data after describing the greening degree, how to use the data to build a more humanized and diversified street environment, and how to live better in the existing environment.

In this paper, we aim to calculate the GVI data by street image analysis and use the GVI data to calculate the optimal GVI paths in the city. The flowchart is shown in Fig.1.

(1) Firstly, the coordinates of all street network road intersections in Osaka City, were collected through the google open source API as the investigation site for the study, with a total of 49770 points. In order to collect more comprehensive street view information, we set the width of view field to 60°, and selected one image at every 60 degrees to get a set of six images containing 0°, 60°, 120°, 180°, 240°, and 300°, which can cover 360-degree panoramic street view (see Fig.2). A total of 298,620 photos of Osaka City were
collected.

(2) The obtained images are segmented by image segmentation model PSPNet \cite{11}. The value of GVI, which is the ratio of green view index (vegetation and terrain) in the landscape elements of the street view image, was calculated to assess the degree of green view of the image.

(3) Create a map of the GVI distribution of Osaka City. The evaluation criterion about the GVI is referred to the survey report of the Ministry of Land, Infrastructure, Transport and Tourism of Japan\cite{14} and the satisfaction criteria of the first phase of the Kyoto Greening Promotion Plan\cite{2} about the GVI. The GVI was divided into four grades: 0~10\%, 10~18\%, 18~25\%, and 25\% or more. Based on this criterion, a general overview of GVI distribution and degree of satisfaction in Osaka City was constructed.

(4) The geographic data software ArcGIS was used to calculate the optimal GVI path in Osaka City. We select two points as start point and end point to calculate the optimal GVI path. Without doubt, the route with the shortest route, i.e., without turning back, is a prerequisite for the calculation of the optimal path. Then the method is optimized in a relatively small area in Osaka City to verify feasibility. The North Senba in Chuo-ku was selected to analyse GVI path, based on GVI value of each intersection in direction to obtain the optimal GVI path.

The following contributions were achieved: (1) Visualized statistics were compiled on the distribution of GVI in Osaka City. (2) We evaluate and visualize the satisfaction of GVI in Osaka City. (3) Several calculation methods of the optimal GVI path were proposed based on GVI distribution in Osaka City. (4) We also analyzed the directional GVI of streets. It can not only be used to analyze the coherence of continuous landscape, but also to provide intuitive guidance on how to complement the streetscape in areas with low GVI value, based on analysis of direction.

2. Related Works

Google Street View is an interactive web map that is highly accessible and has a wide coverage all over the world. It provides a 360° panoramic view of the city, capturing all scenes of a street or neighborhood. Street view shows the basic information of the landscape from the human perspective and offers considerable possibilities for urban visual greening studies. Street view image has now been used extensively in various researches \cite{12,13}. These studies have shown that Street view image datasets are useful and quantitative tools to help policy makers, planners and researchers to understand the streetscape from the human perspective. In recent years, Google Maps\cite{3} and Baidu Maps\cite{4} have provided POI alternatives in their public GIS databases. With the development of these public GIS databases, easier access to raw data (street view image API) facilitates large scale street view image based research \cite{5,14}.

In computer vision field, deep learning uses deeply neural networks for feature extraction and parameter optimization. Semantic segmentation is one of the branches to classify each pixel in an image with different classes based on training data. It is a kind of supervised learning \cite{11}. For natural images, it is achieved by using neural networks to identify landscape categories on pixel-level, rather than identifying their type and position in the overall image. There are many scenarios where semantic segmentation is currently used, for example, in civil engineering \cite{15}, to detect cracks in concrete; in medical field to identify diseases in X-ray images \cite{10} and in autonomous driving for road

\url{https://www.mlil.go.jp/kisha/kisha05/04/040912_3/01.pdf} \url{https://www.city.kyoto.lg.jp/kensetu/cmsfiles/contents/0000102/102008/planhonpen.pdf} \url{https://www.google.com/streetview/} \url{http://lbsyun.baidu.com/}
boundary and object detection [17].

Models for semantic segmentation have developed rapidly in recent years, and representative models include U-Net [18], SegNet [19], Deep-lab V3 [20], PSPNet [11], and so on. All these models have shown good results in terms of segmentation accuracy. In this study, we use the Pyramid Scene Parsing Network (PSPNet) [11] to analyze street view image to calculate GVI value. The basic structure of PSPNet is shown in Fig.3.

PSPNet has been designed with a basic Encoder-Decoder structure to implement semantic segmentation. For the inference process, specifically, in the Encoder part on the left side of the figure, features are extracted from the original input image to integrate and simplify the basic information. The feature is input into different convolution layer with different kernel size in order to get as much as possible features of original image with different level. All the integration processes are completed in a structure named pyramid pooling module. In the Decoder part on the right side of the figure, several feature maps is up-sampled to same size in order to concatenate all feature maps with different level. All pixels are classified based on learned weights. After the process of restoration, we can get colored output image with segmented classes. In PSPNet, the Encoder part uses the Resnet [21] model for complex and advanced feature extraction, because it has perfect performance on feature extraction.

There have existed many pioneering studies on the integration of landscape data processing with Geographic Information System (GIS) research, such as [22] in which used geographic data analysis tools of ArcGIS to analyze and assess the accessibility of urban green spaces. In [23], aerial photographs were implemented on GIS for view area analysis. Three types of landscapes were quantified: open landscapes (visibility), green landscapes (visibility of open spaces), and marine landscapes (visibility of the ocean). It is shown that ArcGIS is a very useful tool in achieving the level of integration of photo information quantification and geographic data [23].

3. Material and Method in Deep Learning

In this study, we constructed a GVI distribution map system according to the following process: (1) Obtain the street view images of Osaka City. (2) Semantic segmentation of streetscape images with PSPNet [11]. (3) Calculate the GVI of different directions and the average GVI for each point. (4) Integrate the geographic data with GIV data using ArcGIS to visualize the greening rate. (5) The route with the largest average GVI was extracted with three different analysis methods.

3.1. Research Area

Since the optimal green view paths in this thesis were analyzed step by step using different methods, the target sites of the study were implemented from the beginning in Osaka City to the North Senba area in Chuo-ku, Osaka City. Osaka City in Japan is the administrative, economic, cultural and transportation center of the Kinki region and western Japan. It has an area of approximately 225.21 km² and is composed of 24 administrative districts. The Osaka Metropolitan Area and the Keihan-Kobe Metropolitan Area are formed with Osaka City as the center. The Keihan-Kobe Metropolitan Area is second only to the Tokyo Metropolitan Area in terms of Gross Domestic Product (GDP) in Japan and ranks among the highest in the world. It was ranked 35th in the world in the "Global Cities Index 2020" ranking of world cities by a U.S. think tank [5]. It is well suited as an object city for studying urban landscapes. Therefore, we set Osaka City as first target area to calculate and visualize GVI distribution as a way to understand the city as a whole. The visualization of the distribution facilitates a more intuitive understanding.

In the part of optimal GVI path, it is unpractical to analyze GVI path of all Osaka City because we could not get a view facing the road at the intersection in Osaka City by Google Map APIs. So in order to verify whether the methods of getting optimal GVI path are feasible, we select a...
representative area in Osaka City where at each intersection, it is possible to get a perfect angle of view facing due east, west, north and south, i.e. the direction of view is parallel to the road. Therefore we set North Senba as the second target area to analyse detailed optimal GVI path. North Senba in Osaka City is the Central Business District (CBD) of Osaka City. It is the center of Osaka’s merchant culture and is a quadrangular area of about 2 km from north to south and 1 km from east to west. The orthogonal pattern of the streets and the north-south alignment of the roads make it easier to automatically download street view images, which is ideal for the study of directional street views. Specifically, North Senba is tends to be square, about 0.97 km$^2$, with 12 roads in east-west direction and 13 roads in north-south direction, with 153 intersections (see Fig.4).

3.2. Get Street View Images of Osaka City

In order to get map networks of Osaka City, we used the Python OSMnx package [24] to obtain the coordinate data of all intersections in Osaka city from the Open Street Map. With the overall networks, we can get coordinates of all the intersections. In recent years, companies such as Google, Amazon, and Twitter have been actively providing data through web service APIs (application programming interfaces) for the purpose of leveraging the various types of big data they have. Google Street View Image API makes it easy to download Google Street View images. Therefore, we collected street view images by Google Map API with obtained coordinates of all intersections in Osaka city (a total of 49770 points). Street view images were collected for six angles (0°, 60°, 120°, 180°, 240°, 300°) in the direction of the street centered at each intersection. The size of them was 640×640, because by Google Map API we can only get the maximum size of 640×640. In order to get more accurate result by PSPNet, we need as much as possible to get high resolution of input image. We totally downloaded 298,620 images of Osaka City. One thing need to be mentioned, the street view images obtained may not be suitable for analysis, such as night-time street scenes, indoor images and blurred images. These points need to be manually filtered and discarded.

3.3. Semantic Segmentation and Calculate GVI

In this study, we use PSPNet [11] which was pre-trained on Cityscapes [25] to generate segmented images into 19 landscape elements. For the training process, we download labeled opensource dataset Cityscapes which contains 5,000 images with correlated fine label in Europe. 75% of them is for training, 10% for validation and 15% for test. Even though the basic composition and details are different, for the composition on class level, we can get clear segmentation result by PSPNet which is pre-trained on Cityscapes. The colored label image is transformed to true label image which use index of class as pixel number in order to improve training efficiency. Because the original training data has large image size, we also need to set crop window with relatively small size on local part of original image, ensuring not to lose much basic information if we straight resize the original image in training dataset. For the evaluation criterion of training result, we use Intersection over Union (IoU) to evaluate the result of segmentation. It is consisted of three part, for each class, the generated region of one class is calculated by IoU with label region. The IoU is defined as below:

$$IoU = \frac{TP}{TP + FP + FN} \quad (1)$$

Where $TP$ is True Positive, $FP$ is False Positive and $FN$ means False Negative. This criterion can by simplified by the ratio of intersection area of each class between generated region and label region to the union area between them. The mean IoU of PSPNet achieved 79.6% in our test set. The inference result is shown in Fig[5]. In fact, the size of the generated segmented image is closely related to the generation efficiency. As Osaka City contains a large number of images, we have reduced the generated image size from 2048×1024 to the original image size, which is
640×640, in order to improve the inference efficiency of segmentation. We spent almost 90 hours to generate segmented street view images of Osaka City. In order to calculate ratio of greenery in a whole street view image, we need to select “vegetation” and “terrain” as target colors to calculate GVI value.

In a segmented image, for each angle $i \in (0^\circ, 60^\circ, 120^\circ, 180^\circ, 240^\circ)$, $G_i$ is the number of greenery pixels. $T_i$ means the total number of this segmented image. $GVI_i \in \{0, 1\}$ means corresponding GVI value of this angle, which is given by

$$GVI_i = \frac{G_i}{T_i} \times 100\% \quad (2)$$

For each intersection point $M$, which contains number of $m$ street view images. The average of $GVI_i$ represents GVI value of this point. Therefore, the $GVI_{avg}$ of each intersection is given by

$$GVI_{avg} = \frac{1}{m} \sum_{i=1}^{m} GVI_i \times 100\% \quad (3)$$

Where $m = 6$ in this paper.

3.4. Integrate GVI with ArcGIS

After getting the geographic data included street line and point network by open street map, we can also get an excel with GVI data and coordinate. In order to make point and line map to visualize the GVI distribution in Osaka City, we use ArcGIS to integrate them manually. The created point map of the distribution of GVI value in Osaka City is shown in Fig.6(a). Each point is represented by $GVI_{ave}$ and its coordinate with various color depths in four classes.

In order to get street GVI value of Osaka City to represent GVI line map, based on GVI point map, we have adopted a criterion that uses the average of the GVI values of the vertices on both sides of the street to represent the value of the whole street. Therefore, the line graph of GVI distribution in Osaka city was obtained as shown in Fig.6(b).

4. Methods of Optimal GVI Path

4.1. Reference Corridor Analysis

By the cell statistics tool, we can get the raster returned by the corridor analysis tool, the sum of the cost distances (cumulative cost) of the two input cost rasters is calculated for each image location. For each raster location, the sum of the two raster costs is used to identify the minimum cost path from one source to another and through that raster location. With this idea, the line graphs of the GVI distribution of Osaka City were exported to raster files according to the satisfaction thresholds of GVI (0~10%, 10~18%, 18~25%, and more than 25%) and assigned resistance values of 200, 150, 100, 50 respectively. The land use data of Osaka City was also assigned with the corresponding
resistance values according to the land attributes. Assign a resistance value greater than 200 to all land outside the road. After integrating the above files into one raster file, the corridor analysis was performed using the cost distance tool. Take Osaka Castle as the starting point and Sumiyoshi Taisha as the destination as an example. Using Osaka Castle as a source, two cost accumulation rasters were created. Using Sumiyoshi Taisha as another source, the process of creating cost surfaces based on individual image locations actually occurs at each image location of the input raster, and the total cumulative cost of the path through the image is calculated. The results obtained are shown in Fig. 7(a).

**4.2. Using Geometric Network Analysis**

Geometric network analysis has many applications in calculating the shortest distance between two points, and this logic can also be applied to try to calculate the optimal GVI path between two points. First, we import the GVI point map and GVI line map of the GVI distribution in Osaka City, and transform them into dataset to generate a geometric network. By setting the weight of GVI in the geometric network to be greater than the weight of road length, the geometric network version of the optimal GVI path of Osaka City can be obtained. The starting point is Osaka Castle and the end point is Sumiyoshi Taisha (see Fig. 7(b)).

**4.3. A Combination of the Above Two Methods**

The criterion for the optimal GVI path is the combined maximum value of the average GVI of the intersections in the selected route. The GVI of each intersection is used as a measure to calculate the optimal GVI path for any location. However, in a continuous landscape of real life, the human eye mainly observes the landscape in the direction of travel. Therefore, a more accurate calculation method for the optimal GVI path should be proposed to select the direction of maximum GVI at each intersection, and finally the GVI of the overall travel direction of the route gets the maximum value.

In the example of the North Senba in Osaka City, we collected street view images from the angle of the road and processed them to obtain the GVI rate of each direction of the intersection, following the method of collecting and analyzing street view images in Osaka City. We obtained 0°, 90°, 180°, and 270° street view images of each intersection. The data were then imported into ArcGIS, and a network dataset consisting of paths and nodes was created using ArcGIS Network Analyst. The best path was calculated based on the street direction GVI as shown in Fig. 8.

**5. Result and Summary**

In the part of GVI analysis in this paper, there are 37,622 intersections with an average GVI between 0-10% of the 49,770 intersections in Osaka City overall, accounting for...
Figure 8. Calculation of the optimal GVI path using the network analysis method.

75.59% of the total; there are 7,914 intersections with an average GVI between 10-18% of the total; there are 2,629 intersections with an average GVI between 18-25% of the total 5.28% of the total; 1,605 intersections (3.22% of the total) had an average GVI of 25%. From the result of statistics with criterion of satisfaction, it can be seen that Osaka City has a low level of satisfaction with the green view in terms of streetscape (see Fig.9).

Based on the distribution data of the GVI in Osaka City, the results of calculating the optimal GVI path can be derived. The following summarizes the results of different three methods: reference corridor analysis, geometric network analysis, and combination of them.

The threshold output generated in the reference corridor analysis can be considered as the least-cost corridor of the image element, rather than the least-cost path. From Osaka Castle to Sumiyoshi Taisha, the path area is planned, not a specific route. Further and more detailed route planning is needed for practical application. For the second method, geometric network analysis, it can calculate the detailed optimal GVI path. The path from Osaka Castle to Sumiyoshi Taisha is also used as an example, the geometric network analysis method can be used to design multiple stopping points, and this calculation method provides a new way of thinking when designing routes for urban travel. This method can be used in urban tourism route design, marathon race route design, and other situations where urban landscape needs to be displayed.

For the third method, it is a combination of first and second methods. The optimal GVI path is calculated based on the directional GVI. Take the example of the northeast point to the southwest point in the North Senba of Osaka City. We can take advantage of traditional method to get the optimal GVI path, because the detailed GVI value of different due direction of each point is calculated, we can get more realistic optimal GVI path as application in real world. However, this method is tedious in GVI statistics. Since the roads in the North Senba in Osaka City are orthogonally distributed, it is possible to set the street view images to be downloaded automatically at 0°, 90°, 180°, and 270°. However, it is time consuming to count street view images of road directions in areas where the angle of road intersections is complicated and uncertain. This method can be better implemented if there is a database about the road angles of the street intersections in the surveyed cities.

In addition, the network analysis method can be used not only for the planning of tourist routes in the area, but also for the evaluation of neighborhood landscapes and the evaluation of business values in the area. At the same time, the directional green view rate study can also analyze the coherence of the continuous landscape. It is an intuitive guide for how to make up for the poor streetscape of the area, and finally a more specific and objective landscape analysis result of the study area can be obtained by summarizing all the data.

6. Conclusion

In this paper, we implement a semantic segmentation network in deep learning to analyse the Green View Index in Osaka City. We implement the GVI analysis in a large scale and get reasonable results and visualize them with some feasible method by ArcGIS Software. After getting the basic GVI distribution of Osaka City, we proposed three different methods for optimal GVI path in order to lead to a better life style with street view.

References

[1] Sjerp De Vries, Sonja Me Van Dillen, Peter P Groenewegen, and Peter Spreeuwenberg. Streetscape greenery and health: Stress, social cohesion and physical activity as mediators. Social science & medicine, 94:26–33, 2013.

[2] E Gregory McPherson, James R Simpson, Qingfu Xiao, and Chunxia Wu. Million trees los angeles canopy cover
and benefit assessment. *Landscape and Urban Planning*, 99(1):40–50, 2011.

[3] Timothy Van Renterghem and Dick Botteldooren. View on outdoor vegetation reduces noise annoyance for dwellers near busy roads. *Landscape and urban planning*, 148:203–215, 2016.

[4] Francesco Ferrini, Alessio Fini, Jacopo Mori, and Antonella Gori. Role of vegetation as a mitigating factor in the urban context. *Sustainability*, 12(10):4247, 2020.

[5] Lun Liu, Elisabete A Silva, Chunyang Wu, and Hui Wang. A machine learning-based method for the large-scale evaluation of the qualities of the urban environment. *Computers, environment and urban systems*, 65:113–125, 2017.

[6] Biao Zhang, Ji-xi Gao, Yang Yang, et al. The cooling effect of urban green spaces as a contribution to energy-saving and emission-reduction: A case study in Beijing, China. *Building and environment*, 76:37–43, 2014.

[7] Jun Yang, Linsen Zhao, Joe McBride, and Peng Gong. Can you see green? assessing the visibility of urban forests in cities. *Landscape and Urban Planning*, 91(2):97–104, 2009.

[8] Nanqi Ye, Bowen Wang, Michihiro Kita, Ming Xie, and Wenyue Cai. Urban commerce distribution analysis based on street view and deep learning. *IEEE Access*, 7:162841–162849, 2019.

[9] Bowen Wang, Liangzhi Li, Yuta Nakashima, Ryo Kawasaki, Francesco Ferrini, Alessio Fini, Jacopo Mori, and Antonella Gori. Role of vegetation as a mitigating factor in the urban context. *Sustainability*, 12(10):4247, 2020.

[10] Liang-Chieh Chen, George Papandreou, Florian Schroff, and James Philion. Deeplabv3+: Encoder-decoder with atrous Separable Convolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(10):2340–2351, 2019.

[11] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. *ICCV, 2015.*

[12] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on pattern analysis and machine intelligence*, 39(12):2481–2495, 2017.

[13] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[14] Mehmet Cetin. Using GIS analysis to assess urban green space in terms of accessibility: case study in Kutahya. *International Journal of Sustainable Development & World Ecology*, 22(5):420–424, 2015.

[15] Yoshiki Yamagata, Daisuke Murakami, Takahiro Yoshida, Hajime Seya, and Sho Kuroda. Value of urban views in a bay city: Hedonic analysis with the spatial multilevel additive regression (SMAR) model. *Landscape and Urban Planning*, 151:89–102, 2016.

[16] Geoff Boeing. OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, 65:126–139, 2017.

[17] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. The cityscapes dataset for semantic urban scene understanding. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3213–3223, 2016.