This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Visualizing Routes With AI-Discovered Street-View Patterns

Tsung Heng Wu, Md Amiruzzaman, Ye Zhao, Deepshikha Bhati, and Jing Yang

***Abstract*—Street-level visual appearances play an important role in studying social systems, such as understanding the built environment, driving routes, and associated social and economic factors. It has not been integrated into a typical geographical visualization interface (e.g., map services) for planning driving routes. In this article, we study this new visualization task with several new contributions. First, we experiment with a set of AI techniques and propose a solution of using semantic latent vectors for quantifying visual appearance features. Second, we calculate image similarities among a large set of street-view images and then discover spatial imagery patterns. Third, we integrate these discovered patterns into driving route planners with new visualization techniques. Finally, we present VivaRoutes, an interactive visualization prototype, to show how visualizations leveraged with these discovered patterns can help users effectively and interactively explore multiple routes. Furthermore, we conducted a user study to assess the usefulness and utility of VivaRoutes.**

***Index Terms*—Driving routes, geo-visualization, street-view imagery, visual appearance.**

I. Introduction

HENRY Miller, an American Author, said that "One's destination is never a place, but rather a new way of seeing things." In this article, we present new computational and visualization methods that can help people "see things" along the rising-up roads. Roadside visual features reveal built environments and play a vital role in understanding a social system involving locations and geo-contexts. Thus, visualizing street views is of interest to applications such as urban and community planning, criminology, social equity, business, and investment. Moreover, visualizing street views is of interest to personal route planning, since they link to many social, economic, and environmental factors that can affect personal route decisions. For example, some people may prefer to drive in greenery while others may want to navigate in urban surroundings.

Manuscript received 26 June 2023; revised 19 September 2023 and 14 February 2024; accepted 14 March 2024. *(Corresponding author: Tsung Heng Wu.)*

Tsung Heng Wu, Ye Zhao, and Deepshikha Bhati are with the Department of Computer Science, Kent State University, OH 44242 USA (e-mail: twu10@kent.edu).

Md Amiruzzaman is with the Department of Computer Science, West Chester University, PA 19383 USA.

Jing Yang is with the Department of Computer Science, University of North Carolina at Charlotte, NC 28223 USA.

Digital Object Identifier 10.1109/TCSS.2024.3382944

We are all familiar with the visualizing routes on maps, mostly as color-coded trajectories, to explore, select, and navigate to destinations. Apparently, visualization of street views can be an elegant complement to existing tools. Unfortunately, street-view information is often represented by a large set of spatially sampled and heterogeneous pictures. Directly visualizing them together with urban structures and maps can easily lead to visual clutter and thus overwhelm users.

To make street-view visualization compatible with route views and enable easy understanding, it is mandatory to find a summarized way to present street-view images. According to "pattern theory," a theoretical model in visual analytics [1], abstraction in data analysis is achieved by finding patterns in data distributions. The model also regards pattern discovery as a fundamental operation in visual analytics processes. In this article, we propose several computational approaches to discovering street-view patterns. Machine learning (ML) tools are employed for data transformation to handle the big size and diversity of the raw images. Afterward, we develop new visualization methods to integrate these visual patterns within an interactive interface. The problems we tackle and our technical approaches include the following.

1) *Finding quantitative representation to compute similarities among street-view images*: It is important to define what are similar styles of street views. There exists no simple equation and optimal solution, but rather an issue related to human perception and experiences. We explore several deep learning (DL) methods to compute quantitative vectors in latent spaces that can represent inherent imagery features.
2) *Extracting area-aware visual patterns from street-view images*: With the latent vectors, we further employ several clustering methods to discover a small group of "visual appearance patterns" (VaPatterns). These clusters provide a succinct set of the aimed "patterns."
3) *Visually exploring the visual patterns over routes*: A set of deliberately designed visualizations present the discovered VaPatterns together with roads and geographical context. Visual interactions enable users to quickly explore different routes and compare them. The system also supports multiresolution exploration with both coarse and fine details over different parts of routes. A visualization prototype, **Vi**sualizing **v**isual **a**ppearance of Routes (VivaRoutes), is implemented. The system can be combined with existing route planners (currently Google Map is used).

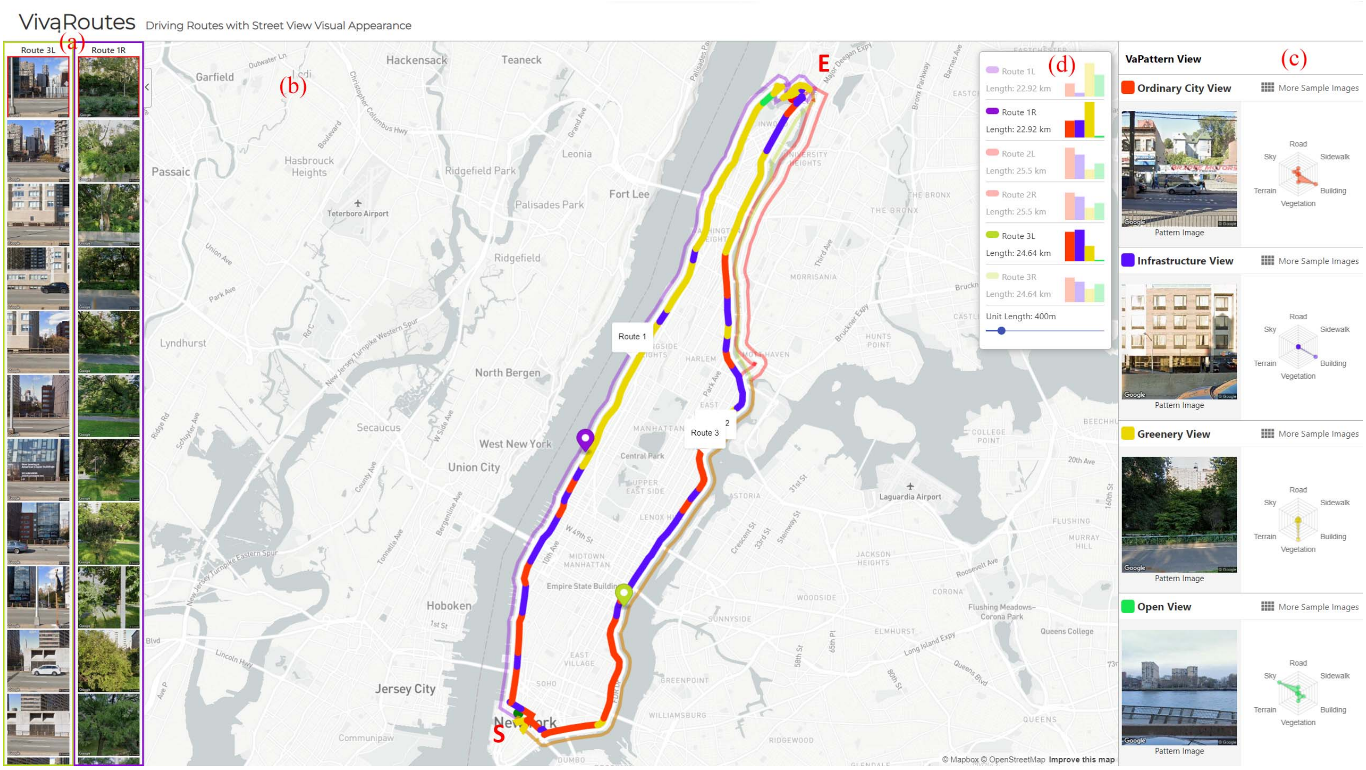


Fig. 1. VivaRoutes interface for visual exploration of driving routes in Manhattan, NY, USA. (a) Route image window shows the captured street-view images [extracted from Google Street View API (GSV)] on two selected routes from a starting location (S) to a destination (E). (b) Map view with the two routes colored according to discovered VaPatterns. (c) VaPatterns are visualized to show the street-view patterns. (d) Map inlet for route information, VaPattern distribution, and user control.

The main contributions of this article can be summarized as follows.

1) We identify the importance of street-level visual appearance and include street-view features in routing services.
2) We discover and quantify visual appearance features by employing DL and ML tools, which enable efficient visualization and interactive exploration.
3) A visualization system, VivaRoutes, is developed for visualizing and comparing visual appearances on routes.
4) A few case studies have been conducted to illustrate the usefulness and effectiveness of VivaRoutes.

VivaRoutes can be used for planning and exploring routes by tourists and commuters. Urban planners and community workers can also use it for their community study. Our approach has the potential to improve visualization systems in fields such as tourism, urban planning, and the social-economical study of communities.

II. Related Work

A. Geographic Routing

Route planning is a research topic in transportation and urban studies, as well as in logistics, autonomous vehicles, and energy saving.

A routing algorithm calculates paths between two locations (e.g., source and destination) [2] with different metrics [3] such as distance, cost, tolls, and time [4]. Context-aware route planning further adds crime, energy, and social information to route computation [5], [6] to find an optimal and feasible set of routes. Several studies (e.g., [7], [8], [9], and [10]) combine multiple criteria into one decision metric, such as weighting crime rate together with distance criteria. Moreover, recent studies of personalized route recommendations use crowdsourcing and/or social media data in finding optimal routes for specific users [11], [12], [13]. User experiences and preferences are mined and used in route computation. For example, Mirri et al. [12] find routes that provide more accessibility for elderly people by collecting data from Foursquare and Yelp. The street-view imagery, which is the focus of this article, can add another dimension to route study and give users new options in route decisions.

B. Street-View Imagery

Street-view visual contents, such as different views of roads, buildings, greenery, and sky openness, form an important environmental and social factor, which has become a critical research topic in landscaping, urban planning, transportation, and social studies [14], [15]. Traditional approaches are often conducted by in-person surveys, mapping, and remote sensing of the built environment [14]. Many researchers have used GSV images in community studies. For example, assessing damage made by tornadoes [16], understanding the association between the built environment and health outcome [17], finding green areas [18], and discovering criminal activities [19], [20] and animal habitats [21].

The recent developments in computer vision and DL technologies have made this process less expensive and faster. They can find objects and extract semantic categories from street-view images and videos (e.g., Segnet [22] and PSPnet [23]). Several studies have taken advantage of the DL models to extract semantic categories from images and use them in social studies (e.g., [24] and [25]). In this article, we extract visual appearance features with DL tools for route information visualization.

C. Spatial and Street-View Data Visualization

Various visualization systems have been developed to make sense of geospatial data in transportation and urban applications [26], [27], [28]. In particular, street-view images are utilized in a few VA systems [14], [29], [30] for visual comparison of spatial distributions and exploration of fine-grained visual details at the street scale. Geo-narratives integrate opinions and descriptions with geo-videos for social geographical research [31]. In this article, we extract and design visualizations of visual appearance features with the new goal of providing information on multiple driving routes.

III. Overview: Rationale and Methodology

Our goal is to add visual information about the street-side landscapes to route visualizations. The landscape's visual appearances can be captured from a large set of street-view images. However, these images need to be summarized as VaPatterns for easy visualization and understanding. It defines our computational goal.

The patterns would preferably match our mental pictures that represent our mind's experiences of perceiving street scenes. Usually, the mental pictures are derived with high-level abstraction and categorization. For instance, we consider one section of a road with open and green views as "country style," and another section with mixed building and road views as "town style."

We then seek ML techniques to discover the patterns from street-view images. This includes two major computational tasks (C1–C2) as follows.

1) *C1: Defining similarities among street view images*: We need to find quantitative representations of the images so that "distance" among different images can be computed. This similarity should be able to reflect the typical visual perception difference from human observers.
2) *C2: Discovering patterns by clustering the images*: We need to use appropriate clustering methods and form a necessary number of clusters. These clusters discover the patterns that can represent the perceiving styles of observers.

In Section V, we show our exploration of multiple computing algorithms in defining similarities among the images and clustering these images.

To explore the visual patterns discovered, we further develop VivaRoutes. It is an interactive route visualization interface that

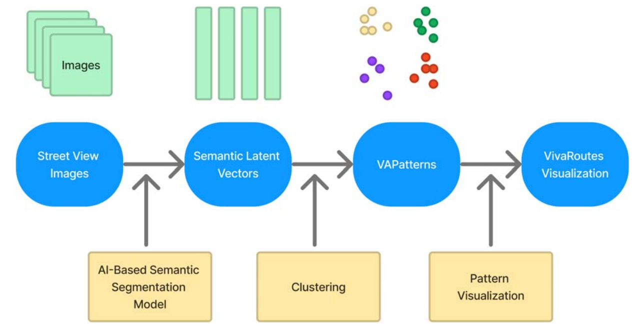


Fig. 2. Illustration of the workflow of VaPattern extraction and visualization.

integrates new visualizations and the computing algorithms in Section VI. The design goal of its visualizations is to allow people to explore alternative street-level visual appearance features along candidate driving routes. The interface should be easy to use and interactive. We identify several visualization design tasks (V1–V4) including the following.

1) *V1: Visualizing street-view patterns intuitively*: We need to summarize the extracted street-view patterns and allow users to easily understand them. The patterns need to be added to existing map views of driving routes. The new information should fit well into the geographical context and not add more recognition burden for users.
2) *V2: Supporting interactive route exploration based on street-view information*: We need to design new visualizations and interactions to help users interactively explore and understand route features and examine their road views.
3) *V3: Facilitating easy comparison of alternative routes with street-view information*: We need to provide a solution for route comparison through street-view patterns, so that users have a new way to examine and select routes.
4) *V4: Visualizing street-views in multiple scales*: We need to allow users to study the VaPatterns in fine scales (i.e., specific areas or street segments) on routes. The study should also be combined with views of street-view images.

Fig. 2 illustrates the workflow of VaPattern extraction and visualization. Through an AI-based semantic segmentation model, street view images are represented by semantic latent vectors. These vectors enable the use of clustering methods to group the images into multiple clusters, i.e., VAPatterns. These patterns are then visualized in an interactive map-based visualization system.

Next, we introduce our data collection process, explore multiple AI methods for pattern extraction, and present our new visualizations and interface that fulfill the specified requirements.

IV. Data Collection

We acquire street geometry data and street-view image data from public sources to help users plan their routes based on visual appearance. There exist many route planner APIs that compute and suggest several alternative routes from a given source to a destination location, such as Open Source Routing

Machine (OSRM), Mapbox, and Google Directions. In this article, we used Google Directions API that recommends three alternative routes.

Next, these routes are matched to street segments which are then used in street-view image retrieval and also in visualization on maps. The road network geometry data of a selected city are downloaded from the open GIS data repository, OpenStreetMap (OSM) [32].

Along these routes, we acquire street-view images from GSV with the available public API [33], [34]. In the implementation, we need to get the heading direction of each street segment on a route. Then, this direction is used to compute the view direction angles toward the left side and right side of the street. This is an essential step as the important visual appearance features are on street sides, but not on the forward (backward) views on the road itself [33].

To refine the accuracy of this computation, for each street segment (which may be a curve), we divide it into small chunks each having a length of about 20 m. Then, the heading direction of the chunk is computed as

$$\theta = \text{atan2}(x, y)$$

where

$$x = \cos(\phi_1) \times \sin(|\lambda_1 - \lambda_2|)$$
$$y = \cos(\phi_1) \times \sin(\phi_2) - \sin(\phi_1)\cos(\phi_2) \times \cos(|\lambda_1 - \lambda_2|). \tag{1}$$

Here, (ϕ_1, λ_1) and (ϕ_2, λ_2) are the latitudes and longitudes of the start and end point of this chunk, respectively. After finding the heading direction (i.e., θ) of each chunk, we get the view angles toward the left side and right side. Then, left-side and right-side street-view images at the mid-location of each chunk are retrieved from GSV API by providing the latitude, longitude, and view angles. Therefore, we acquire the visual appearance images for all roads in a spatial region, with a spatial resolution of 20 m. This resolution may be adjusted for a balance of accuracy and computational load.

Raw street-view images have to be processed to identify visual objects inside them. This is the process of image-based semantic segmentation that extracts street-view object categories for geographical scenes. Moreover, these categories on multiple images in a region are used to retrieve higher level semantic features for meaningful and intuitive representation. We employ state-of-the-art AI and ML algorithms to fulfill these tasks.

V. VaPattern and Computing Methods

A. VaPatterns

In the real world, people often form insights of visual appearances in a built-in environment as abstract and comprehensive patterns in human minds. These "patterns" are distilled from a large set of views by human knowledge and experience. They inherently link human perception impressions with urban settings. For example, one pattern may relate to most greenery and open sky leading to "country style views." Another pattern may associated with most buildings and sidewalks leading to

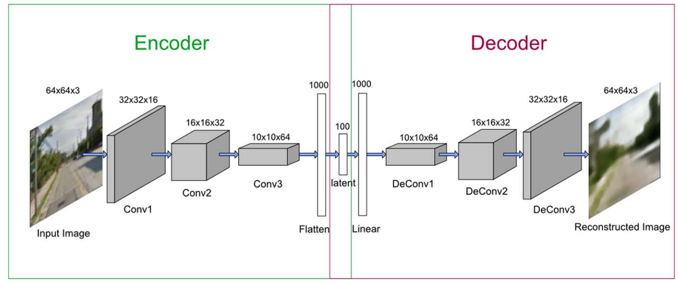


Fig. 3. Architecture of our autoencoder model.

"downtown style views." These patterns may be disparate for different locations due to terrain, weather, architecture, etc.

In a given geo-region, we can discover such patterns from all the street-view images by a visual appearance clustering approach. A small group of "VaPatterns" need to be identified and then contribute to quick visualization of driving route features.

The major challenge to fulfilling this task is the diversity of the streetview images which are affected by the setups and qualities of the cameras, driving conditions, weather, etc. Pixel-based image clustering methods cannot be directly applied. We thus seek help from DL techniques, where the task is divided into two steps: 1) defining quantitative vectors for image similarity computation; and 2) clustering the streetview images into multiple patterns. Next, we show our explorations of multiple approaches.

B. Defining Streetview Image Similarity

We need to find a quantitative representation of each input streetview image. The representation should reflect the stylish similarity and disparity among the given images so as to discover the patterns. We explore three different encoding methods from recent DL techniques to find quantitative vectors to encode the images. Next, we describe and discuss our experiments with these methods to find an optimal solution to our application.

1) Using Autoencoder for Image Encoding: We design an autoencoder to encode street-view images in a 100-dimensional latent space. The model structure is illustrated in Fig. 3. In training, a group of sampled street-view images in a geographical region passed through three convolutional layers and two linear layers to extract the relevant information (i.e., the information makes them distinguishable), and then those were flattened to get the latent vectors. The vectors are further used to reconstruct the street-view image. The activation functions between layers are mostly the rectified linear unit (ReLU) function, except for the output of the encoder, where we use the hyperbolic tangent activation function to normalize the latent vector to $[-1, 1]$. The activation function of the output of the decoder is sigmoid because it has to map the pixel values back to $[0, 1]$. In training, the loss function is implemented as the pixelwise difference between the reconstructed image and the input image. From GSV, the original street image size is 300×300 with three color channels. They are downsampled and resized to 64×64 for the autoencoder to achieve fast and effective training.

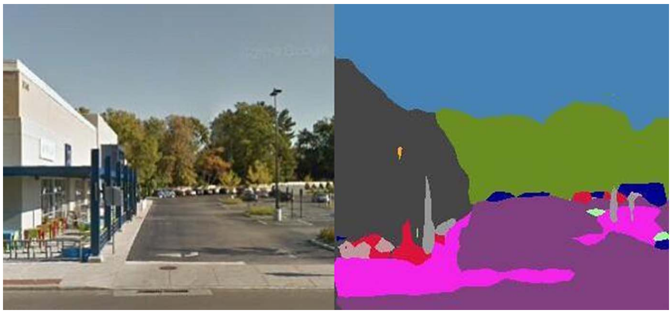


Fig. 4. Street-view image (left) extracted from GSV API and the visual semantic categories extracted by PSPNet (right). The pixels are colored by their categories, such as blue for the sky and grey for buildings. The category distribution is presented in Table I.

TABLE I
SAMPLE VECTOR OBTAINED FROM SEMANTIC SEGMENTATION FROM A STREET-VIEW IMAGE IN FIG. 4

Road	Sidewalk	Bicycle	⋯	Building
0.31	0.03	0	⋯	0.18

This architecture is compact yet efficient for the $64 \times 64 \times 3$ input size. Because our purpose of using the autoencoder is to find a low-dimensional representation rather than image content generation, so we do not choose the more complex variations of the autoencoder.

2) Using Semantic Categories for Image Encoding: Recently, DL models largely improve the effectiveness and accuracy in extracting visual objects and meaningful categories from natural images, which have been used to analyze and explore fine-grained information from street-view images [14], [30], [33]. We employ PSPNet [23], a popular deep neural network model, for extracting visual semantic categories. The PSPNet outputs a 19-dimension vector for an input image while each vector component represents the percentage of pixels belonging to one category in 19 different visual categories: they are road, sidewalk, building, wall, fence, pole, traffic light, traffic sign, vegetation, terrain, sky, person, rider, car, truck, bus, train, motorcycle, and bicycle [23]. An example is illustrated in Fig. 4, and the corresponding result vector is shown in Table I.

However, some categories may not be significant for our purpose (e.g., bicycle). We need to reduce the dimensionality of the vectors. We collect a large number (about 50 000) of street-view images in each of the U.S. regions being studied in VivaRoutes. By studying the percentages for each semantic category in all the images in one region, we identify six major categories—road, sidewalk, building, vegetation, terrain, and sky. To justify this primary observation, we conduct a *dispersion measure analysis*. Dispersion measure is often used in ML to reduce higher dimensional data and select important features [35]. It is shown to be effective in feature extraction from both nominal and categorical data [36]. It computes a dispersion ratio where a high value indicates a high-relevance feature and a low value links to a low-relevance feature in the given dataset. For our study, we use a cutoff threshold of 1.0 to select the top six categories which are the same as our direct selection. These categories form a six-dimensional vector for each image encoding as

$$V = \{\text{Road, Sidewalk, Building, Vegetation, Terrain, Sky}\}. \tag{2}$$

3) Using Semantic Latent Vector for Image Encoding: In addition to direct use of the output semantic categories from the semantic segmentation neural network, we take a deeper look into their architecture.

We realized that the latent vector from the encoder actually ciphers the semantic information and can be utilized in the similarity computation. It uses a higher dimension vector and potentially includes more latent semantic features than the final output with 19-dimensional categories.

In computation, we use a 1000-dimensional latent vector (semantic latent vector) in the ResNet-50 backbone of the DeepLab V3 [37] semantic segmentation model. The model was trained using the Cityscapes dataset, which is best suited for our use cases.

4) Results and Discussion: To show the effects, we select a diverse set of images as an example. It contains different street views in a Mideast region. After mapping these images to three different encoding vectors with the three approaches, these images are visualized in a t-SNE view in Fig. 5. Where the left view is from the autocoder, the middle view is from the semantic category vector, and the right view is from the latent vector in the semantic segmentation network. First, it can be seen that the semantic latent vector method can group the grassland images closely together. However, the other two methods cannot group them very close. For instance, two grassland pictures give users the same visual feelings but in semantic categories, the percentages of the sky in them are quite different so their representative vectors are not considered similar. Second, two pictures in Fig. 6 show two images that have different visual styles. The left one is more like a dense residential view, while the right one presents a small-town business view. These two images are highlighted in Fig. 5. It can be seen that the autoencoder separates them but one of them is made an outlier in the bottom. The semantic category vector method groups them and other similar pictures together. In contrast, the semantic latent vector method separates them and puts them in the vicinity of similar views. From these observations, we found that using the 1000-dimensional semantic vectors to encode images can better represent them for our purpose. A possible explanation is that the autoencoder can detect low-level features but does not consider semantic features, while the semantic categories discover semantic categories but do not include enough low-level features. The semantic latent vectors however can capture both low-level and semantic features since it is from a pretrained semantic segmentation neural network. Therefore, we use the semantic latent vector for the following pattern discovery work.

C. Discovering VaPatterns

ML clustering algorithms are employed to group streetview images into VaPatterns. The appropriate number of clusters (i.e., the number of "patterns") may not be a fixed value. For

(a) (b) (c)

Fig. 5. Using different image encoding approaches for image similarity for a diverse set of sample street-view images extracted from GSV API. (a) Using autoencoder; (b) using semantic categories; and (c) using semantic latent vector.

(a) (b)

Fig. 6. Two example images extracted from GSV API in a Midwestern georegion that represent two different visual patterns. (a) Residential area view. (b) Business district view.

example, in some areas, there may only be 1 or 2 different visual appearance styles while other areas may have more styles. This is depending on the size and social factors of the region.

We study several clustering methods to find a good solution in our experiment datasets including the supervised k-means and agglomerative hierarchical clustering which different numbers of clusters are tried, and an unsupervised meanshift clustering method. They are applied to all the sampled images (e.g., 5000 in the Midwestern area) in a geographical region.

In our experiment, we found that the meanshift methods cannot successfully group the images into meaningful clusters. Then, we adopted the supervised methods by trying different numbers of clusters. The selection of an optimal approach is guided by the Silhouette score, which is a popular metric for clustering result evaluation [38], [39].

In our work with the GSV images in the Manhattan area, New York, the Silhouette score drastically drops while $k = 4$ changes to $k = 5$. Therefore, we choose the four clusters that resulted from the hierarchical clustering as VaPatterns. They discover four VaPatterns in this geographical area. Please note that this work only needs to be done once for a given spatial area in the preprocessing stage.

VI. Visualizing Streetview Patterns

To meet the visualization requirements (V1–V4) in Section III, we design several new visualization tools with the extracted street-view patterns. The new visualization interface of ViVaRoutes integrates these new designs with existing route map view, as illustrated in Fig. 1.

A. Design Overview

1) New Perspective of Route: Our goal is to visualize street-level views along roads. Street views are different on the left and right sides. Therefore, in our system, each traditional route from origin to destination should be considered as two different trajectories. Therefore, in contrast to traditional route planners, we refer to route L (left) and route R (right) and show them as a pair of trajectories. To increase flexibility and easy comparison, we also allow users to select, highlight, and compare two different trajectories such as route 1R with route 3L. For simplicity, in the following, we sometimes refer to a route indeed representing one route trajectory of a specific side.

Our visualization designs are summarized according to the requirements including the following.

1) *For V1:* The discovered VaPatterns are abstracted from various street views. We design a *VaPattern view* to display their semantic contents so that users can quickly understand their key visual components such as open sky and greenery.
2) *For V2 and V3:* We include street pattern visualizations into a traditional map view of driving routes in the *ViVaRoutes interface*. It provides an additional dimension for users to study candidate routes from their origin to destination locations. Users can also study and compare routes with VaPatterns through a map inlet.
3) *For V4:* Users are allowed to explore routes with VaPatterns with zoom-in and out operation. Users can also drag markers on the map, and compare route details with their street-view images.

These views are integrated into the VivaRoutes interface as illustrated in Fig. 1, as follows.

1) VaPattern View [Fig. 1(c)] illustrates VaPatterns to quickly present their representatives of VaPatterns.
2) Map View visualizes the candidate driving routes with VaPatterns in their geographical context [Fig. 1(b)].
3) A map inlet [Fig. 1(d)] where a bar chart shows the VaPattern distribution over each route. It could support users to quickly find the visual appearance information on each side of routes, and also compare them quantitatively. Users can select and highlight two routes for a visual comparison.
4) The street-view images [Fig. 1(a)] of the two highlighted routes are shown based on user interaction by dragging markers (arrows) on the map. It is important that users can check and compare raw street views together with the route view.

Next, we describe the visualization design and functions in detail.

B. VaPattern Visualization

The role of VaPatterns is to present visual features along routes in a transparent and user-friendly way. It is of importance to present their inherent content for users with easy understanding and access. We then find semantic information to represent a pattern. In particular, each image belonging to one pattern has a six-dimensional vector as described in (2). We compute a representative *VaPattern semantic vector* by averaging such vectors of all the images. Moreover, a *VaPattern image* is retrieved by finding one street-view image whose semantic categorical vector is the closest to this pattern vector.

Afterward, in the VaPattern view [Fig. 1(c)], a radar chart [40] is adopted to present one VaPattern semantic vector with its distribution in the six semantic categories [see Fig. 1(c)]. Observers can easily find major visual categories (2) within the corresponding pattern. The corresponding VaPattern image is shown for direct understanding. Users can also open a matrix view of raw images, so as to explore more sample street-views inside this pattern. Moreover, users can name the pattern with preferred name such as "Ordinary City View" and "Infrastructure View." Each pattern is assigned a unique color which is also used in the map and other views for coordination.

C. Integrating VaPatterns With Route Visualization

People are familiar with a map-based interface where multiple candidate (recommended) routes are visualized over map view. Here, we need to couple the VaPattern features over these routes within these interfaces. The design respects the established cognitive framework of users and does not want to add a new burden. Therefore, the basic approach is mapping the VaPatterns to different colors and showing them on the trajectories. In addition, we draw two trajectories for each route representing the left and right sides, respectively.

There exists a key technical challenge: we cannot visualize the VaPatterns in different colors at all image sampling locations along a road. This will cause overwhelming color variation and cannot provide multiresolution control. Therefore, the visualization is implemented based on route segments with the following algorithm: First, the route is divided into "segments" with a user-adjustable length [in Fig. 1(d)]. The segments make it possible for controllable multiresolution information convey. Then, the street-view images of each segment are retrieved. These images belong to different VaPatterns but the dominant one with the majority is used to show the pattern of this route segment. This dominant VaPattern provides the color of the segment on map. We use color-coded line segments on the map [Fig. 1(e)] to show the patterns. It is important and easy for viewers to identify patterns directly on the map. Users can choose preferred colors for different patterns.

For observing real Street-Views, users can interactively select two route trajectories and highlight them on the map inlet [Fig. 1(d)]. The bar chart gives users a summary of VaPattern distribution and the route length. Here, route 1R and route 3L are selected and highlighted with their changing patterns. Users can also drag two markers on them for drill-down study.

In the route image window [Fig. 1(a)], users can choose which side of the routes to be shown in the panel. The raw images around the marker locations are presented. In this window, users can select to show more than two columns for their study.

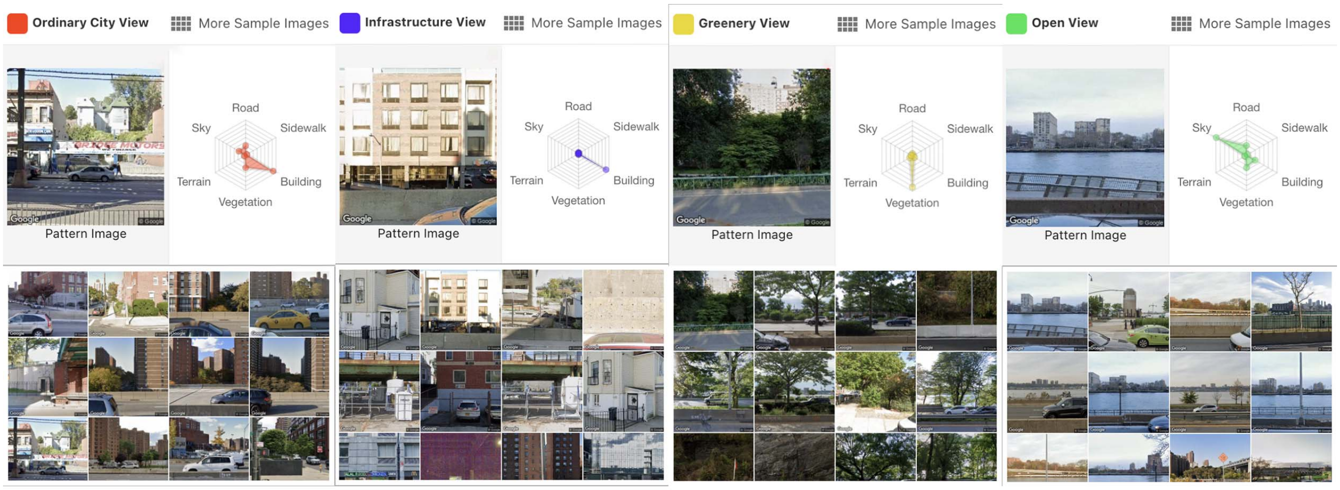
VII. Case Studies

In this section, we use ViVaRoutes in two different geographical areas to show the usage scenarios.

A. Understanding VaPatterns in New York City

VaPatterns provide abstracted city view patterns for quick visual exploration. As shown in Figs. 1(c) and 7, four VaPatterns are discovered from street-view images in Manhattan, New York. These VaPatterns reflect very unique VaPatterns of Manhattan. Users can edit and give the name of each pattern for easy description. Here, in Fig. 7, four patterns are shown side-by-side with the enlarged pattern images, categorical distribution charts, and groups of example representative images. In particular, the four patterns are as follows.

1) Ordinary City Pattern (VaPattern 1): This pattern represents an ordinary street appearance with mixed buildings, sidewalks, bushes, and meadows. It can be seen from Fig. 7(a) that the typical semantic distribution of this pattern: high-level of building with medium vegetation, and some road and sky. Fig. 7(a) shows the pattern image indicating this pattern's visual appearance, also with a set of samples as follows.
2) Infrastructure Pattern (VaPattern 2): This pattern includes the street-views occupied mostly by buildings, bridges, and other transportation infrastructures. It can be seen from Fig. 7(b) that on average, the dominant semantic category is building identified by the AI tool, while other categories are relatively negligible. From the sample images, it can be seen that these street-view images are almost covered by the infrastructure and lack the presence of the sky.



(a) (b) (c) (d)

Fig. 7. Discovering Manhattan visual appearance patterns with four VaPatterns from (a) to (d). The pattern image and a group of sample images (extracted from GSV API) are shown, together with the radar chart that visualizes the semantic category distribution.

3) Greenery Pattern (VaPattern 3): This pattern reflects the green visual appearance of the city with a large portion of trees, plants, and grasslands. It can be seen from Fig. 7(c) that this pattern has a high percentage of vegetation in the chart, while the greenery images present the pattern to users.
4) Open View Pattern (VaPattern 4): This pattern is about the city views with open sky or space, such as faraway views over water surfaces, empty fields, and squares. Fig. 7(d) presents the images usually with open an sky. The sample images show rivers with small buildings at a long distance.

These patterns summarize the city view features in Manhattan, users can quickly recognize and perceive the basic visual appearances in this specific region. Therefore, this abstraction approach helps them explore multiple driving routes in map view with easy understanding. The next case presents a usage scenario of these patterns. For a geographical region with different street-view visual appearances, the patterns can be generated in a similar way to promote visual exploration.

B. Studying Driving Routes With VaPatterns in New York City

With the discovered patterns, users can easily explore the city. Next, we describe an example study.

A user (named Alice for easy description) wants to travel from a starting location in Lower Manhattan to a destination in Upper Manhattan. She employs VivaRoutes to study the driving routes with the interest of street-views. After she inputs the source and destination locations, VivaRoutes recommends three routes she can take on the map [Fig. 1(a)]. Alice can click to select either route so that the route is visualized by colors encoding the street-view patterns. As illustrated in Fig. 1(a), two alternative driving routes (routes 1 and 3) travel along the western and eastern sides of Manhattan Island, respectively. They both use high-speed expressways to avoid going through small streets in the mid-town area. It is of great interest for Alice to find: "What are the different views when driving along Hudson River (route 1 on the western side) or East River (route 3 on the eastern side), respectively?" From the visualization, it can be seen obviously that route 1 has more Greenery Patterns shown in yellow color than route 3. On the other hand, route 3 has more infrastructure patterns with purple color.

Moreover, Alice can check the route comparison panel [Fig. 1(e)] to easily distinguish the two routes. It confirms that route 1 has more yellow segments with beautiful views. In this view, Alice can compare the "left" and "right" sides (based on her driving direction) of these routes. In this example, the left-side view of route 1 has lots of yellow street-view patterns from origin to destination. In contrast, on the right-side Alice will meet many infrastructure views (purple) at the beginning, and after a while, beautiful scenery views (yellow) will appear. On the other hand, on route 3, infrastructure views (purple) can be seen on both the left and right sides. By dragging the marker on the routes, the sampled street-view images are shown in Fig. 1(a). Alice can quickly check the views of interesting locations. Here, she can see the park view on route 1 and the structures on route 3.

It shows that VivaRoutes can help Alice quickly form insights about her potential routes so that to make decisions accordingly. For example, she may choose route 1 for the scene of the Hudson River, but she may also want to choose route 3 for New York buildings and bridges.

C. Exploring Driving Routes in a Midwest Suburban Town

We further show the use of VivaRoutes in a Midwest region of Ohio. This suburban region has very different visual appearance features compared to New York City. Therefore, four different VaPatterns are discovered from street-view sample images. As shown in the VaPattern View of Fig. 8, these patterns are named as follows.

1) Open Pattern: This pattern shows open grassland, sky, and road views.

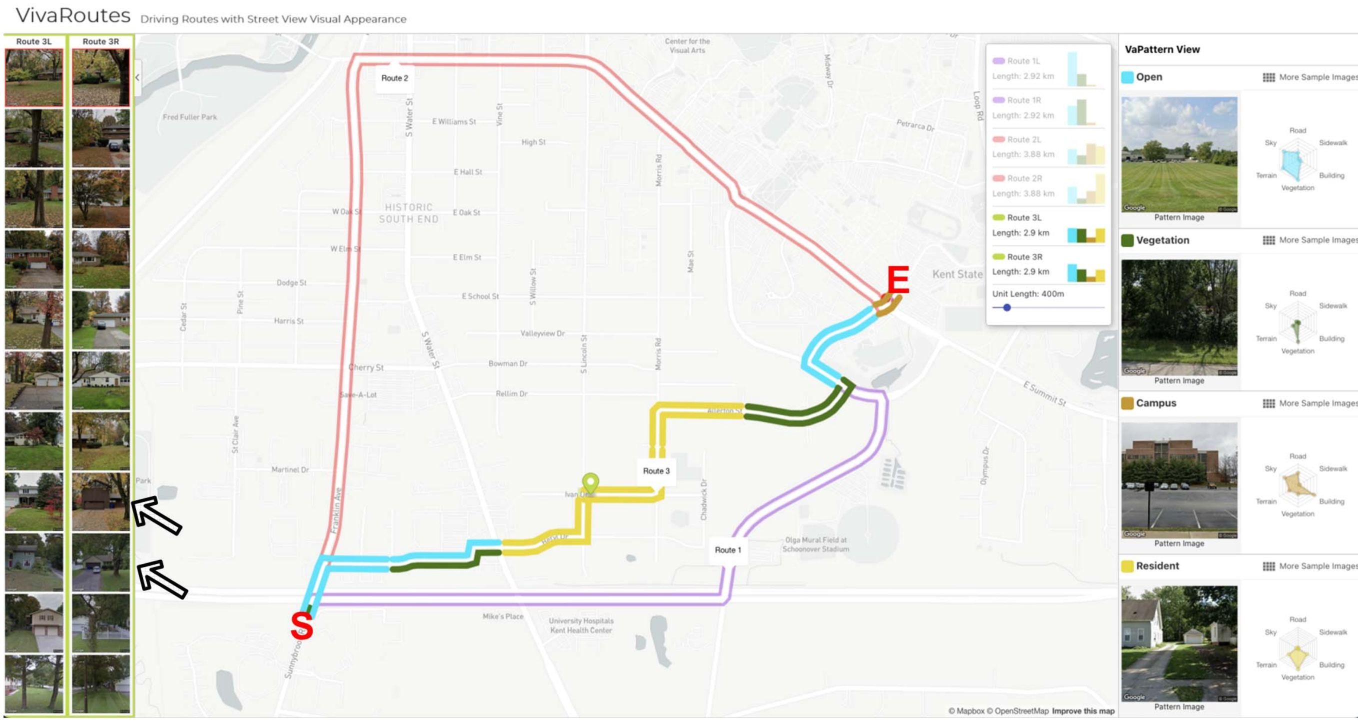


Fig. 8. Studying driving routes in a Midwest university town. The pattern image and a group of sample images (extracted from GSV API) are shown together with the map view.

2) Vegetation Pattern: This pattern reflects views of trees and bushes.
3) Campus Pattern: This pattern mostly includes views from a MidWest university campus in this suburban town.
4) Resident Pattern: This pattern shows mostly the views of resident buildings and neighborhoods.

From them, we can recognize this area as a typical university town with mixed campus buildings, resident houses, trees, and grasses. Next, we show how Alice can use the visualization system for a study of multiple routes.

As shown in Fig. 8, Alice sets the start location (S) from an off-campus resident building to an end location (E) university building on campus. VivaRoutes generates three recommended driving paths from S to E. She can observe the map inlet to quickly discover these routes with left and right side information, respectively. From the bar charts, Alice finds that route 1 has mostly an open view and vegetation view. It does not include a lot of campus views and resident views. In comparison, route 2 has a VaPattern distribution where campus views and resident views are prevalent. Route 3 has a more even distribution of the four patterns. Alice is interested in route 3 as it can give more typical visual views of this university town. In Fig. 8, she selects route 3, and then the VaPatterns of 3L and 3R are highlighted on the map. In this figure, Alice drags the marker along this route, it can be seen that the resident pattern views are highlighted in yellow. On the image view, Alice finds many resident houses along this path. She may drive this route to get more inside information about this town. It can be seen that these images are not sampled in the same season (as indicated by the two arrows), but they are correctly grouped in the resident VaPattern. Similarly, Alice may choose Route 1 where she can find more green views and open views if she does not want to pass through the resident area.

VIII. User Study

We conducted user studies to evaluate our work in two major directions. First, we assessed whether visualizing visual appearance patterns for driving routes is acceptable to people and whether this goal is achieved by our computational and visualization techniques. Second, we evaluated the VivaRoutes system for its usability.

A. User Study of Visualizing VaPatterns

Participants: We invited 45 participants (ages between 19 and 45, 21 males, and 24 females) who are students living in a university town as shown in Fig. 8.

Procedure: We developed an online VivaRoutes Demo System for the participants based on the data shown in Fig. 8. The geographical information and streetview images in the university town were incorporated, and the corresponding visual patterns were discovered and shown to the participants.

In the user study, we scheduled individual interviews lasting 1 to 1.5 hours with each participant. We first discussed our motivation and introduced the work. Then, the participants were guided to explore the demo system.

Finally, they were asked to fill in a form for a group of questions. These questions are shown in Table II with the mean

TABLE II
USER EVALUATION OF VISUALIZING VAPATTERNS

Questions 0 (Poor)–10 (Excellent)	Mean Value	Std. Dev.
Q1: Do you often use map-based navigation services (e.g., Google map, and Apple map) when planning a new route?	9.2	1.2
Q2: Is the new visualization of the streetview information potentially useful for you or others?	7.4	1.5
Q3: How are you familiar with the geo-environment in this example?	7.6	1.9
Q4: Do you think the four patterns can represent the visual appearance of this area?	8.0	1.3
Q5: Do you agree that the pattern distribution shown on the three routes is reasonable, based on your own experience?	7.9	1.4
Q6: Can you quickly get the information from the visualizations?	7.9	1.6
Q7: Does the map view convey useful information?	8.1	1.4
Q8: Does the RouteViewLine convey useful information?	7.4	1.7
Q9: Does the coordinated raw street-view images convey useful information?	8.4	1.4

TABLE III
QUIS EVALUATION OF VIVAROUTES SYSTEM

Questions	Mean Value	Std. Dev.
Reading labels and icons on the screen. 0(very hard)–9(very easy)	8.2	0.8
Selecting and highlighting items/areas. 0(not at all)–9(very much)	7.8	1.0
Organizing information in the interface with positions and layouts. 0(confusing)–9(very clear)	7.3	1.3
Sequential operations on the interface. 0 (confusing)–9 (very clear)	8.1	0.9
Interaction on visual interface 0 (very hard)–9 (very easy)	7.9	1.4
Learning to operate the system. 0 (difficulty)–9 (easy)	8.1	0.8
System response with good speed. 0 (very slow)–9 (fast enough)	8.1	0.8

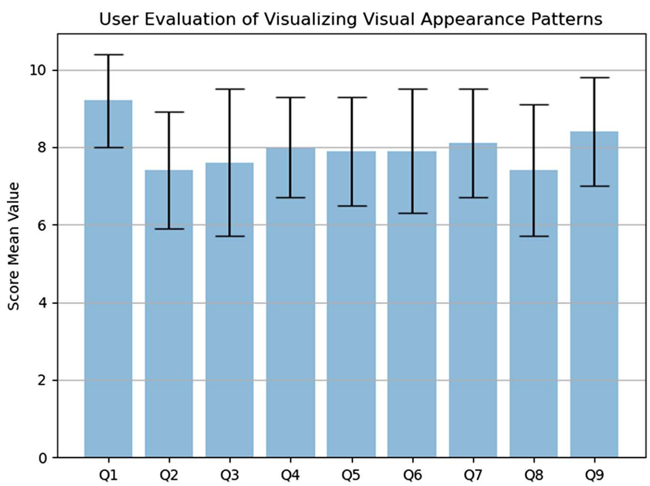


Fig. 9. Bar chart view of user evaluation results in Table II.

values and standard deviations from the participants' scores from 0 to 10. Fig. 9 shows the bar chart view of the results. Next, we analyze the following results.

1) *Visualizing VaPatterns is useful*: As indicated in Table II, the participants were familiar with map-based route services (Q1 with mean 9.2). Then, in Q2 (mean 7.4), they mostly agreed that adding streetview information is potentially useful. We further asked them to list the application areas for the usage. The results indicate: 1) planning trips in a new place as a tourist (87%); 2) route planning in daily life (36%); 3) city planning (38%); 4) exploring maps for fun or knowledge (67%); and 5) community study (27%). It can be seen that people believe the new visual information is useful in multiple directions, while mostly in route planning and exploration. It is worth mentioning that no one replied that the VaPatterns are not useful at all.
2) *The visualization results are reasonable and meaningful*: Next, we asked the participants to evaluate the visual patterns. The answer to Q3 (mean 7.6) shows that they were mostly familiar with the geographical area and landscape environments, since this is the university town where most of them are living. Then, for Q4 (mean 8.0), they thought that the four patterns we discovered by the algorithms (Section V) is meaningful based on their own knowledge of this town. Moreover, they checked the visualized patterns on the three candidates' routes. In Q5 (mean 7.9), they agreed that the patterns shown on the routes are reasonable, which met their experiences on these routes.
3) *The visualization is informative*: In Q6, they evaluated that the visualizations convey the information in a good way (mean 7.9). Furthermore, they gave assessments for three major views. For the map view in Q7, the mean score is 8.1. For the RouteViewLine, the mean score of Q8 is 7.4. They liked the coordinated images in Q9 with a mean of 8.4. The scores indicate that the participants generally find the visualizations informative.

B. QUIS of VivaRoutes System

We further conducted a study to evaluate the interactions in the VivaRoutes system. In this study, we invited 30 participants who are graduate students. Some of them have experience in visualization systems. They conducted this study one by one.

An instructor first introduced the demo system, and then the visualization and interaction functions to each participant. They were guided to freely explore the system for about 15 minutes. Then, they conducted some work on defining routes and visualizing the visual patterns. Afterward, they provided the evaluation by filling out a Questionnaire for User Interaction Satisfaction (QUIS). Table III shows the questions and ratings. Fig. 10 shows the bar chart view of the results. The average score of the visualization and interaction functions in VivaRoutes are very good above 7.0.

C. User Feedback

The participants feel the system is easy to operate and the system response time is very fast. Several participants indicated that the RouteViewLine view may need more time for understanding. This may be attributed to the longer learning curve of this new design, which is different from the map view and image view that most people are familiar with. Some participants also pointed out that the visualizations in the current format may not be easily extended to mobile platforms. Since most people now use mobile phones for route planning,

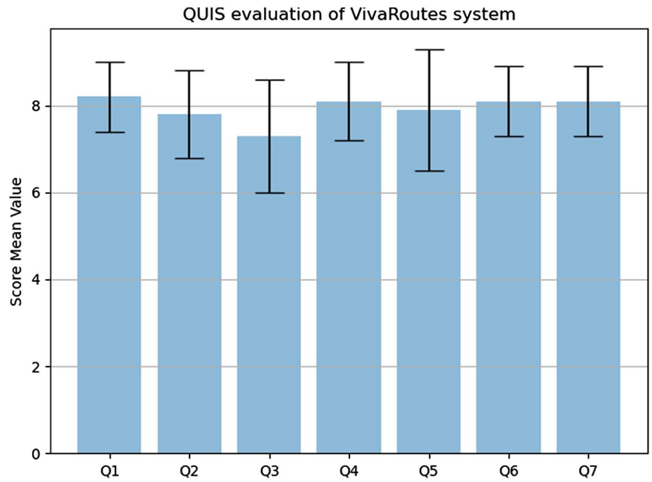


Fig. 10. Bar chart view of QUIS evaluation results in Table III.

the system could be further enhanced for mobile applications. Moreover, it was indicated that when traffic and other information need to be included, cluttering may happen. It could be a challenge since now the route color is used for visual appearance information.

They also provided many suggestions, including reducing the amount of information presented to users, adding more labels and guidance in the system to shorten the learning curve, enlarging the photos for detailed study, hiding and popping up some parts for easy depiction, and integrating with satellite images, among others.

IX. Conclusion and Discussion

In this study, we develop algorithms to discover visual appearance features along driving routes from the street-view imagery dataset. Visual semantic categories are extracted from the image with AI tools, which are also utilized to quantitatively cluster images to find VaPatterns. A VivaRoutes system presents the categories and patterns within an interactive geographical visualization interface, which allows users to explore the discovered features for route planning and decision-making.

The approach identifies and presents street-view features based on the recommended routes from a route planner. These routes may be computed by time, cost, traffic, and other transportation factors. However, visual appearance features are not used in route computation. It is of great interest if routes can be recommended by utilizing these features, for example, finding a route that drives mostly over greenery views. In future work, we will incorporate visual semantic categories and/or VaPatterns in graph-based routing algorithms to further extend this work.

Moreover, this work studies AI-based inductive approaches to discover VaPatterns. However, we realize that it is challenging to theoretically explain and analyze such patterns from diverse visual appearances. We will study more theoretical aspects of representing these street-view pictures. On the other hand, city-level patterns may be an interesting topic for GIS applications. Another future direction is to extend this route-based work to city-level visual appearance extraction and visualization.

Finally, we will also improve VivaRoutes with enhanced visualization and interactions for public deployment.

Acknowledgment

The street view images are extracted from GSV Static API. The map view uses Google Map API. The routes are acquired through Google Directions API. D3.js is used in generating visualizations.

References

[1] N. Andrienko, G. Andrienko, S. Chen, and B. Fisher, "Seeking patterns of visual pattern discovery for knowledge building," in *Proc. Comput. Graph. Forum*, vol. 41. Hoboken, NJ, USA: Wiley, 2022, pp. 124–148.
[2] V. T. N. Nha, S. Djahel, and J. Murphy, "A comparative study of vehicles' routing algorithms for route planning in smart cities," in *Proc. 1st Int. Workshop Veh. Traffic Manage. Smart Cities (VTM)*. Piscataway, NJ, USA: IEEE, 2012, pp. 1–6.
[3] J. Fawcett and P. Robinson, "Adaptive routing for road traffic," *IEEE Comput. Graph. Appl.*, vol. 20, no. 3, pp. 46–53, May/Jun. 2000.
[4] M. Shoaib and W.-C. Song, "Traffic aware optimized zone based hierarchal link state routing protocol for VANET," in *Proc. IEEE 24th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*. Piscataway, NJ, USA: IEEE, 2013, pp. 3117–3122.
[5] B. Malakooti, I. Thomas, S. K. Tanguturi, S. Gajurel, H. Kim, and K. Bhasin, "Multiple criteria network routing with simulation results," in *Proc. IIE Annu. Conf. Proc. Inst. Ind. Syst. Eng. (IISE)*, 2006, p. 1.
[6] M. Musolesi, S. Hailes, and C. Mascolo, "Adaptive routing for intermittently connected mobile ad hoc networks," in *Proc. 6th IEEE Int. Symp. World Wireless Mobile Multimedia Networks*. Piscataway, NJ, USA: IEEE, 2005, pp. 183–189.
[7] H. Samet, J. Sankaranarayanan, and H. Alborzi, "Scalable network distance browsing in spatial databases," in *Proc. ACM SIGMOD Int. Conf. Manage. Data*, 2008, pp. 43–54.
[8] J. Sankaranarayanan and H. Samet, "Query processing using distance oracles for spatial networks," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 8, pp. 1158–1175, Aug. 2010.
[9] B. Yang, C. Guo, Y. Ma, and C. S. Jensen, "Toward personalized, context-aware routing," *VLDB J.*, vol. 24, no. 2, pp. 297–318, 2015.
[10] C. Guo, B. Yang, J. Hu, C. S. Jensen, and L. Chen, "Context-aware, preference-based vehicle routing," *VLDB J.*, vol. 29, no. 5, pp. 1–22, 2020.
[11] C. Chen, D. Zhang, B. Guo, X. Ma, G. Pan, and Z. Wu, "TripPlanner: Personalized trip planning leveraging heterogeneous crowdsourced digital footprints," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 3, pp. 1259–1273, Jun. 2015.
[12] S. Mirri, C. Prandi, P. Salomoni, F. Callegati, and A. Campi, "On combining crowdsourcing, sensing and open data for an accessible smart city," in *Proc. 8th Int. Conf. Next Gener. Mobile Apps, Serv. Technol*. Piscataway, NJ, USA: IEEE, 2014, pp. 294–299.
[13] C. Prandi, P. Salomoni, and S. Mirri, "mPASS: Integrating people sensing and crowdsourcing to map urban accessibility," in *Proc. IEEE 11th Consum. Commun. Netw. Conf. (CCNC)*. Piscataway, NJ, USA: IEEE, 2014, pp. 591–595.
[14] Q. Shen et al., "StreetVizor: Visual exploration of human-scale urban forms based on street views," *IEEE Trans. Visual. Comput. Graph.*, vol. 24, no. 1, pp. 1004–1013, Jan. 2018.
[15] Y. A. Yatmo, "Street vendors as 'out of place' urban elements," *J. Urban Des.*, vol. 13, no. 3, pp. 387–402, 2008.
[16] W. Zhai and Z.-R. Peng, "Damage assessment using Google Street View: Evidence from Hurricane Michael in Mexico Beach, Florida," *Appl. Geogr.*, vol. 123, 2020, Art. no. 102252.
[17] Q. C. Nguyen et al., "Using Google Street View to examine associations between built environment characteristics and US health outcomes," *Preventive Med. Rep.*, vol. 14, 2019, Art. no. 100859.
[18] A. Berland and D. A. Lange, "Google Street View shows promise for virtual street tree surveys," *Urban For. Urban Greening*, vol. 21, pp. 11–15, 2017.

[19] M. Gerell and K. Kronkvist, "Violent crime, collective efficacy and city-centre effects in Malmö," *Brit. J. Criminol.*, vol. 57, no. 5, pp. 1185–1207, 2017.

[20] C. Vandeviver, "Applying Google Maps and Google Street View in criminological research," *Crime Sci.*, vol. 3, no. 1, pp. 1–16, 2014.

[21] P. P. Olea and P. Mateo-Tomás, "Assessing species habitat using Google Street View: A case study of cliff-nesting vultures," *PLoS One*, vol. 8, no. 1, 2013, Art. no. e54582.

[22] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder–decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.

[23] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 2881–2890.

[24] H. Zhou, S. He, Y. Cai, M. Wang, and S. Su, "Social inequalities in neighborhood visual walkability: Using street view imagery and deep learning technologies to facilitate healthy city planning," *Sustain. Cities Soc.*, vol. 50, 2019, Art. no. 101605.

[25] H. Ning, X. Ye, Z. Chen, T. Liu, and T. Cao, "Sidewalk extraction using aerial and street view images," *Environ. Plann. B: Urban Analytics City Sci.*, vol. 49, no. 1, pp. 7–22, 2022.

[26] N. Andrienko and G. Andrienko, "Visual analytics of movement: An overview of methods, tools and procedures," *Inf. Visualization*, vol. 12, no. 1, pp. 3–24, 2013.

[27] G. Andrienko, N. Andrienko, W. Chen, R. Maciejewski, and Y. Zhao, "Visual analytics of mobility and transportation: State of the art and further research directions," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 8, pp. 2232–2249, Aug. 2017.

[28] T. Von Landesberger, S. Bremm, N. Andrienko, G. Andrienko, and M. Tekusová, "Visual analytics methods for categoric spatio-temporal data," in *Proc. IEEE Conf. Vis. Analytics Sci. Technol. (VAST)*. Piscataway, NJ, USA: IEEE, 2012, pp. 183–192.

[29] X. Li, C. Zhang, and W. Li, "Building block level urban land-use information retrieval based on Google Street View images," *GIScience Remote Sens.*, vol. 54, no. 6, pp. 819–835, 2017.

[30] W. Zeng and Y. Ye, "VitalVizor: A visual analytics system for studying urban vitality," *IEEE Comput. Graph. Appl.*, vol. 38, no. 5, pp. 38–53, Sep./Oct. 2018.

[31] S. Jamonnak et al., "GeoVisuals: A visual analytics approach to leverage the potential of spatial videos and associated geonarratives," *Int. J. Geogr. Inf. Sci.*, vol. 34, no. 11, pp. 2115–2135, 2020.

[32] J. Bennett, *OpenStreetMap*. Birmingham, UK: Packt Publishing Ltd, 2010.

[33] M. Amiruzzaman, A. Curtis, Y. Zhao, S. Jamonnak, and X. Ye, "Classifying crime places by neighborhood visual appearance and police geonarratives: A machine learning approach," *J. Comput. Soc. Sci.*, vol. 4, pp. 1–25, 2021, doi: 10.1007/s42001-021-00107-x.

[34] M. Amiruzzaman, Y. Zhao, S. Amiruzzaman, A. C. Karpinski, and T. H. Wu, "An AI-based framework for studying visual diversity of urban neighborhoods and its relationship with socio-demographic variables," *J. Comput. Soc. Sci.*, vol. 6, no. 1, pp. 315–337, 2023.

[35] A. J. Ferreira and M. A. Figueiredo, "Efficient feature selection filters for high-dimensional data," *Pattern Recognit. Lett.*, vol. 33, no. 13, pp. 1794–1804, 2012.

[36] S. Roy, S. Mondal, A. Ekbal, and M. S. Desarkar, "Dispersion ratio based decision tree model for classification," *Expert Syst. Appl.*, vol. 116, pp. 1–9, 2019.

[37] L. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," 2017. [Online]. Available: http://arxiv.org/abs/1706.05587

[38] D.-T. Dinh, T. Fujinami, and V.-N. Huynh, "Estimating the optimal number of clusters in categorical data clustering by silhouette coefficient," in *Proc. Int. Symp. Knowl. Syst. Sci*. Singapore: Springer-Verlag, 2019, pp. 1–17.

[39] A. P. Reynolds, G. Richards, B. de la Iglesia, and V. J. Rayward-Smith, "Clustering rules: A comparison of partitioning and hierarchical clustering algorithms," *J. Math. Modell. Algorithms*, vol. 5, no. 4, pp. 475–504, 2006.

[40] Y. Albo, J. Lanir, P. Bak, and S. Rafaeli, "Off the radar: Comparative evaluation of radial visualization solutions for composite indicators," *IEEE Trans. Visual. Comput. Graph.*, vol. 22, no. 1, pp. 569–578, Jan. 2016.