Automated Planning of Leisure Walks Based on Crowd-sourced Photographic Content

Alexander Kachkaev

PhD research student

giCentre, City University London

Jo Wood

Professor of Visual Analytics

giCentre, City University London

Abstract

All walking trips can be classified into two main groups: functional walks and leisure (or recreational) walks. While the goal of functional walks is moving from one point in space to another, the purpose of leisure walks is the process of walking itself. Unlike functional walking, recreational walking implies a more complex combination of factors that form the selection of a particular route in the mind of a pedestrian, and many of these factors are having a psychological nature being related to human perception of space. One of the most hard-to-formalize factors that a person can be considering when planning a leisure walk is the attractiveness of areas that appear on the way. Conventional map data that are informing existing routing algorithms cannot be used for extracting such measure as attractiveness of streets. Indeed, even a very rich description of all road segments including their type, surface, slope, accessibility, etc. does not contain a subjective component, or in other words, does not tell whether or not the pedestrians enjoy their presence at a particular place. In order to resolve this issue external information sources should be used. This project is focusing on data from 4 photo-sharing services (Flickr, Panoramio, Picasa and Gerograph) and is examining how they can be used for road segments weighting in Central London area.

The idea of using the density of geotagged photographs as a measure of attractiveness of urban streets is based on the peculiarity of the process of photography sharing. In order for an image to appear on a photo-sharing website it must be taken and then uploaded by a user. Both of these actions are voluntary and due to the human psychology often happen when a person finds something interesting that is worth showing to others. When such behaviour is repeated among hundreds of people, this results patterns in distributions of photographs that can be potentially turned into a measure of attractiveness of different places and streets in cities.

Following the discussion of the idea at the last year's UTSG conference, this paper presents the results of the PhD research and covers a number of findings and conclusions.

The first part of the paper is devoted to data analysis and filtering. Because the photographic datasets are not originally collected for the purpose of measuring street attractiveness and thus contain bias, they need to be studied and cleaned in order to increase their reliability and suitability for the chosen purpose. The photographs with different content do not contribute to the measure of street attractiveness equally and the challenge is to classify them into ones that should inform the routing algorithm and those that must be excluded. Because the datasets that this project is working with contain hundreds of thousands of entries and the automated image content classification is unfeasible, the classification is done with the help of an online survey. 900 randomly picked images were shown to a group volunteered participants, who were asked to classify each photograph by a set of criteria: whether an image is a real photograph, is taken outdoors, is taken during daytime, is containing human faces, is featuring something permanent, is made by a pedestrian and is suggesting a nice place for a walk. With 8,434 subjective responses from 359 users (at least

8 subjective responses per photograph), it was possible to suggest filtering methods based on metadata of the photographs as well as their content. The following approaches are discussed in the paper: filtering based on EXIF data, presence of faces in the photographs (involving automated face detection), photo timestamp, tags, title and description, amount of green in the photographs. A combination of successful filtering techniques together with spatiotemporal filtering discussed in the last year's UTSG paper allows reducing bias in the photographic datasets and makes them more suitable for estimating street attractiveness.

The second part of the paper describes the routing algorithm itself. Based on the filtered versions of the photographic datasets and road network data from OpenStreetMap, a methodology for weighting road segments has been proposed. We discuss the work of the algorithm and possible ways if its improvement.

1 Introduction

Routing services are able to provide travel directions for users of all modes of transport. Most of them are focusing on functional journeys (i.e. journeys linking given origin and destination with minimum cost) while paying less attention to recreational trips, in particular leisure walks in an urban context. These walks are additionally predefined by time or distance and as their purpose is process of walking itself, the attractiveness of areas that are passed by can be an important factor in route selection. This factor is hard to be formalised and requires a reliable source of information, covering the entire street network. The data behind maps do not contain any subjective information on street attractiveness, there are only classes of roads, types of land use, etc. These attributes can be used (and actually are widely used) for giving preference to particular streets, however they don't reflect human opinion on the places, and apparently this information is not enough for choosing attractive routes.

We see the solution to the problem in user-generated content, and in particular, shared geotagged images. It is known that some areas are found more popular among photographers than others (Kisilevich et al., 2010), which means that there is something in those places that makes people to (a) take a photograph, (b) upload this photograph on the website for others to see. This peculiarity of such dataset suggest that the spatial density of the photographs can be used for filling the gap in the problem of finding attractive areas and inform an automated routing system about what paths are more attractive than others. So the goal of our research is to look at how we can utilize shared geotagged images for building a routing system that suggests attractive leisure walks. We are aware of some related projects (De Choudhury et al., 2010; Kurashima et al., 2012; Lu et al., 2010; Okuyama, K. and Yanai, K., 2010) and focus on the detailed analysis of the photographic content as the information source for a routing system. We also look into the way of converting the density of photographs into road network scores and suggesting the routes with given constraints.

This paper provides an overview of the project and continues the discussion of the research at the last year's UTSG conference (Kachkaev and Wood, 2013).

2 Geotagged Photographic Content

We selected 4 popular photo-sharing websites that contain large volumes of geotagged images contributed by thousands of volunteers; these are Flickr[1], Panoramio[2], Picasa[3] and Geograph[4]. We chose Central London to be an area for out experiments and downloaded the metadata for all available images in this region using the APIs provided by the services.

Assuming that a geotagged image is a 'vote' for the attractiveness and walkability of the streets nearby, an ideal dataset would only have the photographs that are:

- reasonably accurately georeferenced,

[1] http://www.flickr.com/
[2] http://www.panoramio.com/
[3] http://picasaweb.google.com/
[4] http://www.geograph.org.uk/

- taken outdoors,
- taken during the day,
- not taken during an event,
- taken by a pedestrian,
- not having human faces (except passers'-by), or other moving objects as the main subject.

In order to estimate which of the chosen datasets are suitable for informing a walk planning algorithm, it is necessary to perform their analysis and to find the proportions of photographs that are not matching at lest one of the requirements. If the amounts of photographs not contributing to the attractiveness of the nearby streets are significant, a candidate dataset must be rejected or a corresponding filtering method must be proposed. We split the analysis in two main stages: spatial density evaluation for all items in each dataset (to check the first requirement) and content-based analysis of dataset samples (for all other requirements).

2.1 Spatial Density Analysis and Filtering

To detect potential anomalies in the spatial distribution of the images we chose visual analytics as the main approach. The analysis was done using a software tool that was written for this particular purpose. With use of visual analytics we detected hotspots of different natures that were negatively influencing on the distributions of photographs and also evaluated the proposed filtering methods of incorrectly placed images. We had to reject Picasa dataset as such due to problems in the design of their API that we could not overcome. More details on this stage can be found in our last year's UTSG paper.

2.2 Photo Content Survey

Although it is possible to use the densities of images by converting them into scores of road segments right after performing the spatial filtering, we found it important to do the analysis of the content of the photographs as well. Because not every correctly placed geotagged image may be a sign of the attractiveness of surrounding space, there might be patterns that make some places overrepresented with unwanted photographs of a certain kind, and this may distort the result of a route planner. For example, the dataset may contain be portraits, photographs taken indoors, overnight, during the events, etc.

In order to find the proportions of photographs with unwanted content we ran an online survey asking volunteers to classify a sample of 900 photographs, 300 for each chosen dataset except Picasa. The survey is still available at http://photoassessment.org/; it consists of the following questions about each photograph:

1. Is the image a real photograph?
2. Is it a photograph of something outdoors?
3. At what time of the day is the photograph taken?
4. Is it a photograph of something temporary?
5. Are people the main subject of the photograph?
6. Could the photograph be taken by a pedestrian?
7. Does the photograph suggest this is a nice place to walk?

The participants could reply *yes / no / hard to say* to all questions except the one about the time of day, where the options were *day / twilight / night / hard to say.* We collected 8,434 subjective classifications from 359 participants, giving minimum 8 responses per photograph. Figure 1 shows the aggregated summary of the numbers of different types of photographs.

Panoramio was found to be the dataset with the highest proportion of relevant images (those that depict attractive walkable areas). Nearly a half of Flickr images are taken indoors, and about a half of outdoor photographs are classified as taken during events, which makes the 'street attractiveness rate' in this dataset very low (21%). Although the most of the photographs in the Geograph sample are passing 6 formal criteria of a suitable photograph, only a half of them are depicting attractive places. This peculiarity can be explained by the nature of this photo-sharing website: the contributors are encouraged to evenly fill the geographical grid rather than to upload photographs entirely based on what places they like to be at.

The results of the survey suggest not to recommend Geograph for the use in a routing algorithm and to try further filtering of Flickr and Panoramio datasets in order to reduce the amounts of unrelated images.

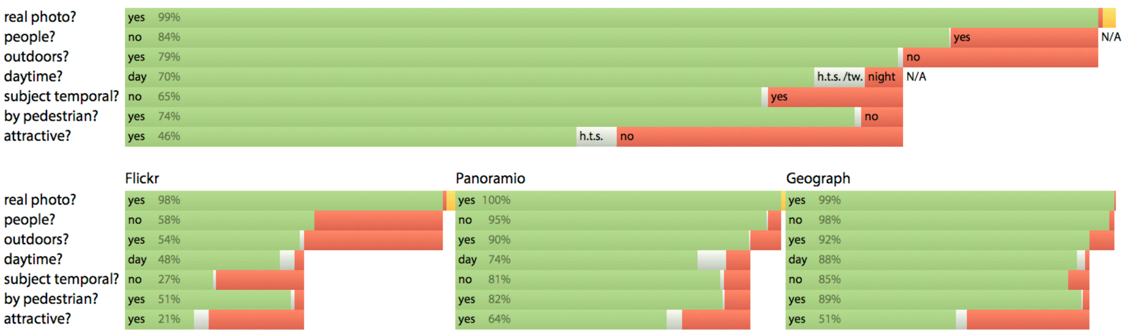

Figure 1: Aggregated results of the survey (mode answers). *Green:* suitable photographs in each classification, *grey:* hard to say, *red:* unsuitable photographs, *yellow:* photographs that were deleted at their origin since the launch of the survey (they do not influence the base for percentages).

2.3 Advanced Content Filtering

Having access to the content of the photographs and some additional metadata attached to the images, we can apply a number of filtering techniques. In this section we describe the methods we used and considered.

Luminance

The most successful filtering method we discovered is based on luminance, that can be extracted from EXIF data attached to the photographs (Jacobson, 2000). Before taking a picture, a camera sets up ISO speed, aperture and explosion according to the amount of light at the scene. Having this information we can restore the value of luminance and compare it against the results of the survey. Assuming that a value of luminance below a certain threshold puts most indoor and night photographs into a separate group, we can find the value of luminance that performs their separation with the smallest number of mistakes. Figure 2 demonstrates that filtering works best when the luminance threshold is set between 3 and 4. We choose 4 to make the filter stricter; an example of image separation by luminance is shown in Figure 3.

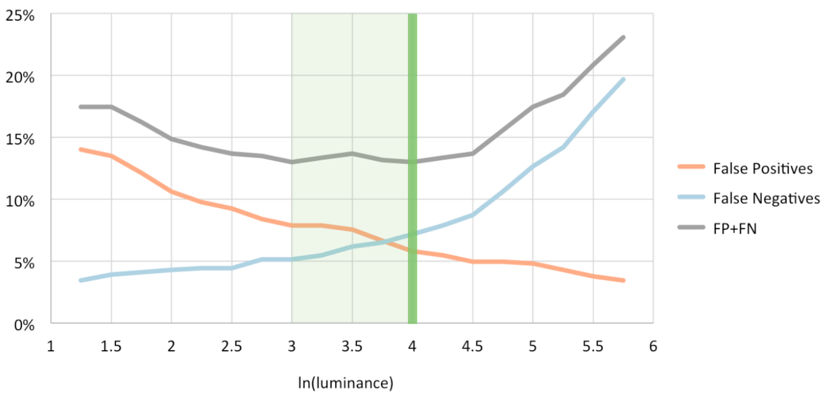

Figure 2: Selection of the luminance filtering threshold based on the results of the survey (Positive = a photograph is classified as outdoors and taken during the day by both survey participants and the luminance filter).

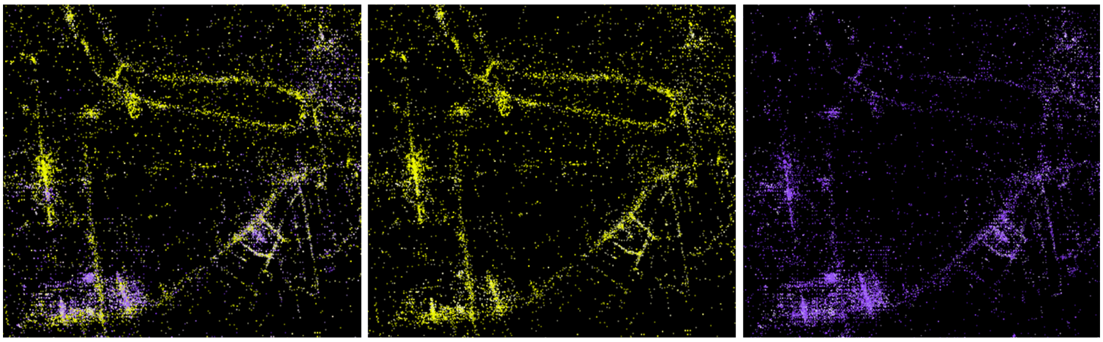

Figure 3: *Left:* Patterns in spatial distribution of luminance in Flickr images; *middle* and *right*: result of segregating Flickr images with luminance ≥ 4 and < 4, respectively.

Event Exclusion based on photo timestamp and spatial clustering

In order to reduce the proportion of photographs taken during events we suggest applying spatial clustering of image locations to find and suppress the peaks of daily activity. We use Voronoi algorithm for this purpose (Andrienko et al., 2010). In each cluster we find up to 30 days with more active users than a threshold of 10 users per cluster a day and reject the photographs from a subset of these users (so that the number of remaining users was equal to the value of the threshold). Although the photographs that are left are still depicting the events, their proportion in the dataset reduces significantly. Examples of event-based filtering is given in Figure 4.

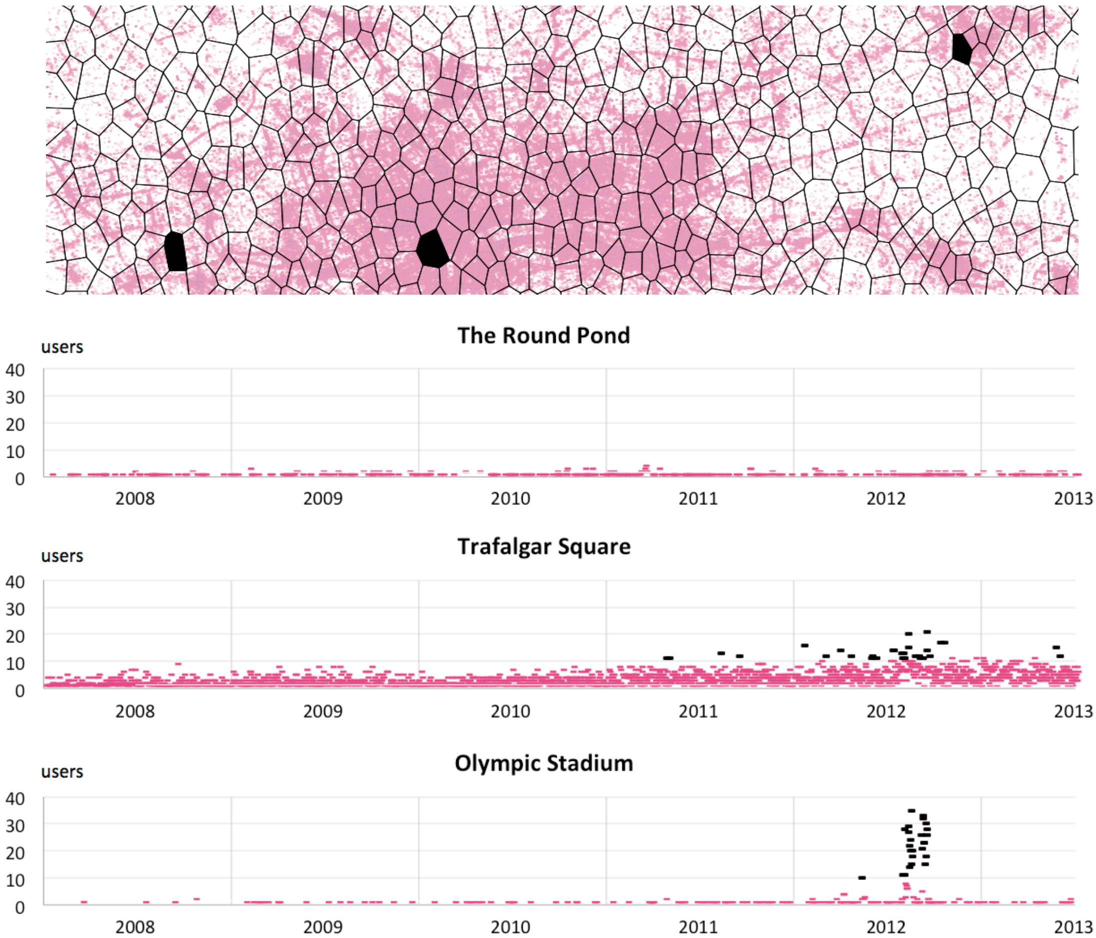

Figure 4: Voronoi spatial clustering for Flickr dataset and 3 examples of event filtering

Exclusion of portraits using face detection
We use automatic face detection to reduce the proportion of photographs with human faces in Flickr dataset and try 2 commonly used libraries: OpenCV[1] and CoreImage[2]. The tests show that the best match between the input from the survey participants and the results of face detection is reached by CoreImage. We use this library to detect faces in about 200,000 images that are not excluded by any other filtering methods.

Filtering based on other parameters of photographs
We looked towards the use other image attributes for filtering of unrelated photographs, such as time extracted from the timestamp, the amount of green in the photographs, tags, title and description, However, the tests showed no or very little correlation between the values of these parameters and the likelihood of a positive answer to the last question in the survey. For example, we were expecting a timestamp to be a better option for filtering night photography compared to use of luminance, but discovered that in many cases this attribute does not reflect the real time of the day. We suppose that this may be because of a high proportion of tourists among the photographers who have different time zones in their camera settings.

By using a combination of the techniques described above together with the spatial filtering explained in the last year's UTSG paper we could significantly refine the distribution of locations of photographs, making Panoramio and Flickr datasets more suitable for the purpose of distinguishing between attractive and non-attractive walkways.

3 Routing

After the source photographic datasets are assessed and filtered, we need a routing system that will work with these data. It should consist of the following components:

1) road network data obtainer and processor,
2) a module for converting spatial densities of photographs into road network scores,
3) a routing algorithm that takes user-defined origin and destination, the calculated scores for the roads and their lengths and finds a path that meets the passed time constraint.

3.1 Road Network Data

In our experiments we chose to use OpenStreetMap as a source for the road network for its free licence[3] and promising accuracy (Haklay, 2010). We obtained the data using OSM XAPI[4] and then convert it into a topology graph with OSM2PO[5]. At this stage we excluded all private roads and segments where pedestrian access is not allowed (e.g. streets with pavements marked as separate lines).

We recommend to check the topology against isolates and to remove them from the graph. Isolates are signs of probable errors in the data, because in real life every public road is linked to the rest of the network. Not removing the isolates will cause an exception in any path finding algorithm when asked to build a route between unlinked nodes.

As of November 2013, the routing graph for pedestrians in Central London that includes publicly accessible walkable roads consisted of 89,032 edges and contained 497 isolated clusters (Figure 5).

[1] http://opencv.org/
[2] https://developer.apple.com/library/mac/documentation/GraphicsImaging/Conceptual/CoreImaging/
[3] http://www.openstreetmap.org/copyright
[4] http://wiki.openstreetmap.org/wiki/Xapi
[5] http://osm2po.de/

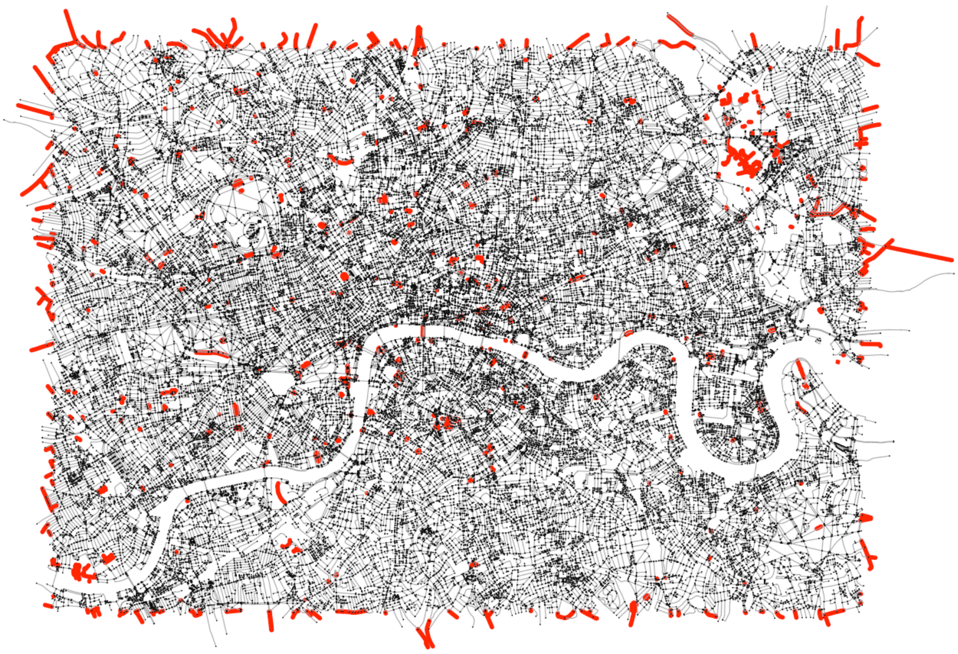

Figure 5: Road network topology for Central London (*NESW: 51.56, 0.02, 51.46, -0.21*). Excluded isolates and edges outside the bounding box are highlighted. *Map data © ODbL OpenStreetMap contributors.*

3.2 Calculation of Network Scores

To inform a walk planner about the attractiveness of streets, all edge segments should be assigned a numeric value representing the density of surrounding photographs if they match re requirements listed in section 2. The simplest thing that can be done is counting photographs located within a given number of meters from each edge and making this a score (Figure 6a). However, in such case some edges get higher scores simply for being longer. In order to avoid this problem we suggest normalizing the scores dividing them by the edge distance (Figure 6b). Another thing we try is adding less to score if a photograph is close to several edges (Figure 6c). Finally, instead of using a window with a fixed size, we suggest assigning the score proportionally to the distance of a photograph to an edge within a window. (Figure 6 d). We are still in the process of comparing these methods and haven't concluded about which is the best yet.

(a) within-window count

(b) edge length dependent within-window count

(c) score effect fission for items located near multiple edges

(d) window blurring

Figure 6: Different methods of score-to-edge assignment based on locations of nearby photographs.

The process of score assignment is resource-intensive as it involves dealing with large datasets and running millions of geography-related function to calculate distances between edges and photographs. We suggest splitting this task into two stages in order to improve the performance. First, for each item in every photographic dataset we find its neighbouring edges, storing the array with pairs of edge ids and distances in a temporary table named ‘photo windows’. We set the size of the window to be the maximum of what may be required later. Then we take photo windows and transpose them into edge windows, but with respect to the window size of the current experiment and photo filtering preferences. Now we can quickly convert a set of neighbouring photographs for every edge into a score depending on what assignment method is chosen. Such approach reduces the numbers of calls of spatial functions and allows the calculated distances between edges and nodes to be recycled.

We find it efficient not to process the entire routing graph at once and split the area into quads (Samet, 1984), working on multiple parts of it in parallel. Doing so makes the approach scalable and less RAM- and CPU-consuming. When only a portion of geotagged photographs is loaded into memory to find out neighbouring edges, it is necessary to extend the bounding box in order to fetch edges that are outside of it, but are still located within the defined maximum distance to some photographs. The same is also important when querying precalculated photo windows to convert them into scores of edges in a given region. In this step the extended region must have a complex shape and include areas around all roads that are partially outside the current bounding box. The process is illustrated in Figure 7 and consists of 4 steps:

1) finding the edges that are partially outside of the bounding box and making a small spatial buffer around them;
2) combining this buffer with the original bounding box;
3) making a buffer around the obtained area;
4) simplifying the result using Douglas-Peucker algorithm (Douglas and Peucker, 1973) with tolerance of not more than 1 meter (the reduction of the complexity of the polygon increases the performance of spatial queries that fetch photo windows).

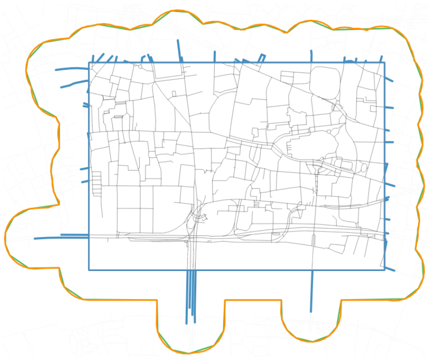

Figure 7: Extension of a bounding box for a given quad when calculating edge scores.

The smaller the quads the less edges and photo windows are loaded into memory, however the more is the proportion of edges, for which the scores are calculated more than once (they are being considered when processing every quad that contains at least one of their nodes). Thus, working with the areas that are too small will not improve the performance.

We choose to have quads not greater than 2 square kilometers, which results 256 of them in the Central London. The process of generation of photo windows for Flickr, the largest dataset we work with, does not take longer than 20 minutes on a personal computer with average characteristics when running 10 threads in parallel. The performance of score calculation depends on the chosen filtering mode, but does not take longer, because no distances between objects need to be calculated during this process.

3.3 Routing Algorithm

Most of the routing systems are based on one of the algorithms that solve the *single-source shortest path problem*, for instance Dijkstra's algorithm, Bellman–Ford algorithm, Floyd's algorithm, etc. (McHugh, 1990). In general case a *shortest path* is a route with minimum cost of moving from the origin to destination; it consists of individual costs of network edges and costs of interchanges. Unlike cycling or driving, where some streets are one way only, there are turn restrictions and speed limits, walking can be done in any direction at a preferred pace with no additional costs for making turns. This simplifies the problem of searching for the *shortest path* bringing it down to finding a route with the smallest sum of edge lengths.

Edge scores that contain the proportions of surrounding densities of geotagged photographs can be used to influence the original costs of the road segments. If a cost of an edge is reduced proportionally to the value of the score, the *shortest path* becomes longer in distance, but 'cheaper' in terms of the total cost, which is achieved by choosing the edges with higher scores. Thus, the preference will be given to roads with higher counts of surrounding photographs, or, in other words, more attractive streets.

The idea can be formalised as following. Let a path *P* be a set of *k* nodes (vertices) *v* or *k-1* edges *e*:

$$P = v_1 \to v_2 \to ... \to v_k, P = (e_1, e_2, ...e_{k-1})$$

then the cost of the a is:

$$\omega(P) = \sum_{i=1}^{k-1} \omega(e_k)$$

where $\omega(e_k)$ is the cost of edge *k*. The cost of an edge depends of its length and the relative normalised score value:

$$\omega(e_k) = l(e_k)(1 - M \times min(1, c\frac{s_k}{S}))$$

where $l(e_k)$ is the length of edge *k*, s_k is the score of edge *k*, *S* is the average score among all edges in the network, *c* is the coefficient of score influence and *M* is a number between 0 and 1 that sets the maximum influence of a weight by a score. Thus, when *c* is equal to 0, the weights of all network edges are equal to their lengths, and the *shortest path* algorithm suggests a route with the smallest distance. As *c* increases, the weights of higher number of road segments become reduced, which makes the *shortest path* algorithm preferring edges with higher values of scores. *M* is introduced to avoid obtaining edges with negative weights, which may lead to infinite loops and freeze the algorithm.

As the goal of our routing system is to find a path that suits a time constraint, we calculate the lengths for each path found by the *shortest path* algorithm with different values of *c* and stop when the time required to complete a suggested walk at a given pace is within 5% of a given constraint. An example of the process is given in Figure 8.

The suggested approach has high performance and may generate an attractive walkable route between given points in a few seconds; however, it has some limitations and can be further improved. The described method does not work for constructing circular routes and also fails to suggest routes when the minimum walking time between the origin and the destination is too small compared to the time specified. We are currently working on resolving this issue and are looking into introducing dynamic weight amplification for edges that are a part of a last calculated *shortest path* and have the smallest value of a score. This will lead to small diversions rather than complete changes of a route and will both allow a more precise adjustment of the resulting time and a more frequent exclusion of the least attractive segments.

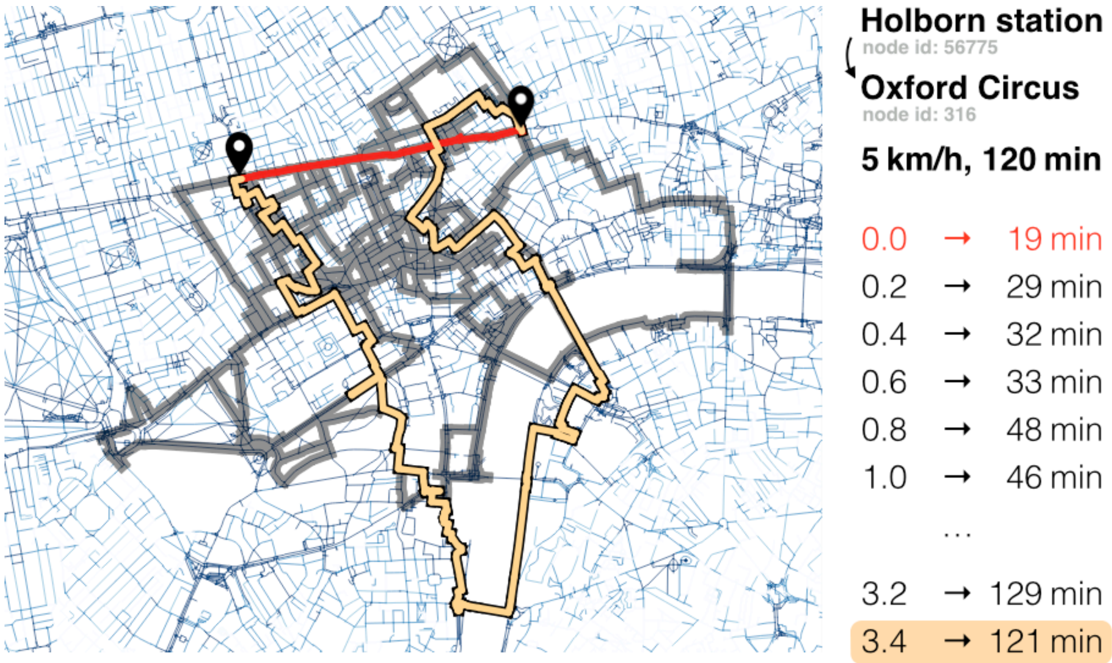

Figure 8: Example of a leisure walk generated by the algorithm between Holborn Station and Oxford Circus with a given time of 120 minutes and pace of 5 km/h, based on locations of Panoramio images, edge length dependent scores and window size of 20 m. *Map data © ODbL OpenStreetMap contributors.*

Having the scores representing street attractiveness according to the opinion of the photographers, a completely different approach to routing can be applied. For example the problem of finding the most attractive path can solved with the use of an ant-colony algorithm (Jain et al., 2010) or a genetic algorithm (Pahlavani et al., 2006). These methods are more computation-intensive, but hay help to overcome the issues that we currently have.

4 Conclusions and Future Work

The studies of photographic content show that spatial densities of shared geotagged images may be utilised for generating routes for leisure walking. The results of the survey we have conducted show that the most suitable photo sharing service for this purpose is Panoramio, where the proportion of images depicting attractive places is the highest. The methodology we propose for filtering unrelated images increases this rate and makes it possible to use Flickr (the largest dataset among those we work with) applicable for this purpose too.

As the result of the experiments with score-to-edge assignment we propose an efficient two-step method of converting spatial density of photographs into measures of attractiveness of road segments in a routing graph. The tests of the approach we choose for path finding demonstrate its applicability, and also help to discover problems, the solutions of which we are currently looking for. The framework we have built makes it possible to set up the photo-based routing system for any city in the world in only a few hours and to integrate it with other software.

We believe that automated planning of leisure walks where attractive streets get preference can be found useful by many people, especially among tourists who are not very familiar with the region of their interest. Increased informativeness and reliability of a journey planner for pedestrians may make leisure walking more inviting and enjoyable.

References

Andrienko, G., Andrienko, N., Mladenov, M., Mock, M., Pölitz, C., 2010. Discovering bits of place histories from people's activity traces, in: IEEE Symposium on Visual Analytics Science and Technology. pp. 59–66.

De Choudhury, M., Feldman, M., Amer-Yahia, S., Golbandi, N., Lempel, R., Yu, C., 2010. Constructing travel itineraries from tagged geo-temporal breadcrumbs, in: Proceedings of the 19th International Conference on World Wide Web. pp. 1083–1084.

Douglas, D.H., Peucker, T.K., 1973. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. Cartographica: The International Journal for Geographic Information and Geovisualization 10, 112–122.

Haklay, M., 2010. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. Environment and planning. B, Planning & design 37, 682.

Jacobson, R.E., 2000. The Manual of Photography: Photographic and Digital Imaging, Media Manual. Focal Press.

Jain, S., Seufert, S., Bedathur, S., 2010. Antourage: mining distance-constrained trips from flickr. ACM Press, p. 1121.

Kachkaev, A., Wood, J., 2013. Crowd-sourced Photographic Content for Urban Recreational Route Planning. Presented at the UTSG, Oxford.

Kisilevich, S., Keim, D., Andrienko, N., Andrienko, G., 2010. Towards acquisition of semantics of places and events by multi-perspective analysis of geotagged photo collections. GeoCart.

Kurashima, T., Iwata, T., Irie, G., Fujimura, K., 2012. Travel route recommendation using geotagged photos. Knowledge and Information Systems.

Lu, X., Wang, C., Yang, J.-M., Pang, Y., Zhang, L., 2010. Photo2Trip: generating travel routes from geo-tagged photos for trip planning. ACM Press, p. 143.

McHugh, J.A., 1990. Algorithmic graph theory. Prentice Hall.

Okuyama, K., Yanai, K., 2010. A Travel Planning System Based on Travel Trajectories Extracted from a Large Number of Geotagged Photos on the Web.

Pahlavani, P., Samadzadegan, F., Delavar, M.R., 2006. A GIS-Based Approach for Urban Multi-criteria Quasi Optimized Route Guidance by Considering Unspecified Site Satisfaction, in: Raubal, M., Miller, H.J., Frank, A.U., Goodchild, M.F. (Eds.), Geographic Information Science. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 287–303.

Samet, H., 1984. The Quadtree and Related Hierarchical Data Structures. ACM Computing Surveys 16, 187–260.