Assisted Journey Recollections from Photo Streams (Demo Paper)

Tyng–Ruey Chuang[*]
Academia Sinica
Taipei, Taiwan

Jheng–Peng Huang[†]
Academia Sinica
Taipei, Taiwan

Hsin–Huei Lee[‡]
Academia Sinica
Taipei, Taiwan

Kae–An Liu[§]
Academia Sinica
Taipei, Taiwan

Huang–Sin Syu[¶]
Academia Sinica
Taipei, Taiwan

ABSTRACT

We extract GPS traces from photo streams and analyze them to reveal movement types. Speed and locale patterns hint about the kinds of activity as captured by the photos of the day. When properly categorized and visualized, the photos and their movement patterns help people in navigating the itineraries of their past, and in retreating images of possible highlights. Our method is tolerant of erroneous and missing positional information in the photos' metadata. External geospatial resources can be further combined and visualized with the itineraries and photos to assist people's recollection of the places they were visiting.

CCS Concepts

•Information systems → Spatial-temporal systems; Global positioning systems; •Human-centered computing → Visualization systems and tools;

Keywords

GPS, Itinerary, Journey, Movement, Photo, Recollection

[*]Institute of Information Science, with joint appointments at the Research Center for Information Technology Innovation, and at the Research Center for Humanities and Social Sciences; Research supported in part by the Ministry of Science and Technology, Taiwan, under grants 103-2627-M-001-008 and 105-2627-M-001-014. (trc@iis.sinica.edu.tw)

[†]Institute of Information Science; Research student from the National Taiwan University of Science and Technology.

[‡]Institute of Information Science; Undergraduate intern from the National Tsing Hua University.

[§]Institute of Information Science; Undergraduate intern from the National Taiwan University.

[¶]Institute of Information Science; PhD student at the National Cheng Kung University. (gipong@iis.sinica.edu.tw)

SIGSPATIAL'16 October 31 - November 03, 2016, Burlingame, CA, USA

ACM ISBN 978-1-4503-4589-7/16/10.

DOI: http://dx.doi.org/10.1145/2996913.2996955

1. MOVEMENT TYPE AND LOCALE

Let event $d_i = (p_i, t_i)$ be a pair of values extracted from a photo's metadata recording the position p_i and time t_i when the photo was taken. We assume t_i is always present and accurate, while p_i may be missing or erroneous. The index i is often associated to an identifier — for example the file name of the photo — from which more information about the photo can be fetched. For all photos taken in a given day, we order them by the time they were captured. That is, $t_i < t_{i+1}$ for all $0 \le i < n$, where d_0 is extracted from the first photo taken on the day, and d_{n-1} is from the last photo. There are in total n photos captured on the day.

We choose a constant w, a time window, to partition the event sequence $D = (d_0, d_1, \ldots, d_{n-1})$ into m consecutive non–empty sequences $D = E_0 E_1 \ldots E_{m-1}$ so that for any pair d_{i-1} and d_i,

$$\begin{array}{rcll} t_i - t_{i-1} & \le & w & \text{iff both } d_{i-1} \text{ and } d_i \text{ are in } E_k \text{, and} \\ t_i - t_{i-1} & > & w & \text{iff } d_{i-1} \text{ is in } E_{k-1} \text{ and } d_i \text{ in } E_k \end{array}$$

for all $0 \le k < m$. Note that, given w, this partition is unique. The constant w is the length of a time interval where if no event ever occurs within it, the interval is used to separate the day's events into those before and those after it. For the travel data we analyze, we use a window of 15 minutes.

We call E_j an episode, and d_i an event. An event $d_i = (p_i, t_i)$ has a coordinate p_i and a timestamp t_i. It is indexed by i. Both events and episodes are ordered, temporally, by their indices. Within an episode, any two neighboring events occur within w in time. With this temporal continuity, we analyze the events in an episode for patterns of movement.

When the position of an event d_i is missing or erroneous, we write $p_i = \uparrow$ to mean d_i has an undefined position. We define two functions *pred* and *succ* on index i such that for any event d_i with a defined position, $d_{pred(i)}$ is the event right before d_i that also has a defined position. Likewise, $d_{succ(i)}$ is the event right after d_i that also has a defined position. If $p_i = \uparrow$ then both $pred(i)$ and $succ(i)$ are undefined. We calculate the (instantaneous) speed s_i at d_i by

$$s_i = \frac{1}{2}(speed(d_{pred(i)}, d_i) + speed(d_i, d_{succ(i)}))$$

when p_i is defined, and the three events $d_{pred(i)}$, d_i, and $d_{succ(i)}$ are all in the same episode. Otherwise, $s_i = \uparrow$ is undefined.

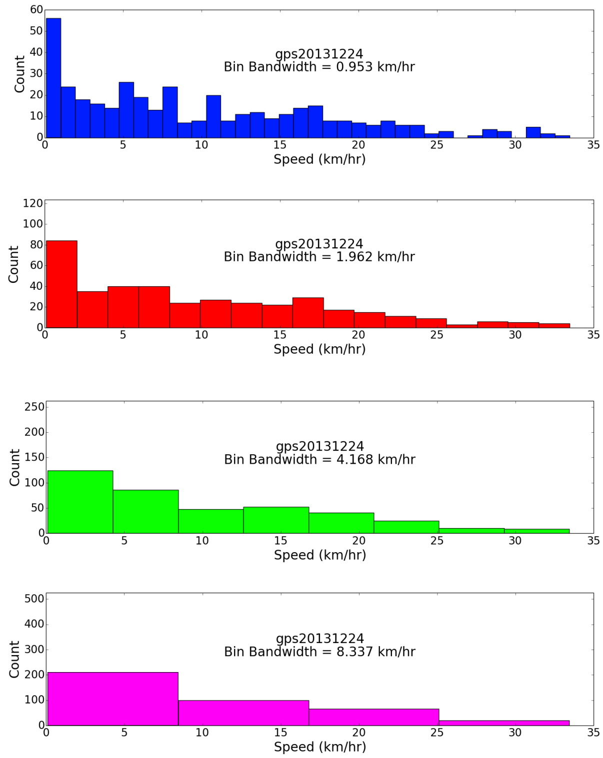

Figure 1: Histograms of Speeds (Biking).

From the speeds so gathered in a day, we can learn a few things about the events such as the maximal speed captured for the day, as well as when and where it occurred. This is an event we may want to highlight. If we find out the speeds are all within the range of say, 0 and 3 km per hour, then the day was probably spent walking or stationary. In the following we look into the distribution of all the event speeds in a day, and use it as a basis to distinguish different movement types.

We can use a histogram to get some ideas of the distribution of speeds. First of all, undefined speeds are not counted at all. We need to decide on a bandwidth, the bin width, to sort the speeds into different bins. If the bandwidth is b km/hr, then bin numbered k will count the number of events whose speeds are within $[kb, (k+1)b)$ km/hr. Bandwidth selection affects the shape of the resultant histogram. Large bandwidths get oversmoothed histograms where the distributions are overly flat. Small bandwidths lead to undersmoothed distributions. Figure 1 shows four histograms plotting the same set of speeds collected for a day gone biking. The respective bandwidths are, roughly, 1, 2, 4, and 8 km/hr. A histogram can be viewed as a probability density function (PDF) when the counts are converted to relative frequencies (with a summary frequency of 1.0 for the total count). This results in a (box) kernel density estimation for the underlining PDF of the speed population. The PDFs in Figure 1 are discontinuous as the kernel (the boxcar function) is discrete. Figure 2 is another kernel density estimation for the same population of event speeds. This estimation uses the Gaussian function as the kernel. With a good estimation, we hold on to a smooth and fitted PDF

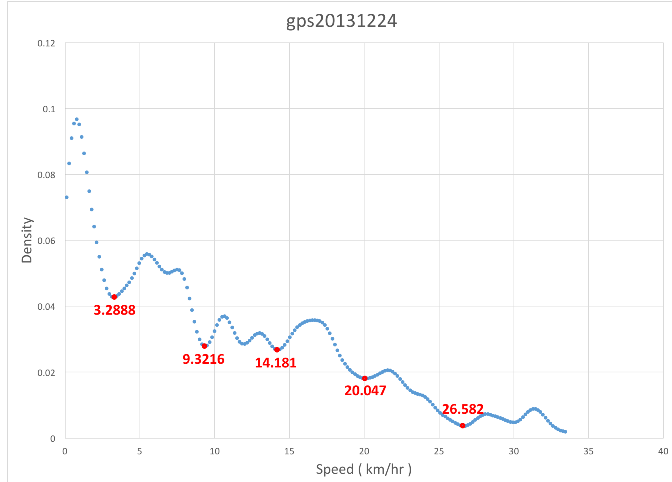

Figure 2: Probability Density of Speeds (Biking).

of the event speeds in the day. More on this in Section 3.

We identify speeds that are relatively rare, and use them to segment the events in an episode into various movement types. That is, we locate those speeds whose densities are minimums in the PDF. Take Figure 2 as an example, the minimum densities occur when the speeds are 3.2888, 6.8079, 9.3216, 12.003, 14, 181, 20.047, and 26.582 km/hr. The respective densities are 0.042754, 0.050081, 0.027836, 0.028571, 0.026724, 0.018075, and 0.003686. As there may be too many minimums to be of practical use, we keep only those speeds whose densities not only are minimum but also decreasing. For this example, we keep only the speeds 3.2888, 9.3216, 14, 181, 20.047, and 26.582. Their densities are minimums and form a decreasing sequence of 0.042754, 0.027836, 0.026724, 0.018075, and 0.003686. The above five speeds can be used as delimiters to categorize events into one of the six speed types. Type I denotes events whose speeds are in the range of $[0, 3.2888)$, type II denotes events whose speeds are in the range of $[3.2888, 9.3216)$, etc.

Within an episode, we collect succeeding events of the same type into a sequence. We use the term movement for such a sequence. Within a movement, the events all have the same speed type. A day's journey can now be reviewed by movement types and by their temporal orders. When overlaid on a map, these movements, in combination of movement directions and the areas of visit, help recall the day's journey.

2. VISUALIZATION AND EXPLORATION

We demonstrate how speeds, movement types, and directions calculated from a day's events can be gathered and visualized for journey recollections. Recall that we often record hundreds of photos for a day's travel. The method we propose allows us to have an overview of the movements in the day, and to look out for places where the interesting movements are. Photos are called up and reviewed after we have located movements of particular interests.

In Figure 3, we visualize and explore a set of 795 photos taken in a day of biking. The photos were captured every 45 seconds by a camera mounted at the handle of a bike. Some of the events extracted from the photos have no location

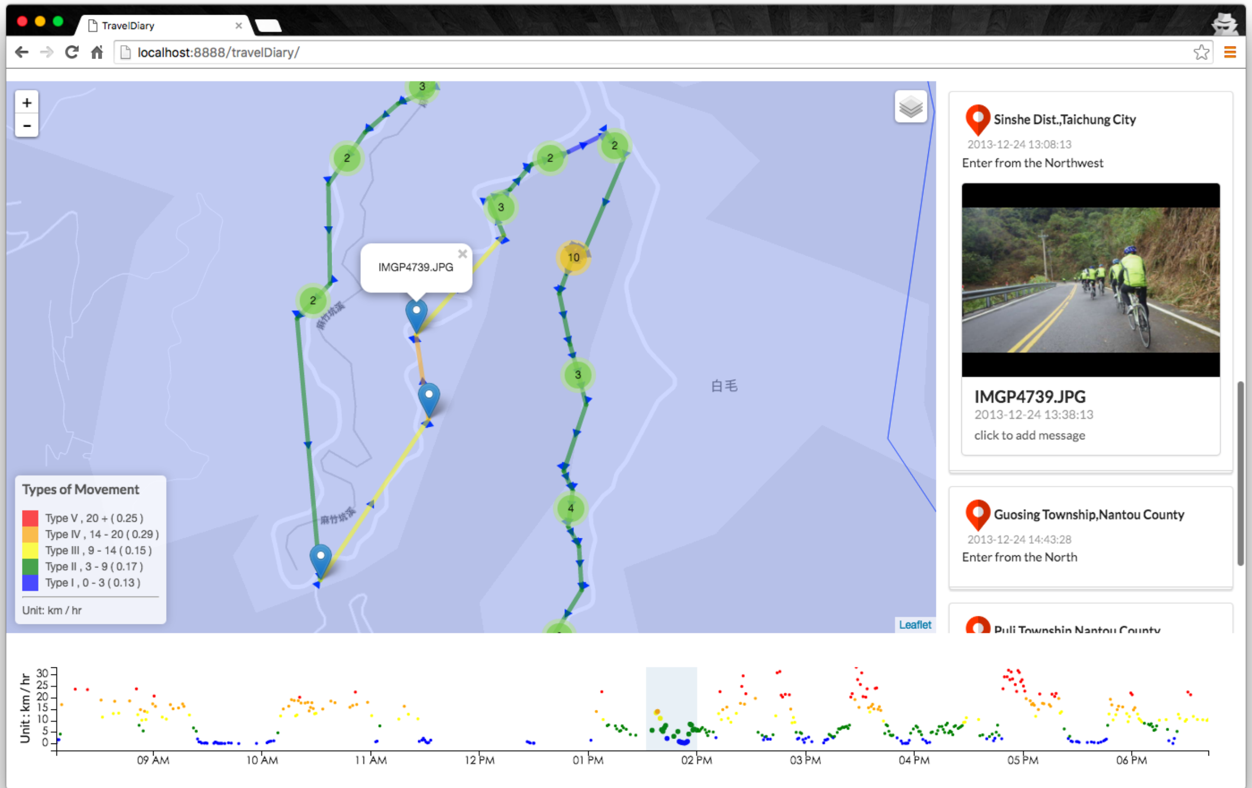

Figure 3: Recalling Events and Movements (Biking).

data (*e.g.* when in doors) or have erroneous locations (*e.g.* when a photo has exactly the same location of the previous photo because the GPS unit couldn't compute a new position in time). Only 395 events have defined speeds; they are displayed along the timeline at the bottom of the window. The speeds are color–coded by their types. Above the timeline is a map, showing the locale of the events within the selected interval of the timeline. These events form movements, and they are displayed on the map with directions. Note that the movements are color–coded too. The color code for movement type is summarized at the lower left corner of the map. Suppose we are searching for a type IV movement where we were going at the speed of 14 to 20 km/hr. The orange segment at the center of the map can be what we are looking for. We click on one of the the markers on the segment. The photo named `IMGP4739.JPG` is retrieved and displayed on the right, so we can refresh our memory. In the photo we were riding down the hill which explains the fairly high speed. The moment was 13:08, and we were in Sinshe District, Taichung City. We also notice the blue segment near the top of the map which indicates a slow movement (0 to 3 km/hr). On the map we can see there is a very sharp turn ahead, so we were slowing down.

3. DISCUSSION

It is obvious the proposed method can be applied to GPS traces whereas each data point in the traces is an event. It is just that there will not be any picture to go along with visualization. We also emphasize there need not be many events for the method to produce reliable and useful speed types for categorizing movements. In Figure 4 and 5 we show the histogram and a kernel density estimation of the PDF of 37 events with defined speeds, taken from a total of 45 (the rest have undefined speed). These were captured by a smartphone's GPS unit when we were making a trip from Taipei to Tainan. We started on a highspeed train, followed by taking a commuter train, and ended by a walk in the Tainan city. In Figure 6, we see that in the map a type V movement (184 km/hr and above) is followed by type II movements (20 to 74 km/hr). The trip ended with type I movements (0 to 20 km/hr).

Also in Figure 5, we witness a scenario where the minimum density reaches 0 when the speed is 127.75 km/hr, but there are other minimums for speeds over that. In such a scenario, we will reset and start to find more minimum densities for the rest of event speeds. This way, we are able to identify two more minimum densities (when the speeds are 184.2 and 209.42 km/hr). As for the general issue of selecting a good bandwidth in kernel density estimation, we resort to an optimization method by Shimazaki and Shinomoto [3]. Their method selects a fixed bandwidth minimizing the integrated squared error for measuring the goodness–of–fit of an estimate to the unknown underlying distribution. The distributions in Figure 2 and 5 are each generated by their method

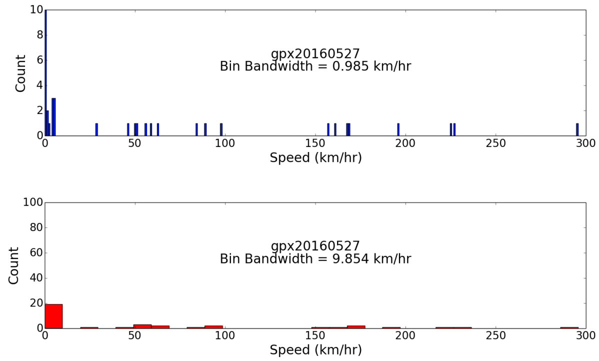

Figure 4: Histograms of Speeds (Trains).

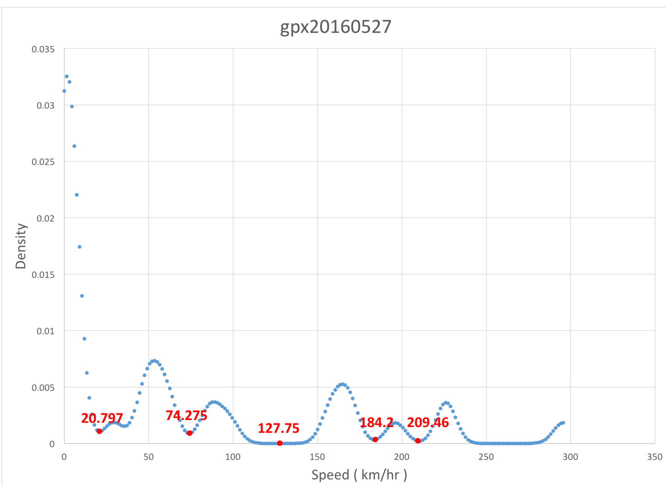

Figure 5: Probability Density of Speeds (Trains).

when supplied with all the events with defined speeds in the given day.

Although we have overlaid a day's movements on a map, we have not started to correlate the day's itineraries with the routes from transportation networks. Also, even though we use an administrative map to learn about where an event was happening (by the name of the administrative area), this is rather dry. Ideally we shall use a spatial database about landmarks, attractions, and various kinds of POI (Point of Interest) to highlight places that were visited. Currently we only extract a few fields from a photo's embedded EXIF record. Altitude information is not used, so is camera orientation. We do not analyze images for content semantics either.

An online demo of our current system can be found at `<http://www.iis.sinica.edu.tw/~gipong/sigspatial16/>`. Several previous works are related to the current one. Lerin et. al. used a pace–based clustering method to infer visit locations and durations from GPS traces [1]. Our use of kernel density estimation seems to be a simpler way in getting similar outcome. There are systems utilizing EXIF metadata for photo album organization and management [2], but they analyze not so much the spatiotemporal relations among the photos. Yan et. al. described a framework for enabling semantic trajectory representation of physical movements as recorded in GPS traces [5]. Their framework is comprehensive, and involves layers of geospatial datasets. Yue et. al. surveyed the field of trajectory–based travel behavior studies [6]. Su et. al. proposed a method to summarize trajectory data in order to generate short text descriptions [4]. These works emphasized the collection and analysis of large trajectory datasets, usually from diverse sources and different users, in order to derive insights into collective behaviors and/or to gain intelligence about specific traces. The persons carrying out the analysis likely are not the people from whom the trajectory datasets are collected. Our method, on the other hand, is intended to help individuals make good sense of their own data, and to assist them recall their whereabouts and actions.

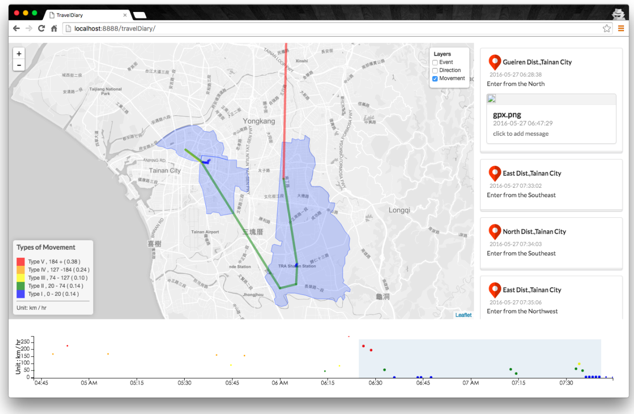

Figure 6: Recalling Events and Movements (Trains).

4. REFERENCES

[1] P. M. Lerin, D. Yamamoto, and N. Takahashi. Pace–based clustering of GPS data for inferring visit locations and durations on a trip. *IEICE Transactions on Information and Systems*, 97(4):663–672, 2014.

[2] D.-S. Ryu, K. H. Kim, S.-Y. Park, and H.-G. Cho. A Web–based photo management system for large photo collections with user–customizable quality assessment. In *Proceedings of the 26th Annual ACM Symposium on Applied Computing*, pages 1229–1236. ACM, 2011.

[3] H. Shimazaki and S. Shinomoto. Kernel bandwidth optimization in spike rate estimation. *Journal of Computational Neuroscience*, 29(1):171–182, 2010.

[4] H. Su, K. Zheng, K. Zeng, J. Huang, S. Sadiq, N. J. Yuan, and X. Zhou. Making sense of trajectory data: A partition–and–summarization approach. In *Proceedings of the the 31st International Conference on Data Engineering*, pages 963–974. IEEE, 2015.

[5] Z. Yan, D. Chakraborty, C. Parent, S. Spaccapietra, and K. Aberer. SeMiTri: A framework for semantic annotation of heterogeneous trajectories. In *Proceedings of the 14th International Conference on Extending Database Technology*, pages 259–270. ACM, 2011.

[6] Y. Yue, T. Lan, A. G. O. Yeh, and Q.-Q. Li. Zooming into individuals to understand the collective: A review of trajectory–based travel behaviour studies. *Travel Behaviour and Society*, 1(2):69–78, 2014.