映像要約技術に基づく都市移動における重要時点の抽出

Key Time Extraction in City Mobility based on Video Summarization

学籍番号　47-196798
氏　　名　尹 航（Yin Hang）
指導教員　宋　軒　教授

1. Introduction

Human mobility mode in range of city provides sufficient information for urban planning and transportation management, which can help improving traffic efficiency and managing public transportation, thus benefiting life of residents. Nowadays, management methods like timeshare fare, speed limit of road, number limit of license plate, etc. have been widely used in transportation management. However, these methods share a similarity that the fare or the management plan are decided beforehand, which means it could not respond to changes in a timely manner.

If we can grasp the dynamic change of human mobility macroscopically, we can find out when the change of management should be adjusted. Furthermore, with the tendency of human mobility change and the detailed degree, specific measures like change of toll, road regulation time zone designation, increase police force, etc. can be carried out.

In another field, video summarization technologies aim to create a concise and complete synopsis by selecting the most informative parts of the video content. Nowadays, when we surf the Internet and watch videos on modern video sites such as YouTube, we can find the extracted key frames are automatically played to tell users their general contents. In this way, users can get a glimpse of certain video summary and quickly decide if it suits their interests just by looking at the extracted summary video.

Back to key time extraction, as we know, frames are arranged in order of time to compose a video. From a dynamic point of view, traffic data is also organized in this way, which is keeping changing with human mobility. If making comparison between video summarization and time

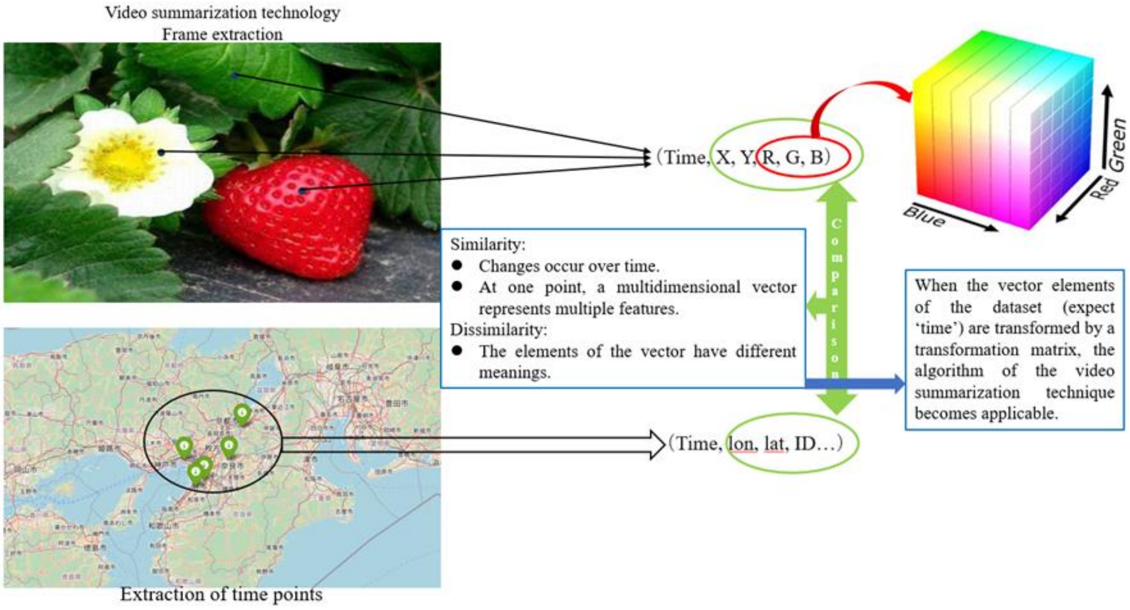

Figure 1. Relationship between key time extraction and video summarization

point extraction, we can find their similarity and dissimilarity:

Similarity: Both contain time in feature vectors, which means change occurs every time.

Dissimilarity: Mobility data vectors contain features such as longitude, latitude, ID, and density, while frame vectors contain two-dimensional coordinates and RGB matrices.

Through matrix transformation, if mobility data set can be treated as one video and the map of the location and movement of people on every time point can be treated as one frame of the video, the algorithm of video algorithm will be applicable for key time extraction.

2. Methodology

We divide the whole process of key time extraction into three steps: (1) data filtering; (2) *Jaccard* distance-based population density dissimilarity matrix computation; (3) Comparison and Improvement based on Submodularity optimization.

2.1 Data Filtering

step 1. Trajectory clustering

SSPD (Symmetrized Segment-Path Distance) from trajectory T_1 to trajectory T_2 is defined as:

$$D_{SSPD}(T_1, T_2) = \frac{D_{SPD}(T_1,\, T_2) +\, D_{SPD}(T_2,\, T_1)}{2}$$

$$D_{SPD}(T_1,\, T_2\,) = \frac{1}{n_1}\sum_{i_1=1}^{n_1} d_{pt}(p_{i_1}^1,\, T_2)$$

$$d_{pt}\big(p_{i_1}^1,\, T_2\big) = min_{i_2\in[0,\dots,n_2-1]}d_{ps}(p_{i_1}^1,\, s_{i_2}^2).$$

Based on SSPD, trajectory distance between every trajectory combination will be roughly computed with no need to do data interpolation. Then we can get to know which group of mode the target ID belongs to.

step 2. Reconstruction of low-rank matrix

IDs belong to the same group means they share some similar mobility tendency. In another word, in the field of matrix computation, it means the vectors of these IDs have linear correlation. Also, the number of groups is surely much less than the number of IDs, so the matrix containing all IDs is low-rank. Based on low-rank matrix reconstruction technology, we can estimate the missing value to recovery the mobility matrix.

step 3. Map-matching

From step 1 and 2, we have got the relatively completed mobility data set, but some abnormal values appear like flash phenomenon, pendulum phenomenon, etc., which are unreasonable due to the difference with real human mobility. We do map-matching following this flow:

(1) Form candidate road list using range query.

(2) Calculate travel likelihood for all roads in candidate road list.

(3) Is Vehicle or human off-road? Yes→ **(2)**; No→ **(4)**.

(4) Map GPS-estimated position to identified road of travel.

(5) Generate successor roads? Yes→ **(6)**; No→**(2)**.

(6) Add successor roads to candidate road list and delete roads with low travel likelihood.

2.2 Population density dissimilarity matrix computation

Based on the common sense of population density computation formula, we can get population density matrix set:

$$\wp = \{P_{t_1}^{m\times n}, P_{t_2}^{m\times n}, \ldots, P_{t_s}^{m\times n}\}$$

$$t_{i+1} - t_i = interpolation\ time\ unit,$$

$$m = \frac{latitude}{\Delta latitude}, n = \frac{longitude}{\Delta longitude},$$

$$s = number\ of\ interpolation$$

According to *Jaccard* distance, we can compute dissimilarity between each population density matrix $P_{t_a}^{m\times n}$ and $P_{t_b}^{m\times n}$ $(a, b \in [1 \ldots s])$, and formula will be transformed as:

$$d_J(a,b) = d_J\left(P_{t_a}^{m\times n}, P_{t_b}^{m\times n}\right)$$
$$= 1 - \frac{\sum_n \sum_m \min\left(P_{t_a}^{i,j}, P_{t_b}^{i,j}\right)}{\sum_n \sum_m \max\left(P_{t_a}^{i,j}, P_{t_b}^{i,j}\right)}$$

2.3 Comparison and Improvement

2.3.1 Key Time 1

Compute *Jaccard* distance only between neighbor time points a and $a+1$ $(a, a+1 \in [0 \ldots s])$ to get the vector $d_{J-neig} = (d_J(1,2), d_J(2,3), \ldots, d_J(s-1,s))$

2.3.2 Key Time 2

Construct the density dissimilarity matrix $D_J^{s\times s} = (d_J(a,b))_{s\times s}$ based on population density dissimilarity between each time point combination.

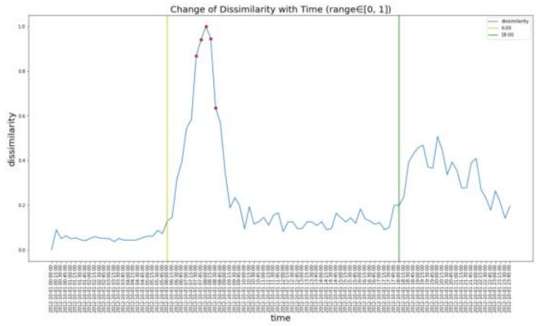

Figure 2 Polyline of Key Time 1

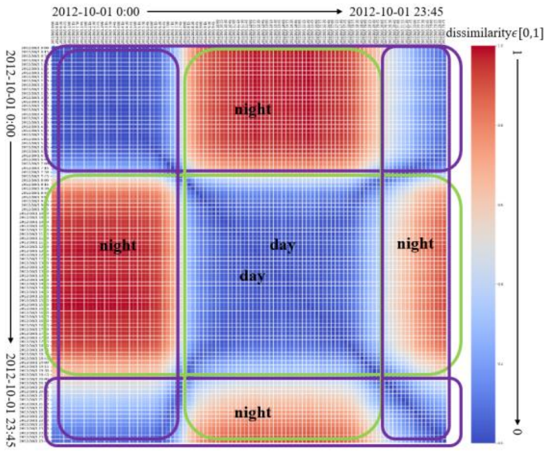

Figure 3. Dissimilarity of Key Time 2

2.3.3 Comparison

- Key Time 1 pays more attention to continuous time, while Key time 2 grasps overall change and adjacent change at the same time
- Key Time 1 extracts time points all fall in the range from 7 am to 9 am, which is the morning peak. During morning peak, it is certainly that mobility situation will vary fast, but it couldn't summarize the situation of mobility mode change.
- Extraction consequence of Key Time 2 is separated overall day, which can roughly represent 5 time segments: sleeping time, morning peak, working time, dinner time, night peak.

2.3.4 Improvement

Due to the information ignorance (ignore the OD (origin-destination) information of ID) of population density, we should add OD similarity into original target function.

Then we use Submodularity Optimization, (which is a greedy algorithm to use the least resource to complete the most things.) to select representative time points. To measure appearance, we transform time as:

$$tansform(t(y,m,d,h,m,s)) = \frac{3600h + 60m + s}{86400}$$

Table 1 shows the standard deviation and red number is coincidentally '成人の日'.

Date	Day of Week	Extracted time1	Extracted time2	Extracted time3	Extracted time4	Extracted time5	standard deviation
2012/1/1	Sun	0.14583	0.35417	0.59375	0.82292	0.9375	5.48%
2012/1/2	Mon	0.11458	0.3125	0.5625	0.80208	0.92708	9.35%
2012/1/3	Tue	0.13542	0.33333	0.58333	0.80208	0.92708	8.76%
2012/1/4	Wed	0.13542	0.34375	0.57292	0.80208	0.92708	8.23%
2012/1/5	Thu	0.13542	0.33333	0.55208	0.79167	0.91667	7.46%
2012/1/6	Fri	0.13542	0.34375	0.57292	0.79167	0.91667	7.83%
2012/1/7	Sat	0.14583	0.35417	0.54167	0.76042	0.90625	4.86%
2012/1/8	Sun	0.125	0.3125	0.57292	0.83333	0.9375	6.42%
2012/1/9	Mon	0.17708	0.4375	0.57292	0.70833	0.84375	13.48%
2012/1/10	Tue	0.13542	0.38542	0.52083	0.73958	0.91667	7.54%
2012/1/11	Wed	0.14583	0.39583	0.51042	0.72917	0.89583	8.36%
2012/1/12	Thu	0.13542	0.39583	0.52083	0.75	0.91667	7.94%
2012/1/13	Fri	0.17708	0.34375	0.5625	0.73958	0.91667	7.68%
2012/1/14	Sat	0.14583	0.34375	0.5625	0.8125	0.92708	4.98%
							7.74%

Table 1. Submodularity and OD matrix based standard deviation