

THUIR at the NTCIR-16 Lifelog-4 Task: Enhanced Interactive Lifelog Search Engine

Task Definition

- > Retrieve a number of specific moments in a lifelogger's life.
- > Input: query of a topic
- > Output: relevant lifelog images
- > Interactive Manner: the user involved in the search process to provide relevance feedback or query reformulation.
- > Dataset: multi-modality dataset of 114 days' records collected in 2015, 2016, and 2018. 183,432 wearable camera images with extracted visual concepts and textual metadata.

*****Feature Extraction

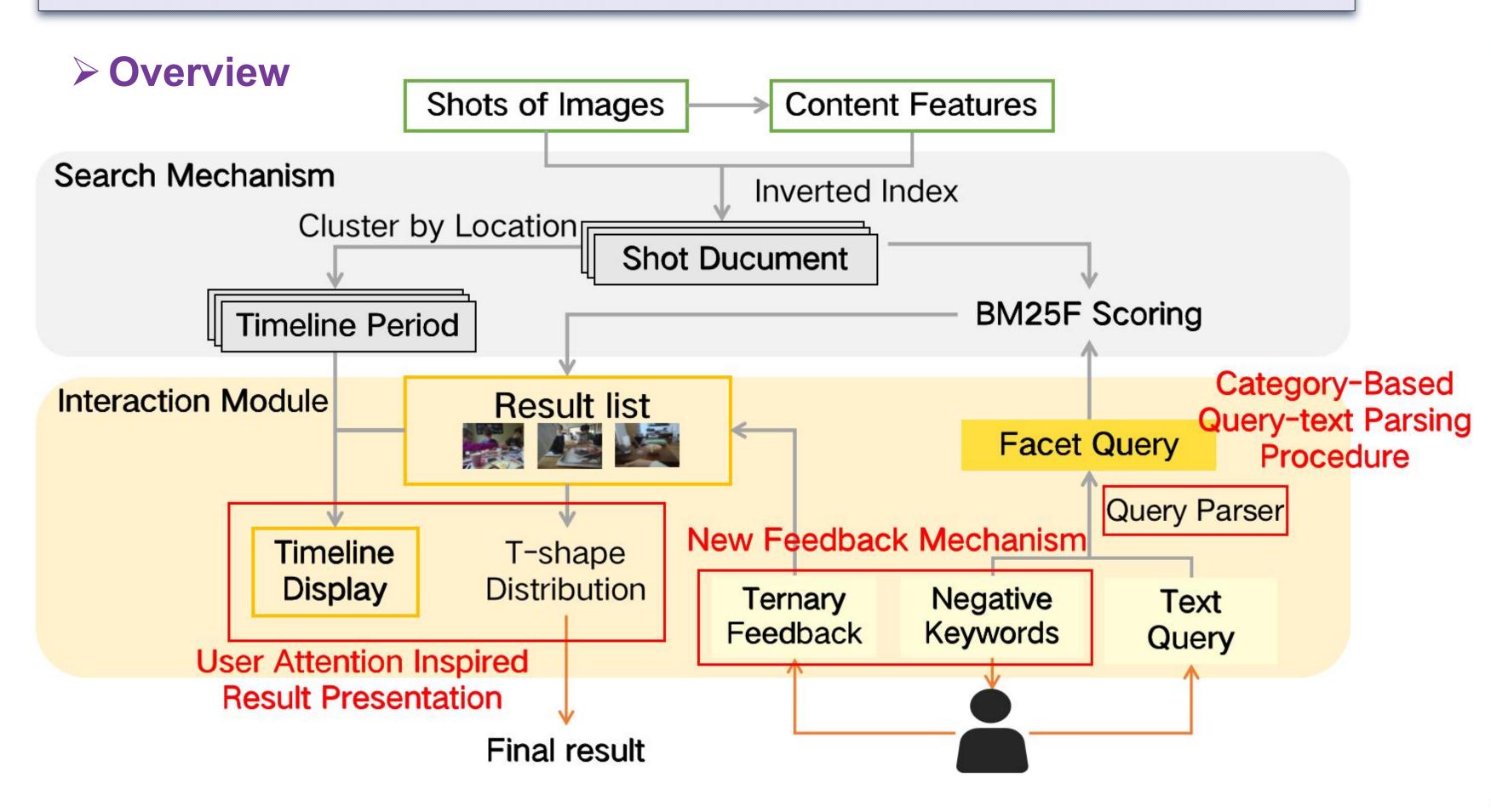
Denoising and Clustering

- Remove blurry and blocked images.
- Consecutive and similar pictures are clustered into shots.
- > 50,233 shots in total.

> Multilevel Feature Extraction

 \succ Visual features, textual features, and behavior features.

Interactive Methods



Information Retrieval Group in Tsinghua University

Zhiyu He, Jiayu Li, Wenjing Wu, Min Zhang, Yiqun Liu, and Shaoping Ma Information Retrieval Group, Department of Computer Science & Technology Tsinghua University

Category-based Query-text Parsing

To extract the practical information of each field from the free text.

match their **JN** type noun phrases.

Type	Grammar	Explanation	Example	
JN	<dt>*(<cc>*<jj>*)*<nn></nn></jj></cc></dt>	noun phrases containing adjectives	a white and blue shirt	
PN	<jn><in>+<jn></jn></in></jn>	noun phrases joined by prepositions	a white cat on a chair	
NJ	<jn><vb><in>*<jn>+</jn></in></vb></jn>	noun phrases joined by verbs	a man sitting on a chair	

> Extract the category of the nouns based on rules.

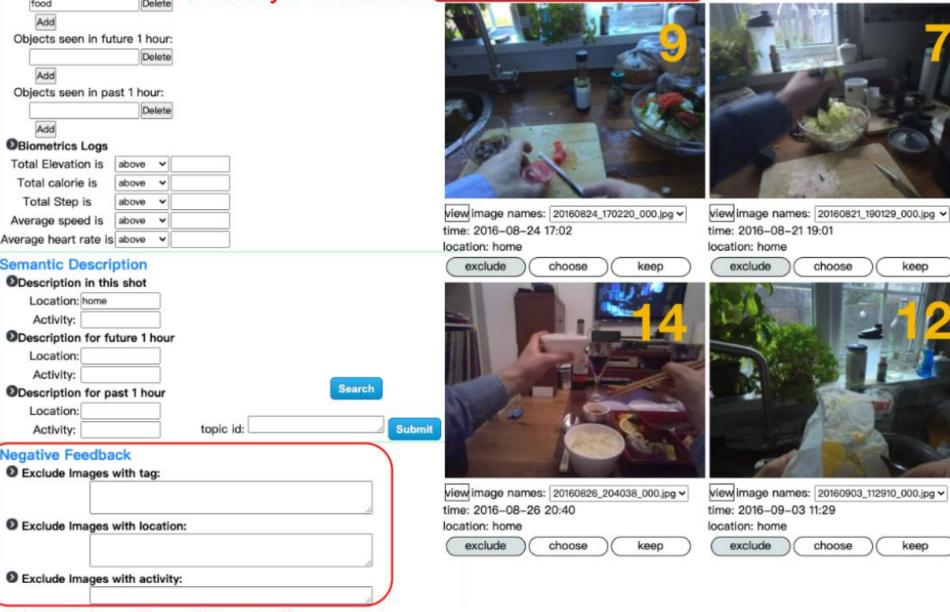
Examples:

Query Text	
Find examples of when I was lifelogged when eating lunch at work in my office.	Tag App Loc
Find examples of when I was in meetings with other people in rooms with red carpet. Before that, i had lived in hotel.	tags root tags loca

> New Feedback Mechanism

> Ternary feedback for each image in result list. feedback).

Timeline Viewing Button ing a example when i was chopping vegetables into a bow ill in Forms time: 2016-09-03 13:47 time: 2016-08-20 16:15 Objects in the scene Delete Ternary Feedback



Negative Feedback Boxes

Approximate time

Range of hour:

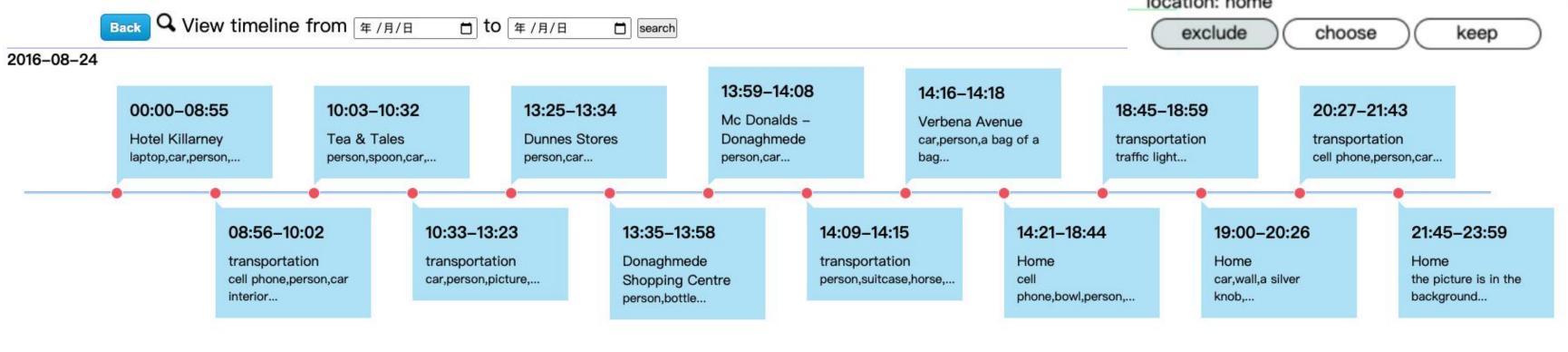
- > Determine the sentence's **tense** and generate **syntactic tree**.
- > Classify the object label descriptions into three categories and

Facet Query

- gs in this shot: [food, office] proximate time: [noon] cation in this shot: [office, work]
- s in this shot: [red carpet, om, rooms, people] s in past 1 hour: [hotel] ation in past 1 hour: [hotel]

> Negative feedback keywords (while query input as positive

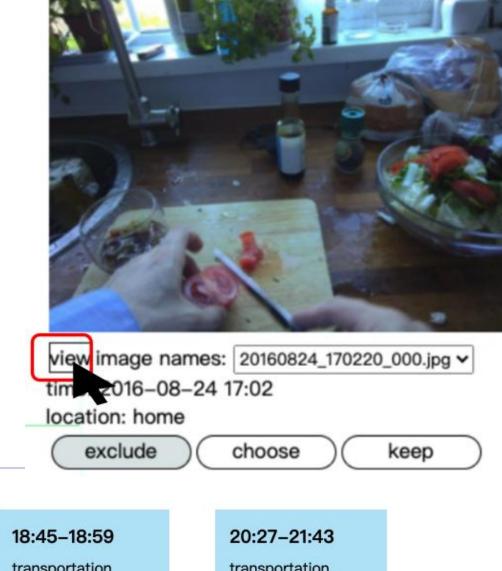
num_q view image names: 20160820 161552 000.jpg v view image names: 20160820 163653 000.jpg > view image names: 20160820_162613_000.jpg v view image names: 20160824_173227_000.jpg time: 2016-08-20 16:36 time: 2016-08-20 16:26 ime: 2016-08-24 17:32 ocation: the sisters hon location: the sisters home num_ret num_rel_ret map Rprec view image names: 20160821_191713_000.jpg ~ view image names: 20160821_190806_000.jpg ~ view image names: 20160821_190129_000.jpg ~ view image names: 20160903_110855_000.jpg time: 2016-08-21 19:17 time: 2016-08-21 19:08 ime: 2016-09-03 11:08 location: home location: home exclude exclude exclude keep keep choose choose choose keep choose recip rank P_5 P_20 view image names: 20160904_063930_000.jpg ~ view image names: 20160824_165158_000.jpg ~ view image names: 20160904 124443 000.jpg > P_100 time: 2016-09-04 12:44 time: 2016-09-04 06:39 time: 2016-08-24 16:51 location: home location: home ocation: home) exclude choose keep exclude choose keep exclude choose keep



constructed topics.

>User Attention Inspired Result Presentation for Interaction > Design **T-shaped fixation distribution** of results that conforms to the user's image browsing habits: the shots diverge from the middle in the horizontal direction.

> Build the **timeline display** and match the shot to the background information of life activities during the time.



*****Experiment and Results

We compare the search engine before and after the interactive methods enhancement by both a novice and an expert user based on the

KNOWNITEM				ADHOC			
Novice		Expert		Novice		Expert	
pre	post	pre	post	pre	post	pre	post
6	6	6	6	5	5	6	6
24	26	37	40	215	233	159	162
19	21	31	36	51	66	118	103
0.2086	0.3753	0.5711	0.6410	0.3047	0.4857	0.4850	0.7160
0.2086	0.3753	0.5842	0.6410	0.2556	0.5220	0.5487	0.7571
0.6667	0.8333	1.0000	1.0000	0.6848	0.6917	0.5931	0.8431
0.3333	0.4000	0.5667	0.6000	0.5600	0.6400	0.5000	0.8000
0.1583	0.1750	0.2583	0.3000	0.4300	0.4300	0.4917	0.6000
0.0317	0.0350	0.0517	0.0600	0.1020	0.1320	0.1967	0.1717

Email:z-m@tsinghua.edu.cn