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

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

❖ 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**
  ➢ Visual features, textual features, and behavior features.

❖ Interactive Methods

➢ **Overview**
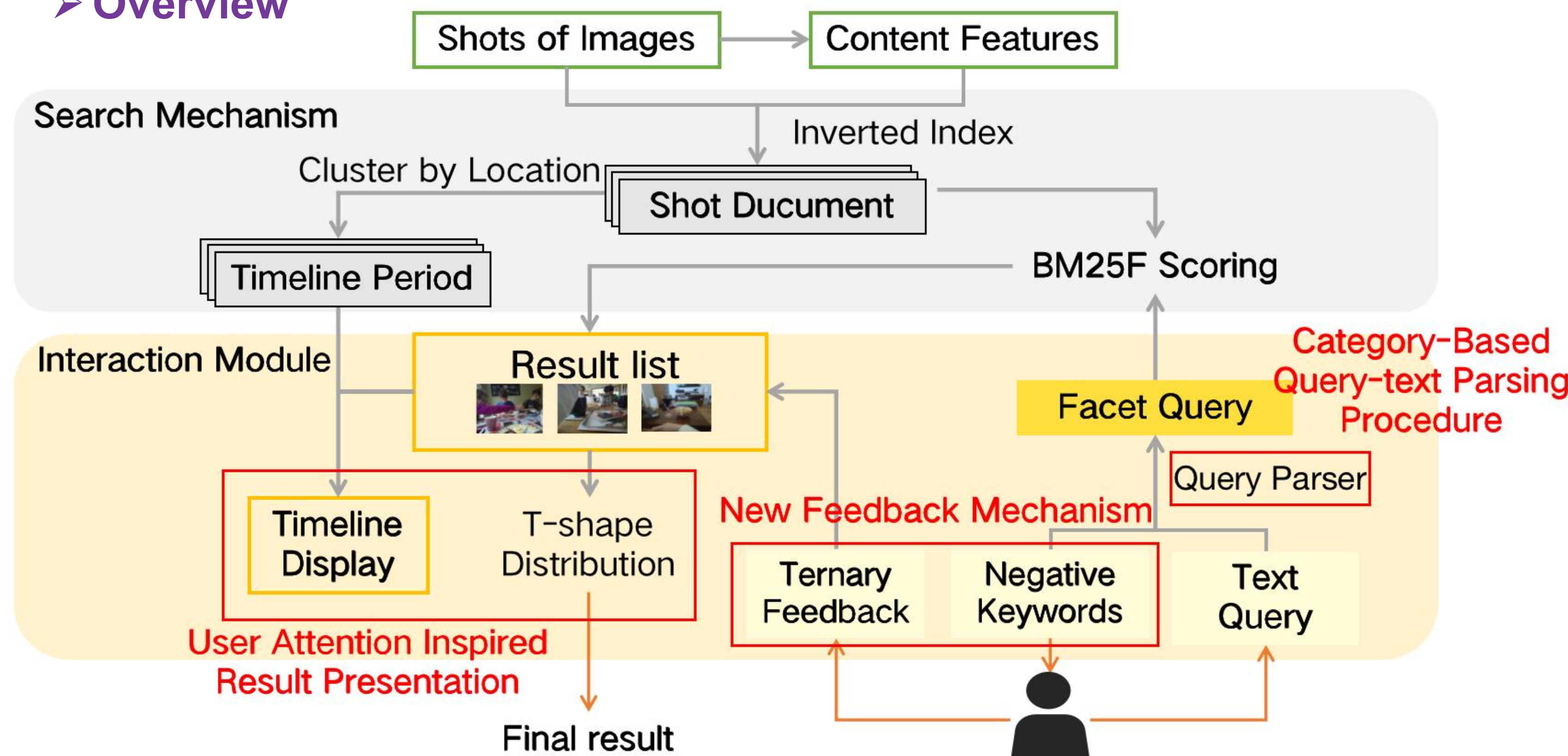
➢ Category-based Query-text Parsing

To extract the practical information of each field from the free text.
➢ Determine the sentence's **tense** and generate **syntactic tree**.
➢ Classify the object label descriptions into three categories and match their **JN** type noun phrases.

Type	Grammar	Explanation	Example
JN	<DT>*(<CC>*<JJ>*)*<NN>	noun phrases containing adjectives	a white and blue shirt
PN	<JN><IN>+<JN>	noun phrases joined by prepositions	a white cat on a chair
NJ	<JN><VB><IN>*<JN>+	noun phrases joined by verbs	a man sitting on a chair

➢ Extract the **category** of the nouns based on rules.

Examples:

Query Text	Facet Query
Find examples of when I was lifelogged when eating lunch at work in my office.	Tags in this shot: [food, office] Approximate time: [noon] Location in this shot: [office, work]
Find examples of when I was in meetings with other people in rooms with red carpet. Before that, i had lived in hotel.	tags in this shot: [red carpet, room, rooms, people] tags in past 1 hour: [hotel] location in past 1 hour: [hotel]

➢ New Feedback Mechanism

➢ **Ternary feedback** for each image in result list.
➢ **Negative feedback keywords** (while query input as positive feedback).

➢ 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.
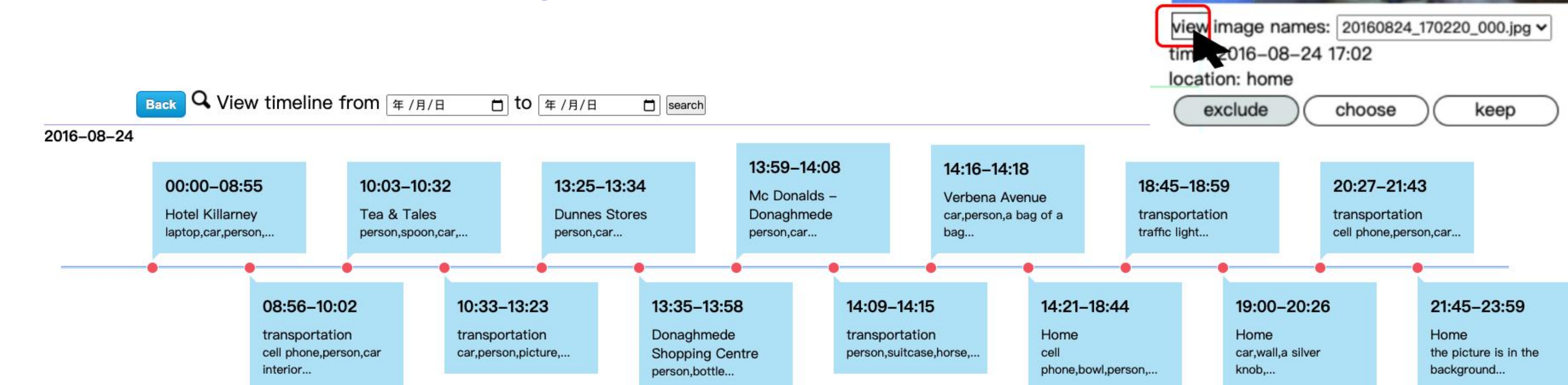
❖ Experiment and Results

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

	KNOWNITEM				ADHOC			
	Novice		Expert		Novice		Expert	
	pre	post	pre	post	pre	post	pre	post
num_q	6	6	6	6	5	5	6	6
num_ret	24	26	37	40	215	233	159	162
num_rel_ret	19	21	31	36	51	66	118	103
map	0.2086	0.3753	0.5711	0.6410	0.3047	0.4857	0.4850	0.7160
Rprec	0.2086	0.3753	0.5842	0.6410	0.2556	0.5220	0.5487	0.7571
recip rank	0.6667	0.8333	1.0000	1.0000	0.6848	0.6917	0.5931	0.8431
P_5	0.3333	0.4000	0.5667	0.6000	0.5600	0.6400	0.5000	0.8000
P_20	0.1583	0.1750	0.2583	0.3000	0.4300	0.4300	0.4917	0.6000
P_100	0.0317	0.0350	0.0517	0.0600	0.1020	0.1320	0.1967	0.1717