

WonderWhat: Real-time Event Determination from Photos

Mingyan Gao^{*}
School of Information and
Computer Science
University of California, Irvine
gaom@ics.uci.edu

Xian-Sheng Hua
Microsoft Research, Asia
Beijing, China
xshua@microsoft.com

Ramesh Jain
School of Information and
Computer Science
University of California, Irvine
jain@ics.uci.edu

ABSTRACT

How often did you feel disappointed in a foreign country, when you had been craving for participating in authentic native events but miserably ended up with being lost in the crowd, due to ignorance of the local culture? Have you ever imagined that with merely a simple click, a tool can identify the events that are right in front of you? As a step in this direction, in this paper, we propose a system that provides users with information of the public events that they are attending by analyzing in real time their photos taken at the event, leveraging both spatio-temporal context and photo content. To fulfill the task, we designed the system to collect event information, maintain dedicated event database, build photo content model for event types, and rank the final results. Extensive experiments were conducted to prove the effectiveness of each component.

Categories and Subject Descriptors

H.4 [Information Systems Applications]: Miscellaneous

General Terms

Design, Experimentation, Human Factors

Keywords

Event Determination, Photo, Context, Content

1. INTRODUCTION

When we are traveling in foreign countries, it is often a big problem for us to understand the events or activities that are going on in the local area. The difficulty prohibits common tourists from better appreciating different places and cultures. Can we create a tool that is capable of determining these events and delivering information about them in real time?

Efforts have been made on utilizing the rich data on the web to provide knowledge to end users. However, no existing approaches can fully satisfy our need. Traditional web search engines rely on keyword indexing. Although many techniques have been proposed to better understand user intentions in search, the gap between concepts and keywords still exists. Work on semantic web and related search have

^{*}This work was performed when the author was visiting Microsoft Research Asia as a research intern.

been advanced to bridge the gap, but the languages and query systems designed so far are still very “expert-oriented” [2] and not usable for common users. Visual search has gained increasing popularity and worked successfully for certain objects and text [3]. However, due to the semantic gap, the approach is insufficient for concepts with richer semantic meaning, like event.

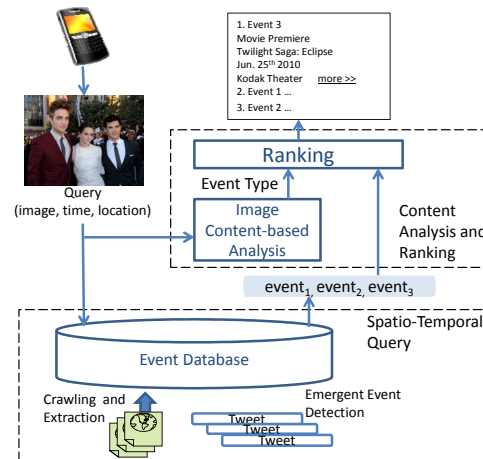


Figure 1: System Architecture

In this work, we propose a system that provides users with information of the public events that they are attending by analyzing in real time their photos taken at the event. Whenever a user wants to know about an event she is currently at, she only needs to take a picture of it. By examining the photo content together with the spatial and temporal data carried with it, our system automatically returns a ranked list of events with which the photo may be associated. Our approach has the following advantages. 1) Use of the system is very intuitive and requires no special efforts; 2) The system keeps a dedicated event database and index, and automatically constructs queries for users, which enables the delivery of exact event information; 3) Our system not only detects planned events, but also tries to discover concurrent events by analyzing real-time micro-blogs; 4) Different types of events do reveal distinct visual characteristics, so visual content is also taken into account to improve search results. As far as we know, there is no previous work that has addressed a similar problem.

2. OUR APPROACH

Problem We define a contextual photo as $p = (img, time, location)$, where img is the image content of the photo p ,

and *time* and *location* = (*latitude*, *longitude*) correspond to the timestamp and geo-coordinate of the photo *p*. We model event as $e = (time, location, title, description, type, media)$. $time = (start, end)$ defines the time interval during which *e* occurs. $location = (lat_1, long_1, lat_2, long_2)$ represents the geo-coordinates of the southwest and northeast corner of the place where *e* takes place. Name of the event *e* is stored in *title*, and the textual explanation to *e* is saved in *description*. Event types, such as performances, exhibitions, sports are stored in *type*. Media data, e.g. posters and photos, is kept in *media* attribute. Given an incoming contextual photo *p*, our problem is to compute a ranked list of events (e_1, \dots, e_n), which are in decreasing order of likelihood that e_i is the event at which the photo *p* was taken.

Approach We present the system architecture in Figure 1. Our system consists of the following major steps.

1. We create an event database, and ingest both planned and emergent events into it. Planned events, which are usually pre-declared online, are extracted from web pages or downloaded via web services that perform event integration. Emergent events, the occurrence of which are impromptu, are detected from Twitter by following the approach in [1].

2. After an user takes a photo for an event, her capturing device creates a contextual photo containing the image content, time and location information. The device then sends this contextual photo to our system as a query.

In our system,

3. First, given the time and location information in the contextual photo, a query is issued to the event database, which returns a list of related events.

4. Then the content analysis component analyzes the image content and returns the event type of the event captured in the photo. In this work, we model the relationship between event types and the raw visual features through a middle layer of visual concepts. We employed a learning based approach to perform the analysis, which consists of four major steps: 1) Train concept detector; 2) Detect concepts from photos associated with different event types; 3) Train event type detector; 4) For each incoming photo, based on the models, decide which type of event the photo is most likely to be.

5. Both the event list from event database and the detected event type are given to the ranking component. The component considers spatial, temporal, and visual distances in the final ranking process.

6. Finally, a ranked list of events and their associated information are returned to the user and presented on her device.

3. EXPERIMENTS

We conducted experiments on both Flickr dataset and a real event photo set shot in New York City.

Flickr Dataset In this experiment, we verify the hypothesis that people do take photos at events. And by making use of the taking time and location of photos, we are able to match them to the corresponding events. We built the event database for events in NYC from year 2008 to 2010. Also, we called the Flickr API and downloaded all the photos shot from year 2008 till June 2010. We matched the photos to the events in the event database. Figure 2 shows examples of matched events and photos. The left column details the events and the urls where these events were extracted, and the right column lists the photos taken at the events.

Real Photo set In this section, we test on real photo

Events	Photos
Title: Phantom of Opera Time: 2008-03-15 14:00:00 Location: Majestic Theater Type: Shows http://newyork.metromix.com/browse/events?date_type=today&date_range=2008-03-15%2C2008-03-15	
Title: Playing the Building: An Installation by David Byrne Time: 2008-06-28 12:00:00 - 18:00:00 Location: 10 South Street, Manhattan Type: Exhibitions http://www.nyc.gov/portal/site/nycgov/menuitem.b4175b51da1d74f472ae1852f8089a0/index.jsp?&epi-content=GENERIC&epi-process=generic_process.jsp&beamID=1622248501&view=D=calendar_process_view&pageop=eventbrowser&range=day&selectedDate=6/28/2008&p=100	
Title: Heritage of Pride Parade Time: 2008-06-29 12:00:00 - 18:00:00 Location: Fifth Ave - Greenwich Street - to Christopher St., Manhattan Type: Parades http://www.nyc.gov/portal/site/nycgov/menuitem.b4175b51da1d74f472ae1852f8089a0/index.jsp?&epi-content=GENERIC&epi-process=generic_process.jsp&beamID=1622248501&view=D=calendar_process_view&pageop=eventbrowser&range=day&selectedDate=6/29/2008&p=50	
Title: New York Yankees vs. Boston Red Sox Time: 2009-05-05 19:05:00 Location: Yankee Stadium Type: Sports http://new.york.eventguide.com/days/90505.htm	

Figure 2: Examples of matched events and photos.

sets collected from 4 volunteers living in NYC. We asked each volunteer to hang around on streets in NYC during their spare time, in August and September 2010, and try attending some events that they discovered. They were advised to take as many pictures as possible at the event, and there were no requirements on the subjects of these photos. The ranking result is depicted in Figure 3. The photo column shows a sample of pictures from each photo set, and the result column lists the top 5 ranking result for most pictures in the photo set. The last column provides the ground truth of these events. For event 1, 3, and 4, we correctly returned the information of the corresponding events on the first place in ranked lists. But for event 2, since the exact event was not stored in our database, our system returned a musical event in the Mitzi Newhouse Theater of Lincoln Center, which was a very close match.

Photos	Ranking Result	Ground Truth
	1. Free Music Fridays 2. Target Free Fridays 3. Film Collection at MoMA 4. Drawings Collection at MoMA 5. Contemporary Art from the Collection	Title: Free Music Fridays Time: 09/10/2010 5:30PM - 7:30PM Location: American Folk Museum Type: Museum, Performance
	1. The Grand Manner	Title: Metropolitan Opera's HD Festival Time: 09/01/2010 8:00PM Location: Lincoln Center Type: Performance
	1. River-to-River Festival	Title: River-to-River Festival Time: 08/29/2010 10:00AM Location: Governors Island Type: Festival/Fair
	1. Street Fairs 2. Improv 4 Kids! 3. Tony n Tinas Wedding 4. Chicago 5. The Quantum Eye - Magic Deceptions	Title: Street Fairs Time: 08/21/2010 10:00AM Location: On 6 th Avenue from 42 nd - 56 th Street Type: Festival/Fair

Figure 3: Result on real photo sets.

4. REFERENCES

- [1] J. Allan, V. Lavrenko, D. Malin, and R. Swan. Detections, bounds, and timelines: Umass and tdt-3. In *Proceedings of Topic Detection and Tracking Workshop (TDT-3)*, 2000.
- [2] E. Prud'hommeaux and A. Seaborne. SPARQL Query Language for RDF. Technical report, W3C, 2006.
- [3] A. W. M. Smeulders, S. Member, M. Worrington, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:1349–1380, 2000.