Contextual Topic Model based Image Recommendation System

Lei Liu

HP Labs, Palo Alto, USA

lei.liu2@hp.com

***Abstract*—With the incredibly growing amount of image data uploaded and shared via the internet, recommender systems have become an important necessity to ease users' burden on the information overload. Existing image recommendation systems are designed for discovering the most relevant images with a given query image or short query composed by a few words. However, none of them considers deal with long query, where the query could in any length and potentially contains multiple query topics. To address this problem, we present a contextual topic model based image recommendation system. Compared to using a search engine such as Google Image, our system has the advantage of being able to discern among different topics within a long text query and recommend the most relevant images for each detected topic with semantic "visual words" based relevance.**

I. Introduction

With the rapid development of social media web sites, such as Facebook, Flickr and Youtube, the amount of multimedia data has been uploaded and shared is tremendous and unprecedented. For example, according to the most recent report, there are 1.8 billion images to Facebook, Instagram, Flickr, Snapchat and WhatApp are uploaded every single day. Figure 1 [1] shows the daily upload images on each social web site. With such

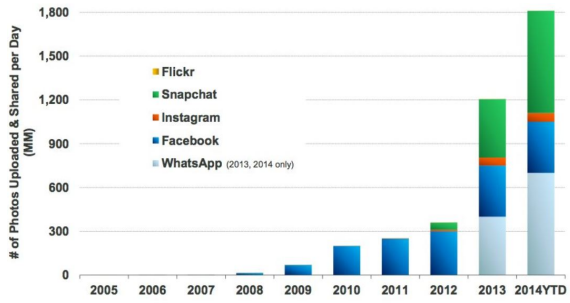

Fig. 1. Daily Number of Photos Uploaded & Shared on Social Platforms

huge amount of image set, it is important to help users quickly find the images that fit their interest, a recommendation system that perform a precise image search is needed. Existing image recommendation systems are designed for short queries that composed by a few words, none of them is designed for long text query, where the query could in any length and potentially contains multiple query topics. One for each query intention, this long text query based image recommendation is critical in many areas. For example, adding illustration images to text for book publishing in education[1][2]. To find images relevant to given long text reading content, a usual practice is to submit the entire text string as a query to an image search engine, like Google Image and Bing Image. However, as existing search engines are designed to only accept a few words as the query, the output from the search engine from a query string will be an error indicating that the query is *too long to process*. The other way is to manually summarize the long query passage to create a query consisting of a few words to find the relevant images. However, this approach is inefficient and may not accurately represent the query passage especially when user do not fully understand the content. Another key disadvantage with current image search or recommendation systems is that although there may be more than one topic underlying the long query content, existing search engines fail to consider this factor and treat all of these concepts as a single topic with which to find the relevant images. In addition, as search engines usually transform the query and candidate resources into bags or vectors of words, the semantic topics underlying the content are totally overlooked. Topics are a better choice for truly understanding both the query and the image.

To address these challenges, we created a contextual topic model based image recommendation system by enabling search with text passages of any length and by recommending a ranked list of images that match the different topics covered within the queried passage. In summary, our system has the following contributions: (1) Our system recommends images for text queries of any length. (2) Our system detects underlying topics from multi-topic content, then recommend image for each topic. (3) Our system introduces a novel semantic "visual words" based image ranking system. Using text content from the web page where the image originated from, we determine "visual words" as the semantic topic features with probabilistic topic modeling technique.

This paper starts with an overview of the system architecture, then describes the features to be demonstrated, and then provides an outline of our planned presentation to WI attendees.

II. Demonstration Features

We address this problem by developing an contextual topic model based image recommendation system. In a nutshell, given query text of any length, our system first detects the underlying topics from the text. For each topic, we generate a set of keywords to represent the meaning of this topic. We

[1]Source: KPCB estimates based on publicly disclosed company data. 2014 YTD data per latest as of 5/14

then recommends a list of images that are relevant for each detected topic. Authors use our system by first submitting text content of any length. These text may be paragraphs, pages or an entire book. The submitted content is fed to **Query Processing** module responsible for preprocessing the query passages, including removing noisy and stop words, stemming, etc. Topics Generation then generates underlying topics from the submitted query content. We represent each topic with set of terms to indicate the concept for each single topic, Topic Compression is designed to remove the redundancy topics that have been detected. **Candidate Image** module fetches relevant images for each extracted topic. We then rank the discovered images based on the user defined criterion in **Image Recommendation** module. Our system can recommend a set of images for any length query content in a ranked list based on "visual words" semantic relevance. Due to the space limit, we only briefly introduce several key components in the paper, but more details are available if requested.

A. Contextual Topic based Image Recommendation

Query Processing: Given a text content as a query, we utilize topic models to discover the abstract topics that underlying the query. Intuitively, provided that a selected text content is about one or more topics, one expects particular words to appear in each topic more or less frequently.

As the number of topics to be generated is given as input to LDA, and this number associated with the queried passage is unknown, it is possible that multiple topics are generated but are associated with similar concepts. In order to remove such redundancy in the topics generated, we propose the idea of topic compression by considering the word distribution of each topic, and then remove duplicate topics if they are discussing the similar concept. To identify if two topics are about similar concepts, we use Pearson correlation to compare the topics.

Image Recommendation: Each topic found from the previous step forms a query that is submitted to candidate images generation module. Specifically, we consider the top n words describing each topic, and these words form a query to be executed over the underlying search engine or image database(we use Bing API in this paper). For each topic query, a set of K candidate images are retrieved. Then we extract image URLs, original URL where the image originated, content from the source webpage, and image size and resolution (height and width) for all the candidate images.

Candidate images can be recommended based on the semantic "visual words" based relevance between query content and content of original page containing the candidate images. This factor is critical to provide a precise image to satisfy user needs. Given that we have K candidate images for each topic in the query, to perform content based image recommendation, we first extract the content from the original page containing candidate images and then generate semantic topic features as "visual words" for each candidate image, then re-rank them based on the content relevance score. To perform this, we present two components: topic feature generator, and relevance discovery and ranking.

For each query topic discovered from the query passage, we have retrieved a set of candidate images by forming a keyword query comprising of the top words that describe a topic. However, traditional search engines match resource to a query based on their term similarity not the semantic "visual words" relevance. Therefore, we need to re-rank the retrieved images based on their topic based semantic similarity.

For this purpose, we treat each set of content of original page containing the candidate images with the original query passage as a content bucket. For each bucket, we generate a set of topics as the semantic "visual words" features with the same hidden topic detection method discussed earlier. provides an illustration of topic feature generator for one content bucket.

We use the topic representations generated for the images and query in each content bucket to re-rank the images of a bucket with respect to the query. Any similarity or distance function could be utilized here, we use cosine similarity in this paper. We can then select the top n ($n < K$) images to show for each query topic discovered from the query passage. More details of topic based content relevance measurement can be accessed in [3].

Users are allowed to customize personal image preferences and filtering out the images that fail to meet her requirements, customization options include image sources, formats, resolutions, size, copyright levels, etc.

III. Demonstration

We have implemented the prototype system using Java. We use Mallet v2.0.6[2] to generate the topics, and in particular the parallel threaded implementation of LDA[4][5].

Our demonstration will start with example content passages with the target to show how our system can recommend images with query text in any length. Then given the query content covers more than one topic, we will show how our system could detect each underlying topic and recommend the most semantic relevant images. Finally, we will start a script to show a comparison analysis between our system and existing image search engine or recommendation system. We will provide interactivity, where participants can interact and experiment with our system over query passages and different customization options.

References

[1] L. Liu, "Semantic topic-based hybrid learning resource recommendation," in *The 24th ACM International Conference on Information and Knowledge Management (CIKM 2015)*, 2015.

[2] J. M. Hailpern, R. Vernica, M. Bullock, U. Chatow, J. Fan, G. Koutrika, J. Liu, L. Liu, S. J. Simske, and S. Wu, "To print or not to print: hybrid learning with METIS learning platform," in *Proceedings of the 7th ACM SIGCHI Symposium on Engineering Interactive Computing Systems, EICS 2015*, 2015, pp. 206–215.

[3] L. Liu, G. Koutrika, and S. Wu, "Learningassistant: A novel learning resource recommendation system," in *International Conference on Data Engineering (ICDE'2015)*, 2015.

[4] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," in *Advances in Neural Information Processing Systems(NIPS'2002)*, 2002.

[5] L. Liu, J. Liu, and S. Wu, "Image discovery and insertion for custom publishing," in *The 9th ACM Conference on Recommender System (RecSys 2015)*, 2015.

[2] http://mallet.cs.umass.edu/download.php