Tag Recommendation System for Photos on Social Websites based on User Preferences and Geo-Location Preferences

Priyanka S K¹ Kusuma S²
¹PG Student ²Assistant Professor
¹,²Department of Information Science
¹,²RNSIT College of Engineering, Bangalore, India

Abstract—Photo tagging is becoming more and more important now-a-days to organize and search large number of photos on social websites. To generate high quality social tags and automatic tag recommendation is the main research topic. In this paper main focus is on the personalized and geo-specific tag recommendation. Consider users and geo-locations have different preferred tags assigned to a photo, a new subspace learning method is proposed to individually discover both the preferences. The goal is learn unified space which is shared by visual domain and textual domain to make visual features and textual features comparable. Visual feature is considered to be lower level representation on semantics than textual feature. Additionally intermediate space is introduced for the visual space and expecting it to have consistent local structure with text space. Unified space is mapped from the textual space and the intermediate space respectively. When an untagged photo with its geo-location is given based on the nearest neighboring search user preferred and geo-location-specific tags are found in the corresponding unified space. Then combine these obtained tags and the visual appearance of the photo to find semantically and visually similar photos, among which the most frequently used tags are suggested to the user and user is allowed to select based on his preference.

Key words: Geo-location preference, subspace leaning, unified space, user preference

I. INTRODUCTION

In modern world uploading and sharing of photo has become very common and the number of photos increases on social websites or on the mobile devices. It is all due to the popularization of GPS-enabled camera devices and mobile phones. So the increased number of photos should be organized in a systematic way so that it becomes easy to search photos. Personal photos are with rich context like tags, geo-locations, and visual attributes like colors and textures [1].

There are many photo-sharing websites, such as Instagram, Corbis, Picasa, and Flicker which will allow millions of users to share and upload their personal multimedia data by their smartphones or other internet access devices. As a result the number of community-contributed photo increases rapidly on both personal devices or on social websites. It is very challenging task to exploit the context data for multimedia applications such as annotation, retrieval or recommendation.

Assigning proper tags to photos is the very important task. Users spend lot of time to organize their photo albums geographically by describing photos with tags related to locations where they were taken, example for this is in fig. 1 the Zone tag[2] a camera phone photo capture and annotation tool which uploads images to flicker, tags are displayed on the right. Visitors to the page can click on a tag, or on the earth icon next to it, to see photos with this tag from this user or from the entire Flicker community, respectively. Fully manual tag assignment is very lengthy process and it is impractical due to large number of photos and limited screen size of mobile phones as result many tag recommendation methods[3],[4],[5] are proposed to suggest relevant tags to the newly uploaded photos and facilitate user to select tags based on their preference. This will not only ease the burden of user to share and upload photos on social websites but it will help to organize and tag their photos on mobile devices. The user and geo-location have different preference towards assigning a tag to a photo. Consider fig 2 it represents the user preference and geo-location preference. Fig 2(a) represents user preference, when a same photo is tagged by two different users have different preference towards assigning a tag to a photo, some users may consider natural landscape and some may consider architecture. Two different users may give two different tags to a same photo. So it is necessary to consider user preference. Fig 2(b) represents the geo-location preference, when a two visually similar photo tagged by two different users both may give different tags if the geo-location preference is consider but if it is not considered then both the photos which are visually similar may get same tag. So a new subspace learning method is proposed which considers both the preferences and suggests tags to a user. It has a unified space which is mapped from the intermediate space and text space. Here it is considered that visual feature is considered to be lower level on representation than text space. Intermediate space is introduced to alleviate semantic gap between text space and visual space.

Fig. 1: A Zone Tag photo page on Flicker
In personalized tag recommendation methods it personalizes the generic method but here the geo-location preference is not considered. In [7], tag co-occurrence for photos is calculated using tags appearing both in the tagging history of a user and in Flicker website, and used to generate recommended tags. Web browsing behavior of a user is exploited to suggest the tags not only to be added to but also to be deleted from the original tags of a photo in Flickr. In [15], image tag recommendation is formulated as a maximum a posteriori problem using a visual folksonomy. With the assumption that favorite images and their associated tags indicate the visual and topical interests of a user, personalized favorite images and their context are used to perform personalized image tag recommendation. A simple personalized image annotation method is designed, which simply annotates an untagged image with the most frequent tags in the user tagging history. However the above methods only focus on photos, users and tags but ignore the geographical information of photos. Some other personalized tag recommendation methods generate candidate tags by exploiting geo-tags [14]. In [14], new photos are tagged using

![Image](image-url)

**Fig. 2:** An illustration of preference of users and geo-locations. (a): Users have their specific preferences towards tags for similar photos; (b): Two visually similar photos corresponding to different geo-locations are tagged with different tags.

## II. RELATED WORK

There are two types of tag recommendation methods proposed previously i.e generic tag recommendation methods and personalized tag recommendation methods.

In generic tag recommendation methods [3], [5], are to predict the same list of tags for the same photo, i.e., in this user factor is not considered. Chen et al. [5] proposed an automatic tag recommendation approach that directly predicts the suitable tags with models which are learned from training data. Shen et al. [3] proposed a multi-task structured SVM algorithm to consider both the inter-object correlations and the loosely-tagged images. Photos are described purely based on image visual content. For a photo, first its top-neighboring images are found from the community photos set and select most frequent tags. In [6], two approaches, based on Poisson Mixture Models and Gaussian process respectively, are proposed to make effective and efficient tag recommendations. In [7], tag concepts derived based on tag co-occurrence pairs are indexed as textual documents. The candidate tags associated with the matching concepts, which are retrieved with the query of user-given tags of an image, are recommended. There are some works focusing on tagging images by exploiting geo-tags [9], [13]. A typical approach as introduced by Moxley et al. [9] and Clean et al. [10] is to annotate a given image by constrained nearest neighbor ( - NN) voting, where the visual neighbors are retrieved from the geo region of the given image. The fundamental idea in [11] is to learn tag semantics, i.e., categorize tags as places, landmarks, and visual descriptors, in order to post-filter tag the results of tag suggestion. Silva et al. [12] annotated geo-referenced photos with descriptive tags by exploring the redundancy over the large volume of annotations available at online repositories with other geo-referenced photos. Geo context is fused with visual concept detection in a concept-dependent manner to improve visual search in [13]. However, the above methods ignore the user preference and suggest same tags to visually similar photos of different users. Different from them, a learning algorithm is proposed to effectively uncover user preference from her tagging history.

![Image](image-url)

**Fig. 3:** The framework of the proposed personalized tag recommendation algorithm. The square and triangle denote photo and tag, respectively. The circle is corresponding to a new photo. The selected related tags are with red boundary.

Users’ own vocabularies by accumulating votes from the candidate images, which are selected in term of three factors: visual features, geographical coordinates and image taken time. Different from the above work, a subspace learning approach is proposed to individually uncover user preference by exploiting user’s tagging history and geo-location preference by exploiting the geographic information of photos, and then jointly explore the learned subspaces assisted with the search scheme to recommend user preferred tags to a photo.

## III. PROPOSED METHOD

This work mainly focuses on how to ease personalized photo tagging process by exploiting the community-contributed multimedia data with rich contextual information. The proposed framework is illustrated in Fig. 3, which contains two parts: the offline process and online processes. The offline process consists of three subdivisions:
Tag Recommendation System for Photos on Social Websites based on User Preferences and Geo-Location Preferences

(IJSRD/Vol. 3/Issue 3/2015/257)

data collection, user preference learning and geo-location preference learning. For data collection, collected a vast amount of photos with their tags, users, geo-locations and some relevant text information from Flickr. With the collected resources, organize the photos according to different users and geo-locations individually. Given a collection corresponding to a user or a geo-location a new subspace learning approach is proposed to uncover the user’s preference and the geo-location’s preference towards tags. Here main goal is to find a unified space for the both visual and textual domains, in which the visual features and the tagging information are comparable, i.e., the correlations between the both heterogeneous representations can be directly constructed.

In the online module, given a new photo with a specific user and a specific geo-location, first find its top-ranked neighboring tags in the user-specific unified space and the geo-specific unified space individually, and combine the both sets of neighboring tags to generate the initial tags, by which semantically relevant photos are chosen from the community-contributed photo set. Then visually similar photos are found by implementing content-based photo retrieval from these semantically relevant photos. Finally, the most frequent tags in the semantically and visually related photos are recommended to the user.

Fig. 4: The illustration of the proposed subspace learning algorithm. The dashed line represents the analogous relationship between these two spaces.

IV. SUBSPACE LEARNING ALGORITHM

The personalized tag recommendation task is addressed with the help of overwhelming community-contributed information such a user, tag and geo-location. There exists a semantic gap making it challenging. The community-contributed photos are associated with rich text information which reduces the semantic gap. The text space and visual space have different structures. To address this problem, it is important and necessary to discover a common structure to link them. The visual and text representation should be consistent. This motivates an approach to discover the unified space, in this the text features and visual features should be same. However, compared with the tag information, visual feature is a much lower level representation on semantics. To reduce the semantic gap, adopted a progressive way as shown in Fig. 4 to first map the visual features into an intermediate space, which is analogous to the tag space, and then perform the transform from the intermediate one to the unified space.

Let \( u \) be a user for whom tags has to be suggested and \( g \) be a geo-location. Main goal is to find latent spaces of user \( u \) and geo-location \( g \) individually.

Let \( X_{u,te} = [x_1, x_2, \ldots, x_n] \) be a set of photos of user \( u \) which are already tagged and \( X_{ute} \) is a set of untagged photos to which tags has to be recommended. \( Y_u = [y_1, y_2, \ldots, y_n] \in \mathbb{R}^{m \times n} \) is the text representation matrix of photo tagged by user \( u \), where \((Y_u)_{ij}\) is used to represent the relevance between tag \( w_i \) and photo \( X_j \) assigned by user \( u \).

With the transformation matrices \( W \) and \( V \) can easily obtain the latent representation \( G = [g_1, g_2, \ldots, g_n] \in \mathbb{R}^{q \times n} \) and \( U = [u_1, u_2, \ldots, u_n] \in \mathbb{R}^{p \times n} \) of the original visual features in the intermediate space and the unified space, respectively.

\[
G = WX \quad (1)
\]
\[
U = VG = VWX \quad (2)
\]

Similarly, the latent representations \( Z = [Z_1, Z_2, \ldots, Z_p] \in \mathbb{R}^{p \times n} \) of the original text features are obtained by

\[
Z = TY \quad (3)
\]

The unified space is expected to connect the visual and text structure.

\[
U = Z \Rightarrow VWX = TY \quad (4)
\]

V. EXPERIMENTAL RESULTS

Whenever a new photo with its geo-location is given fig.4 represents the user uploading new image, based on the nearest neighbor search the relevant tags are searched. Those tags are combined and searched in the website to get the most frequent and semantically relevant tags. Those tags are recommended to the user fig.6 represents the tag suggestion to the user and user is allowed to select based on his preference.
VI. CONCLUSION

As the photo tagging is the important concept in today’s world due to popularization of GPS-enabled camera and mobile phones. People take millions of photos and share with others, to make this task easier subspace learning method is proposed to consider user preference and geo-location preference and suggest tags to the newly updated photo and allow user to select tag based on his preference. This helps users to organize, share and search photos from their mobile phone or on social websites. In future work some new social applications can be developed such as personalized product recommendation; geo-location based travelling suggestion, personalized geo-specific news report.

REFERENCES