doi:10.21311/001.39.7.38

Ranking Social Image Tags via Neighbor Voting and Random Walk

Jing Wang*
School of Telecommunications Engineering, Beijing Polytechnic, Beijing 100106, China

Dawei Chu
Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China
**Corresponding author(E-mail: 25953111@qq.com)*

Abstract
We have witnessed the popularity of photo sharing websites for people to share their own experiences by uploading personal photos. However, due the complexity of social tagging behaviors, user-supplied tags may be imprecise, incomplete, and even irrelevant to the image. Therefore, in this paper, we propose a novel social image tag ranking algorithm to sort tags according to their relevance to the image, which is of great importance for social image retrieval and semantic mining. Firstly, for a seed image, visually similar neighbors are collected from the social image dataset. Secondly, visually similar neighbors vote for each tag of the seed image, and the voting powers are weighted by both visual similarity and tag relevance. Thirdly, based on the neighbor voting results, ranking order of all tags is determined by implementing the random walk on tag graph. To testify the performance of the proposed algorithm, we collect social images from Flickr to construct a dataset which is made up of ten categories. Experimental results demonstrate that the proposed algorithm outperforms exiting methods, and it can effectively boost the performance of social image retrieval as well.

Key words: Social Image, Tag Ranking, Neighbor Voting, Tag Graph, Random Walk.

1. INTRODUCTION

With the rapid development of Web 2.0 platform, there are many photo sharing websites with large-scale multimedia data available online to encourage users to share their personal photos, such as Flickr, Photosig, Facebook, Photobucket, MySpace and zoomr, which are also named as social media (Ballan, 2015). Due to the popularity of multimedia devices and services, and the availability of cheap storage devices and mobile phones, more and more people can easily access the Internet (Zhang, 2016). Therefore, massive user-supplied images are uploaded to photo sharing websites rapidly. These images are frequently shared on online social network websites. Statistical data show that until June 2012, Flickr has been uploaded nearly seven billion images, with more than 2500 new images in each minute (Qian, 2014; Zhang, 2013). Furthermore, these Web 2.0 based photo sharing websites allow users not only upload their photos, but also provide descriptive tags and comments for their photos (Liu, 2011).

Photo sharing websites motivate and allow users to freely provide a set of tags to describe the uploaded images according to their will (Li, 2014). Tags can also be regarded as indexing keywords, which may greatly facilitate social image search and other applications. The massive user-supplied image tags on the photo sharing websites can significantly promote the performance of image retrieval system (Xu, 2014). Unfortunately, because of the subjectivity and diversity of social tagging behaviors, user-supplied tags are often imprecise and incomplete, and many tags are irrelevant to the host image. Therefore, it can be observed that a gap locates between these user-supplied tags and the visual contents of images, and this problem greatly influences the performance of social image search and mining (Sun, 2011).

We aim to effectively promote quality of user-supplied tags by tags ranking process. The main idea of this paper lies in that if a tag has frequently been annotated for its visually similar neighbor images, it is likely to reflect the visual contents of the host image. Inspired by this intuition, we propose a novel social image tag ranking algorithm utilizing a weighted neighbor voting policy.

The main innovations of this paper lie in the following aspects:

1) A weighted neighbor voting scheme is proposed, in which both visual similarity and tag relevance are exploited. That is to say, both visual contents and tag semantics influence the proposed neighbor voting process.

2) To effectively mine the correlations between seed image and its visual neighbors, a tag graph is constructed to integrate all information of neighbors and the seed image.

3) The random walk is executed on the tag graph to optimize tag relevance score by fully mining correlations between tags and visual features of different social images provided by different users.

The rest of the paper is organized as follows. We review related research about social image tag recommendation and tag ranking in Section 2. In Section 3, we provide the overview of the social image tags

ranking problem. Section 4 proposes the social image tag ranking algorithm by neighbor voting and random walk. To prove the effectiveness of our proposed algorithm, experimental results are given in section 5. The last section concludes the whole paper and points out future works.

2. RELATED WORKS

With the boom of social multimedia data available online, such as YouTube, Flickr and Photosig, many researchers have concentrated on how to promote the performance of social image search. Among them, a great deal of works aim to refine and enhance social image tags. Li et al. proposed a comparative survey on image tag assignment, refinement, and retrieval, and several crucial issues about social image retrieval have been discussed in this paper. This survey proposed a two-dimensional taxonomy to organize the growing literature, understand the ingredients of the main works, clarify relationship and difference, and discover their advantages and disadvantages (Li, 2016).

In order to suggest more accurate tags for social images, several researchers have focused on the tag recommendation problem. Sun et al. proposed a new tag recommendation algorithm using the user-given tags associated with images, and each candidate tag is generated from the collective knowledge (Sun, 2011). Lee et al. regarded the image tag recommendation problem as a maximum a posteriori problem exploiting a visual folksonomy (Lee, 2010). Eom et al. proposed a personalized image tag recommendation algorithm exploiting favorite image contexts, and then this work integrated tag statistics and visual similarity together to assign tags to images (Eom, 2011). To achieve tag recommendation with diversity, Cui et al. proposed a new method to recommend tags for images considering both the relevance and semantic diversity (Cui, 2013). Sigurbjörnsson et al. proposed a tag recommendation approach based on tag concurrence analysis (Sigurbjörnsson , 2008).

Apart from the above research, social image tag ranking is of great importance in social image retrieval, because only high quality tags can effectively enhance the accuracy of social image retrieval. Typical social image ranking methods are listed as follows.

Feng et al. proposed a novel method to integrate the strength of tag ranking with the power of matrix recovery, and this paper aggregated the prediction models for various tags into a matrix, and then converted the tag ranking problem to matrix recovery (Feng, 2015).

Tang et al. aimed at properly ranking tags via propagating relevance over community-contributed images and related tags. In this work, the salient region sub-images are extracted using the Itti model, and then establish two graphs with the whole image and the salient region sub-image. Afterwards, the graph-based model is applied to compute the relevance score in terms of the relevance propagation (Tang, 2015).

He et al. proposed an image tag-ranking model by exploiting the ranked social image tag lists as the supervision information. In particular, each ranked tag list is decomposed to several image/tag pairs, and the intrinsic ranking structures can be obtained by pairwise supervision. Furthermore, using the pairwise supervision together with the unsupervised structural information, the proposed method is able to capture the tag ranking list by using the tag relevance information (He, 2015).

Jeong et al. developed an image tag ranking system named as i-TagRanker, which was designed to use the semantic relationships between tags for re-ranking the tags in terms of the relevance with an image. Particularly, two steps are contained in this system, that is, tag propagation and tag ranking (Jeong , 2013).

Liu et al. presented a tag ranking method, which ranked the tags corresponding to a given image in terms of the relevance to visual contents. This work calculated the initial relevance scores for the tags with probability density estimation in advance, and then executed random walk process on a tag graph to re-computere the relevance scores (Liu, 2009).

Xiao et al. propose a coupled probability transition method to calculate the text-visual group relevance from the observed data and utilize it to forecast tag relevance for a new query image. Particularly, in this research the tag-visual group relevance is obtained with mutual reinforcement in both visual space and semantic space (Xiao, 2012).

Zhang et al. proposed a social image tag ranking framework with three modules: 1) low-resolution social images are established from the compressed image data, 2) visual words are generated from SIFT descriptors which are obtained in low-resolution social image, and 3) the neighbor voting policy is used to sort the social image tags (Zhang, 2015).

Sun et al. proposed a tags ranking algorithm according to the relevance to visual contents. In particularly, initial relevance values of the tags about a given image's visual contents are obtained based on the Bayesian framework, in which a fused visual similarity is estimated by integrating both global and local visual similarities (Sun, 2013).

Different from the above works, in this paper, we aim to tackle the problem of social image tag ranking by effectively mining relationships between user supplied tags and social image visual contents.

3. OVERVIEW OF THE SOCIAL IMAGE TAGS RANKING PROBLEM

3.1. Social Image Description

Personal photo sharing website, such as Flickr, allows users to freely provide tags for their uploaded photos. Fig. 1 shows an example of a social image in Flickr with user-supplied tags[1].

Figure 1. Example of an image in Flickr with user-supplied tags

As is shown in Fig.1, a photo which is taken at a beach is provided, and eight tags are supplied for this photo by the author. We can see that most of the tags can reflect semantics of the image, however, a noisy tag (e.g. Snow) is also in the tag list. The main task of this paper lies in two aspects, that is, 1) Pruning noisy tags, and 2) Ranking tags according to the relevance to the image.

3.2. Overview of the Tag Ranking Problem

Suppose that social image dataset is denoted as Θ, and tag dictionary is represented as T. For a social image $I \in \Theta$, and a tag $t \in T$, we define $R(t, I)$ as the tag relevance estimation function. We want to design a tag relevance estimation function which satisfies the following two conditions:

Condition 1: (Ranking social image) Suppose that there are two social images $I_1, I_2 \in \Theta$ and a tag $t \in T$, if the tag t is more relevant to image I_1 than I_2, then the following equation is satisfied.

$$R(t, I_1) > R(t, I_2)$$

(1)

Condition 2: (Ranking social image tag) Suppose that there are two social images $t_1, t_2 \in T$ and a social image $I \in \Theta$, if tag t_1 is more suitable to be tagged for image I than t_2, then the following equation is satisfied.

$$R(t_1, I) > R(t_2, I)$$

(2)

As shown in Fig.2, the framework of the social image tags ranking system is given. In this system, the social image dataset should be constructed in advance, from which visually similar neighbors of the seed image are collected. Afterwards, all visually similar neighbors vote for tags of the seed image. Furthermore, to promote the voting powers of important neighbors, voting score is weighted by both visual similarity and tag relevance. After the neighbor voting process, random walk is carried out on the tag graph, and then the final tag list can be obtained according to tag relevance scores.

3.3 Visual Content Description by Bag of Visual Words Model

Bag of visual Words (BoW) model has been widely used in computer vision, such as video event detection, salient object recognition, image segmentation, multimedia information retrieval, and so on. Bag of words model is inspired by test process of the natural language processing area, in which documents are organized as a

[1] https://www.flickr.com/photos/parismadrid/7615372290/

bag of textual words from a fixed vocabulary dictionary (Kim, 2015). In the bag of visual words model, an image is represented as a visual document including distinctive basic visual elements (Karakasis, 2015).

BoW model assumes that each visual word is not dependent to others, and it is constructed by four parts, including: 1) Local feature extraction: Extracting image descriptors and representing images as a set of local features computed from image patches; 2) Codebook construction: Codebook is used to represent an images as a set of local features; 3) Image representation: After visual feature clustering, the cluster centroids are used to construct the codebook. Finally, an image is described as the visual word histogram from the codebook. The BoW model construction process is illustrated in Fig. 3.

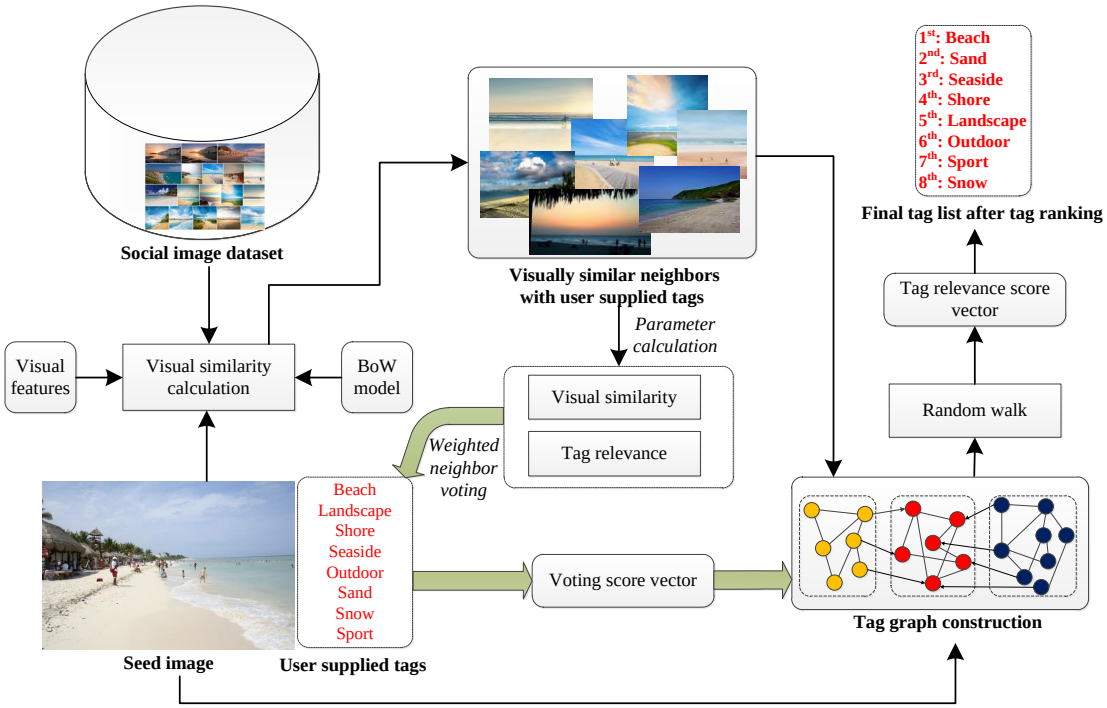

Figure 2. Framework of the social image tags ranking system

SIFT descriptor is used as the visual feature in our work, SIFT is able to detect extrema point in the difference of Gaussian multi-scale space, and constructs a 128D vector (Wang, 2016; Zhao 2016).

To detect SIFT descriptors, the scale space is defined as follows.

$$D(x,y,\delta)=\left(C(x,y,k\delta)-C(x,y,\delta)*I(x,y)\right)=S(x,y,k\delta)-S(x,y,\delta)$$

(3)

where $S(x,y,\delta)$ refers to the scale space function of a social image, which is generated from the convolution of $C(x,y,\delta)$, symbol $*$ refers to a convolution operation, and δ denotes a scale factor.

Traditional codebook construction methods utilize K-means algorithm to cluster the SIFT descriptor. However, disadvantages of these methods are that the seeds of K-means algorithm are random, and there is great randomness in the codebook establishment. Moreover, traditional method has not considered relationships between various SIFT descriptors.

We use Gaussian mixture model (denoted as GMM) to establish the codebook, in which each Gaussian component is regarded as a visual word.

Each Gaussian represents a visual word v_i, and the codebook is defined as $V=\{v_i\}, i\in\{1,2,\cdots,N\}$. σ means the set of parameters of a Gaussian mixture model. Parameter η is equal to $\{\omega_i,\mu_i,\Sigma_i\}, i\in\{1,2,\cdots,N\}$,where

$\omega_i, \mu_i, \Sigma_i$ and N denote the weight, mean vector, covariance matrix of the ith Gaussian, and number of Gaussians respectively. We suppose that h is the hidden mixture variable with respect to an observation x, and $p(x|\eta) = \sum_{i=1}^{N} \omega_i \cdot p(x|h=i,\eta)$ is satisfied.

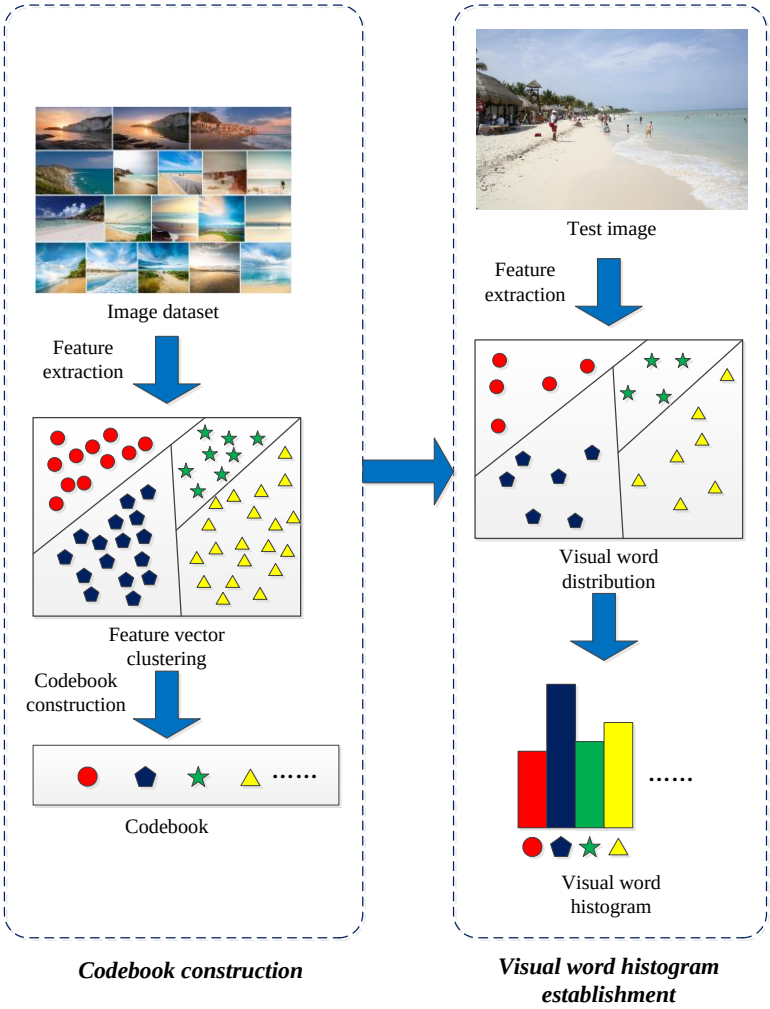

Figure 3. Process of the codebook construction and visual word histogram establishment

After constructing codebook of the visual words, next task is to utilize visual word histogram to represent visual features of images, and term frequency is exploited in our work.

$$TF_{i,j} = \frac{N_{ij}}{\sum_{k=1}^{M} N_{kj}} \tag{4}$$

where N_{ij} refers to the times of visual word w_i appears in the image I_j, and M means the total number of images in the training dataset. Thus, the image I_i is represented as a visual word histogram H_i, and similarity of image I_i and I_j is calculated by estimating the distance between two histograms as follows.

$$Sim(I_i, I_j) = Dis(H_i, H_j) \tag{5}$$

4. TAG RANKING BY NEIGHBOR VOTING AND RANDOM WALK

Tag ranking algorithm is designed in order to calculate the relevance between visual contents of images and semantics of tags, and then sorts tags according to relevance score. To better describe the relationship between visual feature and tag, different visual neighbors should be set different weights. Intuitively, different neighbors may have different voting powers, and the neighbor which is more similar to the seed image has stronger voting power. Therefore, the voting score is weighted by visual similarity and tag relevance. For the seed image I_j, the voting score of tag t_i by its visual neighbor I_k is calculated as follows.

$$Vote\left(t_i, I_j, I_k\right) = Sim\left(I_j, I_k\right) \cdot relevance\left(t_i, T\left(I_k\right)\right) \tag{6}$$

where $relevance\left(t_i, T\left(I_k\right)\right)$ denotes the relevance between tag t_i and tag set $T\left(I_k\right)$ of image I_k. $relevance\left(t_i, T\left(I_k\right)\right)$ is calculated by averaging semantic similarity between tag t_i and each tag of $T\left(I_k\right)$.

$$relevance\left(t_i, T\left(I_k\right)\right) = \frac{1}{\left|T\left(I_k\right)\right|} \cdot \sum_{t_j \in T(I_k)} r\left(t_i, t_j\right)$$

(7)

where $r\left(t_i, t_j\right)$ is the tag correlation between tag t_i and t_j. Inspired by Google distance (Cilibrasi , 2007), in our work, semantic distance $d\left(t_i, t_j\right)$ is estimated as follows.

$$d\left(t_i, t_j\right) = \frac{\max\left(\log N\left(t_i\right), \log N\left(t_j\right)\right) - \log N\left(t_i, t_j\right)}{\log \Gamma - \min\left(\log N\left(t_i\right), \log N\left(t_j\right)\right)}$$

(8)

where $N\left(t_i\right)$ and $N\left(t_j\right)$ denote the numbers of images which are tagged by t_i and t_j respectively, and $N\left(t_i, t_j\right)$ refers to the numbers of images which are tagged by both t_i and t_j. Moreover, Γ is the total images in Google image searching engine. Afterwards, the tag correlation between tag t_i and t_j is computed as follows.

$$r\left(t_i, t_j\right) = \exp\left(\frac{-d\left(t_i, t_j\right)^2}{\sigma_t^2}\right) \tag{9}$$

where σ_t means the empirically set.

However, the above voting score computation method has not fully considered correlations between different tags belonged to different social images. In this work, we try to mine correlations between tags and visual features of images by fully utilizing a group of tags which are provided by various users.

To achieve this goal, we construct a tag graph based tag correlation mining model using random walk policy. In the tag graph, user supplied tags are regarded as graph nodes, and tags are obtained from both seed image and its visual neighbors. Moreover, tag correlation is represented as edge weight. To reduce the computation cost, the edge is deleted from the tag graph if its weight is lower than a threshold. As is shown in Fig. 4, a tag graph corresponding to a seed image and two visual neighbors is given as follows.

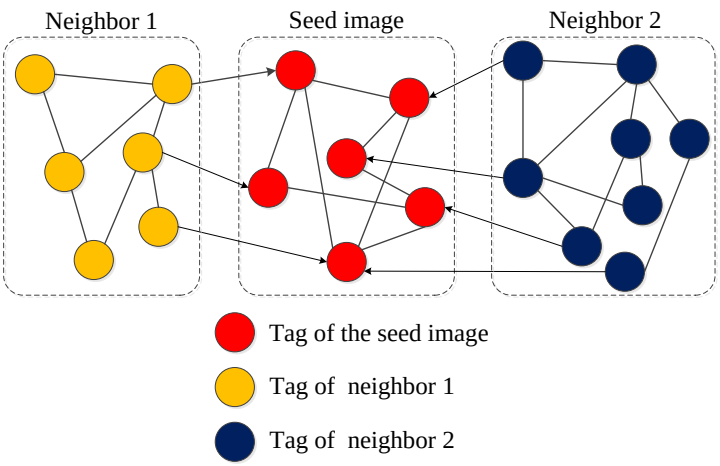

Figure 4. An example of tag graph with two visual neighbors.

Suppose that the tag graph $G=(V,E)$, where V,E denote the nodes set and edges set respectively. Particularly, $V=V_S+V_N$, and V_S,V_N refers to tags of the seed image and its neighbors respectively. To prevent tags of seed image affecting tags of neighbors, the edge connecting seed image and its neighbor is directed, and other edges are undirected.

The transition probability matrix of tag graph G is represented as M, and the element m_{ij} of it is the probability of the transition from tag t_i to t_j. The value of m_{ij} is computed as follows.

$$m_{ij}=\frac{r(t_i,t_j)}{\sum_{k=1}^{|V|} r(t_i,t_k)} \tag{10}$$

The transition probability matrix M is defined as follows.

$$M=\begin{bmatrix} M_S & M_{SN} \\ 0 & M_N \end{bmatrix} \tag{11}$$

where M_S and M_N refer to square matrix describing the transition probability of V_S and V_N, and M_{SN} refers to the matrix which saves the transition probability from V_N to V_S.

Algorithm 1: Tag ranking by neighbor voting and random walk

Input: The seed image I with the user-supplied tag set $T=\{t_1,t_2,\cdots,t_N\}$, Nearest visual neighbors $\{I_1,I_2,\cdots,I_K\}$

Output: Tag list $\hat{T}$

1) Input each user-supplied tag to Wikipedia, the tags which cannot be matched are pruned, and the remaining tags are represented as $\{t_1,t_2,\cdots,t_M\}$.

2) Sum all voting scores from similar neighbors as follows.

$$Vote_i=\frac{1}{K}\cdot\sum_{k=1}^{K} Vote(t_i,I,I_k) \tag{12}$$

where $Vote_i$ is all neighbors voting power for tag t_i.

3) Construct the voting score vector $Vote=(Vote_1,Vote_2,\cdots,Vote_M)$

4) Define the tag relevance score vector as $R=\{r_1,r_2,\cdots,r_M\}$, where r_i is the relevance score of tag t_i.

5) Set the initial state $R^{(0)}=Vote$.

6) The tag relevance score in the lth step is calculated as follows.

$$R^{(l)} = (\delta M)^l R^{(0)} + (1-\delta)(E-\delta M)^{-1}\left(E-(\delta M)^l\right)Vote \tag{13}$$

7) The final state of tag relevance score vector is as follows.

$$R = \lim_{l\to\infty} R^{(l)} = (1-\delta)(E-\delta M)^{-1} Vote \tag{14}$$

8) Sort tags $\{t_1, t_2, \cdots, t_M\}$ to construct the tag list $\tilde{T}$ according to the relevance score vector R in descending order.

where parameter δ (shown in Eq.14) represents the influence of initial state of voting scores and transition matrix on final tag ranking results. In the following experiment, δ is set to be 0.5.

5. EXPERIMENT

In this section, we test the performance of our proposed social image tags ranking algorithm, and 50000 social images are collected from Flickr with ten categories, including 1) cat, 2) automobile, 3) mountain, 4) water, 5) sea, 6) bird, 7) tree, 8) sunset, 9) flower and 10) sky. The above ten tag categories are the most popular tags in Flickr. Afterwards, we put these tags to Flickr to execute the tag based social image searching process according to interestingness, and there are 5000 images in each category.

Performance evaluation metrics used in this paper are NDCG and MAP. Mean average precision (MAP) is designed using average precision (AP), and it denotes mean precision scores after each retrieved relevant item for a query q.

$$MAP = \frac{1}{|Q|} \cdot \sum_{q\in Q} AP(q)$$

(15)

where q is a single query, and Q denotes the query set.

NDCG (normalized discounted cumulative gain) is used to test the tag ranking performance. Each tag in a social image is classified to five grades, that is, 1) Very relevant (five scores), 2) Relevant (four scores), 3) Moderately relevant (three scores), 4) Weakly (two scores), and 5) Not relevant (one score). Assume that a social image is tagged by tag set $\{t_1, t_2, \cdots, t_n\}$, and then NDCG is calculated as follows.

$$N_n = \xi_n \cdot \sum_{i=1}^{n} \frac{2^{r(i)} - 1}{\log(1+i)}$$

(16)

where $r(i)$ is the relevance degree of the tag i, and ξ_n means the normalization constant.

5.1. Influence of Number of Neighbors

The main idea of this work is to exploit the neighbor voting policy in social image tag ranking problem, hence, we should study on how the tag ranking accuracy is influenced by different number of neighbors. As is shown in Eq. 6, the voting score is calculated by integrating two factors: 1) visual similarity, and 2) tag relevance. In the following, we compare the MAP value of our proposed algorithm and other voting methods, and experimental results are illustrated in Fig. 5.

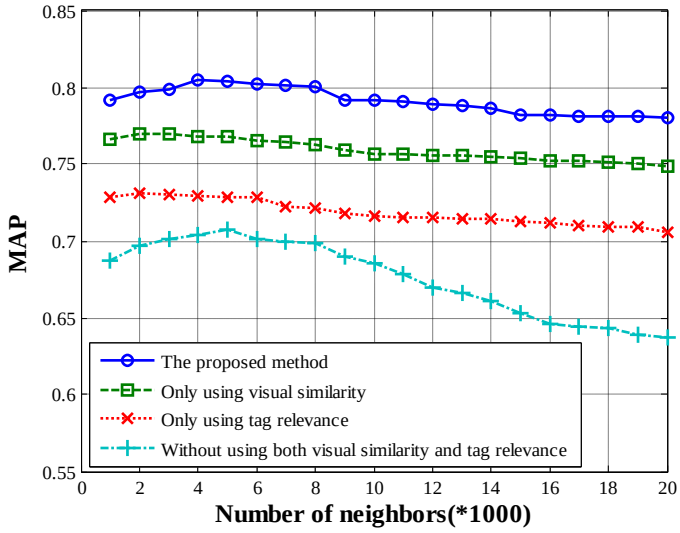

Figure 5. MAP for different methods

In Fig. 5, the neighbor voting method without both visual similarity and tag relevance supposes that all neighbors have equal voting power, and it is regarded as the baseline. Experimental results of Fig. 5 demonstrate that the proposed method can significantly promote MAP value for tags ranking by effectively integrating visual contents and tag semantics of social images. Moreover, it can also be observed that visual contents are more important than tag semantics in the neighbor voting process. As the largest MAP value is obtained when the number of neighbors is equal to 4000, in the following experiment, number of neighbors is set to 4000. Compared with the baseline, performance improvements of other methods are given as follows.

Fig. 6 shows that the proposed method outperforms than neighbor voting schemes which only utilize visual contents and tag semantics. More precisely, our proposed method reaches more than 20% enhancement when the number of neighbors is larger than 16000.

In general, the MAP value of the proposed method significantly performs better than others, the worst performance of our method is still better than the best performance of other methods. Moreover, the rate of performance decline of our method is relatively slow, therefore, our method is insensitive to the number of neighbors and the robustness of our method is better than others.

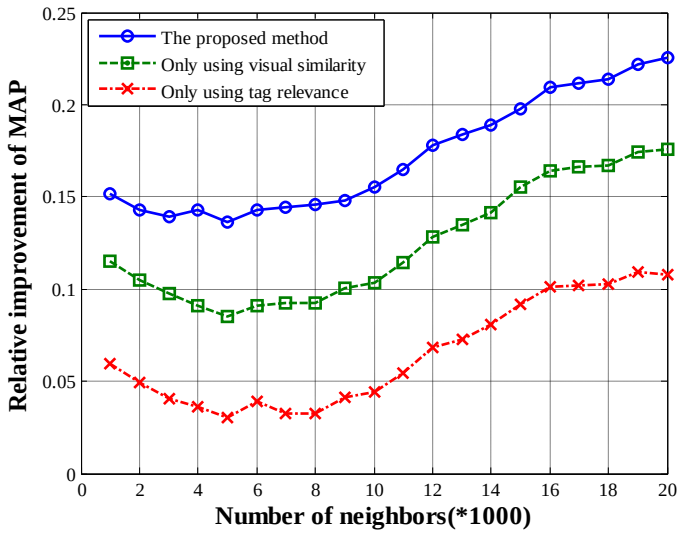

Figure 6. Relative improvement of MAP for different methods

5.2. Tag Ranking Evaluation

To testify tag ranking performance of the proposed method, we use NDCG metric to compare our proposed method and other tag ranking methods:

(1) Neighbor voting based on Low-level Visual feature (Li 2009): The neighbor voting based on low level visual features (Li, 2009).

(2) Tag ranking by probabilistic density estimation and random walk strategy (denoted as Liu 2009) (Liu, 2009).

(3) Tag ranking using visual words in compressed domain (denoted as Zhang 2015) (Zhang, 2015).

The original tags list is regarded as the baseline, and our method is compared with the above related works using average NDCG (experimental results are given in Fig. 7).

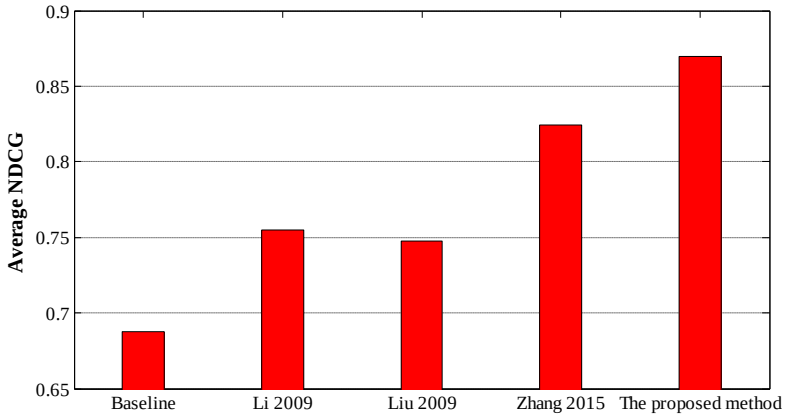

Figure 7. Average NDCG for different methods

From Fig. 7, it can be observed that 1) all methods perform better than original tags, and 2) the proposed method outperforms other methods. Because the proposed method is able to fully take advantages of visual features and effectively mine the tag correlations between different social images.

Specifically, we illustrate the average NDCG for each social image category in this dataset, and experimental results are demonstrated in Fig. 8.

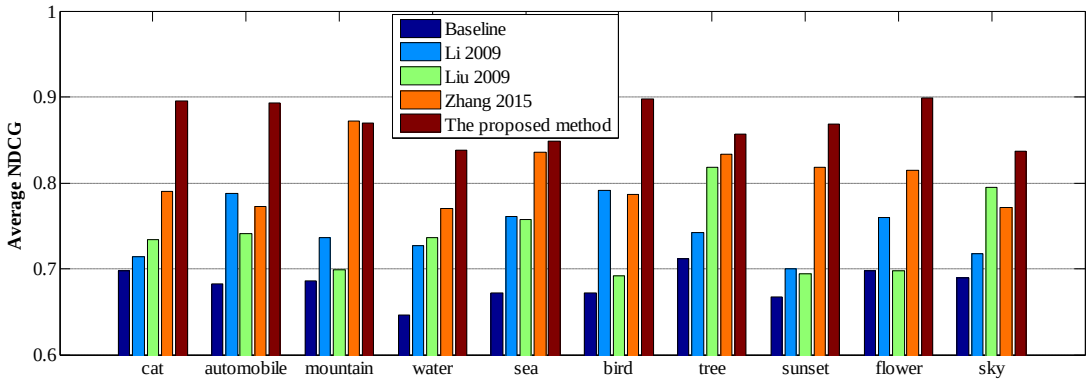

Figure 8. Average NDCG for different social image categories

Fig. 8 further demonstrates the tag ranking results for different image categories. We can see that our proposed method performs better than others in almost all image categories, except for “mountain”. In particular, our proposed method achieve higher NDCD value in categories of cat, automobile, bird, and flower category, and performs worse in water, and sky. The reasons lie in that the proposed method uses local features to describe visual contents of images. Hence, it can achieve better performance when the images contain salient objects, such as cat, automobile, bird, and flower.

To clearly describe the capability of our proposed method, several examples of social image tag ranking results are given in Fig. 9.

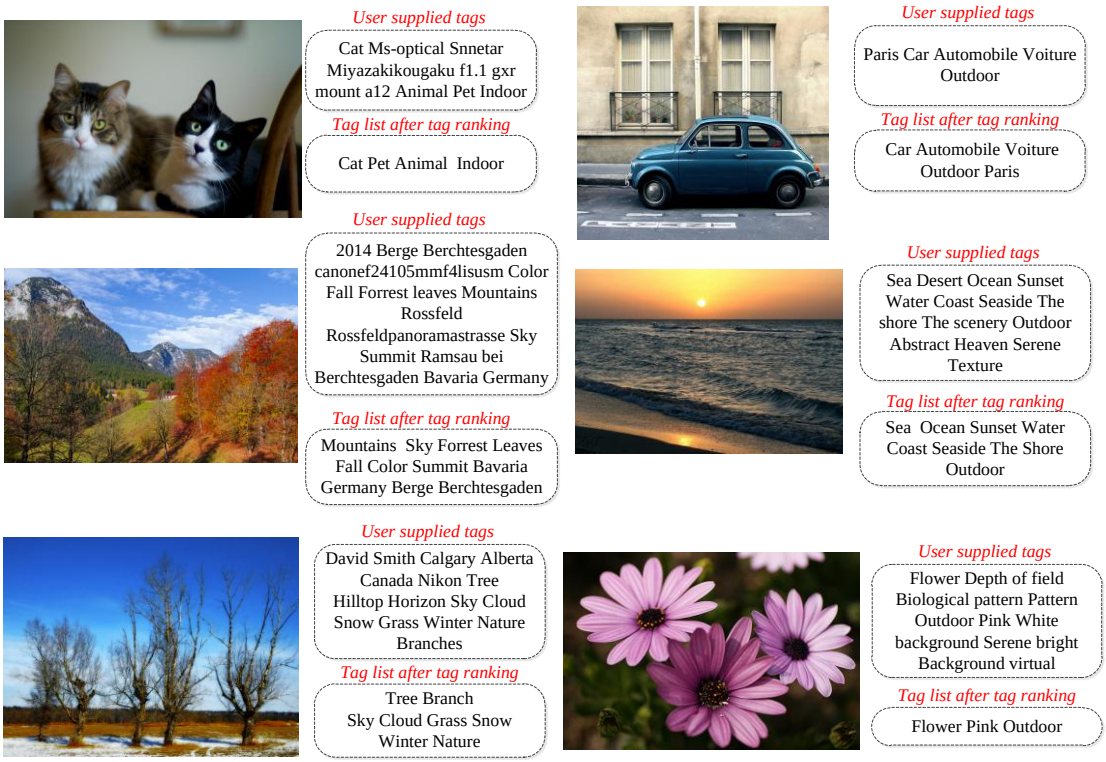

Figure 9. Examples of social image tag ranking results

5.3. Tag based Social Image Retrieval

In order to further explain performance of our proposed method, we exploit it in tag based social image retrieval, which is a potential application scenario of tag ranking. We compare the proposed method with the following three tag ranking strategies: 1) Interestingness based ranking, 2) Uploading time based ranking, and 3) Original tags list provided by user. Flickr provides two tag based image search services for users to rank social images according to interestingness and uploaded time. Utilizing the proposed method, the image ranking list is produced according to the tag position in the ranked tag list.

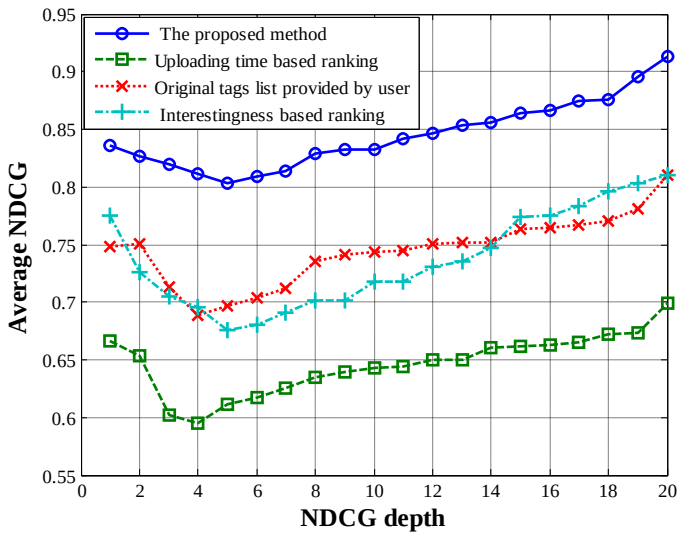

Figure 10. Social image retrieval performance of different tag ranking strategies

Fig. 10 illustrates that under various depths, the average NDCG results obtained by the proposed method are much larger than other strategies. The conclusion can be drawn that high quality tag ranking list is able to promote the capability of social image retrieval.

6. CONCLUSION AND DIRECTIONS FOR FUTURE WORK

This paper proposes a novel social image tag ranking algorithm based on weighted neighbor voting algorithm and random walk process. The main innovations of this paper are to estimate voting powers of neighbors using both visual similarity and tag relevance, and then calculate the tag relevance score by carrying out the random walk on tag graph.

In this work, we exploit the proposed method in tag based social image retrieval. In the future, we will try to use it to other applications, such as social image tag refinement, and social image tag recommendation. Furthermore, other typical social image datasets will be utilized to test the performance of our algorithm, such as NUS-WIDE and MIRFLICKR-25000.

REFERENCES

Ballan, L., Bertini M., Uricchio T., Del B.A. (2015) "Data-driven approaches for social image and video tagging", *Multimedia Tools and Applications*, 74(4), pp. 1443-1468.

Cilibrasi, R.L. and Vitanyi, P. (2007) "The google similarity distance", *IEEE Transactions on knowledge and data engineering*, 19(3), pp.370-383.

Cui C.R., Ma J. (2013) "An image tag recommendation approach combining relevance with diversity", *Chinese Journal of Computers*, 36(3), pp. 654-663.

Eom W.Y., Lee S., De N.W., Yong M. (2011) "Improving image tag recommendation using favorite image context", *In: Proceedings of International Conference on Image Processing*, pp. 2445-2448.

Feng S.H., Feng Z.Y., Jin R. (2015) "Learning to rank image tags with limited training examples", *IEEE Transactions on Image Processing*, 24(4), pp. 1223-1234.

He Y.H., Kang C.C., Wang J., Xiang S.M. Pan Chunhong (2015) "Image tag-ranking via pairwise supervision based semi-supervised model", *Neurocomputing*, 167, pp. 614-624.

Jeong J.W., Hong H.K., Lee D.H. (2013) "I-TagRanker: An efficient tag ranking system for image sharing and retrieval using the semantic relationships between tags", *Multimedia Tools and Applications*, 62(2), pp. 451-478.

Karakasis E.G., Amanatiadis A., Gasteratos A., Chatzichristofis S.A. (2015) "Image moment invariants as local features for content based image retrieval using the Bag-of-Visual-Words model", *Pattern Recognition Letters*, 55, pp. 22-27.

Kim M.U., Yoon K.G. (2015) "Performance evaluation of large-scale object recognition system using bag-of-visual words model", *Multimedia Tools and Applications*, 74(7), pp. 2499-2517.

Lee S.Y., De N.W., Plataniotis K., Ro Y.M. (2010) "MAP-based image tag recommendation using a visual folksonomy", *Pattern Recognition Letters*, 31(9), pp. 976-982.

Li X.R., Cees G., Snoek M., Marcel W. (2009) "Learning Social Tag Relevance by Neighbor Voting", *IEEE Transactions on Multimedia*, 11(7), pp.1310-1322.

Li X.R., Uricchio T., Ballan L., Bertini M., Snoek C.G., Bimbo A.D. (2016) "Socializing the Semantic Gap: A Comparative Survey on Image Tag Assignment, Refinement, and Retrieval", *ACM Computing Surveys*, 49(1), Article No. 14.

Li Z.H., Liu J., Tang J.H., Lu H.Q. "Projective Matrix Factorization with unified embedding for social image tagging", *Computer Vision and Image Understanding*, 124, pp.71-78.

Liu D., Hua X.S., Zhang H.J. "Content-based tag processing for Internet social images", *Multimedia Tools and Applications*, , 51(2), pp.723-738.

Liu D., Hua X.S., Yang L.J., Wang M., Zhang H.J. (2009) "Tag ranking", *Proceedings of the 18th international conference on World Wide Web*, pp. 351-360.

Qian X.M., Hua X.S., Tang Y.Y., Mei T. (2014) "Social Image Tagging With Diverse Semantics", *IEEE Transactions on Cybernetics*, 44(12), pp. 2493-2508.

Sigurbjörnsson B., Zwol R.. (2008) "Flickr tag recommendation based on collective knowledge", *In: Proceedings of the 17th International Conference on World Wide Web*, pp. 327-336.

Sun A.X., Bhowmick S. S., Khanh T.N.N., Bai G. (2011) "Tag-Based Social Image Retrieval: An Empirical Evaluation", *Journal of the American Society for Information Science and Technology*, 62(12), pp. 2364-2381.

Sun A.X., Bhowmick S., Chong J.A., (2011) "Social image tag recommendation by concept Matching", *Proceedings of the 2011 ACM Multimedia Conference and Co-Located Workshops*, pp. 1181-1184.

Sun F.M., Li H.J., Zhao Y.H., Wang X.M., Wang D.X. (2013) "Towards tags ranking for social images", *Neurocomputing*, 120, pp. 434-440.

Tang J.H., Li M.X., Li Z.C., Zhao C.X. (2015) "Tag ranking based on salient region graph propagation", *Multimedia Systems*, 21(3), pp. 267-275.

Wang Y., Du L, Dai H. (2016). "Unsupervised SAR Image Change Detection Based on SIFT Keypoints and Region Information", *IEEE Geoscience And Remote Sensing Letters*, 13(7), pp. 931-935.

Xiao J., Zhou W.G., Li X., Wang M., Tian Q. (2012) "Image tag re-ranking by coupled probability transition", *proceedings of the 20th ACM International Conference on Multimedi*a, pp. 849-852.

Xu Z., Luo X. F. , Liu Y.H., Mei L., Hu C.P. (2014) "Measuring Semantic Relatedness between Flickr Images: From a Social Tag Based View", *Scientific World Journal*, AR 758089.

Zhang J., Liu X., Zhuo L., Wang C. (2015) "Social images tag ranking based on visual words in compressed domain", *Neurocomputing*, 153: 278-285.

Zhang X.M., Ji S.F., Wang S.Z., Li Z.J., Lv X.Q. (2016) "Geographical Topics Learning of Geo-Tagged Social Images", *IEEE Transactions on Cybernetics*, 46(3), pp. 744-755.

Zhang X.M., Zhao X.J., Li Z.J., Xia J.L., Jain R., Chao W.H., "Social image tagging using graph-based reinforcement on multi-type interrelated objects", *Signal Processing*, 93(8), pp.2178-2189.

Zhao Y., Zhai Y.W., Dubois E., Wang S.G. (2016) "Image matching algorithm based on SIFT using color and exposure information", *Journal of Systems Engineering and Electronics*, 27(3), pp.691-699.