Enhancing Query Expansion for Tag-based Social Image Retrieval

Amel Ksibi, Anis Ben Ammar and Chokri Ben Amar

REGIM: REsearch Groups on Intelligent Machines,
Sfax University, BP 1173, Sfax, 3038, Tunisia

Abstract. Recently, extensive research efforts have been dedicated to tag-based social image search which enables users to formulate their queries using tags. However, tag queries are often ambiguous and typically short yielding to retrieve irrelevant images in top ranked list. To overcome this problem, an effective strategy is to produce diverse images in top ranking list covering various aspects of the query. In this context, we propose a Multi-view Concept-based Query Expansion (MVCQE) process, using a predefined list of semantic concepts and following three main steps. First, we harvest social knowledge to capture different contexts related to the query. Second, we perform a Multi-view Concepts weighting by applying concept-based query expansion for the initial query and for each of its contexts. Third, we select the most representative concepts using an adaptive threshold with respect to the dispersion of concept weights. Experiments using ambiguous queries over the NUS-WIDE dataset confirm the effectiveness of our process to improve the diversification compared to well known query expansion approaches . . .

1 Introduction

With the proliferation of Web 2.0, photo-sharing services are hosting a tremendous volume of digital images associated with their users generated tags [11]. Thus, tag-based social image retrieval expect to be an intuitive way to perform search, which presents two specific challenges:

– *Tags Mismatch:* It occurs when tag query fails to appear in tags of relevant images due to either the use of synonyms, or to the incomplete semantic representations (e.g. not containing the tag)[4]
– *Tag Ambiguity:* It occurs when a query is interpreted with several meanings other than user's expectation. [14]

In literature, these two challenges have been well studied separately. In one sight, for the 'tags-mismatch' problem, concept-based approaches have been well intended to overcome it by searching social images based on concept matching rather than tag matching. Indeed, queries and images are transformed into semantic concepts vectors as a standardized representation[6],[7]. In the other sight, to tackle the 'tag ambiguity' problem, an effective approach is to provide diverse results that cover multiple topics underlying a query. To this end, diversity-based approaches can be categorised as either explicit or

implicit [10]. Explicit approaches seek to promote images with maximum coverage of query aspects as *characteristic of the query* itself, while implicit-based approaches rely on *characteristics of the retrieved images* in order to identify diverse images, under the assumption that similar images will cover similar aspects. In current study,we will focus only on explicit approaches. Specifically, we will be interested in query expansion techniques, which aims to alleviate the query ambiguity by adding meaningful terms from a suitable knowledge resource. In literature, query expansion has been shown as a confirmed way for improving retrieval effectiveness in term of Recall value. However, it can generate topic drift problem when too broaden the query. Thus, different challenges are identified:

1. Which knowledge resource should be retained reflecting a sufficient coverage of the dynamic human knowledge?
2. How to optimize the coverage of all query aspects underlying an ambiguous query?
3. How many terms should be added and how to assign weights to each one without hampering the overall precision?

In this paper, we will harness the semantic representation and the social knowledge not only to solve tag-mismatch problem, but also to cope with query ambiguity challenge. Indeed, we will reformulate the tag query using a new concept-based query expansion process called "Multi-view concept-based query expansion" by weighting semantic concepts from different view or context, aggregating the obtained weights and selecting the most representative ones using a dynamic threshold.

This paper is organized as follows: In section 2, we provide an overview of the existing orientations for diversifying search results. In section 3, we present the overall architecture of the proposed tag-based social image retrieval system and we detail the Multi-view concept based query expansion process. In section 4, we give experimental results.

2 Overview of Query Expansion Approaches for Social Image Retrieval

Query expansion process has been considered as an intuitive and promising way to diversify results by adding new meaningful terms from knowledge resources[12]. In literature, different knowledge bases have been exploited for query expansion. For instance, in [9], authors proposed to expand query through an open-source knowledge such as WordNet and ConceptNet based on synonyms and concepts. Myoupo et al.[8] proposed to reformulate queries using Wikipedia Knowledge by adding terms that are closer to the query. Similarly, Hoque et al. [2] explored Wikipedia resources to ensure query expansion. Given an ambiguous query, they attempted to capture its various aspects, and for each aspect, a dynamic number of terms pertaining to the query were discovered from wikipedia. Weinberger et al.[14] introduced a new tool to disambiguate a tag query using a probabilistic framework. In this work, ambiguity is detected when the same tag generates two tags that occur in two divergent contexts.

The aforementioned query expansion approaches are influenced by the number of added terms which affects results diversification. In fact, this number can be considered
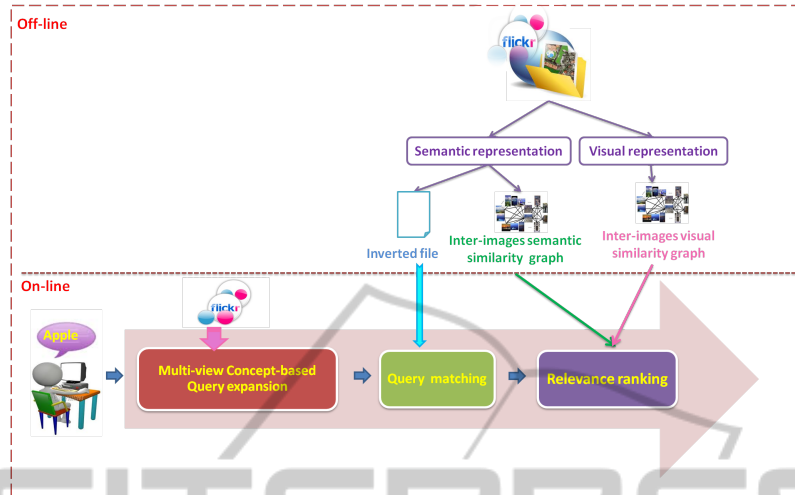
Fig. 1. Flowchart of the proposed tag-based social image retrieval.

as a diversification trade-off where the more added terms are, the higher the diversity is. Typically, this trade-off is uniformly optimized by maximizing the average diversification performance on a set of training queries. However, not all queries are similarly ambiguous. Thus, different queries might benefit from different trade-off since any uniform choice of this trade-off for all queries would be suboptimal. This challenge has been studied in social image retrieval by Hoque et al.[2]. They proposed to automatically estimate the trade-off based on the level of ambiguity of the query itself. this trade-off denotes the number of most related concepts within the query expansion process based on the number of senses of the query as determined by Wikipedia. The main weakness of this approach consists in using lexical resources such as Wikipedia to extract concepts and their weights. Such knowledge resource may only extract the lexical relatedness query tag and extracted concepts and cannot reflect the visual relatedness between them.

3 Mutli-view Concept-based Query Expansion for Tag-based Social Image Retrieval

In this section, we will present the overall architecture of the proposed retrieval process. Then, we will give a preliminary overview about concept-based query expansion approach(CQE). Finally, we will describe the Muti-view concept-based query expansion process.

3.1 Overall Architecture of the proposed Retrieval System

The flowchart of the proposed social image retrieval system is illustrated in figure 1: Given a set of N pre-defined concepts, we model each image x_i in the collection by a

vector $C_i = \{c_1^i, c_2^i, .., c_N^i\}$ containing concept weights using the annotation approach described in [5]. Each vector defines the semantic representation underlying an image.

Take the tag query "Apple" as an example, when "Apple" is submitted to our tag-based social image retrieval system, a step of Multi-view concept based query expansion is performed by aggregating, for each concept, the associated weights obtained from different views. This step is achieved by selecting the most appropriate concepts that capture the different meanings of the query using a dynamic threshold per-query. We note the expanded query C_q by a vector $\{c_1^q, c_2^q, .., c_N^q\}$, An inverted file is, then, constructed to reduce the search space by selecting images having at least one selected concept by the query. We denote by $D_q = \{x_1^q, x_2^q, .., x_{|D_q|}^q\}$ the set of vectors corresponding to images that are associated with the set of query concepts C_q. This collection, which is a part of the large set $D = \{x_1, x_2, .., x_{|D|}\}$, is obtained by the aforementioned inverted file generation.

A step of query-images matching is applied by estimating the cosine similarity between the expanded query vector C_q and each image vector x_i among sub-collection D_q. Once the relevance scores are estimated for all images in the selected collection, these images are ranked by relevance. Generally, query expansion results in a gain in recall often compensated by the corresponding loss in precision, since the integration of some query terms may be less plausible and hence lead to topic drift. To remedy this problem, we apply a relevance re-ranking model using random walk with restart process as such we move relevant images upward assuming that images ,which are visually and semantically similar to highly ranked images, should be upward [1].

Next subsection describes the process of Multi-view concept based query expansion in details.

3.2 Multi-view Concept-based Query Expansion: MVCQE

Concept-based query expansion plays a pivotal role in the overall success of any tag-based retrieval task. Indeed, it can implicitly tackle the query ambiguity problem by expanding a tag query to a list of top related concepts over the semantic space. In other words, a tag query is reformulated by assigning high scores to concepts that overlap different aspects underlying an ambiguous query.

Intuitively, concepts related to the most known sense with respect to the ambiguous query, will have high scores. As a result, not all aspects will be covered. In order to reduce the influence of the most common senses, we propose a new approach called "Multi-view concept-based query expansion" in which we extract different contexts related to the tag in question using social knowledge and we apply concept-based query expansion for the original query and the captured contexts. By doing such, we obtain different query interpretations with respect to different contexts. As such, one concept can have a high weight for a one context and low weight for another. In this situation, we obtain the maximum of weights. As a result, we give high weights to different concepts representing all the query aspects in different contexts.

Figure 2 illustrates the multi-view concept-based query expansion process in details: The first step consists in extracting semantic clusters related to a given tag-query. Each cluster defines a view characterizing a specified context of the query. Indeed,
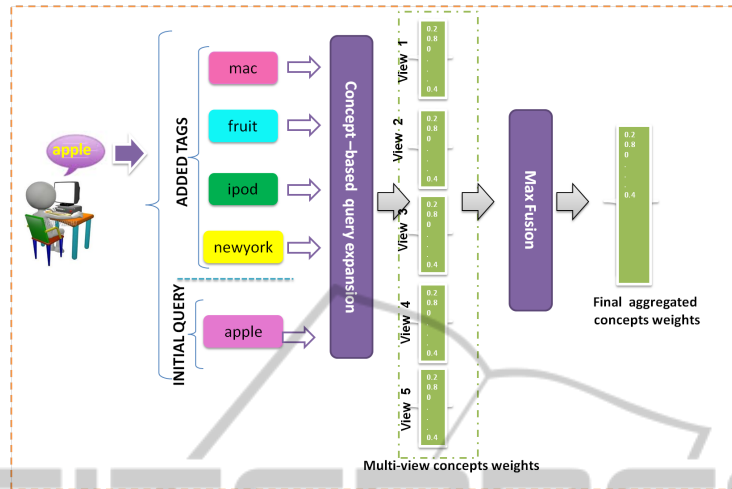
Fig. 2. Our proposed Multi-view Concept-based query expansion process scheme.
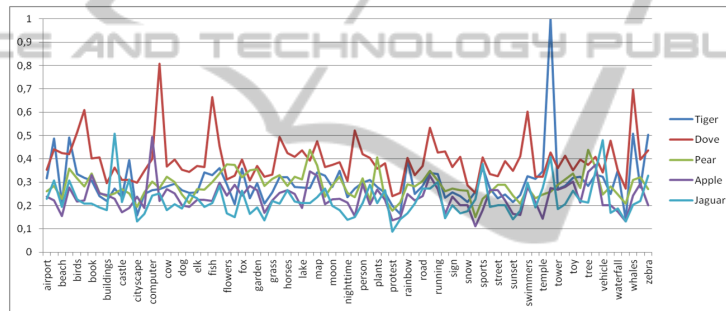

Fig. 3. Dispersion of concept weights for different queries.

The second step consists in performing concept-based query expansion for each view and for the original query. By doing so, we obtain different ranking lists of weighted concepts corresponding to each view or aspect. Then, we aggregate all ranking lists by applying Max-Fusion for each concept to build the multi-view expanded query vector. Thus, we assign for each concept the highest weight among the obtained weights by different views.

The third step is concept selection where the aim is to choose an optimal subset of concepts from the available set that are able to capture the majority of query's aspect and avoid topic drift risk. A challenging question is where to stop selecting concepts from a ranked list of concept weights? In order to select concepts, concept weights ranking can be thresholded at an arbitrary rank or score. This threshold improves the diversification performance of the retrieval process as it can also be considered as a diversification parameter: too tight threshold would extract a limited number of concepts being only the common senses of tag query, while a too loose threshold would produce too query broadening resulting in a topic drift problem.

Typically, this threshold is uniformly optimised so as to maximize the precision on

a set of training queries. As a result, a fixed threshold is estimated for all queries. Figure 3 shows the distribution of concept weights for different queries. From this figure, it is clear that different queries have different degree of dispersion among the weights. So, that any uniform choice of the threshold for all queries would be suboptimal.

A main factor determining what the right threshold is, consists in weights distribution. Since this factor is query dependent, the right threshold should be selected dynamically per query, not statically as most previously proposed methods in the literature. To achieve this objective, we develop a new method defining an optimal trade-off score τ, per query, using concepts scores and their distribution as input. Indeed, we opt for threshold optimization per-query by focusing on the dispersion degree among the scores using the standard deviation σ. The threshold will be estimated using the following formula:

$$\tau = \begin{cases} \frac{1}{N} * \sum_{i=1}^{N} c_i^q & \text{if} \quad \sigma \geq \epsilon \\ \frac{1}{N} * \sum_{i=1}^{N} c_i^q + \sigma & \text{else} \end{cases}$$

where $\epsilon \in [0, 1]$ is a heuristic parameter, c_i^q is the weight of concept i and N is the number of concepts. On one hand, if a query's ranking list has a high value of dispersion among the concepts scores, it could be a clue that the ranking function has been able to discriminate between relevant and irrelevant concepts. So, we estimate the threshold as the average of all scores. On the other hand, if a low level of dispersion appears, because the ranking function has assigned similar weights, it can be interpreted as it was not able to distinguish between relevant and irrelevant concepts. In such case, we add the value of standard deviation to the average to estimate the threshold.

4 Experiments and Results

4.1 Experiments Setup

To validate our proposed retrieval process, we conduct experiments on the challenging real-word NUS-WIDE [1] dataset. It is one of the largest social media datasets which contains 269,648 Flickr images accompanied by their associated tags and their visual features. Each image is also indexed by 81 concepts. In addition, we select a set of 12 common ambiguous tag-queries, including Apple, Jaguar, Dove,Tiger, Pear, Jordan, Eagle, Washington, Flash,...

4.2 Study of MVCQE Effectiveness

We study the effectiveness of MVCQE process by responding to the following questions:

- To which extent does social knowledge improve the detection of contexts performance underlying an ambiguous query?
- What is the impact of using Multi-view on diversifying results?
- What is the impact of using adaptive threshold in MVCQE?
- What is the impact of knowledge resource selection in adaptive threshold in MVCQE?

---

[1] http://lms.comp.nus.edu.sg/research/NUS-WIDE.htm

What is the Impact of using Multi-view on Diversifying Results? In this experiment, we compare our proposed MVCQE approach with the baseline concept-based query expansion (CQE)[13] and tag-based query expansion(TQE)[3]. In TQE, tag query is expanded by its top-k related tags from Flickr that frequently co-occur with the original query.

In the following figure, we illustrate the obtained tag query reformulations for query 'Tiger' using the aforementioned approaches. Top-15 related concepts and tags are extracted. For TQE, the coverage of query aspects is low as all the selected tags are related to one context for query 'Tiger'. CQE outperforms TQE by capturing more aspects and contexts than TQE. This improvement is due to the ability of the predefined list of semantic concepts to cover the comprehensive human world knowledge. However, we note that top-5 selected concepts are belonging to the same context resulting in a lack of capturing all related contexts of a tag query (such as Tiger OS). MVCQE performs the best by detecting more diverse aspects, in top-5 selected concepts, other than in TQE and CQE by selecting 'computer'. The success of our approach MVCQE is due to the diversification of top-k selected concepts yielding to diverse query aspects, yet diverse search results. In fact, the multi-view concepts weighting with respect to different contexts is the responsible of this diversification.

Concepts	CQE	Concepts	MVC-QE	Concepts	TQE
tiger	1	tiger	1	cat	1
whales	0.5091	animal	0.6847	animal	1
zebra	0.5039	computer	0.5689	animals	1
bear	0.4915	whales	0.509	nature	1
animal	0.4861	zebra	0.5039	wildlife	1
rainbow	0.3968	bear	0.4915	bigcat	1
cat	0.3951	leaf	0.4757	white	1
sports	0.3624	rainbow	0.3967	stripes	1
flags	0.3611	cat	0.3951	wild	1
fox	0.3573	book	0.3788	flickrbigcats	1
waterfall	0.351	flowers	0.3707	canon	1
mountain	0.3482	birds	0.3674	nikon	1
temple	0.3471	sports	0.3624	tigre	1
map	0.3429	garden	0.3614	cats	1
fire	0.3417	flags	0.3611	feline	1

Fig. 4. Different interpretations for query "Tiger".

What is the Impact of using Adaptive Threshold in MVCQE? In this experiment, we investigate two types of thresholding: static and adaptive. In static thresholding, the same fixed pre-selected rank threshold τ is applied to all queries. We test different scores of τ at 0.2, 0.3 and 0.4. In dynamic thresholding, we estimate, for each query, the optimal threshold score with respect to the dispersion degree among the concept weights as described above. We obtain the following results in figures 5 and 6.

It can be seen from these figures that defining an unified threshold score for all
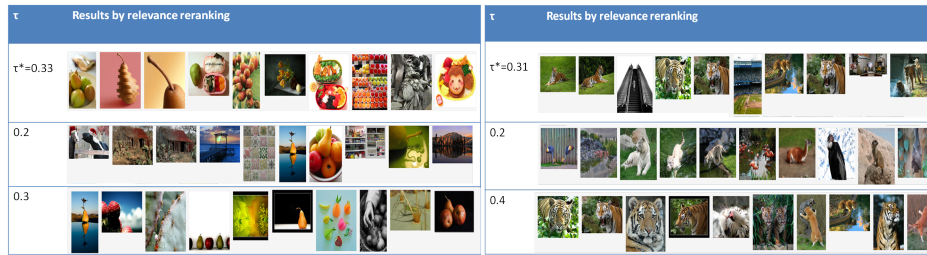
Fig. 5. Obtained results for queries 'Pear' and 'Tiger' using different threshold values.

	σ	τ	mean
dove	0.104	0.39	0.39
pear	0.053	0.33	0.28
apple	0.059	0.29	0.231
tiger	0.137	0.31	0.31

Fig. 6. Optimal thresholds for different queries with respect to dispersion degree among concept weights.

queries would be suboptimal, which prove the need for dynamic thresholding. Actually, static threshold hurts retrieval precision and leads to topic drift risk in case of low value of threshold or under-estimation of query in case of high value of threshold.

What is the Impact of Knowledge Resource Selection in Adaptive Threshold in MVCQE? In this experiment, we study the impact of knowledge resource choice in the threshold estimation. For this purpose, we compare Flickr and Wikipedia resources. Indeed, we estimate the correlation between concepts and the query using FCS measure over Flickr resources and DiscoDistance over Wikipedia resources. Figures 7 and 8 illustrate the results:

From figures 7 and 8, we notice that weights of concepts for queries 'Tiger' and 'jaguar' are more dispersed than queries 'apple' and 'pear'. Thus, we deduce that the dispersion of concept weights is the main factor for determining the optimal threshold. In addition, we note that the interval of weights repartition differ from one another. Thus, the mean of all concept weights for each query is also a factor for determining the focused threshold. These deductions prove the formula of our proposed threshold.

5 Conclusions

In this paper, we have presented a new query expansion process MVCQE using semantic concepts representation in different query contexts. The key advantage of our process is its ability to make effective the use of semantic concepts representation not only for solving tag-mismatch but also for diversifying search results. In fact, we have demonstrated that mapping a tag query in the semantic space guarantee the complete coverage of all query aspects. Further, we have proved the necessity to diversify the captured query aspects. For this purpose, we have analysed the query from different contexts or views. Moreover, we have demonstrated that a step of concepts selection
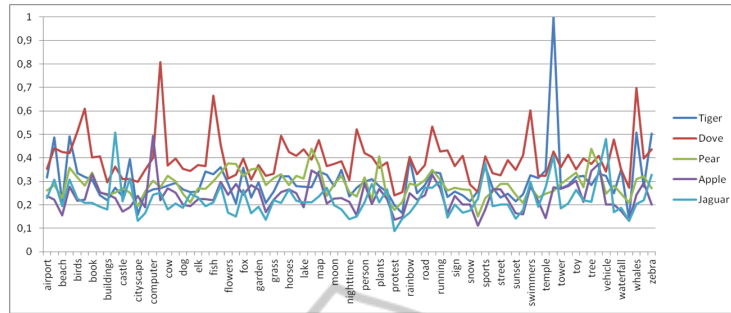
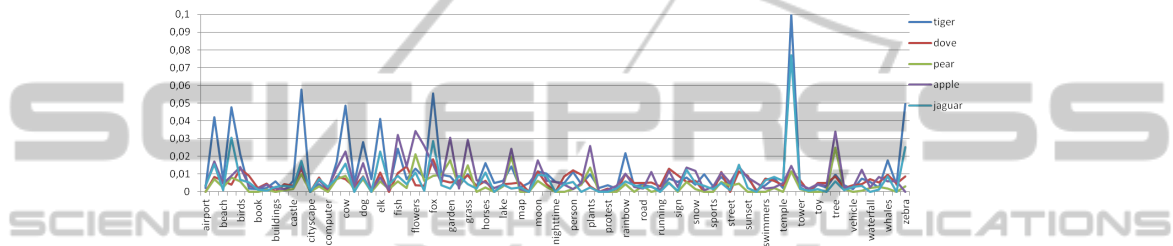**Fig. 7.** Weights of concepts estimation over Flickr resources.

Fig. 8. Weights of concepts estimation over Wikipedia resources.

was required. Therefore, we have proposed an automatic adaptive threshold with respect to the dispersion of concept weights for a given query. Finally, we argued that the proposed adaptive thresholding can be transferable to other applications that need an optimal threshold in a ranking list and having only the items scores.

Our future work will involve investigation into the further refinement of the proposed system. More specifically, we plan to expand our MVCQE approach by capturing hierarchical aspects from a taxonomy of semantic concepts.

References

1. Feki, G., Ksibi, A., Ammar, A. B., Amar, C. B.: Regimvid at imageclef2012: Improving diversity in personal photo ranking using fuzzy logic. In: ImageCLEF12 (2012)
2. Hoque, E., Hoeber, O., Gong, M.: Balancing the trade-offs between diversity and precision for web image search using concept-based query expansion. Journal of Emerging Technologies in Web Intelligence 4(1) (2012)
3. Jin, S., Lin, H., Su, S.: Query expansion based on folksonomy tag co-occurrence analysis. In: Granular Computing, 2009, GRC '09. IEEE International Conference on, pp. 300–305 (2009). DOI 10.1109/GRC.2009.5255110
4. Kato, M., Ohshima, H., Oyama, S., Tanaka, K.: Can social tagging improve web image search? In: Web Information Systems Engineering-WISE 2008, pp. 235–249. Springer (2008)
5. Ksibi, A., Ammar, A. B., Amar, C. B.: Effective concept detection using second order co-occurence flickr context similarity measure socfcs. In: CBMI, pp. 1–6 (2012)

6. Ksibi, A., Dammak, M., Ben Ammar, A., Mejdoub, M., Ben Amar, C.: Flickr-based semantic context to refine automatic photo annotation. In: Image Processing Theory, Tools and Applications (IPTA), 2012 3rd International Conference on, pp. 377–382. IEEE (2012)

7. Ksibi, A., Elleuch, N., Ammar, A. B., Alimi, A. M.: Semi-automatic soft collaborative annotation for semantic video indexing. In: EUROCON, pp. 1–6. IEEE (2011)

8. Myoupo, D., Popescu, A., Borgne, H., Mollic, P. A.: Multimodal Image Retrieval over a Large Database, Lecture Notes in Computer Science, vol. 6242. Springer Berlin Heidelberg (2010)

9. Roohullah, Jaafar, J.: Exploiting the query expansion through knowledgebases for images. In: H. Zaman, P. Robinson, M. Petrou, P. Olivier, T. Shih, S. Velastin, I. Nystrm (eds.) Visual Informatics: Sustaining Research and Innovations, Lecture Notes in Computer Science, vol. 7067, pp. 93–103. Springer Berlin Heidelberg (2011)

10. Santos, R. L., Macdonald, C., Ounis, I.: Exploiting query reformulations for web search result diversification. In: Proceedings of the 19th international conference on World wide web, WWW '10, pp. 881–890. ACM, New York, NY, USA (2010)

11. Sun, A., Bhowmick, S. S., Nguyen, K. T. N., Bai, G.: Tag-based social image retrieval: An empirical evaluation. JASIST 62(12), 2364–2381 (2011)

12. Tang, X., Liu, K., Cui, J., Wen, F., Wang, X.: Intentsearch: Capturing user intention for one-click internet image search. IEEE Trans. on Pattern Analysis and Machine Intelligence 34(7), 1342–1353 (2012)

13. Wei, X. Y., Ngo, C. W., Jiang, Y. G.: Selection of concept detectors for video search by ontology-enriched semantic spaces. Multimedia, IEEE Transactions on 10(6), 1085–1096 (2008)

14. Weinberger, K. Q., Slaney, M., Van Zwol, R.: Resolving tag ambiguity. In: Proceedings of the 16th ACM international conference on Multimedia, pp. 111–120. ACM (2008)