

## Research Article

# Measuring Semantic Relatedness between Flickr Images: From a Social Tag Based View

Zheng Xu,<sup>1</sup> Xiangfeng Luo,<sup>2</sup> Yunhuai Liu,<sup>1</sup> Lin Mei,<sup>1</sup> and Chuanping Hu<sup>1</sup>

<sup>1</sup> *The Third Research Institute of Ministry of Public Security, 339 Bisheng Road, Shanghai 201142, China*

<sup>2</sup> *Shanghai University, 149 Yanchang Road, Shanghai 200472, China*

Correspondence should be addressed to Zheng Xu; xuzheng@shu.edu.cn

Received 29 October 2013; Accepted 19 December 2013; Published 23 February 2014

Academic Editors: J. Shu and F. Yu

Copyright © 2014 Zheng Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Relatedness measurement between multimedia such as images and videos plays an important role in computer vision, which is a base for many multimedia related applications including clustering, searching, recommendation, and annotation. Recently, with the explosion of social media, users can upload media data and annotate content with descriptive tags. In this paper, we aim at measuring the semantic relatedness of Flickr images. Firstly, four information theory based functions are used to measure the semantic relatedness of tags. Secondly, the integration of tags pair based on bipartite graph is proposed to remove the noise and redundancy. Thirdly, the order information of tags is added to measure the semantic relatedness, which emphasizes the tags with high positions. The data sets including 1000 images from Flickr are used to evaluate the proposed method. Two data mining tasks including clustering and searching are performed by the proposed method, which shows the effectiveness and robustness of the proposed method. Moreover, some applications such as searching and faceted exploration are introduced using the proposed method, which shows that the proposed method has broad prospects on web based tasks.

## 1. Introduction

Relatedness measurement especially similarity between multimedia such as images and videos plays an important role in computer vision. The image similarity is a base for many multimedia related applications including image clustering [1], searching [2, 3], recommendation [4], and annotation [5]. The relatedness problem is relevant to two aspects: images representation and relatedness measurement. The former aspect needs an appropriate model to reserve the related information of an image. The latter aspect requires an effect method to compute the relatedness accurately.

In the early stage, relatedness measurement is based on the low-level visual features such as texture [6, 7], shape [8], and gradient [9]. These visual features are used to represent effective information of an image. Some distance metrics including Chi-Square distance [10], Euclidean distance [11], histogram intersection [12], and EMD distance [13] is used. Overall, these methods ignore the high-level features such as semantic information which can be understood by machine

and people easily. These methods are limited to the applications which need semantic level information.

Recently, with the explosion of community contributed multimedia content available online, many social media repositories (e.g., Flickr (<http://www.flickr.com>), Youtube (<http://www.youtube.com>), and Zoomr (<http://www.zoomr.com>)) allow users to upload media data and annotate content with descriptive keywords which are called social tags. We take Flickr, one of the most popular and earliest photo sharing sites, as an example to study the relatedness measurement between images. Flickr provides an open platform for users to publish their personal images freely. The principal purpose of tagging is to make images better accessible to the public. The success of Flickr proves that users are willing to participate in this semantic context through manual annotations [14]. Flickr uses a promising approach for manual metadata generation named “social tagging,” which requires all the users in the social network to label the web resources with their own keywords and share with others. The characteristics of social tags are as follows.

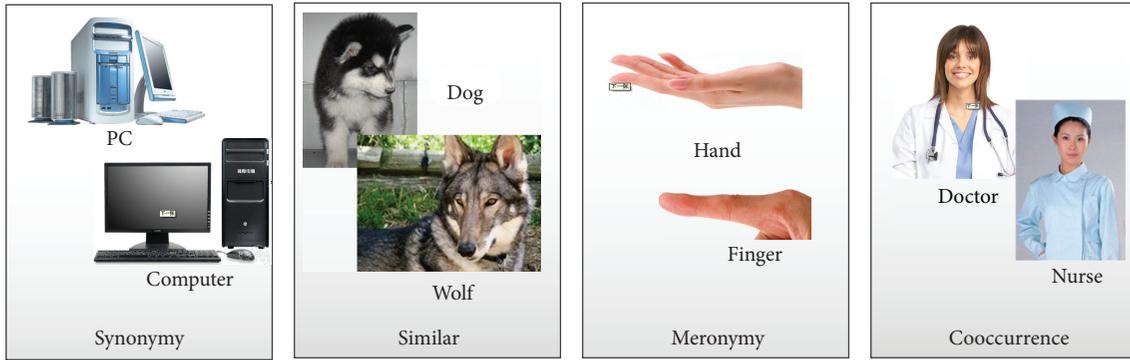


FIGURE 1: The illustration of four kinds of correlations between concepts.

- (1) Ontology free. The ontology based labeling defines ontology and then let users label the web resources using the semantic markups in the ontology. Social tagging requires all the users in the social network to label the web resources with their own keywords and share with others. Different from ontology based annotation, there is no predefined ontology or taxonomy in social tagging. Thus, the tagging task is more convenient for users.
- (2) User oriented. The users can annotate images with their favorite tags. The tags of an image are determined by users' cognitive ability. To a same image, users may give different tags. Each image may be with one tag at least, and each tag may appear in many different images.
- (3) Semantic loss. Irrelevant social tags frequently appear, and users typically will not tag all semantic objects in the image, which is called semantic loss. Polysemy, synonyms, and ambiguity are some drawbacks of social tagging.

Based on the above characteristics, we aim at measuring semantic relatedness between images using social tags. It is observed that the correlations between the concepts of images can be divided into four kinds: synonymy, similarity, meronymy, and concurrence, as illustrated in Figure 1. Synonymy means the same object with different names. Similarity denotes that two objects are similar. Meronymy means that two objects follow part-of relation. Concurrence means that two objects appear frequently. Overall, the above four correlations can be summarized as semantic relatedness [15]. Semantic relatedness is a more generic concept than semantic similarity. Similar concepts are usually considered to be related for their likeness (synonymy); dissimilar concepts can also be semantically related such as meronymy or concurrence. In this paper, we focus on measuring semantic relatedness between images.

- (1) Semantic relatedness follows the cognitive mechanism of people. In [16], the author suggests that the association relation is the basic mechanism of brain. When people know a concept such as "hospital," she/he may index the related concept such as "doctor" for appropriate understanding of the original concept.

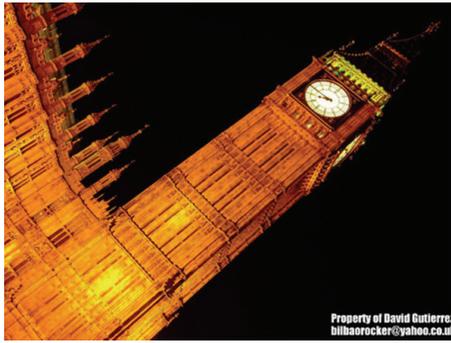
Since the goal of relatedness measurement is to facilitate related applications such as searching and recommendation, the proposed method should follow user's cognitive mechanism.

- (2) Semantic relatedness can be used to organize images based on their associations. In recent literatures, such as Linked Open Data (LOD) [17] and Semantic Link Network (SLN) [18–20], the resources are managed by their semantic relations. The proposed semantic relatedness measures can be used to build semantic links between resources especially images, which can be easily applied in real applications.

The major contributions of this paper are summarized as follows.

- (1) We propose a framework to measure semantic relatedness between Flickr images using tags. Firstly, the cooccurrence measures are used to compute the relatedness of tags between two images. Secondly, we transform the tags relatedness integration to the assignment in bipartite graph problem, which can find an appropriate matching to the semantic relatedness of images. Finally, a decline factor considering the position information of tags is used in the proposed framework, which reduces the noise and redundancy in the social tags.
- (2) A real data set including 1000 images from Flickr with ten classes is used in our experiments. Two evaluation methods including clustering and retrieval are performed, which shows that the proposed method can measure the semantic relatedness between Flickr images accurately and robustly.
- (3) We extend the relatedness measures between concepts to the level of images. Since the association relation is the basic mechanism of brain. The proposed relatedness measurement can facilitate related applications such as searching and recommendation.

The rest of the paper is organized as follows. Section 2 gives the related work of social tags and image similarity measures. The problem definition is introduced in Section 3.



Big Ben at night: London through my lens



London eye

FIGURE 2: The illustration of a pair of images from Flickr.

Section 4 proposes the method for measuring semantic relatedness of images. Experiments are presented in Section 5. Conclusions are made in the last section.

## 2. Related Work

In this section, we give two related aspects of the proposed work. Some researches about social tags are introduced first. Then, we give the related work about image similarity measures.

**2.1. On Social Tags.** In the area about the usage patterns and semantic values of social tags, Golder and Huberman [21] mined usage patterns of social tags based on the delicious ([del.icio.us/post](http://del.icio.us/post)) data set. Al-Khalifa and Davis [22] concluded that social tags were semantically richer than automatically extracted keywords. Suchanek et al. [23] used YAGO (<http://www.mpi-inf.mpg.de/yago-naga/yago>) and WordNet (<http://wordnet.princeton.edu>) to check the meaning of social tags and concluded that top tags were usually meaningful. Halpin et al. [24] examined why and how the power law distribution of tag usage frequency was formed in a mature social tagging system over time.

Beside research on mining social tags, some researches modeled the network structure of social tags. Cattuto et al. [25] investigated the network features of social tags system, which is seen as a tripartite graph using metrics adapted from classical network measures. Lambiotte and Ausloos [26] described the social tags systems as a tripartite network with users, tags, and annotated items. The proposed tripartite network was projected into the bipartite and unipartite network to discover its structures. In [27], the social tags system was modeled as a tripartite graph which extends the traditional bipartite model of ontologies with a social dimension.

Recently, many researchers investigated the applications of social tags in information retrieval and ranking. In [28], the authors empirically study the potential value of social annotations for web search. Zhou et al. [29] proposed a model using latent dirichlet allocation, which incorporates the topical background of documents and social tags. Xu et al. [30] developed a language model for information retrieval

based on metadata property of social tags and their relationships to annotated documents. Bao et al. [31] introduced two ranking methods: SocialSimRank, which ranked pages based on the semantic similarity between tags and pages, and SocialPageRank, which ranked returned pages based on their popularity. Schenkel et al. [32] developed a top- $k$  algorithm which ranked search results based on the tags shared by the user who issued the query and the users who annotated the returned documents with the query tags.

**2.2. On Measuring Images Similarity.** Measuring semantic similarity is a basic issue in computer vision field. Usually some low-level visual features are used for similarity measures. For example, shape features, texture features, and gradient features can be extracted from images. Based on the extracted low-level features, distance metrics such as the Euclidean distance, the Chi-Square distance, the histogram intersection, and the EMD distance are used. In this paper, the proposed method addresses the problem by semantic-level features such as social tags.

Different from the methods using low-level features, recently, a number of papers build image representation based on the outputs of concept classifiers [33]. Our observation is that Flickr provides the related social tags by web users, which reflect how people on the internet tend to annotate images. Several previous methods [34] learn object models from internet images. These methods tend to gather training examples using image search results. Besides, their approaches have to alternate between finding good examples and updating object in order to robust against noisy images. On the other hand, some papers [35] use images from Flickr groups other than search engines, which is claimed to be clean enough to produce good classifiers.

## 3. Problem Definition

In this paper, we study the problem of measuring semantic relatedness between images or videos with manually provided social tags. Here, a social tag refers to some concepts provided by users, which is semantically related to the content of an image or a video. The input of the proposed method is a pair of images or videos with social tags. The goal of the proposed method is to identify the semantic relatedness between two

images or videos. Figure 2 shows the illustration of a pair of images from Flickr with social tags. These two images are about “Big Ben” and “London eye”. These two images may be dissimilar according to the traditional similarity measurement, since they do not share some common low level visual similarity. But, these two images are semantic related since they are both the famous sightseeings of London. In the proposed method, we can compute their semantic relatedness though they may share little similar visual features.

**3.1. Basic Definitions.** We first introduce three important definitions in this paper, the social tags set of an image and the semantic relatedness between two images.

**Definition 1** (social tags set of an image). The social tags (denoted by  $t$ ) set of an image  $f$  (denoted by  $s(f)$ ) is a set of tags provided by users of an image:

$$s(f) = \{t_1, t_2, \dots, t_{|s(f)|}\}. \quad (1)$$

For example, in Figure 2, the tags of the right images are “London” and “eye” other than “London eye”. Since Flickr provides the related tags of each image, we just download the tags by Flickr. We do not perform any NLP operations on the tags.

**Definition 2** (semantic relatedness between tags). The semantic relatedness between tags (denoted by  $sr(t_1, t_2)$ ) is the expected correlation of a pair of tags  $t_1$  and  $t_2$ .

**Definition 3** (semantic relatedness between images). The semantic relatedness between images (denoted by  $sr(f_1, f_2)$ ) is the expected correlation of a pair of images  $f_1$  and  $f_2$ .

The range of  $sr(t_1, t_2)$  and  $sr(f_1, f_2)$  is from 0 to 1. A high value indicates that semantic relatedness between tags or images is more likely to be confidential. Please notice that the definition of  $sr(f_1, f_2)$  can also be extended to videos with social tags.

**3.2. Basic Heuristics.** Based on common sense and our observations on real data, we have five heuristics that serve as the base of our computation model.

**Heuristic 1.** Usually each tag of an image appears only one time.

Different from writing sentences, users usually annotate an image with different tags. For example, the possibility of using tags “apple apple apple” for an image is very low. Therefore, in this paper, we do not employ any weighting scheme for tags such as tf-idf [36].

**Heuristic 2.** The order of the tags may reflect the correlation against the annotated image.

Different tag reflects the different aspects of an image. According to Heuristic 1, the weight of a tag against the image cannot be obtained. Fortunately, the order of the tags can be gotten since user may provide tags one by one.

**Heuristic 3.** The number of tags of an image may not be relevant to the annotation correctness.

Different users may give different tags about the same image. For example, users may give tags such as “apple iPhone” or “iPhone4 mobile phone” for the same image about iPhone. It is hardly to say which tag is better for annotation though the latter annotation has three tags.

**Heuristic 4.** Usually some tags may be redundant for annotating an image.

Of course, users may give similar tags for an image. For example, the tag “apple iPhone” may be redundant since iPhone is very semantic similar to apple.

**Heuristic 5.** Usually some tags may be noisy for annotating an image.

Users may give inappropriate or even false tags for an image. For example, the tags “iPhone” are false for an image about the iPod.

## 4. Computation Model

In this section, we propose the computation model for measuring semantic relatedness between images. Based on the above five heuristics, the social tags provided by users are used in our computation model. Overall, the proposed computation model is divided into three steps.

- (1) Tag relatedness computation. In this step, based on Heuristic 1, all of the tag pairs between two images are computed.
- (2) Semantic relatedness integration. In this step, based on Heuristics 3–5, we measure semantic relatedness between images.
- (3) Tag order revision. In this step, based on Heuristic 2, the image relatedness on step 2 is revised.

Table 1 shows the variables and parameters used in the following discussion. Figure 3 illustrates an overview of the proposed computation model.

**4.1. Tag Relatedness Computation.** According to Definition 1, an image can be represented as a set of tags provided by users. As for the semantic relatedness of a pair of images, we can measure the semantic relatedness between tags of these images. For example, two images with tags “apple iPhone” and “iPod Nano”, we can measure the semantic relatedness between these tags. Since the number of each tag is usually one according to Heuristic 1, the semantic relatedness between tags can be computed without considering their weight.

Many different methods of semantic relatedness measures between concepts have been proposed, which can be divided

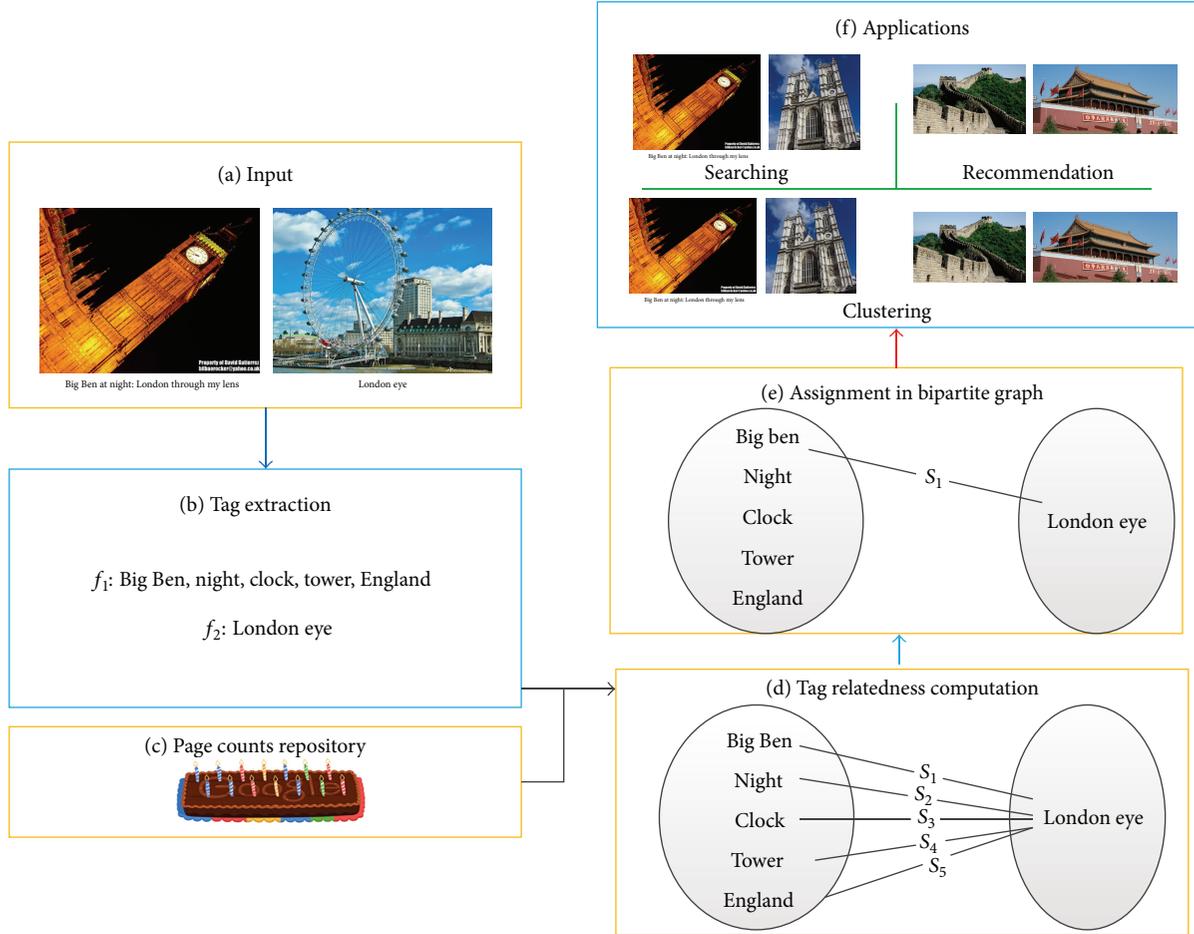


FIGURE 3: The illustration of the proposed method.

TABLE 1: The variables and parameters used in the proposed computation model.

Name	Description
$f$	An image
$t$	A tag
$s(f)$	Tags set of an image
$sr(t_1, t_2)$	Semantic relatedness of two tags
$sr(f_1, f_2)$	Semantic relatedness of two images
$N(t)$	Page counts of a tag
$N(s(f))$	Set of page counts of an image
$pos(t)$	Position information of a tag

into two aspects [37]: taxonomy-based methods and web-based methods. Taxonomy-based methods use information theory and hierarchical taxonomy, such as WordNet, to measure semantic relatedness. On the contrary, web-based methods use the web as a live and active corpus instead of hierarchical taxonomy.

In the proposed computation model, each tag can be seen as a concept with explicit meaning. Thus, we use some equations based on cooccurrence of two concepts to measure their semantic relatedness. The core idea is that “you shall know a word by the company it keeps” [38]. In this section,

four popular cooccurrence measures (i.e., Jaccard, Overlap, Dice, and PMI) are proposed to measure semantic relatedness between tags.

Besides cooccurrence measures, the page counts of each tag from search engine are used. Page counts mean the number of web pages containing the query  $q$ . For example, the page counts of the query “Obama” in Google (<http://www.google.com>) are 1,210,000,000 (the data was gotten in the date 9/28/2012). Moreover, page counts for the query “ $q$  AND  $p$ ” can be considered as a measure of cooccurrence of queries  $q$  and  $p$ . For the remainder of this paper, we use the notation  $N(p)$  to denote the page counts of the tag  $p$  in Google. However, the respective page counts for the tag pair  $p$  and  $q$  are not enough for measuring semantic relatedness. The page counts for the query “ $q$  AND  $p$ ” should be considered. For example, when we query “Obama” and “United States” in Google, we can find 485,000,000 Web pages; that is,  $N(\text{Obama} \cap \text{United States}) = 485,000,000$ . The four cooccurrence measures (i.e., Jaccard, Overlap, Dice, and PMI) between two tags  $p$  and  $q$  are as follows:

$$\text{Jaccard}(p, q) = \frac{N(p \cap q)}{N(p) + N(q) - N(p \cap q)}, \quad (2)$$

$p \cap q$  denotes the conjunction query “ $p$  AND  $q$ ”.

Consider

$$\text{Overlap}(p, q) = \frac{N(p \cap q)}{\min(N(p), N(q))}, \quad (3)$$

$\min(N(p), N(q))$  means the lower number of  $N(p)$  or  $N(q)$ .

Consider

$$\text{Dice}(p, q) = \frac{2 * N(p \cap q)}{N(p) + N(q)}. \quad (4)$$

According to probability and information theory, the mutual information (MI) of two random variables is a quantity that measures the mutual dependence of the two variables. Pointwise mutual information (PMI) is a variant of MI (see (5)):

$$\text{PMI}(p, q) = \frac{\log((N * N(p \cap q)) / (N(p) * N(q)))}{\log N}, \quad (5)$$

where  $N$  is the number of Web pages in the search engine, which is set to  $N = 10^{11}$  according to the number of indexed pages reported by Google.

Through (2)–(5), we can compute the tag relatedness as follows.

- (1) Extracting the tags from two images  $f_1$  and  $f_2$ , which are denoted by

$$\begin{aligned} s(f_1) &= \{t_1, t_2, \dots, t_{|s(f_1)|}\}, \\ s(f_2) &= \{t_1, t_2, \dots, t_{|s(f_2)|}\}. \end{aligned} \quad (6)$$

- (2) Issue the tags from  $f_1$  and  $f_2$  as the query to the web search engine (in this paper, we choose Google for its convenient API (<http://developers.google.com>)), the page counts can be denoted by

$$\begin{aligned} N(s(f_1)) &= \{N(t_1), N(t_2), \dots, N(t_{|s(f_1)|})\}, \\ N(s(f_2)) &= \{N(t_1), N(t_2), \dots, N(t_{|s(f_2)|})\}. \end{aligned} \quad (7)$$

- (3) Computing the semantic relatedness between each tags pair from  $f_1$  and  $f_2$  by (2)–(5). For example, if we use PMI to compute tag semantic relatedness, the equation can be

$$\begin{aligned} \text{sr}(t_i, t_j) &= \frac{\log((N * N(t_i \cap t_j)) / (N(t_i) * N(t_j)))}{\log N}, \\ &t_i \in s(f_1) \wedge t_j \in s(f_2). \end{aligned} \quad (8)$$

From the above steps, the tags relatedness can be computed, which is denoted as a triple  $\langle t_i, t_j, \text{sr}(t_i, t_j) \rangle$ . In the next section, we will give the detailed analysis for choosing the best measures from (2)–(5).

Overall, the page counts of each tag should be issued. Then some cooccurrence based measures are used to compute the semantic relatedness between tags. The reasons for using page counts based measures are as follows.

- (1) Appropriate computation complexity. Since the relatedness between each tag pair of two images should be computed, the proposed method must be with low complexity. Recently, web search engines such as Google provide API for users to index the page counts of each query. The web search engine gives an appropriate interface for the proposed computation model.
- (2) Explicit semantics. The tag given by users may not be a correct concept in taxonomy. For example, users may give a tag “Bling Bling” for an image about a lovely girl. The word “Bling” cannot be indexed in many taxonomy such as WorldNet. The proposed method uses web search engine as an open intermediate. The explicit semantics of the newly emerge concepts can be gotten by web easily.

**4.2. Semantic Relatedness Integration.** In Section 4.1, we compute the tag pair relatedness of two images. Obviously, the tag pair relatedness of two images  $f_1$  and  $f_2$  can be treated as a bipartite graph, which is denoted by

$$\begin{aligned} G &= (V, E), \\ V &= \{f_1, f_2\}, \end{aligned} \quad (9)$$

$$E = \langle t_i, t_j, \text{sr}(t_i, t_j) \rangle, \quad t_i \in s(f_1) \wedge t_j \in s(f_2).$$

Based on (9), we change the semantic relatedness integration of all tag pairs to the problem—assignment in bipartite graph. We want to assign a best matching of the bipartite graph  $G$ .

A matching is defined as  $M \subseteq E$  so that no two edges in  $M$  share a common end vertex. An assignment in a bipartite graph is a matching  $M$  so that each node of the graph has an incident edge in  $M$ . Suppose that the set of vertices are partitioned in two sets  $f_1$  and  $f_2$ , and that the edges of the graph have an associated weight given by a function  $f : (f_1, f_2) \rightarrow [0 \cdots 1]$ . The function  $\text{maxRel} : (f, f_1, f_2) \rightarrow [0 \cdots 1]$  returns the maximum weighted assignment, that is, an assignment so that the average of the weights of the edges is highest. Figure 4 shows a graphical representation of the semantic relatedness integration, where the bold lines constitute the matching  $M$ .

Based on the expression of the assignment in bipartite graphs, we have

$$\begin{aligned} \text{maxRel}(f, f_1, f_2) &= \begin{cases} \frac{\max \sum_{i \in I}^{j \in J} s(t_i, t_j)}{|s(f_1)|}, & |s(f_1)| \leq |s(f_2)|, \\ \frac{\max \sum_{i \in I}^{j \in J} s(t_i, t_j)}{|s(f_2)|}, & |s(f_1)| > |s(f_2)|, \end{cases} \quad (10) \\ I &= [1 \cdots |s(f_1)|], \quad J = [1 \cdots |s(f_2)|]. \end{aligned}$$

Using the assignment in bipartite graphs problem to our context, the variables  $f_1$  and  $f_2$  represent the two images to compute the semantic relatedness. For example, that  $f_1$  and

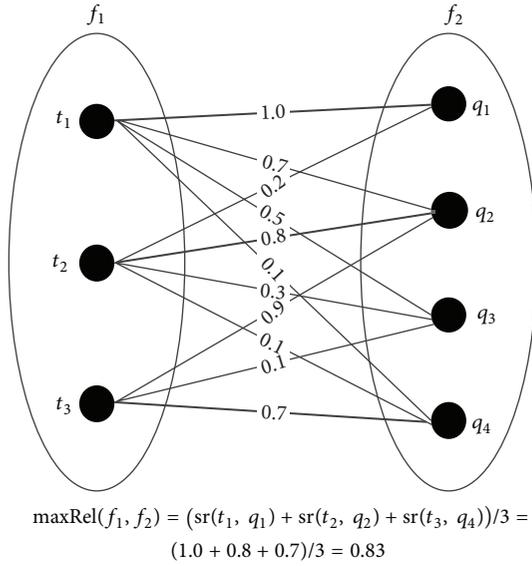


FIGURE 4: Graphical representation of the assignment in bipartite graphs problem.

$f_2$  are composed of the tags  $s(f_1)$  and  $s(f_2)$ .  $|s(f_1)| > |s(f_2)|$  means that the number of tags in  $s(f_2)$  is lower than that of  $s(f_1)$ . According to Heuristic 3, we divide the result of the maximization by the lower cardinality of  $s(f_1)$  or  $s(f_2)$ . In this way, the influence of the number of tags is reduced, and the semantic relatedness of two images is symmetric.

Beside the cardinality of two tags set  $s(f_1)$  and  $s(f_2)$ , the  $\maxRel$  function is affected by the relatedness between each pair of tags. According to Heuristics 4 and 5, the redundancy and noise should be avoided. In  $\maxRel$  function, the one-to-one map is applied to the tags  $s(f_1)$  and  $s(f_2)$ . Thus, the proposed  $\maxRel$  function varies with respect to the nature of two images.

Adopting the proposed  $\maxRel$  function, we are sure to find the global maximum relatedness that can be obtained pairing the elements in the two tags sets. Alternative methods are able to find only the local maximum since they scroll the elements in the first set and, after calculating the relatedness with all the elements in the second set, they select the one with the maximum relatedness. Since every element in one set must be connected, at most, at one element in the other set, such a procedure is able to find only the local maximum since it depends on the order in which the comparisons occur. For example, considering the example in Figure 4,  $t_1$  will be paired to  $q_1$  (weight = 1.0). But when analyzing  $t_3$ , the maximum weight is with  $q_2$  (weight = 0.9). This means that  $t_2$  can no more be paired to  $q_2$  even if the weight is maximum, since this is already matched to  $t_3$ . As a consequence,  $t_2$  will be paired to  $q_3$  and the average of the selected weights will be  $(1.0 + 0.3 + 0.9)/3 = 0.73$  which is considerably lower than using  $\maxRel$  where the sum of the weights was  $(1.0 + 0.8 + 0.7)/3 = 0.83$ .

Overall, the cardinality of two tag sets is used to follow Heuristic 3. The one-to-one map of tags pair is used to follow Heuristics 4 and 5. The  $\maxRel$  function is used to match a best semantic relatedness integration of two images.

**4.3. Tag Order Revision.** According to Heuristic 2, the order of tags should be considered to compute the semantic relatedness between two images. Intuitively, the tags appearing in the first position may be more important than the latter tags. Some researches [39] suggest that people used to select popular items as their tags. Meanwhile, the top popular tags are indeed the “meaningful” ones.

In this section, the  $\maxRel$  function proposed in Section 4.2 is revised considering the order of tags. For example, the relatedness of tags pair with high position should be enhanced, which is summarized as a constrain schema.

*Schema 1* (tag relatedness declining). This schema means that the identical tag pairs of two images  $f_1$  and  $f_2$  should be pruned in  $\maxRel$  function. In other words, the semantic relatedness of the same tag of two images is set as 0.

We add a decline factor to the  $\maxRel$  function, and the detailed steps are as follows.

- (1) According to the  $\maxRel$  function in Section 4.2, the best matching tag pairs are selected, which is denoted by

$$\maxRel(f_1, f_2) = \sum_{t_i \in s(f_1) \wedge t_j \in s(f_2)} sr(t_i, t_j), \quad (11)$$

Of course, the selected tag pairs are the best matching of the bipartite graph between images  $f_1$  and  $f_2$ .

- (2) Computing the position information of each tag, which is denoted by  $Pos(t_i)$ :

$$Pos(t_i) = \frac{|s(f)| + 1 - i}{|s(f)|}, \quad t_i \in s(f). \quad (12)$$

- (3) Add the position information of each tag to (11), which can be seen as a decline factor:

$$sr(f_1, f_2) = \sum_{t_i \in s(f_1) \wedge t_j \in s(f_2)} Pos(t_i) * sr(t_i, t_j) * Pos(t_j), \quad (13)$$

- (4) Of course, similar to  $\maxRel$  function, equation should divide the result of the maximization by

$$sr(f_1, f_2) = \frac{\sum Pos(t_i) * sr(t_i, t_j) * Pos(t_j)}{\sum Pos(t_i) * Pos(t_j)}. \quad (14)$$

We also consider the example in Figure 4. According to (14), the semantic relatedness is revised as

$$\left(1 \cdot 1.0 \cdot 1 + \frac{2}{3} \cdot 0.8 \cdot \frac{3}{4} + \frac{1}{3} \cdot 0.7 \cdot \frac{1}{4}\right) \times \left(1 \cdot 1 + \frac{2}{3} \cdot \frac{3}{4} + \frac{1}{3} \cdot \frac{1}{4}\right)^{-1} = 0.92. \quad (15)$$

Besides adding decline factor to the  $\maxRel$  function, we also add a constrain schema: identical tag pruning.

```

Input: The tags set of two images  $f_1$  and  $f_2$ , which is  $s(f_1)$  and  $s(f_2)$ 
Output: The semantic relatedness of two images  $f_1$  and  $f_2$ 
  for each  $t_i \in s(f_1)$  /* page counts and position initial */
     $N(s(f_1)) \leftarrow N(t_i)$ ;
     $\text{Pos}(s(f_1)) \leftarrow \text{Pos}(t_i)$ ;
    for each  $t_j \in s(f_2)$ 
       $N(s(f_2)) \leftarrow N(t_j)$ ;
       $\text{Pos}(s(f_2)) \leftarrow \text{Pos}(t_j)$ ;
    for each  $t_i \in s(f_1)$ 
      for each  $t_j \in s(f_2)$ 
        if  $(t_i == t_j)$   $\text{sr}(t_i, t_j) = 0$ ; /* pruning */
        else  $\text{sr}(t_i, t_j) = f(N(t_i), N(t_j))$ ; /* relatedness */
  return  $\text{maxRel}(f_1, f_2) = f(\text{Pos}(t_i), \text{Pos}(t_j), \text{sr}(t_i, t_j))$ ;

```

ALGORITHM 1: MaxRel.

*Schema 2* (identical tag pruning). This schema means that the identical tag pairs of two images  $f_1$  and  $f_2$  should be pruned in MaxRel function. In other words, the semantic relatedness of the same tag of two images is set as 0.

The above schema is used to ensure the relatedness measures of two images. If we do not prune the identical tag pairs of two images, the proposed method will be transformed to the similarity measures. For example, the cosine similarity [36] between two tags is to find the number of identical elements of two vectors. The overall algorithm of the proposed computation mode is presented in Algorithm 1.

## 5. Experimental Results

In this section, we evaluate the results of using the proposed method for relatedness measurement. In Section 5.1, we introduce the data set for the evaluation. In Section 5.2, we determine to use the cooccurrence function for tag relatedness measures. In Sections 5.3 and 5.4, clustering and retrieval are used to evaluate the proposed method.

*5.1. The Data Sets.* We choose Flickr groups as the resources for building data sets. Users on online photo sharing sites like Flickr have organized many millions of photos into hundreds of thousands of semantically themed groups. These groups expose implicit choices that users make about which images are similar. Flickr group membership is usually less noisy than Flickr tags because images are screened by group members. We download 1000 images from ten groups. These ten groups can be divided into two classes. The first class includes five groups, which are car, phone, flower, dog, and boat. The second class consists of another five groups, which are Louis Vuitton, Dior, Gucci, Cartier, and Chanel. Of course, these images are selected by humans, which reduce the noise of the data set. The reason why we choose two classes of groups is that we want to test the accuracy of the proposed method against the semantic relatedness of data set. The semantic relatedness of the second set is higher than the first set since the second class is all about the luxury brands. For example, almost all these brands produce handbags. Thus, if the proposed method can do well in these groups, we may say

TABLE 2: The detailed information of the data set.

Group 1	Average tags per image	Group 2	Average tags per image
Car	4.4	Louis Vuitton	3.1
Phone	3.5	Dior	3.2
Flower	2.2	Gucci	2.9
Dog	5.6	Cartier	2.8
Boat	3.1	Chanel	2.6

that it can measure the semantic relatedness between Flickr images accurately and robustly. Table 2 gives the detailed information of the data set. Table 3 gives some selected tags from group 2.

*5.2. Relatedness Function Selection.* In Section 4.1, four cooccurrence measures (i.e., Jaccard, Overlap, Dice, and PMI) are given for relatedness measures between tags. In [40], Rubenstein and Goodenough proposed a data set containing 28 word pairs rating by a group of 51 human subjects, which is a reliable benchmark for evaluating semantic similarity measures. The higher the correlation coefficient against R-G ratings is, the more accurate the methods for measuring semantic similarity between words are. Figure 5 gives the correlation coefficient of four functions against R-G test set. From Figure 5, we can say that PMI performs best on relatedness measures for its highest correlation coefficient. Thus, in the later experiments, we select PMI as the relatedness measures between tags.

*5.3. Evaluation on Image Clustering.* In this section, we evaluate the correctness of using tag order. In Section 4.3, we add the position information of each tag to the semantic relatedness measures. The tags with high position are treated as the major element for semantic relatedness measures. We evaluate the using of tag order by the clustering task. We employ the proposed semantic relatedness of images into  $K$ -means [41] clustering model. Since the  $K$ -means model depends on the initial points, we random select core points 100 times. We evaluate the effectiveness of document clustering with three quality measures:  $F$ -measure, Purity, and Entropy [41]. We treat each cluster as if it were the result of the proposed

TABLE 3: The selected tags of group 2 from Flickr.

Group 2	Tags	Tags	Tags	Tags	Tags
Louis Vuitton	“Louis Vuitton” “Keepall”	“Louis Vuitton” “Alma”	“Louis Vuitton” “Tivoli”	“Louis Vuitton” “Bolsas”	“LV” “Multicolore”
Dior	“DIOR” “lipstick” “makeup”	“Dior” “Diorskin Nude” “Tan Sun Powder”	“Dior” “Makeup” “Palette”	“Dior” “Addict 2”	“Dior” “Jadore” “Perfume”
Gucci	“Gucci” “Leather Belts”	“Gucci” “Trainers”	“Gucci” “Jolie Leopard” “Orange”	“Replica” “Gucci” “Handbags”	“Gucci” “Cruise”
Cartier	“Cartier” “Pasha” “Chronograph”	“CARTIER” “Love Bracelet”	“Cartier” “Santos Galbee”	“Calibre” “Cartier”	“Cartier Watch” “Tank Francaise”
Chanel	“Chanel” “Coco Noir”	“Chanel” “Chanel Riva” “Chanel nail polish”	“Coco Mademoiselle”	“Chanel” “No 5”	“Chance” “Chanel”

method and each class as if it were the desired set of images. Generally, we would like to maximize the  $F$ -measure and Purity and minimize the Entropy of the clusters to achieve a high-quality document clustering. Moreover, we compare the clustering results between the proposed method using tag order or not. Figures 6 and 7 give the clustering results of group 1 and group 2 data sets. From Figures 6 and 7, we can conclude the following.

- (1) The proposed method performs better than cosine based clustering. This result can be obtained from Figures 6 and 7. The three metrics including  $F$ -measure, purity, and entropy of the proposed method are better than cosine based clustering. This may be caused by the inherent feature of the proposed method. The proposed method is based on the semantic relatedness other than the cooccurrence of the cosine based clustering. If the tags of two images are not overlapped, the cosine based clustering may be unavailable.
- (2) The schema on using of tag order is effective. This result can also be obtained from Figures 6 and 7. The three metrics including  $F$ -measure, purity, and entropy of using tag order are the highest. The position information reflects the importance of each tag. The proposed method emphasizes the tags with high order, which raises the performance on images clustering.
- (3) The proposed method is robust in different data sets. The proposed method performs well in group 1 and group 2 data set. It is worth noting that the difference between the proposed method and cosine method of group 2 is higher than that of group 1. The reason of that is due to the semantic correlation of group 2 being stronger than group 1. In other words, the performance of the proposed method relies on the semantic correlation of classes in data sets. The stronger the semantic correlation between classes of data, the better the proposed method performance.

5.4. *Evaluation on Image Searching.* In this section, we evaluate the proposed method query-based image searching task. Five queries from group 2 are selected as the test set including “Louis Vuitton,” “Gucci,” “Chanel,” “Cartier,” and “Dior”. These queries are searched in Flickr. The top 50 images are obtained as the data set. Moreover, we remove the queries on the tags of each image. For example, the tag “Cartier” of the top 50 images is removed of the query “Cartier”. The reason for that operation is that the proposed method is based on the semantic relatedness other than cooccurrence. We choose cut-off point precision to evaluate the proposed method on image searching. The cut-off point precision ( $P^n$ ) means that the percentage of the correct result of the top  $n$  returned results. We compute the  $P^1$ ,  $P^5$ , and  $P^{10}$  of the group 2 test set. Table 4 lists the comparison of the cut-off point precision between the proposed method and Flickr. From the experimental results, we can conclude the following.

- (1) The proposed method performs better than Flickr. In Table 4, the  $P^1$ ,  $P^5$ , and  $P^{10}$  of the proposed method are higher than Flickr. The experimental results prove the correctness of the proposed method on image searching task.
- (2) The proposed method can handle the relatedness searching problem. The proposed method can measure the semantic relatedness of two images robustly and correctly.
- (3) The proposed method can support the faceted exploration of image search. Faceted exploration of search results is widely used in search interfaces for structured databases. Recently the faceted exploration is also appearing in online search engine in the form of search assistants. The proposed method can measure the semantic relatedness of two images. Given the search queries, we can select the related images for faceted search.

TABLE 4: The comparison of the cut-off point precision between the proposed method and Flickr.

Cut-off point	Louis Vuitton	Gucci	Dior	Chanel	Cartier
$P^1$	100%	100%	100%	100%	100%
$P^1$ (Flickr)	100%	100%	0	100%	100%
$P^5$	100%	100%	100%	100%	100%
$P^5$ (Flickr)	80%	60%	60%	60%	80%
$P^{10}$	100%	100%	100%	100%	100%
$P^{10}$ (Flickr)	90%	70%	70%	80%	80%

## 6. Conclusions

This paper mainly discusses the semantic relatedness measures systematically, puts forward a method to measure the semantic relatedness of two images based on their tags, and justifies its validity through the experiments. The major contributions are summarized as follows.

- (1) We propose a framework to measure semantic relatedness between Flickr images using tags. Firstly, the cooccurrence measures are used to compute the relatedness of tags between two images. Secondly, we transform the tags relatedness integration to the assignment in bipartite graph problem, which can find an appropriate matching to the semantic relatedness of images. Finally, a decline factor considering the position information of tags is used in the proposed framework, which reduces the noise and redundancy in the social tags.
- (2) A real data set including 1000 images from Flickr with ten classes is used in our experiments. Two evaluation methods including clustering and searching are performed, which shows that the proposed method can measure the semantic relatedness between Flickr images accurately and robustly.
- (3) We extend the relatedness measures between concepts to the level of images. Since the association relation is the basic mechanism of brain. The proposed relatedness measurement can facilitate related applications such as searching and recommendation.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work was supported in part by the National Science and Technology Major Project under Grant no. 2013ZX01033002-003, in part by the National High Technology Research and Development Program of China (863 Program) under Grant nos. 2013AA014601 and 2013AA014603, in part by

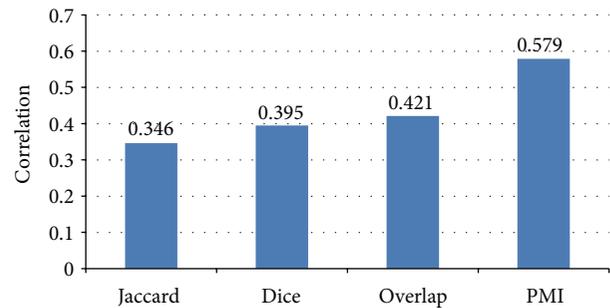


FIGURE 5: The correlation of four selected functions.

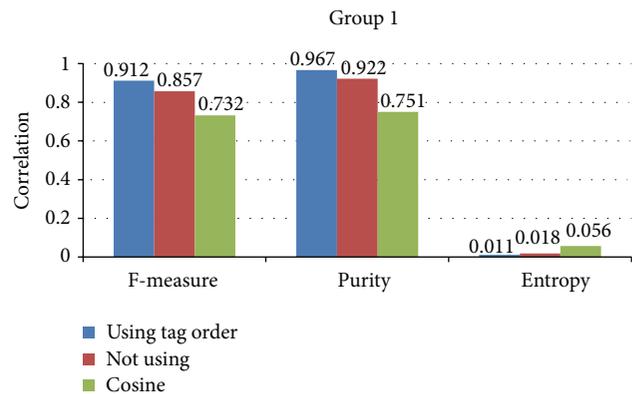


FIGURE 6: The clustering results of group 1 data sets.

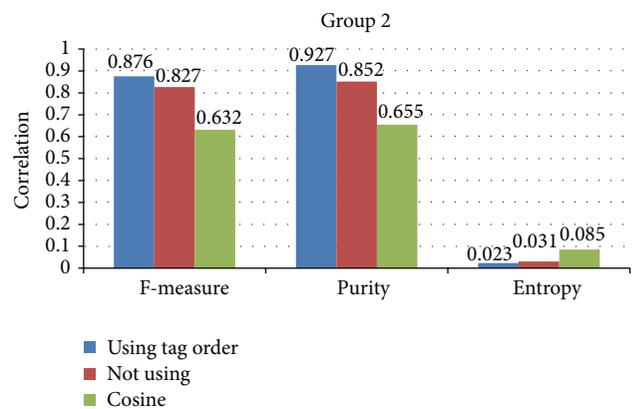


FIGURE 7: The clustering results of group 2 data sets.

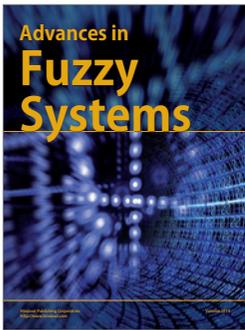
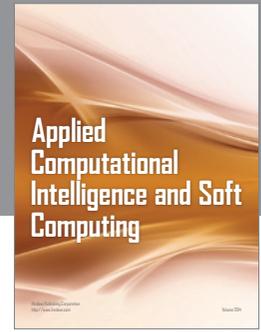
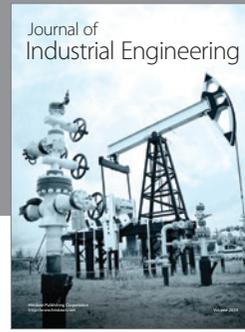
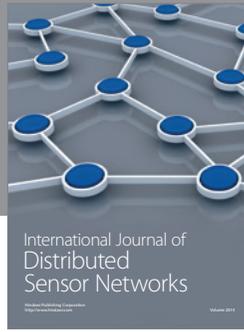
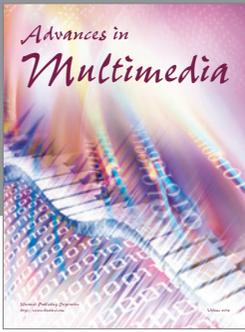
National Key Technology Support Program under Grant no. 2012BAH07B01, in part by the National Science Foundation of China under Grant no. 61300202, and in part by the Science Foundation of Shanghai under Grant no. 13ZR1452900.

## References

- [1] J. Goldberger, S. Gordon, and H. Greenspan, "Unsupervised image-set clustering using an information theoretic framework," *IEEE Transactions on Image Processing*, vol. 15, no. 2, pp. 449–458, 2006.
- [2] T. Evgeniou, M. Pontil, C. Papageorgiou, and T. Poggio, "Image representations and feature selection for multimedia database search," *IEEE Transactions on Knowledge and Data Engineering*, vol. 15, no. 4, pp. 911–920, 2003.

- [3] R. Ji, H. Yao, X. Sun, B. Zhong, and W. Gao, "Towards semantic embedding in visual vocabulary," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 918–925, June 2010.
- [4] J. Fan, D. A. Keim, Y. Gao, H. Luo, and Z. Li, "JustClick: personalized image recommendation via exploratory search from large-scale Flickr images," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 2, pp. 273–288, 2009.
- [5] T. Gong, S. Li, and C. L. Tan, "A semantic similarity language model to improve automatic image annotation," in *Proceedings of the 22nd International Conference on Tools with Artificial Intelligence (ICTAI '10)*, pp. 197–203, October 2010.
- [6] C. Schmid and R. Mohr, "Local grayvalue invariants for image retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pp. 530–535, 1997.
- [7] M. Varma and A. Zisserman, "A statistical approach to texture classification from single images," *International Journal of Computer Vision*, vol. 62, no. 1-2, pp. 61–81, 2005.
- [8] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 509–522, 2002.
- [9] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, vol. 1, pp. 886–893, June 2005.
- [10] D. Huang, M. Ardabilian, Y. Wang, and L. Chen, "Asymmetric 3D/2D face recognition based on LBP facial representation and canonical correlation analysis," in *Proceedings of the 16th IEEE International Conference on Image Processing (ICIP '09)*, pp. 3325–3328, November 2009.
- [11] L. Wang, Y. Zhang, and J. Feng, "On the Euclidean distance of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1334–1339, 2005.
- [12] W. Jia, H. Zhang, X. He, and Q. Wu, "Gaussian weighted histogram intersection for license plate classification," in *Proceedings of the 18th International Conference on Pattern Recognition (ICPR '06)*, pp. 574–577, August 2006.
- [13] Y. Rubner, C. Tomasi, and L. J. Guibas, "A Metric for distributions with applications to image databases," in *Proceedings of the IEEE 6th International Conference on Computer Vision*, pp. 59–66, January 1998.
- [14] L. Wu, X.-S. Hua, N. Yu, W.-Y. Ma, and S. Li, "Flickr distance: a relationship measure for visual concepts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 5, pp. 863–875, 2012.
- [15] D. Cai, "An information-theoretic foundation for the measurement of discrimination information," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 9, pp. 1262–1273, 2010.
- [16] P. van den Broek, "Using texts in science education: cognitive processes and knowledge representation," *Science*, vol. 328, no. 5977, pp. 453–456, 2010.
- [17] C. Bizer, T. Heath, and T. Berners-Lee, "Linked data—the story so far," *International Journal on Semantic Web and Information Systems*, vol. 5, no. 3, pp. 1–22, 2009.
- [18] H. Zhuge, "Communities and emerging semantics in semantic link network: discovery and learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 6, pp. 785–799, 2009.
- [19] H. Zhuge, "Semantic linking through spaces for cyber-physical-socio intelligence: a methodology," *Artificial Intelligence*, vol. 175, no. 5-6, pp. 988–1019, 2011.
- [20] X. Luo, Z. Xu, J. Yu, and X. Chen, "Building association link network for semantic link on web resources," *IEEE Transactions on Automation Science and Engineering*, vol. 8, no. 3, pp. 482–494, 2011.
- [21] S. A. Golder and B. A. Huberman, "Usage patterns of collaborative tagging systems," *Journal of Information Science*, vol. 32, no. 2, pp. 198–208, 2006.
- [22] H. S. Al-Khalifa and H. C. Davis, "Measuring the semantic value of folksonomies," in *Proceedings of the Innovations in Information Technology (IIT '06)*, pp. 1–5, November 2006.
- [23] F. M. Suchanek, M. Vojnović, and D. Gunawardena, "Social tags: meaning and suggestions," in *Proceedings of the 17th ACM Conference on Information and Knowledge Management (CIKM '08)*, pp. 223–232, October 2008.
- [24] H. Halpin, V. Robu, and H. Shepherd, "The complex dynamics of collaborative tagging," in *Proceedings of the 16th International World Wide Web Conference (WWW '07)*, pp. 211–220, May 2007.
- [25] C. Cattuto, C. Schmitz, A. Baldassarri et al., "Network properties of folksonomies," *AI Communications*, vol. 20, no. 4, pp. 245–262, 2007.
- [26] R. Lambiotte and M. Ausloos, "Collaborative tagging as a tripartite network," in *Computational Science*, vol. 3393 of *Lecture Notes in Computer Science*, pp. 1114–1117, 2006.
- [27] U. Maulik, S. Bandyopadhyay, and I. Saha, "Integrating clustering and supervised learning for categorical data analysis," *IEEE Transactions on Systems, Man, and Cybernetics A*, vol. 40, no. 4, pp. 664–675, 2010.
- [28] D. Ramage, P. Heymann, C. D. Manning, and H. Garcia-Molina, "Clustering the tagged web," in *Proceedings of the 2nd ACM International Conference on Web Search and Data Mining (WSDM '09)*, pp. 54–63, February 2009.
- [29] D. Zhou, J. Bian, S. Zheng, H. Zha, and C. L. G. C. Lee Giles, "Exploring social annotations for information retrieval," in *Proceedings of the 17th International Conference on World Wide Web (WWW '08)*, pp. 715–724, April 2008.
- [30] S. Xu, S. Bad, Y. Cao, and Y. Yu, "Using social annotations to improve language model for information retrieval," in *Proceedings of the 16th ACM Conference on Information and Knowledge Management (CIKM '07)*, pp. 1003–1006, November 2007.
- [31] S. Bao, G. Xue, X. Wu, Y. Yu, B. Fei, and Z. Su, "Optimizing web search using social annotations," in *Proceedings of the 16th International World Wide Web Conference (WWW '07)*, pp. 501–510, May 2007.
- [32] R. Schenkel, T. Creclius, M. Kacimi et al., "Efficient top-k querying over social-tagging networks," in *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (ACM SIGIR '08)*, pp. 523–530, July 2008.
- [33] N. Rasiwasia, P. J. Moreno, and N. Vasconcelos, "Bridging the gap: query by semantic example," *IEEE Transactions on Multimedia*, vol. 9, no. 5, pp. 923–938, 2007.
- [34] R. Fergus, L. Fei-Fei, P. Perona, and A. Zisserman, "Learning object categories from Google's image search," in *Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV '05)*, vol. 2, pp. 1816–1823, October 2005.

- [35] G. Wang, D. Hoiem, and D. Forsyth, "Learning image similarity from flickr groups using fast kernel machines," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2177–2188, 2012.
- [36] G. Salton, A. Wong, and C. S. Yang, "A vector space model for automatic indexing," *Communications of the ACM*, vol. 18, no. 11, pp. 613–620, 1975.
- [37] Z. Xu, X. Luo, J. Yu, and W. Xu, "Measuring semantic similarity between words by removing noise and redundancy in web snippets," *Concurrency Computation Practice and Experience*, vol. 23, no. 18, pp. 2496–2510, 2011.
- [38] R. Firth, "A synopsis of linguistic theory 1930–1955," in *Studies in Linguistic Analysis*, Philological Society, Oxford, UK, 1957.
- [39] M. Vojnović, J. Cruise, D. Gunawardena, and P. Marbach, "Ranking and suggesting popular items," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 8, pp. 1133–1146, 2009.
- [40] H. Rubenstein and B. Goodenough, "Contextual correlates of synonymy," *Communications of the ACM*, vol. 8, no. 10, pp. 627–633, 1965.
- [41] M. Steinbach, G. Karypis, and V. Kumar, "A comparison of document clustering techniques," in *Proceedings of the KDD Workshop on Text Mining*, 2000.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

