



# Multi-level diversification approach of semantic-based image retrieval results

Mariam Bouchakwa<sup>1</sup> · Yassine Ayadi<sup>1</sup> · Ikram Amous<sup>1</sup>

Received: 2 March 2019 / Accepted: 8 June 2019 / Published online: 22 June 2019  
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

## Abstract

With the increasing popularity of social photograph-sharing Web sites, a huge mass of digital images, associated with a set of tags voluntarily introduced by amateur photographers, is daily hosted and consequently, the *Tag-based social Image Retrieval* technique has been widely adopted. However, tag-based queries are often too ambiguous and abstract to be considered as an efficient solution for the retrieval of the most relevant images that meet the users' needs. As an alternative, the *Semantic-based social Image Retrieval* technique has emerged for the purpose of retrieving the relevant images covering as much possible the topics that a given ambiguous query ( $q$ ) may have. Actually, the diversification strategies are a great challenge for researchers. In this context, we jointly investigate two processes at the ambiguous query preprocessing and postprocessing levels. On the one hand, we propose a *Tag-based Query Semantic Reformulation* process, which aims at reformulating the tag-based users' queries, according to multiple semantic facets of the different images' views, by using a set of predefined ontological semantic rules. On the other hand, we propose a *Multi-level Image Diversification process* that can first perform a two-level-based image clustering offline, and second, filter and re-rank the image cluster retrieval results according to their pertinence versus the reformulated query online. The experimental results and statistical analysis performed on a collection of 25.000 socio-tagged images shared on *Flickr* demonstrate the effectiveness of the proposed technique, which is compared with the research technique based on one-level-based image clustering, tag-based image research technique and recent CBIR techniques.

**Keywords** Socio-tagged images · Semantic annotation · Semantic retrieval · Semantic rules ·  $k$ -means clustering · Search result diversification

## 1 Introduction

Information is considered one of the most valuable assets in our time. Massive amounts of information are daily exchanged between people through the *World Wide Web*. In other words, we can say that this is the era of digital communication. In the beginning, the exchange of information was mainly used through text documents. Nevertheless, with the

technological progress, the exchange of information covers various other forms such as audio, video and image.

Due to the advancements of digital image acquisition devices, image capturing is no longer a difficult task. In fact, images have been increasingly used since they are one of the best ways of expressing, sharing and memorizing knowledge. Moreover, image databases of artworks, satellites, remote sensing, tourism, biology, medicine, etc., have attracted more and more users in various professional and amateur fields, which has created an urgent need for effective and efficient systems that help retrieve as quickly as possible the relevant visual information from a gigantic amount of image collections.

*Content-based Image Retrieval* (CBIR) systems [1–9] are intended to extract the low-level features of images (texture, color, shape, etc.) to describe their visual contents. The retrieval process consists in matching the visual features of a given query image with those of the image collection in order

---

✉ Mariam Bouchakwa  
mariam.bouhekwa@gmail.com

Yassine Ayadi  
ayadi.yassine@gmail.com

Ikram Amous  
ikram.amous@isecs.rnu.tn

<sup>1</sup> MIRACL Laboratory, Technopole of Sfax, University of Sfax, P.O. Box 242, 3031 Sfax, Tunisia



**Fig. 1** Two images with very close visual appearance, but two different semantic meanings

to get results that are visually similar. Nevertheless, extensive experiments on CBIR systems have proved that the low-level content often fails to describe the top-level semantic concepts in the user's mind [10]. The discrepancy between the limited description of the low-level image features and the richness of the users' semantics is referred to as “*semantic gap*” [4, 11]. Indeed, human eyes discriminate images according to their visual contents. When we apply a feature extraction technique to the images that have a similar visual appearance, it can produce close feature vectors values, which reduce the performance of the adopted CBIR system [12]. To illustrate the “*semantic gap*” problem is the example of the two images in Fig. 1. Although these images have a very close visual appearance, they have two different semantic meanings: The foreground of the first image includes flowers, while the second image shows a man playing golf [10].

Recent research studies have addressed the “*semantic gap*” problem on the basis of visual words fusion of various types of descriptors [12–16]. Other important research studies are based on the addition of image spatial attributes to the image retrieval [17, 18]. Extensive experiments have demonstrated the effectiveness of these techniques in improving the performance of CBIR systems by reducing the “*semantic gap*” problem. However, the common problems faced by CBIR systems consist of the time complexity and the extensive computational cost. In addition, since it is impractical for users to use CBIR technology because they need to provide query images, most of them prefer to perform a *Text-based Image Retrieval* (TBIR) [19].

In fact, the TBIR technology has been widely adopted in various real-world applications, like biomedical [20], radiology [21], data mining [22], art painting [23], image research engines on the Web (such as *Google*<sup>1</sup> and *Yahoo*<sup>2</sup>), social media (such as *Flickr*<sup>3</sup> and *Instagram*<sup>4</sup>), etc. More precisely, this technique focuses on the return of potentially relevant images compared to a textual research query described by a

user [24]. As consequence, it becomes necessary to associate a set of keywords with each image stored in the database. Moreover, a set of formulas must be defined to reformulate the request initially sent by a user. The retrieval process consists in matching the image keywords with the reformulated query in order to get results that meet the user's needs. However, the major drawback of the TBIR-based methods is that the large amounts of image data require annotation, which is a hard and a time-wasting process for the experts in the domain. On the other hand, with the spread of Web 2.0, large-scale image-sharing services emerged. In fact, the collaborative aspect of this type of Web services has given users the opportunity not only to share images with their families, friends and the online community at large, but also to associate them with user-contributed data called tags. Just as an illustration, 350 million tagged images are daily added to *Facebook*,<sup>5</sup> 80 million to *Instagram*<sup>6</sup> and 4.5 million to *Flickr*<sup>7</sup> service. Thus, the tag-based query, which enables the users to formulate their queries using the tagging information, has become the intuitive way to perform a *Text-based Image Retrieval*.

Extensive research studies proved that the tag-based queries are often ambiguous and typically short [25, 26]. In fact, two major reasons are behind this ambiguity: First, the rich nature of the image content can be tough to be described by the social community using the limited expressiveness of tags. Second, the tag-based queries can be interpreted with several meanings other than users' expectations. Due to the users' tag ambiguity issue, image search engines often suffer from the inability to understand the context in question and consequently to provide images that do not meet the users' needs. The heterogeneity of the semantic contents of the images, which are provided as results, makes the browsing performed by the users to select the desired images a tedious and a time-consuming task. For instance, for the tag-based query “*apple*,” images of *apple fruit* are intermixed with images of *fiona apple*, *apple ipod*, *apple computer*, *apple pie*, *apple logo*, and so on.

To deal with these issues, researchers have oriented their efforts toward a *Semantic-based Image Retrieval* (SBIR). Indeed, coverage-based search result diversification is an effective solution in case of an absence of exact knowledge about the users' intentions and their contexts. This solution consists in retrieving the relevant images covering as much possible the topics that a given ambiguous query (*q*) may have. Then, the retrieved documents can be re-ranked. Two postprocessing strategies are distinguished [26]: clustering-based strategy and diversification-based strategy.

<sup>1</sup> <https://www.google.com/>.

<sup>2</sup> <https://fr.yahoo.com/>.

<sup>3</sup> <https://www.flickr.com/>.

<sup>4</sup> <https://www.instagram.com/>.

<sup>5</sup> <http://www.culturecrossmedia.com/reseaux-sociaux/evolution/>.

<sup>6</sup> <https://www.instagram.com/press/>.

<sup>7</sup> <http://www.kullin.net/2012/01/instagram-now-growing-faster-than-flickr/>.

In this paper, we propose a novel *Semantic-based Social Image Retrieval* technique in order to deal with the aforementioned issues. This technique is based on two processes at ambiguous query preprocessing and postprocessing levels. At the ambiguous query preprocessing level, we propose to reformulate the tag-based queries using a process called “*Tag-based Query Semantic Reformulation process*” (TQSR). This process is approached as follows: First, we deduce the different semantic meanings that an ambiguous user’s query can have based on a domain ontology. Thereafter, we reformulate the initial query by executing a set of predefined ontological semantic rules. As a result, we obtain a semantic concept-based query that expresses the semantic content of the desired images according to different semantic facets (*taxonomic, temporal, spatial and qualification*) of the different semantic views (*structural, behavioral and event*). This query is then automatically transformed into SPARQL language to interrogate the target images. On the other hand, at the ambiguous query postprocessing level, we propose a process called “*Multi-level Image Diversification process*” (MLID). The main idea behind this process is gathering the semantically similar images in a same cluster. This is approached at both levels: *semantic facet-based clustering level* and *semantic concept-based clustering level*. The interrogated clusters are then filtered and re-ranked according to their semantic similarity degrees versus the reformulated query. To support a lightweight image search result clustering system, we propose to precompute time-consuming computation tasks, including image similarity measures, image clustering and representative image selection, offline.

The performance of the proposed technique is compared to the image research technique adopting a one-level-based image clustering, the tag-based image research technique and the recent state-of-the-art CBIR techniques. The proposed image research technique adopting a two-level-based image clustering significantly improves the performance of Tag-BIR compared to the image research technique adopting a one-level-based image clustering and the tag-based image research technique. Moreover, the proposed technique outperforms the state-of-the-art CBIR methods on standard image databases.

The main contributions of this research paper are as follows:

1. A novel technique at the ambiguous query preprocessing level enables to move from a tag-based query to a semantic-based one. Hence, the reformulated query can express the content of the desired images according to different semantic views and facets.
2. A novel technique at the ambiguous query postprocessing level enables to relieve and accelerate the task of finding the most relevant image for a user. This technique aims at:

- first, clustering the images of a collection according to their semantic content. The two-level-based image clustering (*semantic facet-based* and *semantic concept-based clustering levels*) improves the accuracy of the clustering results given the importance of the semantic vocabulary size;
  - second, filtering and re-ranking the image cluster retrieval results in order to reorder the retrieved results according to their pertinence versus the reformulated query, and to achieve higher accuracy rate.
3. Supporting the image search result clustering system by precomputing offline the time-consuming computation tasks, like image similarity measures, image clustering and representative image selection.

This paper is organized as follows: In Sect. 2, we provide an overview of the different directions of the research studies dealing with the query ambiguity challenge. In Sect. 3, we describe the general architecture of the proposed “*Semantic-based Social Image Retrieval System*” and provide an overview of the *Semantic Modeling* and *Automatic Annotation* processes. The *Clustering* and *Research processes* are detailed in Sects. 4 and 5, respectively. In Sect. 6, we provide experimental details and performance measurements of the proposed S2IR technique followed by an analysis of the computational complexity. Finally, Sect. 7 concludes the proposed technique and draws our future research studies.

## 2 Related works

In the literature, *query ambiguity problem* (QAP) has been recognized as a hard issue in tag-based image retrieval. To deal with this problem, an effective approach is to provide numerous results covering as many semantic meanings as possible that a query ( $q$ ) may have [27]. However, the diversifying search for results is declared as an NP-hard optimization problem [28], where the objective is to find a ranking of a set of documents  $R(q)$  with a maximum relevance to a given query ( $q$ ) and a minimum redundancy, which makes it cover all its possible underlying aspects. Within this context, two strategies can be considered to diversify the results following an ambiguous query [29]: a preprocessing strategy and a postprocessing strategy. In this section, we detail and discuss the approaches of diversifying research results in information retrieval (IR).

### 2.1 Ambiguous query preprocessing strategy

Sometimes, the user does not have a clear idea about what he/she wants, hence a failure to precisely describe the

intended search. Thus, the system can be faced with an ambiguous query. In such situations, poor-quality result lists are provided. Hence, it will be difficult for the user to find his/her needs. Therefore, it is indispensable to rely on an interactive search enabling to automatically reduce the ambiguity of the users' queries and send more satisfying search results for the browsers.

Expanding queries has been a promising solution for a long time, which offers the possibility of refining the initial queries in order to diversify the research results. Indeed, the diversification of the results increases the opportunity to find the desired documents by the browsers. The policy of the query expansion consists in adding additional meaningful terms to the users' queries, which reduces their ambiguity. These terms can be selected from an unstructured vocabulary or extracted from the informational richness of external knowledge resources, like ontologies and corpus. In this context, we explore different ambiguous query expansion techniques that have been addressed by the *Information Retrieval* (IR) approaches.

Query expansion based on *Pseudo-Relevance Feedback* (PRF) is a simple and effective method enabling to address the problem of users' query ambiguity in information retrieval [30]. The basic assumption of (PRF) is that the top-retrieved documents are often associated with many useful terms, in particular, users' tags, which can help distinguish between the relevant documents from irrelevant ones. Therefore, the policy of (PRF) is to refine the query initial terms by adding these feedback terms [31–34]. For example, Kitanovski et al. [34] demonstrate a complete implementation of three medical image retrieval types, namely: text-based, content-based and mixed retrieval together with a modality classification. In particular, text-based retrieval subsystem uses query expansion based on (PRF) technique, which has significantly boosted the performance of the main retrieval system. Indeed, the application of (PRF) technique consists in taking the top  $n$  number of initially retrieved documents and calculating the  $m$  most informative terms within them. The identified terms are thereafter added to the original query to execute the retrieval process according to the modified query. Lioma et al. [31] hypothesize the collaborative tags representing the semantic information that can make queries more informative and hence improve the performance of the retrieval systems. They adopt PRF as a technique of expansion keyword queries by using semantic annotations found in the freely available *Del.icio.us* collaborative tagging system. Experimental studies with three different techniques of enriching queries (based on individual terms, phrases and on whole queries) with *Del.icio.us* tags, as well as varying the number of tags used for expansion between 1 and 10, show improvement in retrieval precision, on a baseline of short keyword queries.

To summarize, (PRF) assumes that the target collections provide enough feedback information to select the effective expansion terms. This is often not the case in the image retrieval since most images have only short annotation metadata provoked by the *Incomplete Annotation Problem* (IAP). Contrariwise, some of the expansion terms extracted from feedback documents are irrelevant for the query and thus may affect the retrieval performance.

Other research studies have solved the problem of the query ambiguity by enriching the initial users' queries with new concepts that are semantically highly correlated with them. These concepts are derived from tags assigned to the resources shared on the Web. Therefore, it is interesting to focus on how researchers have treated the different semantic relationships between tags. Indeed, in [25, 35–40] the authors adopt some techniques for the measurement of the co-occurrence between concepts. For example, Jin et al. [36] suggest a method in which the used expansion terms are selected from a large volume of social tags in folksonomy. A tag co-occurrence method for similar term selection is adopted to choose the best expansion terms from the candidate tags according to their potential impact on the research effectiveness. For their part, Ksibi et al. [38] propose an approach for query expansion based on the contextual correlations of interconcepts. The main idea consists in exploring *Flickr* resources in order to extract such correlations, which are then presented as an interconcept graph. A random walk process is performed on this graph to discover implicit concepts, which are relevant for a given query, and hence to optimize analysis results. As for Biancalana et al. [40], they present an approach for the personalization of Web search using query expansion. In fact, they extend the family of co-occurrence matrix models by using a new way enabling to explore socio-tagging services. Indeed, the proposed user model consists of a three-dimensional correlation matrix in which each term is linked to an intermediate level that contains the relative belonging classes, each of which is accompanied by a relevance index. Therefore, each term is contextualized before being linked to all the other terms present in the matrix and guided toward the specific semantic categories, identified by tags. This approach has been proved useful particularly in the case of disambiguation of the word contexts.

Based on the newly added terms that are derived from the calculation of the co-occurrences between terms, disambiguate users' queries have enabled to diversify the returned documents, and therefore, enhance the quality of the performed research studies. Nevertheless, in the context of image retrieval, the expansion of the query terms, according to their semantic relationships with tags assigned to the social resources, is not considered an effective solution for a semantic content-based document retrieval due to this type



of query expansion being based on static calculations not founded on a logical reasoning.

As an alternative, many research studies have proposed to disambiguate the users' queries by exploiting the informational richness of the structured knowledge resources. For example, some lexical open-source resources, like the *WordNet*, are used to expand the users' queries with concepts semantically related to them [41–43]. In [44–46], the authors propose to explore *Wikipedia* knowledge repository in order to reformulate the users' requests while in [47, 48], the users' queries have been expanded based on domain ontologies. In the same context, other research studies propose to combine the semantic information derived from various knowledge resources [49].

The refinement of the ambiguous users' queries based on external semantic knowledge resources has greatly improved the quality of the provided results in numerous areas. However, semantic content-based image retrieval presents a real challenge. In fact, an image can be semantically viewed from several angles, such as structural, behavioral and event views, where each view can be expressed according to four semantic facets, namely: taxonomic, temporal, spatial and qualification facets [50–52]. Therefore, an incomplete identification of the different semantic facets of the images' views by a given query makes it difficult to provide all the semantically relevant images.

## 2.2 Ambiguous query postprocessing strategy

In the literature, numerous postprocessing techniques have been proposed to solve the result diversity challenge. These techniques can be categorized in two classes: *Search Result Diversification* (SRD) and *Search Result Clustering* (SRC). While the SRD techniques consider the pairwise similarity inter-documents in order to iteratively select the document that is not only relevant for the request but also different from the previously selected documents, the SRC techniques aim to gather similar results in the same cluster.

- *Search result diversification* The policy of the most research approaches relying on the SRD techniques consists first in retrieving a set of relevant documents compared to the sent query and then re-ranking them in order to ensure diversity. In this context, many researchers have devoted considerable efforts to diversify the highest ranked results. For instance, Leuken et al. [53] study three visually diverse ranking methods for the re-ranking of the search results. On the other hand, Wang et al. [54] propose a diverse relevance ranking algorithm to maximize the average diverse precision in the optimization framework by mining the semantic similarities of the social

images according to their visual characteristics and associated tags. As for Ksibi et al. [26], they suggest re-ranking the search results by promoting both visual and semantic diversity. For their part, Sun et al. [55] have also explored the visual and semantic information of images in order to obtain a social image-ranking scheme. They have then suggested retrieving the images by meeting the relevance, diversity and typicality criteria. Since most people have confirmed that they prefer to retrieve results with interesting and broad topics, as reported in [56], many literature studies about topic coverage have emerged [57–62]. For example, Agrawal et al. [57] propose to classify the taxonomy over queries to represent the different aspects of a query. This approach promotes the documents, which share a high number of classes with the query and demotes those with classes already well represented in the ranking. Qian et al. [61] propose a topic diverse re-ranking method for tag-based image retrieval with the promotion of the topic coverage performance. In this topic diverse re-ranking method, a tag graph construction and a community detection are two effective adopted ways to ensure the diversity. Indeed, a tag graph based on the similarity between each tag is constructed. The community detection method is then conducted to explore the community topic of each tag. After that, intra-community ranking and inter-community ranking are introduced to gain the final retrieved results.

Although the search result diversification techniques have enabled to disambiguate the users' queries by providing results that enable to reflect the various meanings that can have a given request, this type of technique suffers from two major limitations: First, the provided image results are visually very heterogeneous so that the selection of the desired image is a tedious and a time-consuming task. Second, re-ranking the images according to their relevance, as compared to a query, often requires online heavy calculations.

- *Search result clustering* Thanks to their discriminative power, the SRC techniques have been proved effective to promote diversity ranking. For instance, [63–65] create a visually diverse ranking of the search results by clustering the images based on their visual characteristics, such as color, texture and shape. Nevertheless, this type of approach is inefficient for practical use because the online visual clustering is a highly time-consuming process. As a solution, in [66, 67] the authors suggest the implementation of the low-level feature-based image clustering on a *MapReduce* architecture. Indeed, *MapReduce* is a parallel programming model that helps increase the speed of execution of the iterative clustering algorithms. However,

visual-based image ranking algorithms usually neglect the “*semantic gap*” problem between the low-level visual information captured by the imaging devices and the high-level semantic information perceived by humans, which largely degrades the clustering performance, and therefore that of the research process. As an alternative, *Deep-Learning* algorithms (DL) have emerged. For instance, Qayyum et al. [68] propose the modeling of a *Conventional Neural Network* (CNN) [69] to automatically extract the features of multimodal medical images for classification and retrieval tasks. Lyndon et al. [70] also design a CNN architecture for the classification of biomedical images. The proposed architecture consists of two convolution layers where each is followed by a pooling layer. The output of the second pooling layer is connected to a fully connected layer with 500 neurons, which, in turn, is connected to a logistic regression classifier. Despite the fact that DL classification techniques have offered the possibility to perform an automatic feature extraction and bridge the “*semantic gap*” problem, the choice of the network structure is today an open issue in DL approaches. Furthermore, when adopting a *Deep Neural Network* algorithm, it is necessary to design a great number of images for the training stage. Moreover, recent research studies have addressed the “*semantic gap*” problem based on visual words fusion of various types of descriptors, like SIFT and LIOP [12], SURF and HOG [13], SURF and FREAK [14], SIFT and BRISK [15], LIOP and LBPV [16], HOG and ORB [71], etc. For instance, in [14], the SURF and FREAK features are extracted separately from each image in the training and testing image sets. Thereafter, the *k*-means clustering algorithm is applied individually on each descriptor’s extracted features. The high-dimensional feature space of each descriptor is reduced to clusters, which is also known as visual words, in order to separately formulate the SURF and FREAK feature descriptor dictionaries. The visual words of both descriptors are then fused together by integrating the dictionaries of both feature descriptors. A histogram is then constructed using the fused visual words of each image. The set of histograms is used to train the SVM classifier. Euclidean distance is used to measure the similarity between the query image and the images stored in the image collection. In [71], the authors propose to apply a KPCA method [72] on the extracted features of HOG and ORB descriptors to select the best features and reduce the computational cost. Extensive experiments demonstrate the effectiveness of these techniques and their ability to reduce the “*semantic gap*” problem. However, fused features are not enough to achieve significant clustering and research performances, if they are not fused properly. Moreover, the construction of a single histogram from the whole image based on the traditional BoVW methodology suffers from the submergence of feature

salience that can decrease the image clustering performance and therefore that of the research process. In [17, 18], spatial information about salient objects within each image is retained in order to improve the performance of the image clustering and retrieval by reducing the semantic gap issue. However, the common issue in front of the CBIR systems is the time complexity. Finally, tag-based clustering algorithms have emerged as another important research direction. For example, in *IGroup* [73], an image search engine that presents the results in semantic clusters according to the associated users’ tags is proposed. Ramage et al. [74] explore the use of tags in *k*-means clustering in an extended vector space model, which includes both tags and page text. Yin et al. [75] transformed the problem of Web object classification into a problem of optimization on a graph of objects and tags. However, the use of social tags in their raw form is not an effective solution for a semantic content-based image classification (or clustering) since users’ tags can often contain ambiguous, disturbing and even missing information [76].

### 2.3 Summary

Tag ambiguity is the common issue in *Tag-based Social Image Research approaches*. An effective solution in case of an absence of precise knowledge about the users’ intentions and their contexts consists in diversifying the search results to cover as much possible the topics that a given ambiguous query may have. Many interesting techniques at the ambiguous query preprocessing and postprocessing levels were proposed. However, semantic-based query reformulation, diversity-based re-ranking for reordering image results, high cost of online heavy computations and time complexity requirement are the fundamental issues in diversity-based image research approaches. In this paper, we propose a novel technique where we jointly investigate two processes at the ambiguous query preprocessing and postprocessing levels. The proposed *Tag-based Query Semantic Reformulation* process enables to ensure a complete coverage of all the query aspects by reformulating the ambiguous users’ queries, according to multiple semantic facets of the different images’ views, using a set of predefined ontological semantic rules. The *Multi-level Image Diversification process* introduces efficient semantic-based image clustering where the computing-intensive tasks are executed offline. Given the importance of the semantic vocabulary size, the proposed two-level-based image clustering technique enables to improve the accuracy of the clustering results compared to the one-level-based image clustering technique. The filtering and re-ranking of the returned image cluster results according to their pertinence compared to the reformulated query enables to achieve higher accuracy of the search results.

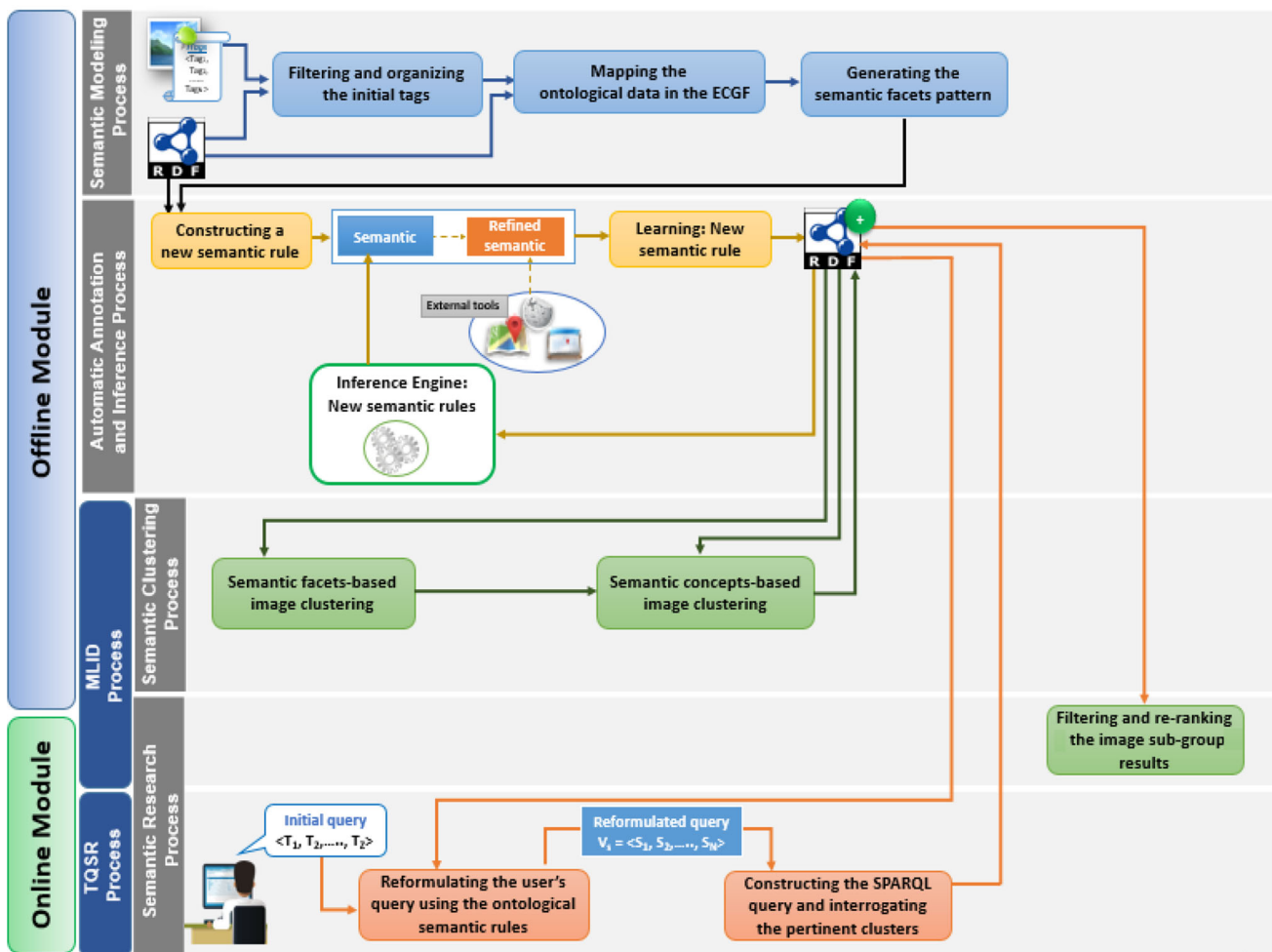


Fig. 2 General architecture of the proposed *Semantic-based Social Image Retrieval System*

### 3 General architecture of the proposed Semantic-based Social Image Retrieval System (S2IRS)

*Semantic-based Image Retrieval* (SBIR) is a dynamic research field, which aims at developing a set of methods enabling to express the high-level of the semantic meanings within images in order to be the keystone of the image retrieval process. Within this context, we focus in our studies on annotating, clustering and querying the socio-tagged images based on their different semantic views and facets. As illustrated in Fig. 2, the proposed system consists of an offline and online modules.

In the offline module, three successive processes occur:

- *Semantic modeling process* It consists in treating the tag information associated with the images shared on the social networks by projecting them on a domain ontology to generate a semantic pattern. This pattern enables to identify the different semantic views and facets that can

have an image based on which its semantic content can be described.

- *Automatic annotation and inference process* It consists of two stages: First, defining a set of semantic rules, enriching the used ontology by these rules and annotating the images using the new information of the enriched ontology. The main aim behind this stage is therefore to know the structure of these rules; second, automatically deducing new semantic rules. Indeed, the inference process can be defined as an action enabling a human or a machine to widen his/its knowledge.
- *Clustering process* It consists in separating and assigning the images to a set of subgroups according to the semantic annotations associated with them. The main goal behind this process is to accelerate the semantic-based image research process by enabling the users to identify the required group at a glance. To design a lightweight image search system, the time-consuming computation tasks including image similarity measures, image cluster-

ing and representative images selection are precomputed offline.

In the online module, ambiguous requests sent by the Internet users are treating and reformulating based on the predefined ontological semantic rules. Thereafter, the new semantic terms of the reformulated request will be automatically transformed into SPARQL language in order to interrogate the pertinent clusters. The interrogated clusters are filtered and re-ranked according to their relevance compared to the reformulated query.

The following subsections briefly describe the modeling, annotating and inference processes. Then, the research and clustering processes will be detailed in Sects. 4 and 5, respectively.

### 3.1 Overview of the semantic modeling process

The images shared on the social networks, such as *Flickr*, are usually associated with a set of keywords called tags, which are assigned to them by the social community. Nevertheless, the users' tags do not constitute an effective way for annotating these images because they often contain ambiguous, disturbing and even missing information. The suggested solution consists in attributing new terms to the socio-tagged images, which can reflect their semantic content. However, a good description of the semantic content of an image requires first the definition of these different semantic facets to be described. Therefore, the main aim behind the modeling process is to generate a semantic pattern that enables to present the different semantic facets of an image in order to make a basic support for the semantic annotation process. In fact, in our previous research studies [51, 77], we have first classified the tags associated with the social images. Moreover, these tags often express “*who*” and “*what*” and can be seen according to several views:

- *The structural view* It presents the set of objects depicted on the image and which are noticeable in the real world (tower, monument, etc.).
- *The behavior view* It presents the activities performed by the characters depicted on the image (sport, adventure, etc.).
- *The event view* It presents the events unfolding within the image (festival, party, etc.).

Moreover, the concepts of each view can have spatial and temporal contexts as well as a set of specifications that characterize them and a set of taxonomies from which they are derived. Thus, we have defined four facets from which the different concepts can be seen:

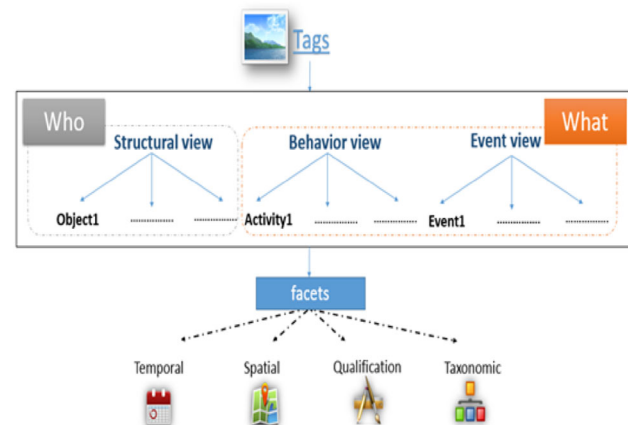


Fig. 3 Different semantic views and facets of an image

- *The temporal facet* It enables to indicate the epoch of construction of the historical objects and the unfolding season of the events and activities.
- *The spatial facet* It enables to indicate the location of the visual objects and the unwinding location of the events and activities.
- *The qualification facet* It enables to describe the nature and the specifications of an object, activity or event that distinguish them from other objects, activities and events.
- *The taxonomic facet* It enables to indicate the type of a visual object, an activity or event taking place.

Figure 3 shows the different semantic views and facets of an image.

Therefore, we have used the image tagging information to define semantic patterns that represent the different semantic facets of these images. Our solution is essentially based on a domain ontology, and it consists of three steps:

- *Filtering and organizing the initial tags* This consists in projecting each tag attached to an image on the approved domain ontology in order to verify its membership. Thus, the tags belonging to the ontology, as classes or instances, have been considered as domain concepts and have been the focus of our work.
- *Mapping the ontological data using the extended conceptual graphs formalism* This consists in moving from the abstract level of an image toward its conceptual description level. In fact, the image conceptual description level is achieved by mapping the ontological data, in particular, the semantic relationships and the specific properties of the filtered and organized tags, on the extended conceptual graph formalism *ECGF*. Subsequently, the various conceptual models derived from an image are merged and drawn on the same graph. It should be mentioned that we have extended the formalism of the conceptual graphs defined in [78], so



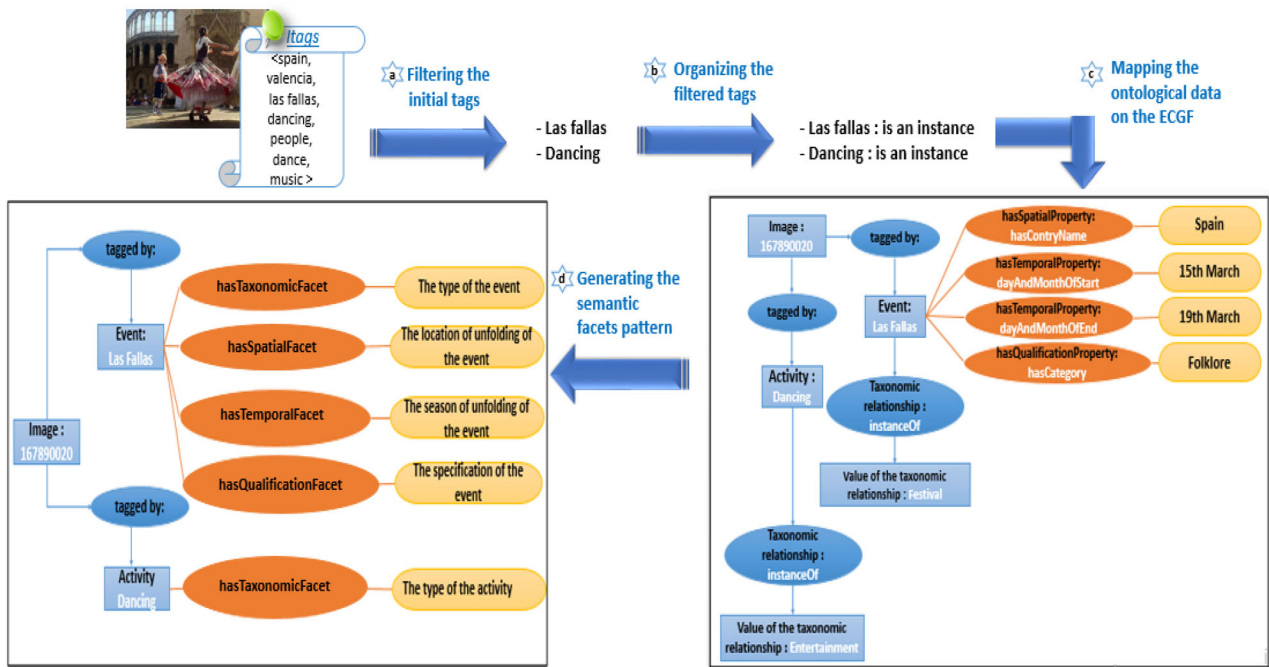


Fig. 4 An example of the semantic facets modeling process

that the concepts and the conceptual relationships as well as the conceptual features can be presented [51].

- *Generating the semantic facet pattern of an image* This consists in deducing the corresponding semantic facet pattern of an image from its conceptual description model. This pattern enables to identify the different semantic facets of such an image and serves as support for the future annotation process. For example, we can deduce from the start and the end date of an event its unfolding season. We can deduce also the name of a monument location from its GPS information.

Figure 4 gives an example of a semantic facet modeling process corresponding to a socio-tagged image.

The generated semantic facet patterns have been saved in XML files as presented in Fig. 5.

### 3.2 Overview of the semantic annotation and inference process

After the modelization of the image semantic facets, a semantic annotation process has been defined. The aim of this process consists in assigning semantic information to the socio-tagged images, so that afterward, the new knowledge becomes the focus of the research process. For this reason, three steps have been defined:

- *Semantic annotation* This consists in constructing a set of semantic rules that use the tags associated with the

```

1 <?xml version="1.0" encoding="ISO-8859-1" ?>
2 <image id="167890020">
3   <TaggedBy Event="Las Fallas" Type="Ontological_Instance">
4     <hasTaxonomicFacet> The type of the event </hasTaxonomicFacet>
5     <hasSpatialFacet> The unwinding location of the event </hasSpatialFacet>
6     <hasTemporalFacet> The season of unfolding of the event </hasTemporalFacet>
7     <hasQualificationFacet> The specification of the event </hasQualificationFacet>
8   </TaggedBy>
9   <TaggedBy Activity="Dancing" Type="Ontological_Instance">
10    <hasTaxonomicFacet> The type of the activity </hasTaxonomicFacet>
11  </TaggedBy>
12 </image>
    
```

Fig. 5 An example of a semantic facet pattern saved in XML file

social images as basic information and rely on the knowledge derived from the adopted ontology to generate taxonomic, qualification, spatial and temporal semantics, which express the different semantic facets of an image. These semantic rules have been presented via the logic of predicates.

- *Semantic inference* This consists in generating new semantics by defining new semantic rules. Its policy is based on the front chaining. For more details on the semantic annotation and inference rules, readers can refer to [77].
- *Enrichment of the used ontology* This consists in enriching the initial adopted ontology with the generated semantic rules, where a semantic rule is characterized by its type (taxonomic, spatial, temporal or qualificative rule). In addition, each resulted semantic value has been defined

```

<rdf:Class rdf:about="http://www.owltouring.org#Activity">
  <rdf:Concept_Name rdf:datatype="string"> Activity </rdf:Concept_Name>
  <rdf:comment rdf:datatype="string"> Occupation of a person or a group of persons that has a specific goal. </rdf:comment>
</rdf:Class>
<rdf:Class rdf:about="http://www.owltouring.org#Adventure">
  <rdf:Concept_Name rdf:datatype="string"> Adventure </rdf:Concept_Name>
  <rdf:subClassOf rdf:resource="http://www.owltouring.org#Activity"/>
</rdf:Class>
<rdf:Class rdf:about="http://www.owltouring.org#Diving">
  <rdf:Concept_Name rdf:datatype="string"> Diving </rdf:Concept_Name>
  <rdf:subClassOf rdf:resource="http://www.owltouring.org#Adventure"/>
  <rdf:Rule> Rule36 </rdf:Rule>
</rdf:Class>
<rdf:Description rdf:about="http://www.owltouring.org#SemanticFacet">
  <rdf:Facet_Name rdf:datatype="string"> Semantic Facet </rdf:Facet_Name>
</rdf:Description>
<rdf:Description rdf:about="http://www.owltouring.org#TaxonomicFacet">
  <rdf:Facet_Name rdf:datatype="string"> Taxonomic Facet </rdf:Facet_Name>
  <rdf:subFacetOf rdf:resource="http://www.owltouring.org#SemanticFacet"/>
</rdf:Description>
<rdf:Description rdf:about="http://www.owltouring.org#TheTypeOfTheActivity">
  <rdf:Facet_Name rdf:datatype="string"> The type of the activity </rdf:Facet_Name>
  <rdf:subFacetOf rdf:resource="http://www.owltouring.org#TaxonomicFacet"/>
</rdf:Description>
<rdf:Description rdf:about="http://www.owltouring.org#SemanticRule">
  <rdf:Rule_Name rdf:datatype="string"> Semantic Rule </rdf:Rule_Name>
</rdf:Description>
<rdf:Description rdf:about="http://www.owltouring.org#TaxonomicRule">
  <rdf:Rule_Name rdf:datatype="string"> Taxonomic Rule </rdf:Rule_Name>
  <rdf:subRuleOf rdf:resource="http://www.owltouring.org#SemanticRule"/>
</rdf:Description>
<rdf:Description rdf:about="http://www.owltouring.org#Rule36">
  <rdf:Rule_Name rdf:datatype="string"> Rule36 </rdf:Rule_Name>
  <rdf:subRuleOf rdf:resource="http://www.owltouring.org#TaxonomicRule"/>
  <rdf:Predicats>
    <rdf:Predicat>
      <rdf:Concept_Name> Diving </rdf:Concept_Name>
      <rdf:Relation_Name> Is-a </rdf:Relation_Name>
      <rdf:Concept_Name> Adventure </rdf:Concept_Name>
    </rdf:Predicat>
    <rdf:Predicat>
      <rdf:Concept_Name> Adventure </rdf:Concept_Name>
      <rdf:Relation_Name> Is-a </rdf:Relation_Name>
      <rdf:Concept_Name> Activity </rdf:Concept_Name>
    </rdf:Predicat>
  </rdf:Predicats>
  <rdf:Semantic rdf:semanticfacet="http://www.owltouring.org#TheTypeOfTheActivity"> Adventure activity </rdf:Semantic>
  <Multimedia> 24989418481.jpg </Multimedia>
  <Multimedia> 66897307633.jpg </Multimedia>
</rdf:Description>

```

Fig. 6 An extract of the enriched ontology

by its semantic facet. An extract of the RDF of the enriched ontology is presented in Fig. 6.

Indeed, in our work we have adopted «*Touring*<sup>8</sup>» ontology, which presents the concepts used in the tourism domain, to model and annotate a set of socio-tagged images. Formally, this ontology is presented by  $On: (C_{On}, R_{On}, X_{On}, I_{On}, P_{On})$ , where  $C_{On}$ ,  $R_{On}$ ,  $X_{On}$ ,  $I_{On}$  and  $P_{On}$  are, respectively, the set of concepts, relationships between the concepts, axioms, instances and properties. To take into account the semantic rules as well as the semantic facets, we proposed to enrich this ontology with new components called «*SemanticRule*»

<sup>8</sup> <http://www.owltouring.org/> Last visited 08/08/2016.

and «*SemanticFacet*». Thus, the new formal definition of the enriched ontology is:  $On: (C_{On}, R_{On}, X_{On}, I_{On}, P_{On}, SR_{On}, SF_{On})$ , where  $SR_{On}$  is the set of rules and  $SF_{On}$  is the set of semantic facets that can have the semantic values provided following the run of these semantic rules. Figure 7 represents the meta-model of the initial ontology and the performed extension.

We define the ontology properties as follows:

- *HasSet property* For each concept  $c$ , relation  $r$ , axiom  $a$ , instance  $i$  and property  $p$  of the ontology  $On (C_{On}, R_{On}, X_{On}, I_{On}, P_{On}, SR_{On}, SF_{On})$ , with  $c \in C_{On}$ ,  $r \in R_{On}$ ,  $a \in X_{On}$ ,  $i \in I_{On}$  and  $p \in P_{On}$ , the *HasSet property* is generated. For each generated *HasSet property*,  $\exists c \in C_{On}$ ,  $r \in R_{On}$ ,



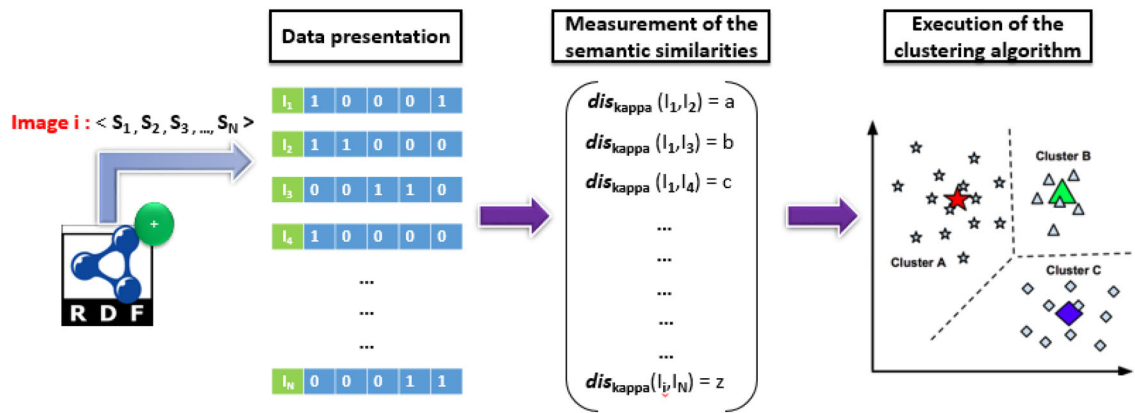


Fig. 8 Semantic-based image clustering process

play retrieved images as long and paged image list does not enable users to obtain overall view of retrieved images and make the localization of the target image a hard and a time-consuming task. To help users browse and consult images, an image summarization technique is an urgent need. Image clustering and representative image selection (called canonical view) are efficient and effective techniques used for summarizing image data. Within this context, we propose in our work an image search system that summarizes the interrogated images in semantic groups. The policy of the grouping process consists in using socio-tagged images after executing the modeling and annotation processes. To support the online semantic-grouped image search process, we propose to precompute image clusters offline. Figure 8 presents the different stages of the proposed semantic-based image clustering process.

- Data presentation

The semantic data attributed to the images after executing the modeling and annotation processes are presented based on a qualitative vectorial model so that if a semantics is associated with an image, the corresponding value is 1 and 0 otherwise. Thus, each image is characterized by a Boolean vector that indicates the presence or absence of a semantic value.

- Measurement of the semantic similarities between images

To estimate the semantic similarity between two images, the distance of *Cohen's kappa* [79] is applied. Indeed, *Cohen's kappa* coefficient, which is a measure created by *Jacob Cohen*, is still considered as the most accurate measure enabling to quantify the relationship between two items with qualitative representations. The general expression of *Cohen's kappa* is presented by Eq. (1):

$$k = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)} \quad (1)$$

- Execution of the clustering algorithm

To cluster the annotated images into different sub-topics, *k*-means algorithm [80] is applied. Indeed, *k*-means clustering is a method frequently used for automatically partitioning a set of data into *k* groups. The following pseudocode describes the general iterations of the *k*-means algorithm:

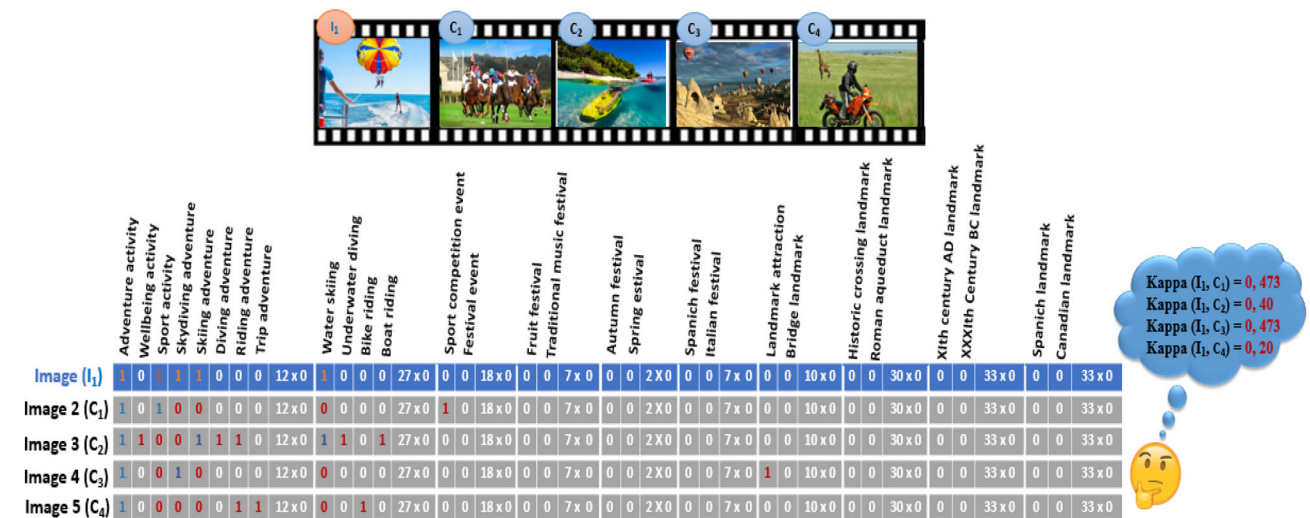
1. Choose the number of clusters
2. Choose and assign the initial centroids
3. For  $d_i \leq \text{Nb instances}$ 
  - (i) Calculate the distance of  $d_i$  to each centroid
  - (ii) Assign  $d_i$  to the more closest cluster's centroid
4. For  $C_i \leq \text{Nb clusters}$   
Recalculate the position of the cluster's centroid
5. Repeat steps 3 and 4.
6. The algorithm converges when there is no further change in assignment of instances to the clusters.

Three challenges are encountered at this level: The first is about the choice of the number of clusters and the second concerns the selection of the initial centroids, while the third is about the calculation of the clusters' centroids when it consists of a binary data stream.

- *How to choose the optimal number of clusters?* Determining the optimal number of clusters when applying *k*-means algorithm is a fundamental step because this number has a considerable influence on the performance of the clustering results. In this work, we use three popular methods, namely: *Elbow* [81], *Average Silhouette* [82] and *Gap Statistic* [83] methods, to determine the optimal number of clusters. The *k* value that we choose is the one which



**Fig. 9** Iterations of calculation of a new cluster’s centroid (an example)



**Fig. 10** Measure of semantic similarities between images with very important semantic data flow (example)

is obtained by at least two methods. In the situation where the three methods generate three different results, we randomly choose one of these values.

- *How to select the best initial centroids?* To optimize the selection of the initial centroids, we rely on the *improved Pillar algorithm* [84] according to which the first centroid is the closest one to all the other instances of the data space. The data point with the maximum distance from the previously selected centroid is considered the next centroid. Likewise, the other centroids are selected in an iterative manner by calculating the distance metrics and choosing the data point with the maximum distance from the previously selected centroids.
- *How to calculate the new gravity centers?* To calculate the gravity center when the data flow is binary, we propose to select the closest one to all the other instances of the data space. This consists first in calculating the distance between all the instances of the data space. Then, add the distances of each instance compared to the others and calculate the averages. The instance with the highest average distance will be the new centroid of the cluster as it is the closest to all the other instances. Figure 9 illustrates an example of calculation of a new cluster’s centroid.

Despite the fact that we have optimized the choice of the number of clusters and started from the best initial centroids when applying *k*-means algorithm, unsatisfactory clustering results have occurred. This is due to the application of *kappa* distance to measure the semantic similarities between a large number of images with very important semantic data flow.

In the example shown in Fig. 10, the candidate image I<sub>1</sub> is semantically more similar to the centroid C<sub>2</sub> than the centroids C<sub>1</sub>, C<sub>3</sub> and C<sub>4</sub>. However, the semantic measures generated by *kappa* distance enable to associate I<sub>1</sub> with the centroids C<sub>1</sub> or C<sub>3</sub>. Therefore, noisy clustering results can be generated.

To resolve this issue, a solution consists in dividing the set of images into various semantic interest fields in order to restrain the semantic data space. In this context, we recall that the semantic content of an image can be described according to different semantic facets (*taxonomic, temporal, spatial and qualification*) of the different semantic views (*structural, behavioral and event*), as referred in our previous work [51]. Thus, we propose to execute the image clustering process at two successive levels (two-level-based image clustering process):

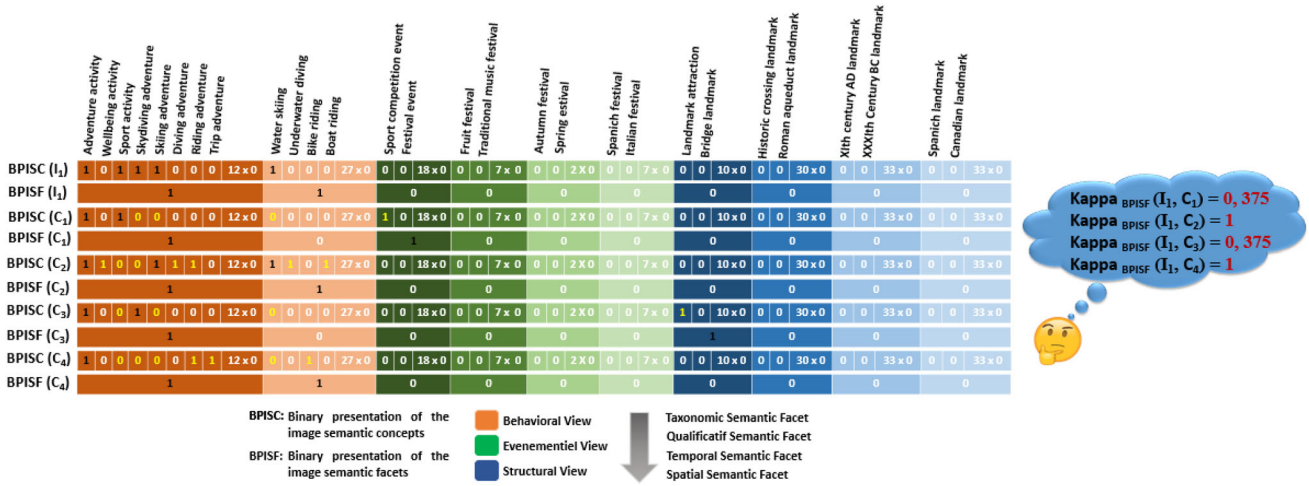


Fig. 11 Clustering images using the similarity measures of the binary presentation vectors of the image semantic facets (example)

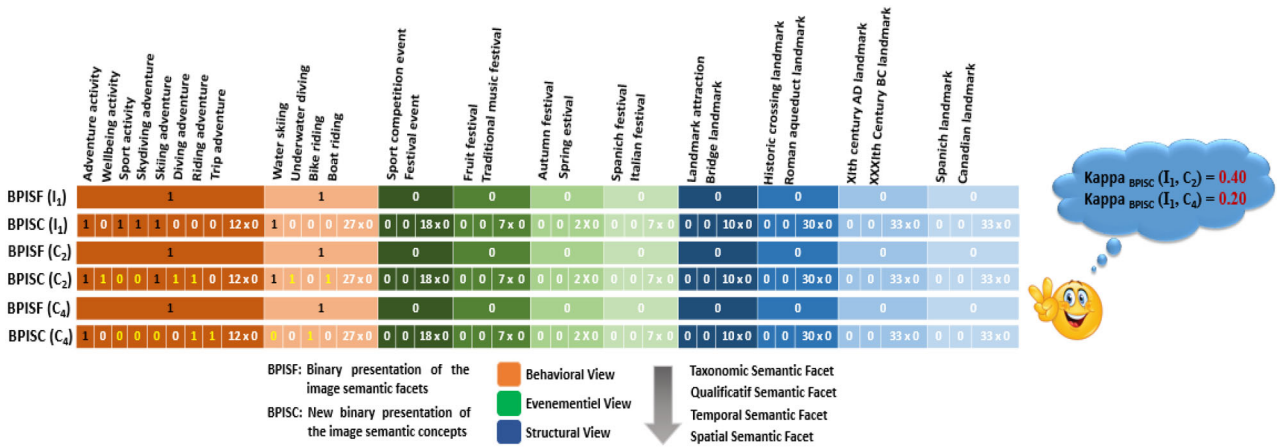


Fig. 12 Clustering images using the similarity measures of the binary presentation vectors of the image semantic concepts (example)

- *Semantic Facet-based Image Clustering* This consists in separating the images that are semantically very distant, by applying *k*-means algorithm and using the semantic similarity measures of the binary presentation vectors of the image semantic facets. The binary presentation vectors of the image semantic facets are generated according to the different semantic facets of the image semantic concepts. Thus, if an image has at least one semantic concept for a semantic facet, the corresponding value is 1 and 0 otherwise. In the example shown in Fig. 11, the image I<sub>1</sub> is semantically more similar to the images C<sub>2</sub> and C<sub>4</sub>, based on the results of the semantic similarity measures of the image semantic facets. Therefore, there is a great chance so that these three images become included in the same cluster, while the images C<sub>1</sub> and C<sub>3</sub> become classified in other different clusters because they have weak semantic similarities with I<sub>1</sub>.
- *Semantic Concept-based Image Clustering* This consists in partitioning each group of images resulted from the

semantic facet-based image clustering into a set of sub-groups, by applying *k*-means algorithm and using the semantic similarity measures of the binary presentation vectors of the image semantic concepts. In the example shown in Fig. 12, the candidate image I<sub>1</sub> is semantically more similar to the centroid C<sub>2</sub>, based on the results of the semantic similarity measures of the image semantic concepts.

The two-level-based image clustering process can be presented as shown in Fig. 13.

The results of the two-level-based image clustering process are saved in an XML file as presented in Fig. 14.

The centroids are inserted in the RDF file as representative images of the clusters. An extract of the RDF of the enriched ontology is presented in Fig. 15.

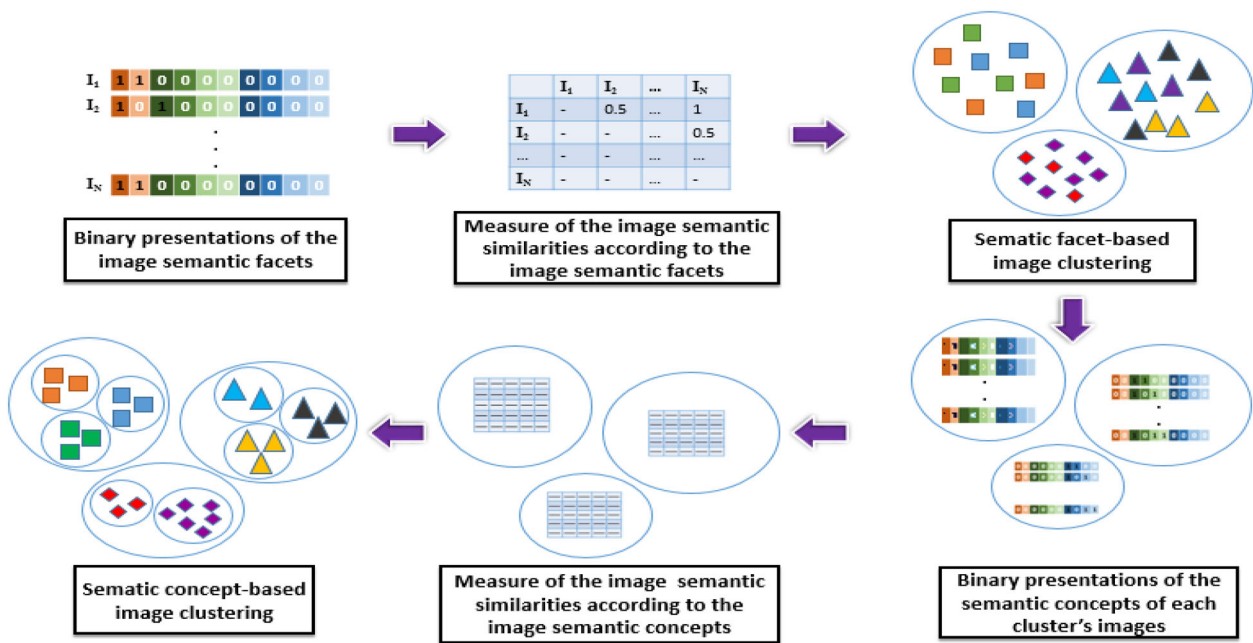


Fig. 13 Two-level-based image clustering process

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<Centroid id="676403793.jpg">
  <Image> 679403840.jpg </Image>
  <Image> 674034030.jpg </Image>
  <Image> 674933943.jpg </Image>
  <Image> 674673927.jpg </Image>
  <Image> 676730211.jpg </Image>
  <Image> 674933902.jpg </Image>
</Centroid>
<Centroid id="8930378309.jpg">
  <Image> 893563876.jpg </Image>
  <Image> 893763509.jpg </Image>
  <Image> 892583651.jpg </Image>
  <Image> 893528030.jpg </Image>
  <Image> 892528943.jpg </Image>
</Centroid>
</xml>
```

Fig. 14 Results of the semantic concept-based image clustering saved in an XML file

### 5 Semantic research process

Tag-based image retrieval has not been proved effective in the field of Information Retrieval (IR). A major limitation of this axis is that it is semantically very weak because of the ambiguity of the tags, which makes it unsuitable for the IR’s needs. Thus, the semantic-based reformulation of the tag-based user’s queries is a promising solution, which

helps improve the research by providing semantically relevant results. Figure 16 presents the different stages of the proposed semantic-based image research process including the tag-based query semantic reformulation as well as the filtering and the re-ranking of the interrogated image sub-groups.

#### 5.1 Ambiguous query reformulation

The interrogation process that we suggest consists in automatically transforming a tag-based query from the textual form toward an interrogation language, in our case SPARQL. Thus, we propose first to create an intermediate level that enables to move from an ambiguous query toward a semantic-based query by using the properties of the ontology previously presented in Fig. 7.

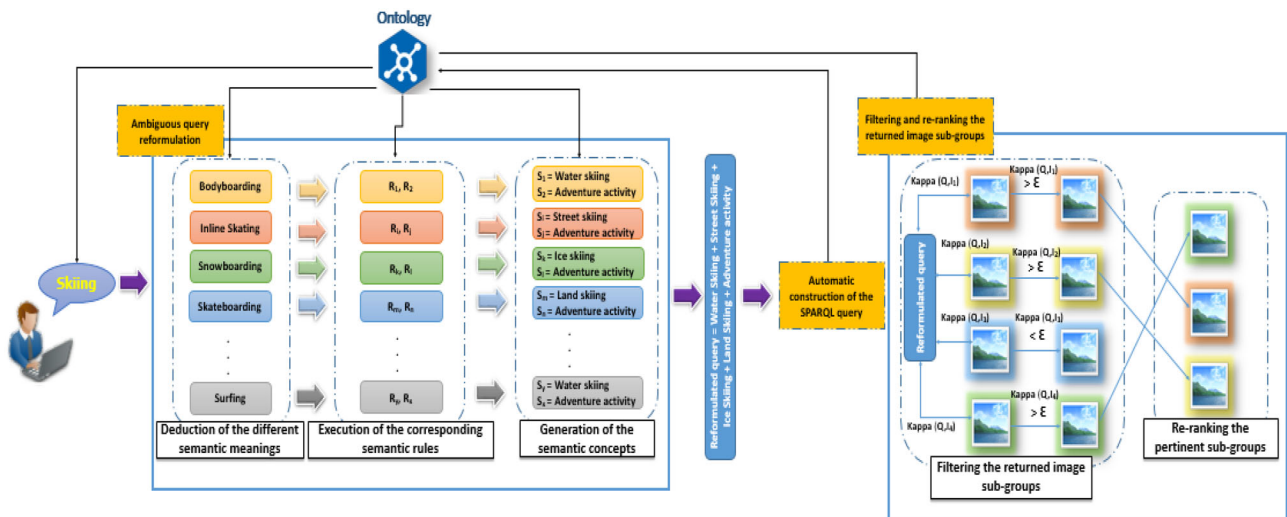
Given that the input of the interrogation process is described by a set of terms sent by a user, any query carrying on the search of the concept  $c_i$ , such as  $c_i \in C_{On}$ , is translated by the property *HasSet* (). Therefore, an initial query carrying on the search of the concepts  $\{c_1, c_2, \dots, c_N\}$ , such as  $\{c_1, c_2, \dots, c_N\} \in C_{On}$ , can be formulated as demonstrated by Formula 1:

$$HasSet(c_1, C_{On}) \wedge HasSet(c_2, C_{On}) \wedge \dots \wedge HasSet(c_N, C_{On}) \tag{1}$$

**Fig. 15** Enrichment of the ontology by the clustering information

```

<rdf:Description rdf:about="http://www.owltouring.org#Rule36">
  <rdf:Rule_Name rdf:datatype="string"> Rule36 </rdf:Rule_Name>
  <rdf:subRuleOf rdf:resource="http://www.owltouring.org#TaxonomicRule"/>
  <rdf:Predicates>
    <rdf:Predicat>
      <rdf:Concept_Name> Diving </rdf:Concept_Name>
      <rdf:Relation_Name> Is-a </rdf:Relation_Name>
      <rdf:Concept_Name> Adventure </rdf:Concept_Name>
    </rdf:Predicat>
    <rdf:Predicat>
      <rdf:Concept_Name> Adventure </rdf:Concept_Name>
      <rdf:Relation_Name> Is-a </rdf:Relation_Name>
      <rdf:Concept_Name> Activity </rdf:Concept_Name>
    </rdf:Predicat>
  </rdf:Predicates>
  <rdf:Semantic rdf:semanticfacet="http://www.owltouring.org#TheTypeOfTheActivity"> Adventure activity </rdf:Semantic>
  <Multimedia> 24989418481.jpg </Multimedia>
  <Multimedia> 66897307633.jpg </Multimedia>
  <Centroid> 24989418481.jpg </Centroid>
</rdf:Description>
  
```



**Fig. 16** Semantic-based image research process

The second step consists in determining the different semantic meanings that can have each concept  $c_i$  of the initial query by deducing their final descendants (sub-concepts or instances). Thus,  $\exists c_i \in C_{On}$ , such as  $HasSet(c_i, C_{On})$ ,  $\exists c_j \in C_{On}$ , such as  $HasSet(c_j, C_{On})$  and  $\exists r \in R_{On}$ , such as  $HasSet(r, R_{On})$ , the property  $ComposedBy(c_j, r, c_i)$  is generated to indicate that  $c_j$  is a direct sub-concept of  $c_i$ , such that  $r$  represents the relation  $\langle is-a \rangle$ . On the other hand,  $\exists c_i \in C_{On}$ , such as  $HasSet(c_i, C_{On})$  and  $\exists i_a \in I_{On}$ , such as  $HasSet(i_a, I_{On})$ , the property  $InstancedFrom(i_a, c_i)$  is generated to indicate that  $i_a$  is an instance of  $c_i$ . The new reformulation based on the semantic meanings can be therefore demonstrated by Formula 2 such as  $\exists c_{ij} \in C_{On}$  is a final sub-concept of the concept  $c_i$  and  $\exists i_{ak} \in I_{On}$  is an instance of the concept  $c_a$ ,  $c_{ij}$  and  $i_{ak}$  are translated by the property  $HasSet()$ .

$$\begin{aligned}
 &HasSet(c_{i1}, C_{On}) \wedge HasSet(c_{i2}, C_{On}) \\
 &\wedge \dots \wedge HasSet(c_{NM}, C_{On}) \\
 &\wedge HasSet(i_{a1}, I_{On}) \wedge HasSet(i_{a2}, I_{On}) \\
 &\wedge \dots \wedge HasSet(i_{XY}, I_{On})
 \end{aligned} \tag{2}$$

The third step consists in determining the different semantic rules associated with the semantic meanings of the reformulated query. In general, a query carrying on the search of the concepts and instances  $\{c_1, c_2, \dots, c_N, i_1, i_2, \dots, i_X\}$ , with  $c_i \in C_{On}$  and  $i_a \in I_{On}$ , can have a new reformulation demonstrated by Formula 3 such as:  $\exists sr_{ij} \in SR_{On}$  is a semantic rule granted to the concept  $c_i$ , the property  $HasRule(c_i, sr_{ij})$  is generated, and  $\exists sr_{ak} \in SR_{On}$  is a semantic rule granted to the concept  $i_a$ , the property  $HasRule(i_a, sr_{ak})$  is generated. The rules  $sr_{ij}$  and  $sr_{ak}$  are translated by the property  $HasSet()$  such as:

$$\begin{aligned}
 &HasSet(sr_{i1}, SR_{On}) \wedge HasSet(sr_{i2}, SR_{On}) \\
 &\wedge \dots \wedge HasSet(sr_{NM}, SR_{On}) \\
 &\wedge HasSet(sr_{a1}, SR_{On}) \\
 &\wedge HasSet(sr_{a2}, SR_{On}) \wedge \dots \\
 &\wedge HasSet(sr_{XY}, SR_{On})
 \end{aligned} \tag{3}$$



The fourth step consists in deducing the different semantic concepts generated following the run of the semantic rules associated with the semantic meanings of the reformulated query. Thus, for each semantic rule  $sr_i \in SR_{On}$ , the property  $HasSemantic(sr_i, s_i)$  is generated to indicate that the application of the semantic rule  $sr_i$  has a semantic value as results. Formula 4 demonstrates the semantic concept-based reformulation of the initial user's query.

$$\begin{aligned} & HasSemantic(sr_1, s_1) \\ & \wedge HasSemantic(sr_2, s_2) \wedge \dots \\ & \wedge HasSemantic(sr_N, s_N) \end{aligned} \quad (4)$$

The fifth step consists in building a final semantic concept-based query by avoiding any redundancy of semantic concepts, as shown in Formula 5.

$$S_i \wedge S_j \wedge S_k \wedge \dots \wedge S_z \quad (5)$$

The steps of reformulation of an ambiguous user's query by semantic concepts are defined by the Algorithm 1 such as:

- **NbOfTermes ()**: enables to count the number of ontological concepts of a query.
- **getConcept ()**: enables to return a concept according to its order within a query.
- **hasSubConcepts (), hasInstances (), hasSemantic-Rules ()**: enable to verify whether a concept has sub-concepts, instances and semantic rules, respectively.
- **getFinalConcepts (), getInstances (), getSemantic-Rules ()**: enable to return the final sub-concepts, instances and semantic rules of a concept, respectively.
- **getSemanticConcept ()**: enables to return the semantic generated following the run of a semantic rule.
- **AvoidRedundancy ()**: enables to avoid any redundancy of semantic concepts of a query.

---

**Algorithm 1** : Ambiguous user's query reformulation
 

---

```

1: Input :
2: Vector  $UReq [c_1, c_2, \dots, c_n]$  : Chain // Initial user's query
3: Query  $R_1, R_2, R_3, R_4, R_5$  : Chain

4: Begin :
   // User's tag-based query reformulation
5: for (int  $i \leftarrow 1$  to  $Size(UReq)$ ,  $i++$ ) do
6:   if ( $HasSet(UReq[i], Con) \leftarrow True$ ) then
7:      $R_1 \leftarrow R_1 + UReq[i]$ 
8:   end if
9: end for

   // Semantic meaning-based query reformulation
10: for (int  $i \leftarrow 1$  to  $NbOfTermes(R_1)$ ,  $i++$ ) do
11:   if ( $hasSubConcepts(getConcept(i)) \leftarrow True$ ) then
12:     Vector  $V_{FinalConcepts} \leftarrow getFinalConcepts(getConcept(i))$ 
13:     for (int  $j \leftarrow 1$  to  $Size(V_{FinalConcepts})$ ,  $j++$ ) do
14:       if ( $hasInstances(V_{FinalConcepts}[j]) \leftarrow False$ ) then
15:          $R_2 \leftarrow R_2 + V_{FinalConcepts}[j]$ 
16:       else
17:         Vector  $V_{Instances} \leftarrow getInstances(V_{FinalConcepts}[j])$ 
18:         for (int  $k \leftarrow 1$  to  $Size(V_{Instances})$ ,  $k++$ ) do
19:            $R_2 \leftarrow R_2 + V_{Instances}[k]$ 
20:         end for
21:       end if
22:     end for
23:   else
24:     if ( $hasInstances(getConcept(i)) \leftarrow True$ ) then
25:       Vector  $V_{Instances} \leftarrow getInstances(getConcept(i))$ 
26:       for (int  $j \leftarrow 1$  to  $Size(V_{Instances})$ ,  $j++$ ) do
27:          $R_2 \leftarrow R_2 + V_{Instances}[j]$ 
28:       end for
29:     else
30:        $R_2 \leftarrow R_2 + getConcept(i)$ 
31:     end if
32:   end if
33: end for

   // Semantic rule-based query reformulation
34: for (int  $i \leftarrow 1$  to  $NbOfTermes(R_2)$ ,  $i++$ ) do
35:   if ( $hasSemanticRules(getConcept(i)) \leftarrow True$ ) then
36:     Vector  $V_{Rules} \leftarrow getSemanticRules(getConcept(i))$ 
37:     for (int  $j \leftarrow 1$  to  $Size(V_{Rules})$ ,  $j++$ ) do
38:        $R_3 \leftarrow R_3 + V_{Rules}[j]$ 
39:     end for
40:   end if
41: end for

   // Semantic concept-based query reformulation
42: for (int  $i \leftarrow 1$  to  $NbOfTermes(R_3)$ ,  $i++$ ) do
43:    $R_4 \leftarrow R_4 + getSemanticConcept(getConcept(i))$ 
44: end for

   // Avoid any redundancy of query's concepts
45:  $R_5 \leftarrow AvoidRedundancy(R_4)$ 
46: End

```

---

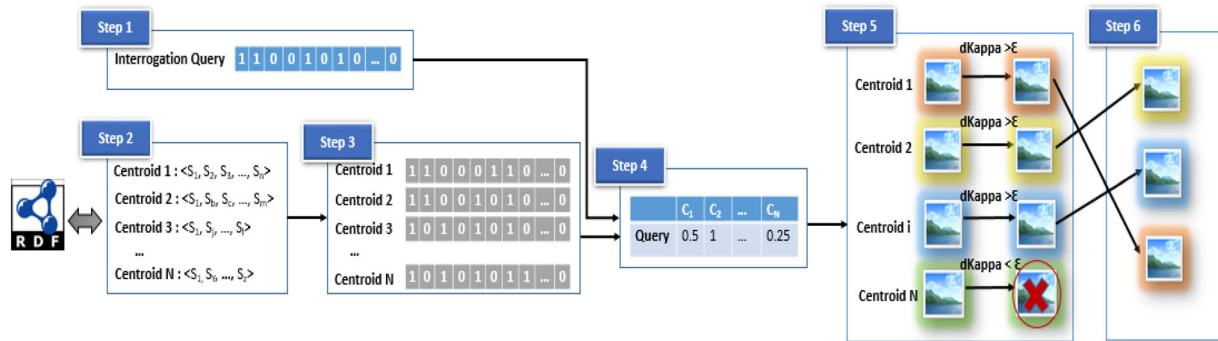


Fig. 17 Filtering and re-ranking steps of the returned image subgroups

## 5.2 Automatic construction of the SPARQL query

Once the ambiguous user's query is reformulated, we move to the construction of the interrogation query. This consists in automatically translating the semantic concept-based query into interrogation language (such as SPARQL language, recommended by the W3C since 2008). The SPARQL interrogation consists in searching in the *Semantic nodes* of the RDF of the enriched ontology the literals:  $\{S_i, S_j, S_k, \dots, S_z\}$  to interrogate the representative images of the clusters that are semantically correlated to the reformulated query. Once the target centroids are selected, the images of each cluster are interrogated from the XML file. The general expression of the SPARQL query is presented as follows:

```
PREFIX ONTOLOGY : <http://owltouring.org/>
SELECT ? image
WHERE {?x, rdf: SemanticRule, SemanticName. ?x, rdf:
SemanticRule, Centroid.
      ?y, rdf: SemanticRule, SemanticName. ?y, rdf:
SemanticRule, Centroid.
      ?t, rdf: SemanticRule, SemanticName. ?t, rdf:
SemanticRule, Centroid.
      ...
      ...
      ...
      ?z, rdf: SemanticRule, SemanticName. ?z, rdf:
SemanticRule, Centroid.
      ?x, «Si». ?y, «Sj». ?t, «Sk». ...?z, «Sz»}
```

## 5.3 Filtering and re-ranking the returned image subgroups

Following the run of an SPARQL interrogation, a set of image clusters semantically correlated with the reformulated query is returned. However, these clusters have various semantic similarity degrees compared to the interrogation query. Therefore, it is interesting to filter the clusters that have weak semantic correlations with the query. In this regard, we propose to fix a  $\epsilon$  value so that the clusters that have semantic

similarities less than  $\epsilon$  are considered irrelevant and therefore must be filtered. The relevant image clusters are then re-ranked so that the ones that have high semantic similarity degrees are ranked before the ones that have lower semantic similarity degrees. The filtering and re-ranking steps of the returned image subgroups consist in:

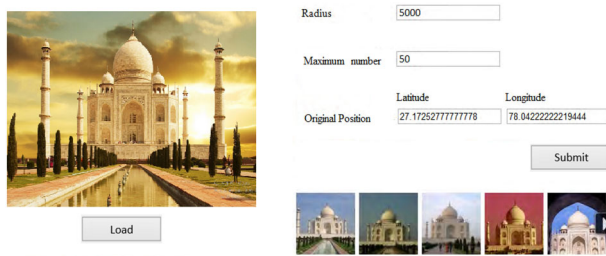
- *Step 1* Presenting the semantic concepts of the interrogation query by a binary vector.
- *Step 2* Searching in the RDF file, the semantic concepts correlated with each cluster's representative image.
- *Step 3* Presenting the semantic concepts of each cluster's representative image by a binary vector.
- *Step 4* Measuring the semantic similarity between each centroid's binary presentation vector and the query's binary presentation vector.
- *Step 5* Filtering the irrelevant interrogated image clusters.
- *Step 6* Re-ranking the pertinent clusters according to the semantic similarity measures between their centroids and the interrogation query.

Figure 17 shows the different steps of the filtering and re-ranking of the returned image subgroups.

## 6 Experimental evaluations

To endorse the efficiency of the suggested *Semantic-based Social Image Retrieval System*, we designed a prototype called "*Socio-touring tool*," which consists of three models, namely: administrator, expert and user's models.

These models help ensure the modelization, annotation and grouping of a set of images, to be semantically interrogated, and help conduct a series of experiments to evaluate the output results of each functionality. Since there are no specific preprepared image collections for each domain of interest, we have conducted this work by constructing a collection of test images. These images are taken from the social network *Flickr*, and they are related to the tourism filed. After



**Fig. 18** Web interface developed to extract the test image collection

modeling and annotating the socio-tagged images, we have carried out a series of experiments to evaluate the clustering and research processes.

## 6.1 Construction of a test image collection

With over millions of users and billions of images, *Flickr* is a social media service that has rapidly evolving in recent decades. The most important reason for the interest *Flickr* may be because of its policy of tagging. Furthermore, *Flickr* provides to the developers a PHP API (*FlickrPHP*), which is available for no-commercial use.

Within this context and for constructing a tourist image collection, we have used the model provided by the RI work of [85, 86]. In terms of implementation, we have constructed the test collection following two steps:

- First, we have collected  $n = 180$  photograph queries that illustrate popular tourist attractions (e.g., churches, bridges, historic monuments, etc.) to ensure their identification later on, by a large number of users. They are also chosen to fit in the spatial neighborhoods of other images and vice versa.
- Second, each query photograph is loaded on the Web interface that we have developed, to extract a set of *socio-tagged* photographs from the free database of *Flickr*. The location coordinates of each query image are automatically extracted from its format. In addition, the Web interface enables to introduce the radius of the photograph recovery as well as the maximum number of photographs to download (for example, in Fig. 18, a maximum number of 50 images and a distance of 5 km are chosen).

As a result, we have ended up with a test collection that includes 25.000 separate images and 15.562 unique tags.

## 6.2 Implementation of the prototype “Socio-touring tool”

As previously mentioned, “*Socio-touring tool*” consists of three modules such as:

- *Administrator’s model* It helps automatically model the semantic facet patterns of the test collection images, initially tagged by the social community. These images are then automatically annotated by applying a set of semantic and inference rules. The next step consists in classifying the images based on the semantic information associated with them. This model also offers the possibility to consult the evaluation results generated following the run of the suggested experiments.
- *Expert’s model* It enables the experts in semiotics and tourism domain to express their opinions about the modeling, annotation, clustering and querying of the images of the test collection.
- *User’s model* It makes it possible to enter the user’s query, reformulate it, interrogate the pertinent image subgroups and visualize the returned results.

Figure 19 shows the global architecture of the *Socio-touring tool*.

Indeed, the prototype “*Socio-touring tool*” consists of many interfaces among which:

- An interface that helps browse the test image collection and generate for each image its corresponding semantic facet pattern that is saved in an XML file (Fig. 20).
- An interface that helps automatically annotate the modeled images and save the resulted annotation information in the RDF file of the used ontology (Fig. 21).
- An interface that helps cluster the annotated images by specifying the clustering process parameters, saving the resulted clustering information and consulting the results of the application of *Elbow*, *Silhouette* and *Gap Statistic* results (Fig. 22).
- An interface that helps the experts in tourism domain manually classify the test images (Fig. 23).
- An interface that helps users enter their queries, automatically reformulate them to interrogate the pertinent image subgroups and visualize the returned results (Fig. 24).

## 6.3 Evaluation of the image semantic clustering process

In order to evaluate the proposed image semantic clustering process, we have relied on the method suggested by Im et al. [87], which is adopted for testing the large collections of data. In fact, this method consists in selecting from each 1.000 images of the test collection 100 sample images. These images are considered as exemplary images and constitute the focus of the experiments that will be performed.

It should be noted that we have constructed a collection of 25.000 images among which 2.500 images are selected as samples to evaluate the semantic clustering process. Three

Fig. 19 General architecture of Socio-touring tool

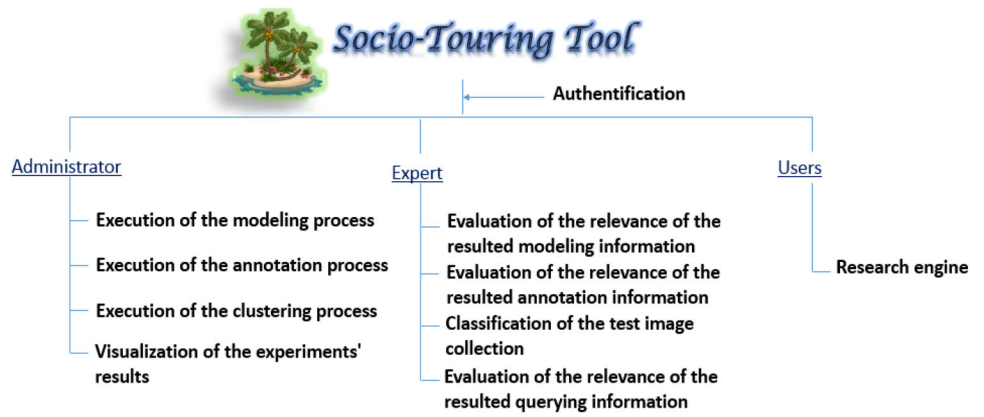


Fig. 20 Administrator interface: semantic modeling process

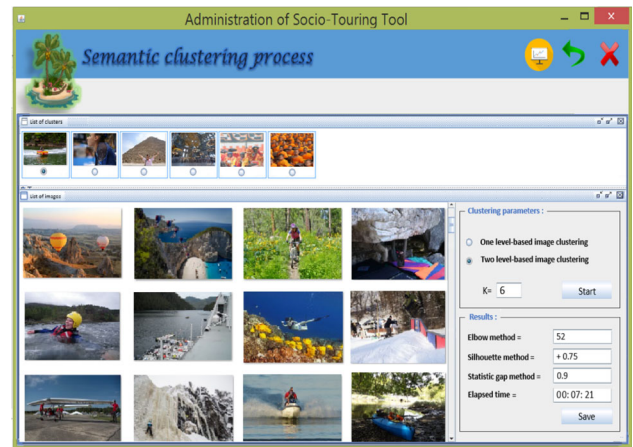


Fig. 22 Administrator interface: semantic clustering process

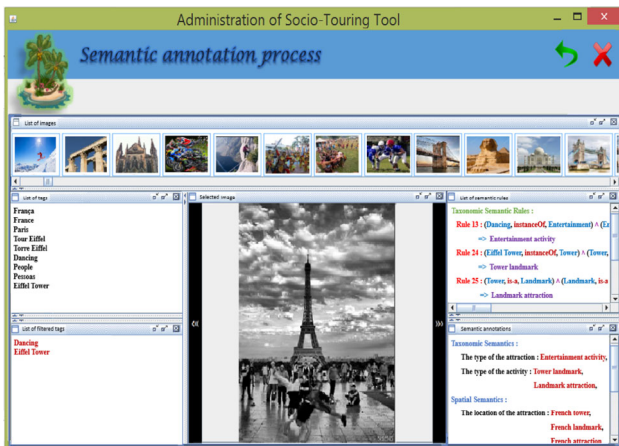


Fig. 21 Administrator interface: semantic annotation process

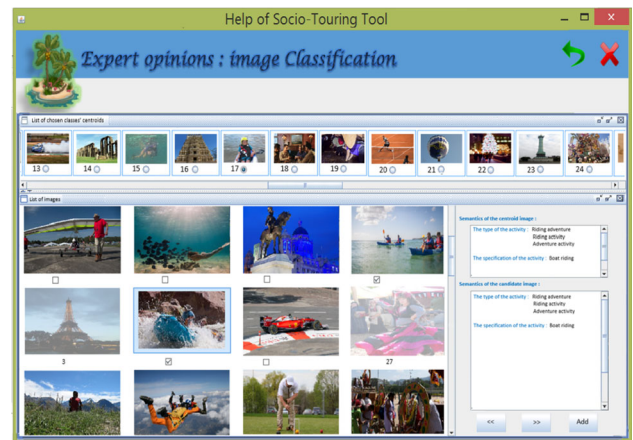


Fig. 23 Expert opinion interface: image classification

experts in semiotics have first chosen and named a set of 49 images from the sampled images. These images have been considered as centroids of the classes built by the experts in the tourism domain during the manual classification. Moreover, it should be mentioned that the optimum numbers of clusters that have been considered after applying *Elbow*, *Sil-*

*houette* and *Gap statistic* methods in the one-level-based and the two-level-based image clustering processes are 30 and 51, respectively. Therefore, we conclude that the optimum number of clusters obtained when applying the two-level-based image clustering process is closer to the real value chosen by



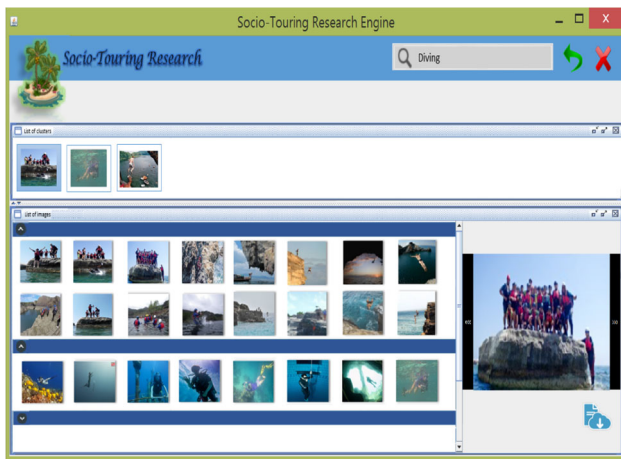


Fig. 24 User research interface

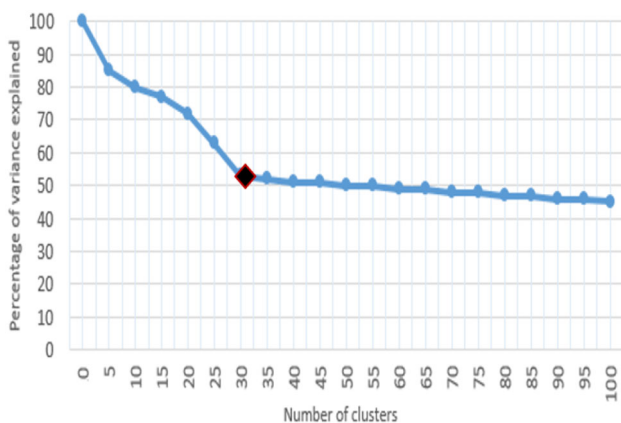


Fig. 25 Results of the Elbow method to determine the optimum number of clusters when applying the one-level-based image clustering process

the experts in semiotics than the one obtained when applying the one-level-based image clustering process.

The results generated following the application of the methods of optimization of the number of clusters in the two image clustering processes are presented as follows (Figs. 25, 26, 27, 28, 29, 30):

After choosing the centroids of the clusters by the experts in semiotics, five experts in the tourism domain have classified the 2.451 remaining test images based on their semantics. Indeed, the manual image classification is intended to construct a basic reference adopted when evaluating the suggested image semantic clustering process.

The procedure of evaluating the image semantic clustering process is guided by the technique suggested by Ricardo et al. [88], which is intended to evaluate the performance of IR systems. Thus, the *precision*, *recall* and *F-measure* are computed as indicated by Eqs. (2), (3) and (4):

$$\text{Precision} = \sum_{k=1}^n \frac{I_r}{I_t} \tag{2}$$

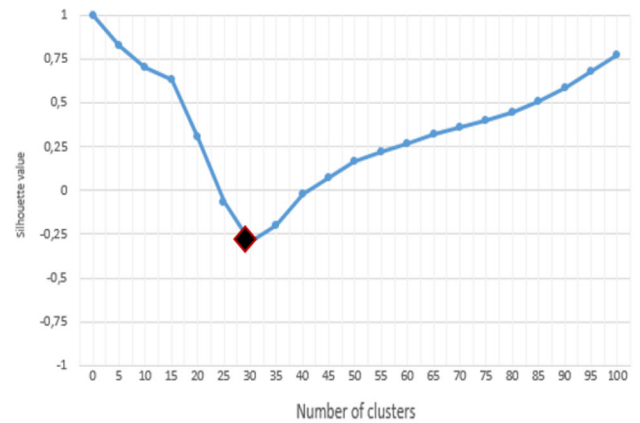


Fig. 26 Results of the Silhouette method to determine the optimum number of clusters when applying the one-level-based image clustering process

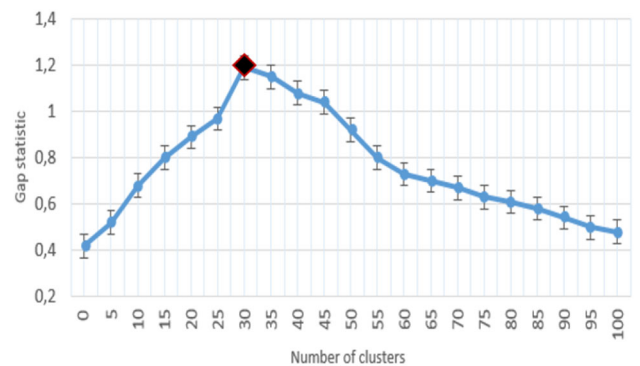


Fig. 27 Results of the Gap statistic method to determine the optimum number of clusters when applying the one-level-based image clustering process

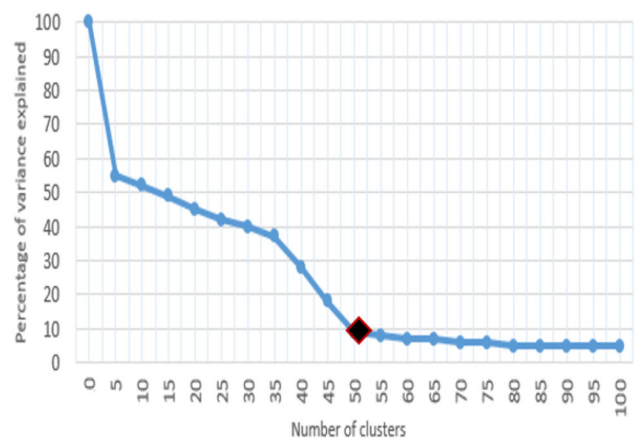
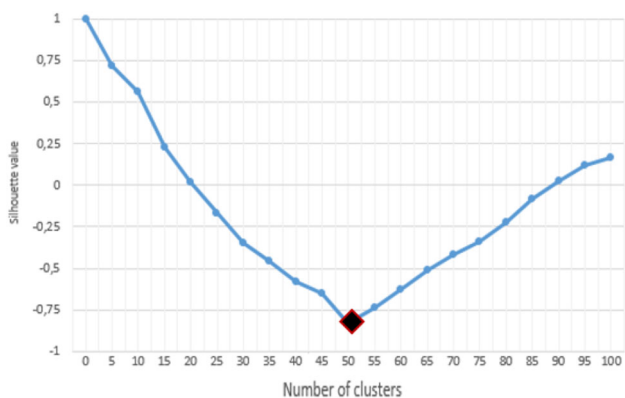


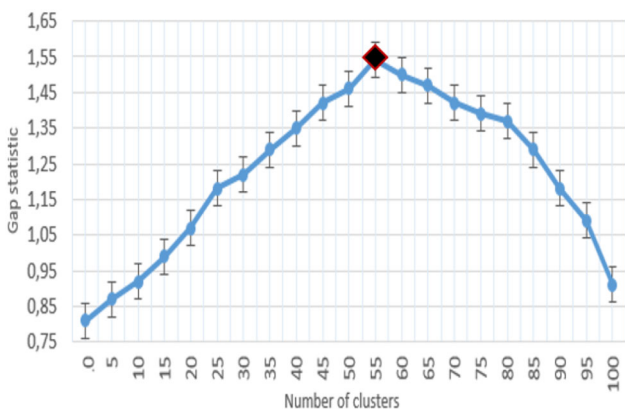
Fig. 28 Results of the Elbow method to determine the optimum number of clusters when applying the two-level-based image clustering process

$$\text{Recall} = \sum_{k=1}^n \frac{I_r}{I_s} \tag{3}$$

$$F\text{-measure} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$



**Fig. 29** Results of the Silhouette method to determine the optimum number of clusters when applying the two-level-based image clustering process



**Fig. 30** Results of the Gap statistic method to determine the optimum number of clusters when applying the two-level-based image clustering process

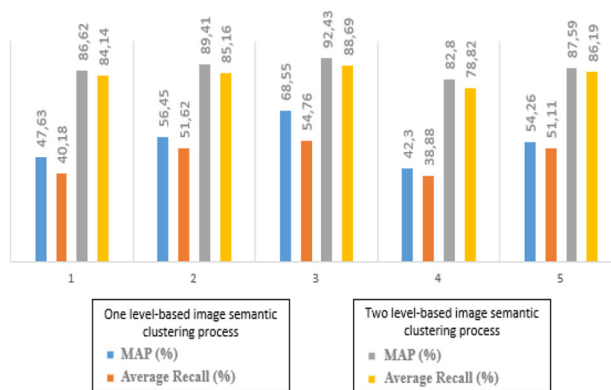
where  $I_r$  represents the number of relevant images assigned to the cluster  $k$ ,  $I_t$  represents the total number of images assigned to the cluster  $k$  and  $I_s$  represents the number of images that must be really assigned to the cluster  $k$ .

The evaluation results of the one-level-based image semantic clustering process are in the order of 53.83% in terms of *mean average precision* (MAP) and 47.31% in terms of *average recall* (AveR). However, the evaluation results of the proposed two-level-based image semantic clustering process are in the order of 87.77% in terms of MAP and 84.60% in terms of AveR. Therefore, the image semantic clustering process at two levels outperforms in terms of *average F-score* (AvgF) performance compared to the AvgF performance of the image semantic clustering process at one level. Table 1 summarizes the different obtained evaluation results.

Figure 31 shows the different evaluation results of the one-level-based and the two-level-based image semantic clustering processes. It should be mentioned that the evaluation results of the 2.500 sampled images are divided into

**Table 1** Average precision, recall and  $F$ -measure: performance comparison (in %) of the one-level-based and the two-level-based image semantic clustering processes

	MAP	AvgR	AvgF
One-level-based image semantic clustering process	53.83	47.31	50.35
Two-level-based image semantic clustering process	87.77	84.60	86.15



**Fig. 31** Evaluation results of the one-level-based and the two-level-based image semantic clustering processes

five groups of ten classes each, with the exception of the last group that consists of nine classes.

The provided evaluation results prove that the application of the image semantic clustering at two levels is more efficient than that of the application of the image semantic clustering at one level.

### 6.4 Evaluation of the image semantic research process

To evaluate the performance of the proposed semantic research process, we have conducted experimental studies on the image collection that we have constructed. Thus, all the 25.000 images of the collection are employed as the database images for retrieval. In addition, we have chosen a set of 12 common ambiguous tag-queries, including skiing, diving, festival, etc. Five experts in the tourism domain have used these queries to indicate among the interrogated images the ones that are relevant and the ones that are missing.

In order to find the best performance of the proposed research process adopting the two-level-based image semantic clustering, different sizes of the semantic vocabulary (20, 50, 100, 150 and 200) using different semantic vocabulary percentages (10, 25, 50, 75 and 100%) per query are formulated. Indeed, the vocabulary size is an important parameter that affects the performance of the proposed technique. Increasing the size of the semantic vocabulary at a certain level for compact representation of the semantic contents of

**Table 2** Performance comparison and statistical analysis of different sizes of the semantic vocabulary and different percentages of semantic vocabulary per query on the constructed collection of tourist images

% of semantic vocabulary per query	Performance analysis in terms of the MAP performance (in %) on the different sizes of the semantic vocabulary				
	20	50	100	150	200
10%	76.71	78.02	80.37	82.64	85.65
25%	76.88	78.42	80.77	82.92	86.03
50%	77.38	78.26	80.98	83.31	86.22
75%	77.60	79.55	81.47	84.58	86.53
100%	77.87	79.88	81.76	84.85	<b>86.86</b>
MAP	77.29	78.83	81.07	83.66	86.26
SD	0.486	0.832	0.553	0.997	0.464
SE	0.217	0.372	0.247	0.445	0.207
Confidence interval	76.68–77.89	77.79–79.86	80.38–81.75	82.42–84.89	85.68–86.83
Nonparametric statistical analysis using Wilcoxon matched-pairs signed-rank test					
Z value	2.023	2.023	2.023	2.023	2.023
P value	0.043	0.043	0.043	0.043	0.043

Bold value indicate the best performance

the images increases the performance of the image retrieval, while larger sizes of the semantic vocabulary result in overfitting that also decreases the image retrieval performance. Query size can also affect the quality of the results. Increasing the size of the semantic query increases the retrieval performance, but the large size of the query tends to overfit.

The procedure of evaluating the image research process is also guided by the technique suggested by Ricardo et al. [88]. Thus, the *precision*, *recall* and *F-measure* are computed as indicated by Eqs. (5), (6) and (4):

$$\text{Precision} = \sum_{k=1}^n \frac{J_r}{J_t} \tag{5}$$

$$\text{Recall} = \sum_{k=1}^n \frac{J_r}{J_s} \tag{6}$$

where  $J_r$  represents the number of relevant images interrogated as a response to the query  $q$ ,  $J_t$  represents the total number of images interrogated as a response to the query  $q$  and  $J_s$  represents the number of images that must be really interrogated as a response to the query  $q$ .

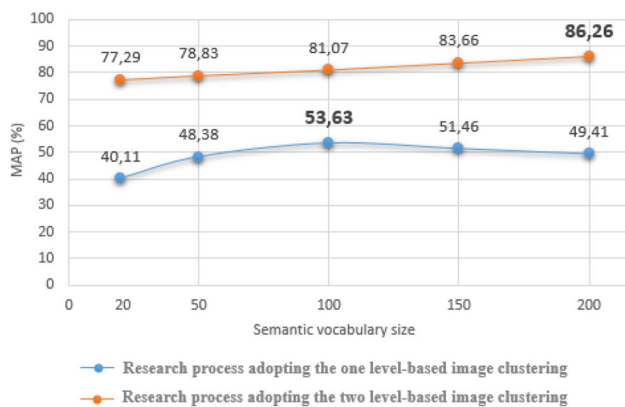
Table 2 presents the experimental details of the proposed research process adopting the two-level-based image semantic clustering on different sizes of the semantic vocabulary using different percentages of semantic vocabulary per query. The best MAP performance of 86.86% is achieved on a vocabulary size of 200 semantics using 100% semantic vocabulary per query. In order to verify the statistical significance of the experimental results of the proposed research process adopting the two-level-based image semantic clustering, the results of the statistical analysis are also reported in Table 2. The statistical results of the nonparametric Wilcoxon

matched-pairs signed-rank test are reported by comparing the obtained MAP performance on a vocabulary size of 200 semantics with other reported semantic vocabulary sizes (20, 50, 100 and 150) using the standard 95% confidence interval value. According to the statistical results of the nonparametric Wilcoxon matched-pairs signed-rank test, the proposed research process adopting the two-level-based image semantic clustering is statistically effective because the value of  $P$  is less than the level of the significance (i.e.,  $\alpha \leq 0.05$ ) for all the reported semantic vocabulary sizes.

The proposed research process adopting the two-level-based image semantic clustering is also statistically more accurate than the state-of-the-art CBIR techniques [18, 89] using visual vocabulary sizes of 200 visual words on the Corel-A image Benchmark.

In order to demonstrate the robustness of the proposed research process adopting the two-level-based image semantic clustering, the performance analysis in terms of MAP versus different sizes of the semantic vocabulary is also compared to the MAP performance of the research process adopting the one-level-based image semantic clustering, whose experimental details are shown in Fig. 32. According to the experimental details, the proposed research process adopting the two-level-based image semantic clustering outperforms in terms of the MAP performance compared to the research process adopting the one-level-based image semantic clustering, on all the reported semantic vocabulary sizes.

It should be noted that we have obtained a MAP performance of 86.90% when applying the proposed research process adopting the two-level-based image semantic clustering on a vocabulary size of 208 semantics using 100% semantic vocabulary per query. We conclude so that the proposed research system adopting the two-level-based image



**Fig. 32** Performance comparison (in %) in terms of MAP performance between the proposed research process adopting the two-level-based image semantic clustering and the research process adopting the one-level-based image semantic clustering on different sizes of the semantic vocabulary on the constructed image collection

semantic clustering is scalable, and we can therefore expand the semantic vocabulary size, unlike the competitor research system adopting the one-level-based image semantic clustering, where the best MAP performance of 54.17% is achieved on a vocabulary size of 100 semantics using 100% semantic vocabulary per query, as well as the state-of-the-art CBIR techniques [18, 89].

Moreover, it should be recalled that the interrogated image clusters can include a set of clusters with a weak semantic correlation versus the reformulated query, which should be filtered. Within this context, we have conducted a series of experiments in order to vary the  $\varepsilon$  value to obtain the optimal harmonic mean in terms of  $F$ -score that combines MAP and average recall (AvgR). Indeed, when fixing  $\varepsilon$  value, the clusters that have semantic similarities less than  $\varepsilon$  are considered irrelevant and therefore are filtered. Tables 3 and 4 summarize the different evaluation results obtained by varying  $\varepsilon$  value, when applying the research process adopting the one-level-based image semantic clustering on a vocabulary size of 100 semantics using 100% semantic vocabulary per query, and when applying the research process adopting the two-level-based image semantic clustering on a vocabulary size of 200 semantics using 100% semantic vocabulary per query.

Based on the results illustrated, respectively, in Tables 3 and 4, we conclude that the research process adopting the one-level-based image semantic clustering on a vocabulary size of 100 semantics using 100% semantic vocabulary per query provides the optimum  $F$ -score when fixing  $\varepsilon$  value at 0.46 and that the research process adopting the two-level-based image semantic clustering on a vocabulary size of 200 semantics using 100% semantic vocabulary per query provides the optimum  $F$ -score when fixing  $\varepsilon$  value at 0.38. Therefore, when applying the research process adopting the

**Table 3** Average precision, recall and  $F$ -measure: performance comparison (in %) when varying  $\varepsilon$  value and applying the research process adopting the one-level-based image semantic clustering on a vocabulary size of 100 semantics using 100% semantic vocabulary per query

	Research process adopting the one-level-based image semantic clustering		
	MAP.	AvgR.	AvgF.
$\varepsilon = 0$	24.30	62.42	34.98
$\varepsilon = 0.25$	28.64	58.83	38.52
$\varepsilon = 0.35$	34.56	56.96	43.01
$\varepsilon = 0.46$	<b>54.17</b>	<b>50.73</b>	<b>52.39</b>
$\varepsilon = 0.5$	57.66	33.24	42.16
$\varepsilon = 0.7$	66.73	26.13	37.55

Bold values indicate the best performance

**Table 4** Average precision, recall and  $F$ -measure: performance comparison (in %) when varying  $\varepsilon$  value and applying the research process adopting the two-level-based image semantic clustering on a vocabulary size of 200 semantics using 100% semantic vocabulary per query

	Research process adopting the two-level-based image semantic clustering		
	MAP.	AvgR.	AvgF.
$\varepsilon = 0$	27.82	83.34	41.71
$\varepsilon = 0.25$	35.24	83.27	49.52
$\varepsilon = 0.35$	77.87	83.16	80.42
$\varepsilon = 0.38$	<b>86.86</b>	<b>83.10</b>	<b>84.93</b>
$\varepsilon = 0.5$	87.27	42.02	56.72
$\varepsilon = 0.7$	87.41	36.17	51.16

Bold values indicate the best performance

one-level-based image semantic clustering, the clusters that have semantic similarities less than 0.46 versus the reformulated query are considered irrelevant and therefore are filtered. While when applying the research process adopting the two-level-based image semantic clustering, the clusters that have semantic similarities less than 0.38 versus the reformulated query are considered irrelevant and therefore are filtered.

On the other hand, and in order to observe the effect of the suggested users' query semantic reformulation on the quality of the returned search results, we have conducted experimental studies on the constructed collection of tourist images using both the users' tags associated with these images by the social community and the semantic annotations associated with them after executing the semantic annotation process. According to the experimental results shown in Table 5, the evaluation results of the research process of the images associated with the users' tags are in the order of 70.15% in terms of MAP and 23.63% in terms of AvgR. While the evaluation

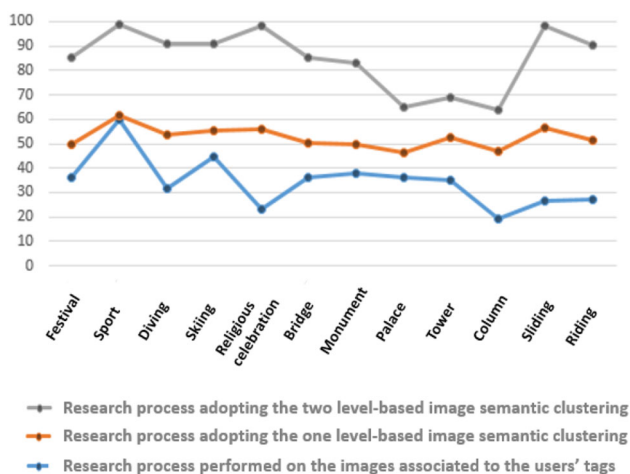


**Table 5** Comparative analysis (in %) in terms of the evaluation metrics of the proposed research process adopting the two-level-based image semantic clustering with its competitors on the constructed image collection

Ambiguous tag-queries	Performance measures	Research process performed on images associated with the users' tags	Research process adopting the one-level-based image semantic clustering	Proposed research process adopting the two-level-based image semantic clustering
Festival	Precision	61.65	51.40	88.45
	Recall	25.40	47.90	82.70
	<i>F-measure</i>	<i>35.97</i>	<i>49.58</i>	<i>85.47</i>
Sport	Precision	92.25	62.23	100
	Recall	44.10	60.80	98.00
	<i>F-measure</i>	<i>59.67</i>	<i>61.50</i>	<i>98.98</i>
Diving	Precision	70.58	56.50	92.33
	Recall	20.40	50.80	89.10
	<i>F-measure</i>	<i>31.65</i>	<i>53.49</i>	<i>90.68</i>
Skiing	Precision	69.13	57.32	93.33
	Recall	32.70	53.20	88.20
	<i>F-measure</i>	<i>44.39</i>	<i>55.18</i>	<i>90.69</i>
Religious celebration	Precision	88.81	56.87	100
	Recall	13.50	54.60	96.70
	<i>F-measure</i>	<i>23.43</i>	<i>55.71</i>	<i>98.32</i>
Bridge	Precision	64.26	52.46	87.35
	Recall	25.00	47.90	83.60
	<i>F-measure</i>	<i>35.99</i>	<i>50.07</i>	<i>85.43</i>
Monument	Precision	60.97	50.95	85.08
	Recall	27.50	48.20	81.60
	<i>F-measure</i>	<i>37.90</i>	<i>49.53</i>	<i>83.30</i>
Palace	Precision	57.64	47.67	66.74
	Recall	26.40	45.10	63.80
	<i>F-measure</i>	<i>36.21</i>	<i>46.34</i>	<i>65.23</i>
Tower	Precision	60.05	55.03	70.58
	Recall	24.50	50.30	66.90
	<i>F-measure</i>	<i>34.80</i>	<i>52.55</i>	<i>68.69</i>
Column	Precision	61.08	48.08	65.61
	Recall	11.30	45.30	62.20
	<i>F-measure</i>	<i>19.07</i>	<i>46.64</i>	<i>63.85</i>
Sliding	Precision	91.22	58.11	100
	Recall	15.60	54.80	96.20
	<i>F-measure</i>	<i>26.64</i>	<i>56.40</i>	<i>98.06</i>
Riding	Precision	64.16	53.42	92.85
	Recall	17.20	49.90	88.30
	<i>F-measure</i>	<i>27.12</i>	<i>51.60</i>	<i>90.51</i>
	MAP	70.15	54.17	<b>86.86</b>
	AvgR.	23.63	50.73	<b>83.10</b>
	AvgF.	35.35	52.39	<b>84.93</b>

Bold values indicate the best performance

Italic values indicate the results in term of F-measure

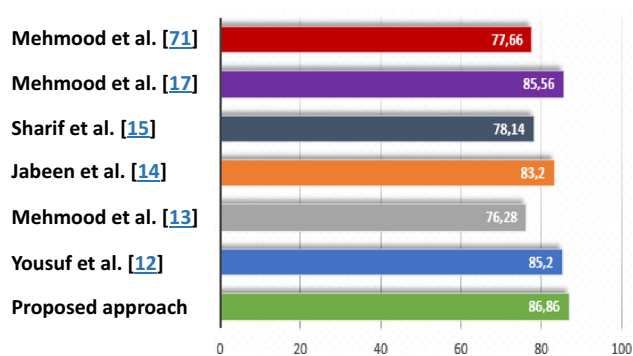


**Fig. 33** Performance comparison in terms of *F-measure* on the tourist image collection

results of the semantic research process of the images are in the order of 54.17% in terms of MAP and 50.73% in terms of AvgR when adopting the one-level-based image semantic clustering process on a vocabulary size of 100 semantics using 100% semantic vocabulary per query, and they are in order of 86.86% in terms of MAP and 83.10% in terms of AvgR when adopting the two-level-based image semantic clustering on a vocabulary size of 200 semantics using 100% semantic vocabulary per query. Therefore, we conclude that the suggested semantic-based image research process outperforms in terms of the *Average F-score* (AvgF) performance compared to the users' tags-based image research process. The provided evaluation results also prove that the application of an image research process is more efficient when we jointly investigate the two processes at ambiguous query preprocessing and postprocessing levels, especially the proposed two-level-based image semantic clustering process.

A comparative analysis of performance of the proposed research process adopting the two-level-based image semantic clustering in terms of *F-measure* is performed with the research process adopting the one-level-based image semantic clustering and the research process performed on the collection of tourist images associated with the users' tags, whose experimental details are shown in Fig. 33. Figure 33 clearly indicates that the proposed research process adopting the two-level-based image semantic clustering yields better performance compared with its competitor techniques, on all the reported tag-based queries.

A comparative analysis of the proposed research process adopting the two-level-based image semantic clustering on the constructed collection of tourist images is also performed with recent CBIR techniques [12–15, 17, 71] on the Corel-B image repository, whose details are presented in Fig. 34. According to the presented experimental results, the suggested approach on a vocabulary size of 200 semantics using



**Fig. 34** Analysis of MAP performance comparison of the proposed research process adopting the two-level-based image semantic clustering on the constructed collection of tourist images with recent CBIR techniques on the Corel-B image repository

100% semantic vocabulary per query also outperforms compared to the recent CBIR techniques.

## 6.5 Computational complexity

The performance of the proposed research process adopting the two-level-based image semantic clustering is measured on a computer, whose hardware specifications are as follows: RAM with 8 GB storage capacity, GPU with 2 GB storage capacity and Intel Pentium (R) Core i3 microprocessor with 1.7 GHz clock frequency.

The required software resources for the implementation of the proposed technique are Microsoft Windows 7 operating system (64-bits) and Eclipse Tool. The average computational complexity required to reformulate an ambiguous user's query consisting of a single word is 0.9809 (s). A comparative analysis is performed with the computational complexity required by the recent CBIR techniques for feature extraction only [13–15, 17, 90], whose details are presented in Table 6.

Computational complexity of the proposed research process adopting the two-level-based image semantic clustering (i.e., complete framework) is reported by selecting the representative images of the clusters. A comparative analysis, presented in Table 7, is also performed with the computational complexity of the recent CBIR techniques [13, 17, 91], reported by selecting the image benchmark of the Corel-1 K, which comprises each image resolution of  $384 \times 256$  or  $256 \times 384$ .

## 7 Conclusion and future research studies

In this paper, we present a novel approach to diversify the research results interrogated on the social image research systems in order to reduce the tag-based query ambiguity issue that affects Tag-BIR performance. We jointly investi-

**Table 6** Comparative analysis of the average computational complexity (time in seconds) of the proposed technique required to reformulate an ambiguous user's query consisting of a single word with the com-

putational complexity required by recent CBIR techniques for feature extraction only

Proposed approach	Mehmood et al. [13]	Jabeen et al. [14]	Sharif et al. [15]	Mehmood et al. [17]	Tian et al. [90]
0.9809	0.1456	0.1730	0.1422	0.0788	5.6

**Table 7** Comparative analysis of the computational complexity (time in seconds) of the proposed image retrieval framework with recent CBIR techniques

Retrieved (representative) images	Proposed approach	Mehmood et al. [13]	Mehmood et al. [17]	Ali et al. [91]
05 (representative) images	1.6783	0.2880	0.3601	0.3726
10 (representative) images	2.2892	0.3980	0.4829	0.5178
15 (representative) images	2.8007	0.6475	0.6901	0.7050

gate two processes at the ambiguous query preprocessing and postprocessing levels. The *Tag-based Query Semantic Reformulation* process is intended to reformulate the tag-based users' queries according to multiple semantic facets using a set of predefined ontological semantic rules. Besides, the *Multi-level Image Diversification process* aims first at clustering images based on their semantic content offline, and second filtering and re-ranking the returned image cluster results according to their pertinence versus the reformulated query online.

In fact, the key advantage of the suggested approach is its ability to make effective the semantic reformulation of an ambiguous tag-query in order to guarantee a complete coverage of all the query aspects and ensure relevance of the returned results. The proposed approach also enables to identify at a glance the required images thanks to a semantic clustering of the diversified results. Indeed, the image clustering process that we propose is executed at two successive levels: a *Semantic Facet-based Image Clustering level* and a *Semantic Concept-based Image Clustering level*. The proposed image semantic clustering at two levels improves the accuracy of the clustering results, when it consists of an important semantic vocabulary size on a large semantic data space, compared to the research process adopting a one-level-based image semantic clustering. The filtering and re-ranking of the image cluster retrieval results enables to reorder the retrieved results according to their pertinence versus the reformulated query, and to achieve higher accuracy rate. Time-consuming computation tasks, including image similarity measures, image clustering and representative image selection, are precomputed offline to support a lightweight image search result clustering system.

Experimental studies on a collection of 25.000 images shared on *Flickr* prove the effectiveness of the suggested research process adopting a two-level-based image seman-

tic clustering compared to the research process adopting a one-level-based image semantic clustering and the research process based on the users' tags associated with the images.

The experimental studies also demonstrate that the proposed research process adopting the two-level-based image semantic clustering on the constructed collection of tourist images outperforms compared to the recent CBIR techniques on the Corel-B image repository.

In our future research studies, we are interested in incorporating the suggested image semantic clustering process at two levels on a *MapReduce* pattern to scale up for large-scale image data.

## References

1. Chaturani, N.W.U.D., Geva, S., Chandran, V., Cynthujah, V.: An effective content based image retrieval system based on global representation and multi-level searching. In: Proceedings of the 10th International Conference on Industrial and Information Systems (ICIIS), pp. 158–163 (2015)
2. Wang, X.Y., Zhang, B.B., Yang, H.Y.: Content-based image retrieval by integrating color and texture features. *Multimed. Tools Appl.* **68**(3), 545–569 (2014)
3. Atlam, H.F., Attiya, G., El-Fishawy, N.: Integration of color and texture features in CBIR system. *Int. J. Comput. Appl.* **164**(3), 23–29 (2017)
4. Liu, G.H., Yang, J.Y.: Content-based image retrieval using color difference histogram. *Pattern Recogn.* **46**(1), 188–198 (2013)
5. Chang, B.M., Tsai, H.H., Chou, W.L.: Using visual features to design a content-based image retrieval method optimized by particle swarm optimization algorithm. *Eng. Appl. Artif. Intell.* **26**(10), 2372–2382 (2013)
6. Liu, G.H., Li, Z.Y., Zhang, L., Xu, Y.: Image retrieval based on micro-structure descriptor. *Pattern Recogn.* **44**(9), 2123–2133 (2011)
7. ElAlami, M.E.: A new matching strategy for content based image retrieval system. *Appl. Soft Comput.* **14**, 407–418 (2014)

8. Jiji, G.W., DuraiRaj, P.J.: Content-based image retrieval techniques for the analysis of dermatological lesions using particle swarm optimization technique. *Appl. Soft Comput.* **30**, 650–662 (2015)
9. Wang, X.Y., Yu, Y.J., Yang, H.Y.: An effective image retrieval scheme using color, texture and shape features. *Comput. Stand. Interfaces* **33**(1), 59–68 (2011)
10. Pass, G., Zabih, R.: Comparing images using joint histograms. *Multimed. Syst.* **7**(3), 234–240 (1999)
11. Wang, X.Y., Li, Y.W., Yang, H.Y., Chen, J.W.: An image retrieval scheme with relevance feedback using feature reconstruction and SVM reclassification. *Neurocomputing* **127**, 214–230 (2014)
12. Yousuf, M., Mehmood, Z., Habib, H.A., Mahmood, T., Saba, T., Rehman, A., Rashid, M.: A novel technique based on visual words fusion analysis of sparse features for effective content-based image retrieval. *Math. Probl. Eng.* **2018**, 1–13 (2018)
13. Mehmood, Z., Abbas, F., Mahmood, T., Javid, M.A., Rehman, A., Nawaz, T.: Content-based image retrieval based on visual words fusion versus features fusion of local and global features. *Arab. J. Sci. Eng.* **43**(12), 7265–7284 (2018)
14. Jabeen, S., Mehmood, Z., Mahmood, T., Saba, T., Rehman, A., Mahmood, M.T.: An effective content-based image retrieval technique for image visuals representation based on the bag-of-visual-words model. *PLoS ONE* **13**(4), e0194526 (2018)
15. Sharif, U., Mehmood, Z., Mahmood, T., Javid, M.A., Rehman, A., Saba, T.: Scene analysis and search using local features and support vector machine for effective content-based image retrieval. *Artif. Intell. Rev.* **2018**, 1–25 (2018)
16. Sarwar, A., Mehmood, Z., Saba, T., Qazi, K.A., Adnan, A., Jamal, H.: A novel method for content-based image retrieval to improve the effectiveness of the bag-of-words model using a support vector machine. *J. Inf. Sci.* **45**(1), 117–135 (2019)
17. Mehmood, Z., Gul, N., Altaf, M., Mahmood, T., Saba, T., Rehman, A., Mahmood, M.T.: Scene search based on the adapted triangular regions and soft clustering to improve the effectiveness of the visual-bag-of-words model. *EURASIP J. Image Video Process.* **2018**(1), 48 (2018)
18. Mehmood, Z., Anwar, S.M., Altaf, M.: A novel image retrieval based on rectangular spatial histograms of visual words. *Kuwait J. Sci.* **45**(1), 54–69 (2018)
19. Smeulders, A.W., Worring, M., Santini, S., Gupta, A., Jain, R.: Content-based image retrieval at the end of the early years. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(12), 1349–1380 (2000)
20. Adebayo, S., McLeod, K., Tudose, I., Osumi-Sutherland, D., Burdett, T., Baldock, R., Parkinson, H.: PhenImageShare: an image annotation and query infrastructure. *J. Biomed. Semant.* **7**(1), 35–44 (2016)
21. Kurtz, C., Rubin, D.L.: Utilisation de relations ontologiques pour la comparaison d'images décrites par des annotations sémantiques. In: *Proceedings of the (EGC)*, pp. 609–614 (2014)
22. Lingutla, N.T., Preece, J., Todorovic, S., Cooper, L., Moore, L., Jaiswal, P.: AISO: annotation of image segments with ontologies. *J. Biomed. Semant.* **5**(1), 50–54 (2014)
23. Hollink, L., Nguyen, G., Schreiber, G., Wielemaker, J., Wielinga, B., Worring, M.: Adding spatial semantics to image annotations. In: *Proceedings of the 4th International Workshop on Knowledge Markup and Semantic Annotation (ISWC)*, pp. 31–40 (2004)
24. Zhang, C., Chai, J., Jin, R.: User term feedback in interactive text-based image retrieval. In: *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 51–58 (2005)
25. Weinberger, K.Q., Slaney, M., Van Zwol, R.: Resolving tag ambiguity. In: *Proceedings of the 16th ACM International Conference on Multimedia (MM)*, pp. 111–120 (2008)
26. Ksibi, A., Ammar, A.B., Amar, C.B.: Adaptive diversification for tag-based social image retrieval. *Int. J. Multimed. Inf. Retr.* **3**(1), 29–39 (2014)
27. Van Leuken, R.H., Garcia, L., Olivares, X., Van Zwol, R.: Visual diversification of image search results. In: *Proceedings of the 9th International Conference on World Wide Web (WWW)*, pp. 341–350 (2009)
28. Radlinski, F., Bennett, P.N., Carterette, B., Joachims, T.: Redundancy, diversity and interdependent document relevance. *SIGIR Forum.* **43**(2), 46–52 (2009)
29. Hoque, E., Hoeber, O., Gong, M.: Evaluating the tradeoffs between diversity and precision for web image search using concept-based query expansion. In: *Proceedings of the 11th International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, vol. 3, pp. 130–133 (2011)
30. Carpineto, C., Romano, G.: A survey of automatic query expansion in information retrieval. *ACM Comput. Surv.* **44**(1), 1–50 (2012)
31. Lioma, C., Moens, M.F., Azzopardi, L.: Collaborative annotation for pseudo relevance feedback. In: *Proceedings of Exploiting Semantic Annotation in Information Retrieval (ESAIR)*, pp. 25–35 (2008)
32. Naouar, F., Hlaoua, L., Omri, M.N.: Relevance feedback in collaborative information retrieval based on validated annotation. In: *Proceedings of the 1st International Conference on Reasoning and Optimization in Information Systems*, vol. 13 (2013)
33. Guo, Q., Liu, W., Lin, Y., Lin, H.: Query expansion based on user quality in folksonomy. In: *Proceedings of Asia Information Retrieval Symposium (AIRS)*, pp. 396–405 (2012)
34. Kitanovski, I., Strezoski, G., Dimitrovski, I., Madjarov, G., Loskovska, S.: Multimodal medical image retrieval system. *Multimed. Tools Appl.* **76**(2), 2955–2978 (2017)
35. Garg, N., Gatica-Perez, D.: Tagging and retrieving images with co-occurrence models: from corel to flickr. In: *Proceedings of the 1st ACM Workshop on Large-Scale Multimedia Retrieval and Mining*, pp. 105–112 (2009)
36. Jin, S., Lin, H., Su, S.: Query expansion based on folksonomy tag co-occurrence analysis. In: *Proceedings of the International Conference on Granular Computing*, pp. 300–305 (2009)
37. Broder, A., Ciccolo, P., Gabrilovich, E., Pang, B.: Domain-specific query augmentation using folksonomy tags: the case of contextual advertising. In: *Proceedings of the 1st Workshop on Information Retrieval for Advertising (2008)*
38. Ksibi, A., Ammar, A.B., Amar, C.B.: Enhanced context-based query-to-concept mapping in social image retrieval. In: *Proceedings of the 11th International Workshop on Content-Based Multimedia Indexing (CBMI)*, pp. 85–89 (2013)
39. Wang, J., Davison, B.D.: Explorations in tag suggestion and query expansion. In: *Proceeding of the 2008 ACM Workshop on Search in Social Media (SSM)*, pp. 43–50 (2008)
40. Biancalana, C., Micarelli, A.: Social tagging in query expansion: a new way for personalized web search. In: *Proceedings of the 12th IEEE International Conference on Computational Science and Engineering (CSE)*, pp. 1060–1065 (2009)
41. Abbasi, R.: Query expansion in folksonomies. In: *International Conference on Semantic and Digital Media Technologies (SAMT)*, pp. 1–16 (2010)
42. He, Y., Li, Y., Lei, J., Leung, C.H.: A framework of query expansion for image retrieval based on knowledge base and concept similarity. *Neurocomputing* **204**, 26–32 (2016)
43. Otegi, A., Arregi, X., Ansa, O., Agirre, E.: Using knowledge-based relatedness for information retrieval. *Knowl. Inf. Syst.* **44**(3), 689–718 (2015)
44. Hoque, E., Hoeber, O., Gong, M.: Balancing the trade-offs between diversity and precision for web image search using concept-based query expansion. *J. Emerg. Technol. Web Intell.* **4**(1), 26–34 (2012)
45. Leong, C.W., Hassan, S., Ruiz, M.E., Mihalcea, R.: Improving query expansion for image retrieval via saliency and picturability. In: *International Conference of the Cross-Language Evaluation Forum for European Languages (CLEF)*, pp. 137–142 (2011)



46. Yuvarani, M., Iyengar, N.C.S., Kannan, A.: Improved concept-based query expansion using Wikipedia. *Int. J. Commun. Netw. Distrib. Syst.* **11**(1), 26–41 (2013)
47. Ayadi, Y., Amous, I., Gargouri, F.: Toward an automatic annotation approach based on ontological enrichment for advanced research. *Int. J. Eng. Technol.* **13**(2), 80–89 (2013)
48. Alromima, W., Moawad, I.F., Elgohary, R., Aref, M.: Ontology-based query expansion for Arabic text retrieval. *Int. J. Adv. Comput. Sci. Appl.* **7**(8), 223–230 (2016)
49. Devi, M.U., Gandhi, G.M.: Wordnet and ontology based query expansion for semantic information retrieval in sports domain. *J. Comput. Sci.* **11**(2), 361–371 (2014)
50. Charhad, M., Zrigui, M., Quénot, G.: Une approche conceptuelle pour la modélisation et la structuration sémantique des documents vidéos. In: Proceedings of the 3rd International Conference: Sciences of Electronic, Technologies of Information and Telecommunications (SETIT), pp. 1–7 (2005)
51. Bouchakwa, M., Ayadi, Y., Amous, I.: Modeling the semantic content of the socio-tagged images based on the extended conceptual graphs formalism. In: Proceedings of the 14th International Conference on Advances in Mobile Computing and Multi Media (MOMM), pp. 35–39 (2016)
52. De Andrade, D.O.S., Maia, L.F., De Figueirêdo, H.F., Viana, W., Trinta, F., De Souza Baptista, C.: Photo annotation: a survey. *Multimed. Tools Appl.* **77**(1), 423–457 (2018)
53. Van Leuken, R.H., Garcia, L., Olivares, X., Van Zwol, R.: Visual diversification of image search results. In: Proceedings of the 18th International Conference on World Wide Web (WWW), pp. 341–350 (2009)
54. Wang, M., Yang, K., Hua, X.S., Zhang, H.J.: Towards a relevant and diverse search of social images. *IEEE Trans. Multimed.* **12**(8), 829–842 (2010)
55. Sun, F., Wang, M., Wang, D., Wang, X.: Optimizing social image search with multiple criteria: relevance, diversity, and typicality. *Neurocomputing* **95**, 40–47 (2012)
56. Carbonell, J., Goldstein, J.: The use of MMR, diversity-based re-ranking for reordering documents and producing summaries. In: Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 335–336 (1998)
57. Agrawal, R., Gollapudi, S., Halverson, A., Ieong, S.: Diversifying search results. In: Proceedings of the 2nd ACM International Conference on Web Search and Data Mining (WSDM), pp. 5–14 (2009)
58. 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), pp. 881–890 (2010)
59. Santos, R.L., Macdonald, C., Ounis, I.: Selectively diversifying web search results. In: Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM), pp. 1179–1188 (2010)
60. Song, K., Tian, Y., Gao, W., Huang, T.: Diversifying the image retrieval results. In: Proceedings of the 14th ACM International Conference on Multimedia (ICM), pp. 707–710 (2006)
61. Qian, X., Lu, D., Wang, Y., Zhu, L., Tang, Y.Y., Wang, M.: Image re-ranking based on topic diversity. *IEEE Trans. Image Process.* **26**(8), 3734–3747 (2017)
62. Yan, Y., Liu, G., Wang, S., Zhang, J., Zheng, K.: Graph-based clustering and ranking for diversified image search. *Multimed. Syst.* **23**(1), 41–52 (2017)
63. Lashari, S.A., Ibrahim, R.: A framework for medical images classification using soft set. *Procedia Technol.* **11**(2013), 548–556 (2013)
64. Wen, W., Hao, Z., Cai, R., Shao, Z.: Gaussian process learning for image classification based on low-level features. In: Proceedings of the 8th International Conference on Natural Computation (CNC), pp. 237–241 (2012)
65. Cai, D., He, X., Li, Z., Ma, W.Y., Wen, J.R.: Hierarchical clustering of WWW image search results using visual, textual and link information. In: Proceedings of the 12th Annual ACM International Conference on Multimedia (MULTIMEDIA), pp. 952–959 (2004)
66. Sabarad, A.K., Kankudti, M.H., Meena, S.M., Husain, M.: Color and texture feature extraction using Apache Hadoop framework. In: Proceedings of the 2015 International Conference on Computing Communication Control and Automation (ICCCBEA), pp. 585–588 (2015)
67. Zhu, H., Shen, Z., Shang, L., Zhang, X.: Parallel image texture feature extraction under Hadoop cloud platform. In: Proceedings of the International Conference on Intelligent Computing (ICIC), pp. 459–465 (2014)
68. Qayyum, A., Anwar, S.M., Awais, M., Majid, M.: Medical image retrieval using deep convolutional neural network. *Neurocomputing* **266**, 8–20 (2017)
69. Li, Y.D., Hao, Z.B., Lei, H.: Survey of convolutional neural network. *Int. J. Comput. Appl.* **36**(9), 2508–2515 (2016)
70. Lyndon, D., Kumar, A., Kim, J., Leong, P.H. W., Feng, D.: Convolutional neural networks for medical clustering. In: CLEF (Working Notes) (2015)
71. Mehmood, Z., Rashid, M., Rehman, A., Saba, T., Dawood, H., Dawood, H.: Effect of complementary visual words versus complementary features on clustering for effective content-based image search. *J. Intell. Fuzzy Syst.* **2019**, 1–14 (2019)
72. Cha, G.H.: Kernel principal component analysis for content based image retrieval. In: Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), pp. 844–849 (2005)
73. Wang, S., Jing, F., He, J., Du, Q., Zhang, L.: Igroup: presenting web image search results in semantic clusters. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 587–596 (2007)
74. Ramage, D., Heymann, P., Manning, C.D., Garcia-Molina, H.: Clustering the tagged web. In: Proceedings of the 2nd ACM International Conference on Web Search and Data Mining (WSDM), pp. 54–63 (2009)
75. Yin, Z., Li, R., Mei, Q., Han, J.: Exploring social tagging graph for web object classification. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), pp. 957–966 (2009)
76. Wang, X.J., Zhang, L., Jing, F., Ma, W.-Y.: Annosearch: image auto-annotation by search. *Proc. Comput. Vis. Pattern Recogn.* **2**, 1483–1490 (2006)
77. Bouchakwa, M., Ayadi, Y., Amous, I.: Semantic pattern-based automatic annotation process of images shared on social networks. In: Proceedings of the 30th IBIMA Conference (IBIMA), pp. 19–38 (2017)
78. Sowa, J.F.: Conceptual structures: information processing in mind and machine. Addison-Wesley publishing Company, Reading (1984)
79. Cohen, J.: A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* **20**(1), 37–46 (1960)
80. MacQueen, J.: Some methods for classification and analysis of multivariate observations. In: Proceedings of the 5th Berkeley Symposium On Mathematical Statistics and Probabilities, vol. 1, pp. 281–296 (1967)
81. Thorndike, R.L.: Who belongs in the family? *Psychometrika* **18**(4), 267–276 (1953)
82. Rousseeuw, P.J., Kaufman, L.: Finding groups in data. *Ser. Probab. Math. Stat.* **34**(1), 111–112 (1990)
83. Tibshirani, R., Walther, G., Hastie, T.: Estimating the number of clusters in a data set via the gap statistic. *J. R. Stat. Soc.* **63**(2), 411–423 (2001)

84. Bhusare, B.B., Bansode, S.M.: Centroids initialization for K-means clustering using improved pillar algorithm. *Int. J. Adv. Res. Comput. Eng. Technol.* **3**(4), 1317–1322 (2014)
85. Chris, B., Ellen, M.: Evaluating evaluation measure stability. In: *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 33–40 (2000)
86. Ellen M.V.: The philosophy of information retrieval evaluation. In: *CLEFEL*, pp. 355–370 (2001)
87. Im, D.H., Park, G.D.: Linked tag: image annotation using semantic relationships between image tags. *Int. J. Multimed. Tools Appl.* **74**(7), 2273–2287 (2015)
88. Ricardo, B., Berthier, R.: *Modern Information Retrieval*. ACM Press, New York (1999)
89. Mehmood, Z., Anwar, S.M., Ali, N., Habib, H.A., Rashid, M.: A novel image retrieval based on a combination of local and global histograms of visual words. *Math. Probl. Eng.* **2016**, 12 (2016)
90. Tian, X., Jiao, L., Liu, X., Zhang, X.: Feature integration of EODH and Color-SIFT: application to image retrieval based on codebook. *Signal Process. Image Commun.* **29**(4), 530–545 (2014)
91. Ali, N., Bajwa, K.B., Sablatnig, R., Mehmood, Z.: Image retrieval by addition of spatial information based on histograms of triangular regions. *Comput. Electr. Eng.* **54**, 539–550 (2016)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.