Expert Systems With Applications 91 (2018) 12–26

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Semantic Guided Interactive Image Retrieval for plant identification

Filipe Marcel Fernandes Gonçalves, Ivan Rizzo Guilherme, Daniel Carlos Guimarães Pedronette*

Department of Statistics, Applied Mathematics and Computing (DEMAC), State University of São Paulo (UNESP), Av. 24-A, 1515, Rio Claro, SP 13506-900, Brazil

ARTICLE INFO

Article history:
Received 1 March 2017
Revised 18 August 2017
Accepted 19 August 2017
Available online 21 August 2017

Keywords:
Interactive image retrieval
Unsupervised learning
Semantic gap
Ontology

ABSTRACT

A lot of images are currently generated in many domains, requiring specialized knowledge of identification and analysis. From one standpoint, many advances have been accomplished in the development of image retrieval techniques based on visual image properties. However, the semantic gap between low-level features and high-level concepts still represents a challenging scenario. On another standpoint, knowledge has also been structured in many fields by ontologies. A promising solution for bridging the semantic gap consists in combining the information from low-level features with semantic knowledge. This work proposes a novel graph-based approach denominated Semantic Interactive Image Retrieval (SIIR) capable of combining Content Based Image Retrieval (CBIR), unsupervised learning, ontology techniques and interactive retrieval. To the best of our knowledge, there is no approach in the literature that combines those diverse techniques like SIIR. The proposed approach supports expert identification tasks, such as the biologist's role in plant identification of Angiosperm families. Since the system exploits information from different sources as visual content, ontology, and user interactions, the user efforts required are drastically reduced. For the semantic model, we developed a domain ontology which represents the plant properties and structures, relating features from Angiosperm families. A novel graph-based approach is proposed for combining the semantic information and the visual retrieval results. A bipartite and a discriminative attribute graph allow a semantic selection of the most discriminative attributes for plant identification tasks. The selected attributes are used for formulating a question to the user. The system updates similarity information among images based on the user's answer, thus improving the retrieval effectiveness and reducing the user's efforts required for identification tasks. The proposed method was evaluated on the popular Oxford Flowers 17 and 102 Classes datasets, yielding highly effective results in both datasets when compared to other approaches. For example, the first five retrieved images for 17 classes achieve a retrieval precision of 97.07% and for 102 classes, 91.33%.

1. Introduction

The increasing image availability accessible through different technologies has demanded the development of effective retrieval and recognition methods. In this scenario, various image processing techniques have been developed and applied to digital media content (Arvor, Durieux, Andres, Laporte, 2013). Many recent advances have been made through the development of techniques that use quantitative features extracted by visual descriptors, capable of retrieving and indexing images. Most of these approaches are based on Content-Based Image Retrieval (CBIR) systems, which retrieve images by taking into account their visual content. The CBIR approaches consider various visual properties such as shape, texture, and color, extracted through global and local low-level features (Datta, Joshi, Li, & Wang, 2008; Kurtz, Depeursinge, Napel, Beaulieu, & Rubin, 2014; Lew, Sebe, Djeraba, & Jain, 2006). Recently, Convolutional Neural Networks (CNNs) have also been applied towards this goal with significant results (Hoi, Liu, & Chang, 2010; Jia et al., 2014; Razavian, Azizpour, Sullivan, & Carlsson, 2014). Therefore, the main aspects of such retrieval methods are based on feature extraction techniques by visual descriptors.

Besides the visual features, advances have been achieved in other stages of the retrieval pipeline. Approaches which exploit the user feedback through supervised learning methods have been integrated to CBIR techniques, improving the image retrieval effectiveness and adaptability to user inputs (Cheng, Jing, & Zhang, 2009; Liu, Liu, Qin, Ma, & Li, 2007b; Thomee & Lew, 2012). More recently, unsupervised learning has also attracted a lot of attention

* Corresponding author.
E-mail addresses: filipemfg@gmail.com (F.M.F. Gonçalves), ivan@rc.unesp.br (I.R. Guilherme), daniel@rc.unesp.br (D.C.G. Pedronette).

http://dx.doi.org/10.1016/j.eswa.2017.08.035
0957-4174/

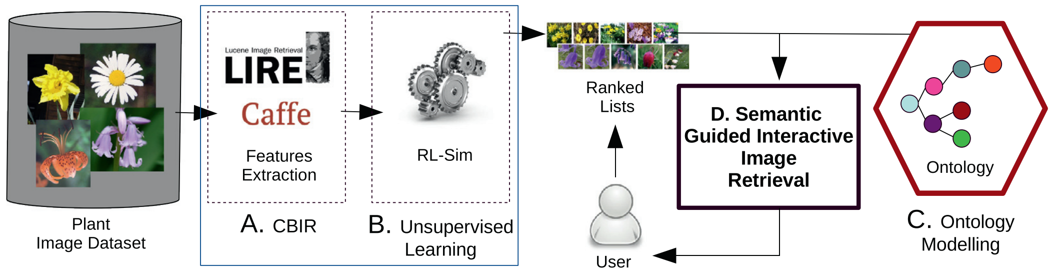

Fig. 1. Proposed retrieval approach for Plant Image Retrieval.

of the research community, once such methods exploit the dataset structure for improving the retrieval effectiveness, dispensing user interventions. In this scenario, unsupervised rank-based methods have been proposed achieving significant effectiveness gains (Bai & Bai, 2016; Bai, Bai, & Wang, 2015; Pedronette, Gonçalves, & Guilherme, 2017; Pedronette & da S. Torres, 2013; 2014).

Despite the continuous development of visual features, supervised, and unsupervised learning methods, retrieving relevant images based on the user needs still is a challenging task. The main challenge is to relate the semantic information of an image domain with the numerical values of low-level features recovered by pattern recognition algorithms. This problem refers to the semantic gap, which is defined as a lack of coincidence between the information that can be extracted from the visual content and the interpretation that the same data present to the user in a given situation (Datta et al., 2008; Smeulders, Worring, Santini, Gupta, & Jain, 2000). The semantic gap remains one of the most challenges of CBIR approaches, directly affecting the retrieval effectiveness.

On the other hand, ontologies have been widely used as a representation technique, allowing the reuse of knowledge since they transcribe a common understanding of a specific area. Ontologies declare explicit semantic, realizing significant statements and supporting the information sharing of attributes and relationships (Gruber, 1993; Guarino, 1998; Lacy, 2005). However, despite the recent advances, there is still a challenge to integrate techniques that use quantitative features with the semantics of structured knowledge representation in ontologies.

In addition, image analysis and identification tasks require specialized knowledge in many research fields, such as Systematic Botany. Traditionally, plant samples and field photographs are analyzed with many systematic descriptions, that allow the identification of organisms and their classification into groups. The identification of Angiosperms (plants with flowers and fruits) requires a vast knowledge of structures and properties of a specimen subject (Souza & Lorenzi, 2007). The identification task is even more challenging when performed solely from image sources since some plant regions are hidden. The image may not show, for example, internal structures in vegetable organs, such as the ovary. Nilsback and Zisserman state in Nilsback and Zisserman (2006) that image classification of flower branches is difficult even for humans, who need a complete knowledge of the domains. In this scenario, it is imperative the development of approaches for better representing the knowledge of many research fields in ontology structures, such that it can be interpreted and processed by both humans and machines.

In this paper, a novel interactive image retrieval approach is proposed aiming at bridging the semantic gap in plant identification tasks. The proposed approach, entitled *Semantic Interactive Image Retrieval* (illustrated in Fig. 1), consists of an automatic interactive system which combines Content Based Image Retrieval (CBIR) techniques, Unsupervised Learning, knowledge representation structured in Ontologies and interactive retrieval mechanisms. Given an image input defined by the user, the system extracts low-level features (Fig. 1A) and executes an unsupervised learning algorithm (Fig. 1B) in order to improve the retrieval results. Additionally, the system uses as an input the structured knowledge given by the ontology (Fig. 1C), which is defined by a domain specialist. The integration between the image retrieval results and the ontology knowledge constitutes the most relevant contribution of the proposed approach (Fig. 1D). The system exploits both information in order to establish a better interaction with the user, defined in terms of textual questions. A bipartite ontology graph and a discriminative attribute graph are proposed to select the most informative attributes from the ontology, capable of better discriminating the plant in the query image from those retrieved based on low-level features.

The proposed approach involves various research challenges of different areas. The main contributions of the paper are summarized in the following, enumerated according to Fig. 1:

- *A. CBIR and features extraction:* (*i*) extraction of visual features using recent CBIR and deep-learning frameworks (LIRE; Lux, 2013 and Caffe; Jia et al., 2014); (*ii*) evaluation of several features and identification of the most effective features for plant image retrieval tasks;
- *B. Unsupervised learning:* (*iii*) use and evaluation of a recent rank-based unsupervised learning method (RL-Sim; Pedronette & da S. Torres, 2013) in plant image retrieval to improve the effectiveness of initial retrieved results;
- *C. Ontology modeling:* (*iv*) the development of a systematic botany ontology, which describes and conceptualizes properties and structures of Angiosperm families;
- *D. Semantic Guided Interactive Image Retrieval:* (*v*) a graph-based integration approach which combines the retrieval results information with the structured knowledge given by the ontology; (*vi*) the proposal of a semantic-guided interactive image retrieval system, in which the questions presented to the user are defined according to the most discriminative attributes of the ontology.

The proposed approach was experimentally evaluated on the two popular datasets: Oxford Flowers with 17 and 102 Classes. Experimental results demonstrated that significant effectiveness gains can be obtained through the interactive retrieval process, indicating the decrease of effects of the semantic gap. The proposed method also yields very high effectiveness results in both datasets when compared to other approaches.

The paper is organized as follows: Section 2 discusses related work and Section 3 presents the CBIR techniques used in plant image retrieval (Fig. 1A). Section 4 discusses the unsupervised distance learning method (Fig. 1B) while Section 5, the ontology modeling (Fig. 1C). Section 6 presents in details the Semantic Guided Interactive Image Retrieval (Fig. 1D). Section 7 presents the experimental evaluation and Section 8 discusses the proposed approach. Finally, Section 9 presents the conclusion and directions for future work.