Neurocomputing 000 (2017) 1–17

Contents lists available at ScienceDirect

Neurocomputing

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

Diversity-based interactive learning meets multimodality

Rodrigo Tripodi Calumby[a,c,*], Marcos André Gonçalves[b], Ricardo da Silva Torres[c]

[a] Department of Exact Sciences, University of Feira de Santana, Avenida Transnordestina, s/n, Novo Horizonte, Feira de Santana, Bahia 44036-900, Brazil
[b] Departamento de Ciência da Computação, Av. Antônio Carlos, 6627 - Prêdio do ICEx Pampulha, Belo Horizonte, Minas Gerais, 31270-901, Brasil
[c] RECOD Lab, Institute of Computing, University of Campinas, Av. Albert Einstein, 1251, Campinas, 13083-852, Brazil

ARTICLE INFO

Article history:
Received 23 February 2016
Revised 31 July 2016
Accepted 6 August 2016
Available online xxx

Keywords:
Diversity
Multimodal retrieval
Relevance feedback
Machine learning

ABSTRACT

In interactive retrieval tasks, one of the main objectives is to maximize the user information gain throughout search sessions. Retrieving many relevant items is quite important, but it does not necessarily completely satisfy the user needs. When only relevant near-duplicate items are retrieved, the amount of different concepts users are able to extract from the target collection is very limited. Therefore, broadening the number of concepts present in a result set may improve the overall search experience. Diversifying concepts present in the retrieved set is one possibility for increasing the information gain in a single search iteration, maximizing the likelihood of including at least some relevant items for each possible intent of ambiguous or underspecified queries. Relevance feedback approaches may also take advantage of diverse results to improve internal machine learning models. In this context, this work proposes and analyses several multimodal image retrieval approaches built over a learning framework for relevance feedback on diversified results. Our experimental analysis shows that different retrieval modalities are positively impacted by diversity, but achieve best retrieval effectiveness with diversification applied at different moments of a search session. Moreover, the best results are achieved with a query-by-example approach using multimodal information obtained from feedback. In summary, we demonstrate that learning with diversity is an effective alternative for boosting multimodal interactive learning approaches.

1. Introduction

Given the technological advances of the last decades, the resources for capturing, processing, storing, and retrieving digital content have tremendously improved. Currently, with the convergence of multimedia devices and the broad access to the Internet, the production and sharing of images is an easy and popular activity. Consequently, we witness the emergence of heterogeneous image repositories and the increasing demand for effective retrieval techniques.

For exploring such collections, image retrieval engines have usually relied on textual information for relevance computation and ranking. In a different paradigm, content-based image retrieval (CBIR) systems exploit low-level features from the images for measuring visual similarity. While CBIR is very effective for some tasks [1], the adequate encoding of high-level visual concepts through low-level visual features is a hard issue. Quite frequently, for complex queries on heterogeneous collections, the low-level features cannot accurately encode the visual concepts of the images (the well-known *semantic gap* problem [2]). In addition, several works have reported the complementarity of textual and visual information [3–7]. These methods have attracted great attention from the research community as the joint usage of various information sources could be useful for attenuating the semantic gap.

Another important issues are that different users quite often have disparate interpretations of the same image or even the same user may have different perceptions at different times, making the retrieval task much more difficult. Additionally, users are not always able to properly express their information needs, meaning that retrieval systems have to process poorly defined queries. In order to make systems adaptive to different users, Relevance Feedback (RF) has been exploited [8] to aid per-user system optimization. In this context, users can help the system to refine results by providing feedback about the relevance of the items retrieved. The system exploits this information to expand queries, enhance internal learning models, and adjust features for constructing new results that are supposed to better satisfy the user needs. Several studies have shown the ability of relevance feedback to improve both, retrieval effectiveness and user satisfaction [9–12]. On the

* Corresponding author at: Department of Exact Sciences, University of Feira de Santana, Avenida Transnordestina, s/n, Novo Horizonte, Feira de Santana, Bahia 44036-900, Brazil.
E-mail addresses: rtcalumby@ecomp.uefs.br, rtripodi@gmail.com (R.T. Calumby), mgoncalv@dcc.ufmg.br (M.A. Gonçalves), rtorres@ic.unicamp.br (R.d.S. Torres).

http://dx.doi.org/10.1016/j.neucom.2016.08.129
0925-2312/

other hand, retrieving several highly relevant items may not completely satisfy the user need, if most of them correspond to near duplicates [13].

Moreover, multimedia retrieval engines have to be robust enough for handling ambiguous or underspecified queries. In such context, ambiguous queries may be decomposed into some, possibly several, search intents. For tackling these issues, several works in the literature have proposed the use of explicit [14–16] or implicit [13,17,18] diversity promotion techniques. Such studies have shown that promoting diversity is helpful for enhancing user satisfaction and optimizing the search experience by retrieving more items for the most likely intent [19] or, as a post-retrieval process, at least some relevant items for the maximum number of the possible search intents [20]. In turn, some complex queries inherently demand diverse results to be properly answered.

For instance, let us focus on a search scenario in which there is no single "correct answer," i.e., several different items may be considered as satisfactory, with each one carrying its particularities. In this scenario, the retrieval system may not be able to select the best item from the group of possible answers and should then provide the user with a set of possibilities. Hence, instead of biasing the result towards a, although correct, single concept, it allows the user to browse and pick the most suitable items. In a different scenario, a given query may only be properly answered not by a single target item but by a complementary group. More specifically, these items may share common aspects whereas each one brings extra information. For example, in an image retrieval context, let us assume the user is looking for "a modern architecture building." This is a broad query whose answer items may differ in several aspects such as their location in the world, construction materials, usage objectives, colors, or even the perspective. Here, the system should provide the user with a set of relevant items, allowing her to choose the most suitable ones. A different user may wish to collect images of buildings designed by a specific architect which means gathering a set of items representing the architect career portfolio and not several pictures from a famous building.

Diversification methods can avoid the generation of highly relevant result sets related only to a single search intent, caused by the imbalance in the number of relevant items for each intent or as a consequence of poorly defined queries [21]. However, a major diversification drawback (and a research challenge) to be considered is that the diversity promotion may erroneously promote non-relevant items to the top positions of the retrieved list. Therefore, finding the proper balance between relevance and diversity is still a hard task, as optimizing one may be detrimental to the other [22].

Interactive learning has been exploited in the information retrieval field for decades with the purpose of tackling several of the aforementioned issues. The possibility of including the user in the retrieval loop has allowed significant effectiveness enhancements over time. By taking advantage of all available data and the collected user preferences, learning-to-rank models [23] leveraged online adaptiveness and consequently improved user search experience.

In this context, in [24], we started a preliminary investigation on the possibility of improving diversity with small or even no detrimental impact on relevance. Such study introduced an approach for enhancing the user experience by interactively learning with user feedback over diversified results produced by a multimodal image retrieval engine. The experiments demonstrated that learning with diverse items helps improving overall diversity during a search session while simultaneously retrieving more relevant items in fewer feedback iterations.

However, the alternatives investigated in [24] were very limited. That work presented only a preliminary study of the many possible scenarios that are exploited in the current investigation. For instance, in that study, only a text-based initial query was considered and, consequently, the visual information was integrated only after the first round of feedback. That work also used a single learning method and a single diversification algorithm. In this work, we propose different methods for simulating alternative search scenarios and the integration of different modalities (see Section 4). Furthermore, in this work, we also assess the diversification impact over several learning methods as well as the effectiveness of an alternative diversification method. Moreover, in [24], the impact of the diversification process and its parameters was not evaluated. We address these issues in this paper. In summary, we present a much comprehensive investigation, considering several novel research intents and significantly extended experiments, as well as a more thorough analysis of the results. In particular, we investigate new research questions not exploited in [24], such as:

- How the "strength/intensity" of the diversification promotion process affects the overall results, considering the complex balance between diversity and relevance?
- Considering the scenario in which different sources of information are available for ranking, how the diversification approaches behave when used along with different retrieval modalities and vice versa?
- Since alternative methods may differently benefit from diversity promotion, what is the impact of diversification on our learning approach (and on the baselines) in the context of multimodal relevance feedback?

Understanding the behavior and possible benefits of how and when to use different retrieval modalities is an important aspect when designing and developing multimedia retrieval systems. At the same time, evaluating the capability of learning-to-rank methods to exploit diversity-oriented feedback and its impact on the learning rate will allow the development of better retrieval systems by boosting the internal models when dealing with diverse-oriented queries. Moreover, since the online adjustment of adaptive ranking methods is directly affected by the available training data and their features/modalities, a great opportunity emerges when the system can automatically help the user on expressing and fulfilling her information needs.

To the best of our knowledge, no previous work has ever proposed introducing diversity into a multimedia interactive learning scenario with explicit user feedback. Moreover, although diversity promotion has gained attention in the last years, there is still a lack of understanding about how to better exploit its benefits, while attenuating the possible drawbacks. Moreover, we also provide an extensive experimental evaluation and discussions on the impact of modalities, features, diversification methods, and learning strategies, when the diversity criteria are considered.

The remainder of this paper is organized as follows. Section 2 focuses on related work. Section 3 describes the diversity-driven learning method that is the base for this work, while Section 4 formally introduces our research questions. Section 5 details the experimental settings. Section 6 presents our experimental results, analysis, and discussions. Finally, Section 7 concludes the paper.

2. Related work

State-of-the-art diversity-promotion methods, such as IA-Select [14] and xQuAD [15], explore different query intents that are explicitly stated or detected at runtime. These techniques are usually employed over traditional information retrieval (IR) engines including probabilistic and language modeling methods.

In more details, IA-Select is a probabilistic method that assumes an existing information taxonomy, which is used to model user interpretations of a query and categorize the documents in the

Please cite this article as: R.T. Calumby et al., Diversity-based interactive learning meets multimodality, Neurocomputing (2017), http://dx.doi.org/10.1016/j.neucom.2016.08.129