Airo International Research Journal
Volume VIII, ISSN: 2320-3714

December, 2016

Study of Software Retrieval Based On Semantic Properties and Its Benefits

Patel Muhammad Umar Mukhatar Ahmad
University - Sri Satya Sai University
Guide Name - Dr. Suresh Chand Tyagi
Guide Designation - Executive Director, IDC Foundation

Abstract

The semantic representation of resources is a domain of intense research at present. In this domain, several directions can be taken. In the present work we are exactly interested in the conceptual analysis of the textual documents. In our approach, we suggest studying the variation of the concept relevance within a text to identify the major theme and all the minor themes evoked in the text. This allows us at the second level of analysis to create a semantic map of concepts and terms extracted from resources. Through this map, we propose to the users a proximity champ allowing them to find, in the one time, pertinent resources according to their query and in the second time to discover a new knowledge by navigating in semantic proximity relations. of resources is a domain of intense research at present. In this domain, several directions can be taken. In the present work we are exactly interested in the conceptual analysis of the textual documents. In our approach, we suggest studying the variation of the concept relevance within text to identify the major theme and all the minor themes in the text. This allows us at the second level of analysis to create a semantic map of concepts and terms extracted from resources. Through this map, we propose to the users a proximity champ allowing them to find, in the one time, pertinent resources according to their query and in the second time to discover a new knowledge by navigating in semantic proximity relations.

Keywords: semantic, proximity, research, navigating, navigating

Introduction

Available information on Internet develops at an exponential rate. Data in these information frameworks is becoming more unpredictable and increasingly dynamic. As users with various backgrounds, traits, abilities, dispositions, and intentions increase dramatically, users' needs also become more various and complicated. Therefore the demand for a more successful and proficient means for managing and exploring data became a pressing issue. This represents a challenge to the traditional approaches and strategies used in current information retrieval frameworks. These frameworks use a catchphrase based search process which is discontinuous because users have no power over the internal matching process which is not transparent to users. Besides, the yield of search frameworks as result list presentation is linear and has a restricted display capacity. Relationships and associations among documents are rarely illustrated. The retrieval condition lacks an interactive mechanism for users to peruse. These inherent weaknesses of traditional information retrieval frameworks avoid them from coping with the sheer unpredictability of information needs and the huge number of data dimensional.

Information retrieval visualization is proposed as a new approach to improve information retrieval. In this degree, we propose a semantic navigation in request to help users in finding relevant resources according to their questions. In the present work, we deal about intelligent retrieval. The remaining of this paper is organized as following. In the second area, we present a study of the state of the art concerning information retrieval visualization. In the third area we describe the general principle of our approach of semantic navigation which is made out of two parts. The first concerns term and concepts extraction. The second part concerns semantic map generation. Finally, in the fourth area, we present and discuss the obtained experimental outcomes.

In the literature, information retrieval is realized by two major classes of approaches. The top of the line of approaches is based on the mapping of watchwords representing a query of users and terms representing the substance of resources.

The second class of approaches is based on the discovery of the concepts. This second class of approaches presents the advantages to be independent of terms representing questions and substance resources. In these two approaches, users have usually a list of the relevant resources. Along these lines, he must check the pertinence of result. From our investigation of the state of

the art, we saw the lack of interest on studying the semantic of documents. We are interested rather at identifying concepts evoked in textual resources, and studying their pertinence which allow us annotating documents. Then, we associate each archive to the additionally representing concepts in request to create a map of navigation. From our point of view, through the concept annotation, we aim to establish a global perspective on the textual corpus providing users with an information search framework to localize their interest focus according to their concepts of search. Indeed, such a module allows enriching archive indexing process. This facilitates users in finding similar documents to their report of interest. Visualization is required at whatever point humans need to discover and reason about complex combinations of high volumes of data (e.g.,). Information visualization and visual data mining is not constrained to the display, yet aims at supporting human perceptual abilities during the data exploration process. A vast literature exists on the topic. Cluster visualization has been used in such various fields as intelligence (for example to show correlation between individuals) and image assortment access (for example to show similarity in images). Alternative visualizations have been used to make easy to identify patterns in homogeneous data (for example in geospatial data); numerous visualizations, instead, map the quality of relationships between components. In text retrieval, much research has investigated the visualization of search results (see and for a review), the visualization of the entire archive assortment (e.g., Treemaps) or a large text corpus (for example Jigsaw). Information exploration, an open-finished procedure that is iterative and multi-tactical is right now gaining interest and stimulating new user interactions beyond traditional text search. The issue of visualization of the substance of resources has been perceived since the publication of the seminal book. It focuses on the formal and even the meaning achieved through rigorously defined structures. In contrast, information visualization emphasizes the semantics and the meaning that can be passed on by visual-spatial models to the users. Much research exertion in semantic-based visualization has been spent on finding ways of visualizing complex graphs that get from the interlinking of semantic data, the relation between various concepts, the various granularities, and (dis)connections. The result is a large number of ontology-based visualization frameworks.

We propose a retrieval framework that endeavors a half and half feature space (HFS) that is worked by integrating low-level image features and elevated level semantic terms, through rounds of relevance feedback (RF) and performs similarity-based retrieval to help self-loader image interpretation. The oddity of the proposed framework is that it can credit the semantic features of the query image by reformulating the query vector representation in the HFS via user feedback. We actualized our framework as a prototype that plays out the retrieval over a database of 811 radiographic images that contains 69 interesting sorts of bone tumors.

Medical information retrieval is important for research and potentially for clinical care, however finding similar cases is largely an unassisted and time-consuming procedure, and precision is

established through many years of training and experience. Indeed, even in spite of this training, substantial inter-reader variation in determining case similarity is challenging. In addition, the volume of medical information is growing faster than the ability of professionals to carry out this responsibility themselves without the help of automated search mechanisms. In radiology, image retrieval has particular importance because the radiologist regularly stands up to rare abnormalities for which diagnosis is troublesome. Finding similar images from large imaging archives, for example, picture archiving and communication framework (PACS), can potentially assist in suggesting diagnoses of many similar cases, and the proof provided by the similar cases can assist the radiologist to improve interpretation of rare abnormalities and may help in determining diagnosis.

In this work, relevance feedback was utilized as a mechanism to achieve great accuracy in CBIR by incorporating semantic and quantitative image features, however not requiring the user to give the semantic features in the query image. The implications of this approach are important in that by removing the prerequisite of collecting semantic features, it may be practical to introduce our approach to CBIR into the clinical workflow, since busy radiologists rarely have time to input such data into the framework. Furthermore, it is relatively easy to pass judgment on visual similarity seeing an alternate arrangement of images, and this procedure, combined with existing knowledge in previously annotated images, enables our framework to perform well in the CBIR task with relevance feedback. While the ultimate goal of machine learning algorithms and artificial intelligence may be to automatically learn from the data with restricted or no human interaction, it should be perceived that achieving accurate outcomes for complex image interpretation tasks such as medical images may require more elevated levels of psychological processing. Our framework shows promising outcomes in this regard by retrieving visually similar images in the challenging area of bone tumor radiography by incorporating a human master as part of the iterative procedure. This kind of "human-in-the-circle" integration or alleged "interactive machine learning" has promise for other complex interpretation tasks in radiology.

A third limitation is that the benefits of relevance feedback are most certainly not guaranteed to be realized for all iterations because of the randomness of human interaction, and the user may need to continue providing relevance feedback until the framework reaches a saturation point. It is also conceivable that the reformulated query vector may actually move farther from the relevant vectors in the feature space after numerous iterations and the optimal saturation point will never be reached, since we as it were thought about the positive relevance for this investigation. However, in the user interface, we have incorporated a sliding control bar that in future testing will allow the users to tune the weights of positive and negative relevance.

Finally, the relevance feedback approach may not work as well if the relevant vectors create several disconnected clusters within the half breed feature vector space. Thus, the revealed

performance may not generalize to other kinds of datasets or with various arrangements of users. Further examinations on larger image assortments will be required. In spite of these limitations, we believe our outcomes show the potential value of our approach to enhancing CBIR with relevance feedback of visual similarity as a new heading for helping radiological image interpretation. Looking toward immediate future work, one of our goals is to test the ability of our framework to improve diagnostic accuracy across readers with varying levels of clinical experience, with the aim of improving the performance of less experienced users through powerful similar image retrieval.

The workflow representation in the CAKE framework uses an article situated representation (originally produced for cases) for ontology representation and metadata annotation. Tasks can be organized in a hierarchy of classes, each of which contains certain properties of a task, which can be inherited from the super class. For example, the Assign room task includes a job depiction stating that the assignment must be performed by the capable administrative department. Another property may state the average duration for task execution. The semantic workflow representations in Wings augment traditional workflow representations with semantic constraints that indicate metadata properties of workflow tasks and datasets. Each workflow constituent has a workflow variable associated with it, and the semantic constraints express constraints over those variables. In the case of data variables, the constraints are communicated in terms of metadata properties. For example, an input dataset may be limited to be a discrete dataset, and in the event that in this way, then its variable would have a constraint stating that it has the property of being discrete. In the case of components, the constraints express how the execution of the software component changes the metadata properties of the input datasets into yield datasets.

We generalize and expand previous approaches for workflow similarity using graph-based representations in several ways. We expressly spread data driven as well as control-stream driven workflows. We link with semantic representations and introduce knowledge intensive similarity measures according to the local/global principle. Further expansions are desirable to represent hierarchical workflows. The correct treatment of workflow agility requires representing workflow instances (rather than templates as in the present approach) including the execution state. Also, increasingly practical involvement in applications of the model is required, involving semantic models of a larger scale. This also raises the inquiry of approachs for developing appropriate similarity models. New techniques for learning similarity measures are demanded as a tool for this reason. For image retrieval, content based approach with semantic features broadly used. The natural language processing approach integrated with lexical libraries and low-level features used for the retrieval of images. It gives the semantic base for ontology creation of the image databases and

improves retrieval accuracy. The half and half model which combines ontology and Bayesian Network aim to improve the quality of image retrieval. Be that as it may in this model, the reasoning and ranking of the concepts has to be improved for better recall rate. Multi-Modality ontology approach gives scalability of retrieval framework and ranking mechanism has to be improved for large domain concepts. The area based approach narrows the semantic gap while retrieving the images and defeats the shortcomings of image-image matching. Image semantic and SIFT features used to improve the accuracy of retrieval. The fluffy domain ontology for image narrows the search range and diminishes the useless information retrieval.

This approach improves retrieval framework productivity and accuracy with the use of fluffy set. In video retrieval, shot choice and object acknowledgment plays the most important job. Automatic shot determination and annotation used in pictorially enhanced ontology to perform automatic annotation of video for retrieval. In video ontology approach, complex occasions can be identified by combining the semantic patterns in various measurements. Semantic-based tolerance image representation improves the viability of retrieval through automatic annotation of image/video according to its semantics. In rule learning approach, rule based annotation gives better precision and recall.Computer-assisted image retrieval applications can assist radiologists by identifying similar images in archives as a means to providing decision support. In the classical case, images are described using low-level features extracted from their substance, and an appropriate distance is used to find the best matches in the feature space. However, using low-level image features to completely capture the visual appearance of diseases is challenging and the semantic gap between these features and the significant level visual concepts in radiology may impair the framework performance. To deal with this issue, the use of semantic terms to give elevated level portrayals of radiological image substance has as of late been advocated. Nevertheless, the majority of the existing semantic image retrieval strategies are constrained by two factors: they require manual annotation of the images using semantic terms and they disregard the intrinsic visual and semantic relationships between these annotations during the comparison of the images. Based on these considerations, we propose an image retrieval framework based on semantic features that depends on two main strategies: (1) automatic "soft" forecast of ontological terms that describe the image substance from multi-scale Riesz wavelets and (2) retrieval of similar images by evaluating the similarity between their annotations using a new term dissimilarity measure, which takes into account both image-based and ontological term relations. The combination of these strategies gives a means of accurately retrieving similar images in databases based on image annotations and can be considered as a potential answer for the semantic gap issue. We validated this approach in the context of the retrieval of liver injuries from processed topographic (CT) images and annotated with semantic terms of the RadLex ontology. The relevance of the retrieval results was assessed using two protocols: evaluation relative to a dissimilarity reference standard defined for pairs of images on a 25-images dataset,

and evaluation relative to the diagnoses of the recovered images on a 72-images dataset. A normalized discounted cumulative gain (NDCG) score of more than 0.92 was obtained with the primary protocol, while AUC scores of more than 0.77 were obtained with the second protocol. This automatically approach could give real-time decision backing to radiologists by showing them similar images with associated diagnoses and, where available, reactions to therapies.

We propose to consider the similarity between semantic terms during the retrieval of similar database images. To this end, we propose a new term dissimilarity measure enabling us to quantify both the image-based relations among the terms, which are given by their visual signatures, and the semantic relations among the terms, which are automatically evaluated from the structure of the ontology. In this measure, the commitments of the image-based and the ontological dissimilarities are automatically evaluated from data examples using a learning strategy based on maximizing the agreement between the perceptual image dissimilarity gave by the radiologist and the global term dissimilarity value. This strategy gives a potential answer for the limitations of BOW approaches that assume that each term describing an image is independent of other features. We propose to consider the similarity between semantic terms during the retrieval of similar database images. To this end, we propose a new term dissimilarity measure enabling us to quantify both the image-based relations among the terms, which are given by their visual signatures, and the semantic relations among the terms, which are automatically evaluated from the structure of the ontology. In this measure, the commitments of the image-based and the ontological dissimilarities are automatically evaluated from data examples using a learning strategy based on maximizing the agreement between the perceptual image dissimilarity gave by the radiologist and the global term dissimilarity value. This strategy gives a potential answer for the limitations of BOW approaches that assume that each term describing an image is independent of other features. Agricultural digital library was the carrier of agricultural information resources. The current digital library retrieval query administration there are a few disadvantages, including low query intelligence, independent outcomes and low level of shared information data. The traditional pattern of information query was based on the catchphrase query. This kind of query can't mirror the profound meaning of user query demand. Semantic web query mode was a hot research topic as of late. In semantic query mode, user's semantics of natural language can be understood by the machine to a certain degree and the semantic level of retrieval was actualized finally. Step by step instructions to assemble the information query model based on semantic innovation was a liable to be explained in which the most key issue was the semantic similarity computation. Traditional semantic similarity algorithms include algorithm based on semantic distance, algorithm based on the concept features and algorithm based on the amount of information. The disadvantage of the above algorithms was easy to lead to errors. Some other researchers contemplated the information retrieval based on semantic retrieval. In this paper, aiming at the shortcomings of the similarity

calculation in traditional semantic algorithm, the traditional semantic algorithm was improved. Using the similarity algorithm based on semantic distance, the model of information query based on semantic innovation was fabricated and information retrieval precision was increased.

GRAPHICAL-BASED KNOWLEDGE REPRESENTATION FORMALISMS

In this segment, we described the main existing graphical based Knowledge representations. These representations are a decent way to understand a subject or how extraordinary topics are associated, enabling a fast and viable way to learn and discuss a given model. These representations went through many years of research and several models have risen to represent these frameworks. These formalisms are relevant for this work according to the need to represent Knowledge graphically. As the intended platform will be used by several users willing to collaborate in the development of conceptual models, it is important that they have access to a graphical representation that reflects, as intently as conceivable, the conceptual model.

Semantic Networks

Semantic Networks are a kind of Knowledge Representation and were created in the beginning of the 60's by Robert F. Simmons at System Development Corporation and improved by Allan M. Collins and M.R. Quillian among others. It's represented by a graph, made by concepts, which are associated by relations. There are three important components in this representation: concepts, relationships and instances. Semantic networks are a powerful knowledge representation framework. They are easy to understand by humans and can be used in automated processing frameworks. This means that they can also be come a vehicle to archive company knowledge.

Conceptual Graphs

Conceptual Graphs (CG) were devised by Sowa from philosophical, psychological, linguistic, and artificial intelligence foundations, and are a powerful knowledge representation and inference condition. CG capture nuances in natural language while being able to be actualized in computer software. Like the author suggests, CG are easy to actualize and understand by humans and are capable of being actualized as a structure accessible by machines. CGs are comprised by two kinds of components:

– Elements

– Concepts - Usually represented by boxes.

– Conceptual Relations - Usually represented by circles

RDF Graph

Resource Description Framework (RDF) is a standard model for data interchange on the Web. RDF has features that facilitate data merging regardless of whether the underlying schemas contrast, and it specifically underpins the advancement of schemas after some time without requiring all the data customers to be changed.RDF expands the linking structure of the Web to use URIs to name the relationship between things as well as the two parts of the bargains (usually alluded to as a "triple"). Using this straightforward model, it allows organized and semi-structured data to be blended, uncovered, and shared across various applications. In a progressively visual point-of-see, RDF can be characterized has a labeled graph (RDF Graph), framed by resources (graph hubs) and labeled links (edges) between them. This data model is represented through triples which are framed by three components: Subject (or Resource), Object (or on the other hand Value) and Predicate (or Property).

Conclusion

This paper presents the analysis of semantic feature based image retrieval. In analysis, proposed approach tangled low-level features and use three layers of tangled features space pooling by non-linear manner. In proposed approach find the linear and non-linear semantic features mapping by sigmoid capacity.

References

[1] R. Amar and J. Stasko. (2004) A Knowledge Task-Based Framework for Design and Evaluation of Information Visualizations. In Proc. of IEEE InfoVis, pages 143–149, Los Alamitos, USA, IEEE Press.

[2] Bederson, B.B. (2000), Fisheye menus, Proceedings of the 13th annual ACM symposium on User interface software and technology, San Diego, California, United States, ACM Press

[3] Bertin J. (1967) La sémiologie graphique. Paris : GauthierVillars.

[4] Lyn Bartram, Albert Hot, John Dill, Frank Henigman, (1995) The Continuous Zoom: A Constrained Fisheye Technique for Viewing and Navigating Large Information Spaces, Proc. ACM Symposium on User Interface Software and Technology (UIST'95), pp. 207-215.

[5] Riccardo A. Cava, Paulo R. G. Luzzardi, Carla M. D. S. Freitas, (2002) The Bifocal Tree: a Technique for the Visualization of Hierachical Information Structures, IHC'2002.

[6] Chen, H., Lynch, K.J., Basu, K., & Ng, T. D. (1993). Generating, integrating, and activating thesauri for conceptbased document retrieval. IEEE Expert, 8(2), 25-34.

[7] John V. Carlis and Joseph A. Konstan. (1998) Interactive visualization of serial periodic data. In Proceedings of the 11th Annual ACM Symposium on User Interface Software and Technology, pages 29–38. ACM Press

8. Su, J.H., Huang, W.J., Philip, S.Y. and Tseng, V.S., 2011. Efficient relevance feedback for content-based image retrieval by mining user navigation patterns. IEEE transactions on knowledge and data engineering, 23(3), pp.360-372.

9. Clinchant, S., Ah-Pine, J. and Csurka, G., 2011, April. Semantic combination of textual and visual information in multimedia retrieval. In Proceedings of the 1st ACM international conference on multimedia retrieval (p. 44). ACM.

10. Akakin, H.C. and Gurcan, M.N., 2012. Content-based microscopic image retrieval system for multi-image queries. IEEE transactions on information technology in biomedicine, 16(4), pp.758-769