UNIVERSITI TEKNOLOGI MARA

MULTI-MODALITY ONTOLOGY SEMANTIC IMAGE RETRIEVAL WITH USER INTERACTION MODEL

MOHD SUFFIAN SULAIMAN

Thesis submitted in fulfilment
of the requirements for the degree of
**Doctor of Philosophy
(Computer Science)**

Faculty of Computer and Mathematical Sciences

April 2020

AUTHOR'S DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Teknologi MARA. It is original and is the results of my own work, unless otherwise indicated or acknowledged as referenced work. This thesis has not been submitted to any other academic institution or non-academic institution for any degree or qualification.

I, hereby, acknowledge that I have been supplied with the Academic Rules and Regulations for Post Graduate, Universiti Teknologi MARA, regulating the conduct of my study and research.

Name of Student   :   Mohd Suffian Sulaiman

Student I.D. No.   :   2011858298

Programme   :   Doctor of Philosophy in Information Technology
                                and Quantitative Science – CS990

Faculty   :   Computer & Mathematical Sciences

Thesis Tittle   :   Multi-Modality Ontology Semantic Image Retrieval
                                with User Interaction Model

Signature of Student :   …………………………………………..

Date   :   April 2020

ABSTRACT

Interest in the production and potential of digital images has increased greatly in the past decade. The extensive use of digital technologies produces millions of digital images daily. However, the capabilities of technologies equipment manifest the difficulty and challenge for the user to retrieve or search the visual information especially in a large and varieties of a collection. The issues of time consuming for tagging the image, often subject to individual interpretation and lack of ability for a computer to understand the semantic high-level human understanding of image become the former approaches unable to provide an effective solution to this problem. In addressing this problem, this research explores the techniques developed to combine textual description with visual features to form as multi-modality ontology. This semantic technology is chosen due to the ability to mine, interpret and organise the knowledge. Ontology can be seen as a knowledge base that can be used to improve the image retrieval process with the aim of reducing the semantic gap between visual features and high-level semantics. To achieve this aim, multi-modality ontology semantic image retrieval model is proposed. Four main components comprising resource identification, information extraction, knowledge-based construction and image retrieval mechanism are the main tasks need to be implemented in this model. In order to enhance the retrieval performance, the ontology is combined with user interaction by exploiting the ontology relationship. This approach is proposed based on an adaptation from a part of relevance feedback concept. To realise this approach, the semantic image retrieval prototype is developed based on the existing foundation algorithm and customised to provide the ability for user engagement in order to enhance the retrieval performance. To measure the retrieval performance, the ontology evaluation needs to be done first. The correctness of ontology content between the referred corpus and the notation of the ontology is important to make sure the reliability of the proposed approach. Twenty samples of natural language queries are used to test the retrieval performance through the generating of the SPARQL query automatically to access the metadata in the ontology. The graphical user interface is designed to display the image retrieval results. Based on the results, the retrieval performance is measured quantitatively by using precision, recall, accuracy and F-measure techniques. An experiment shows that the proposed model has an average accuracy 0.977, precision 0.797, recall 1.000 and F-measure 0.887 compared to text-based image retrieval, 0.666 (accuracy), 0.160 (precision), 0.950 (recall) and 0.275 (F-measure); textual ontology, 0.937 (accuracy), 0.395 (precision), 0.900 (recall) and 0.549 (F-measure); visual ontology, 0.984 (accuracy), 0.229 (precision), 0.300 (recall) and 0.260 (F-measure); multi-modality ontology, 0.920 (accuracy), 0.398 (precision), 1.000 (recall) and 0.569 (F-measure). In conclusion, results of the proposed model demonstrated better performance in order to reduce the semantic gap, enhance the semantic image retrieval performance and provide the easy way for the user to retrieve the herbal medicinal plant images.

ACKNOWLEDGEMENT

TABLE OF CONTENTS