شبكة المعلومات الجامعية

التوثيق الإلكتروني والميكروفيلم

بسم الله الرحمن الرحيم

HANAA ALY

شبكة المعلومات الجامعية

التوثيق الالكتروني والميكروفيلم

HANAA ALY

شبكة المعلومات الجامعية

التوثيق الإلكتروني والميكروفيلم

جامعة عين شمس

التوثيق الإلكتروني والميكروفيلم

قسم

نقسم بالله العظيم أن المادة التي تم توثيقها وتسجيلها
علي هذه الأقراص المدمجة قد أعدت دون أية تغيرات

يجب أن

تحفظ هذه الأقراص المدمجة بعيدا عن الغبار

HANAA ALY

Knowledge-based Image Representation and Retrieval

A thesis submitted to the Department of Information Systems, Faculty of Computer and Information Sciences, Ain Shams University. In partial fulfilment of the requirements for the degree of Master of Science in Computer and Information Sciences

By:

Haitham Samih Atea

Supervised by:

Prof. Tarek F. Gharib

Head of Information Systems Department
Faculty of Computer and Information Sciences Ain Shams University

Prof. Manal Ahmed Ismail

Computer Engineering and Systems Department
Faculty of Engineering
Helwan University

Dr. Sherine Rady Abdel Ghany

Associate Professor in Information Systems Department
Faculty of Computer and Information Sciences
Ain Shams University

Ain Shams University, Cairo, Egypt
December, 2020

Knowledge-based Image Representation and Retrieval

A thesis submitted to the Department of Information Systems, Faculty of Computer and Information Sciences, Ain Shams University. In partial fulfilment of the requirements for the degree of Master of Science in Computer and Information Sciences

By:

Haitham Samih Atea

Supervised by:

Prof. Tarek F. Gharib

Head of Information Systems Department Faculty of Computer and Information Sciences
Ain Shams University

Prof. Manal Ahmed Ismail

Computer Engineering and Systems Department
Faculty of Engineering
Helwan University

Dr. Sherine Rady Abdel Ghany

Associate Professor in Information Systems Department
Faculty of Computer and Information Sciences
Ain Shams University

December, 2020

To my dear family...

Acknowledgements

In the beginning, I bow head thanking to ***"Allah"*** who paved the way and only by HIS will everything can be achieved.

I would like to express my deeply grateful and respect to ***Prof. Tarek F. Ghraib*** *for giving me this opportunity to work under his esteemed guidance, motivation and support. I also wish to express profound gratitude for his constant encouragement and critical discussions throughout this research program and during the preparation of this thesis.*

Deep thanks and sincere gratitude with appreciation to ***Dr. Sherine Rady****.* This work would not have been possible without her guidance, support and encouragement. For her effective supervision, helpful comments, and the extensive time she devoted to this work, providing many facilities during preparation, experimental measurements, and scientific supervision that helped to accomplish this research.

Deep thanks and sincere gratitude with appreciation to ***Prof. Manal Ismail*** *for her continuous advices, useful comments and scientific supervision throughout this research.*

Also, thanks to my friends and my colleagues in the Egyptian E-Learning University *for their fruitful help and support.*

Finally, no words can describe how thankful I'm to my family, my lovely wife for their help, encouragement and inspiration to accomplish all this work.

Abstract

Analyzing complex natural language queries through image/multimedia search engines remains a big challenge. Traditional text-based retrieval systems associate textual descriptions with each image, based on subjective human perception. These descriptions are next matched lexically against the user interrogated queries. Such annotation-based paradigm does not achieve the best results, since the lexical comparison is not sufficient for matching sentences in a semantic manner. Combining image retrieval processing with rich semantics and knowledge-based modeling provide promising solutions towards better image search engines.

This thesis proposes a knowledge-based image representation and retrieval which integrates external knowledge sources for obtaining a higher-level inference that can both handle complex natural language queries and increase the number of relevant retrievals for image search engines.

The thesis presents two solution approaches for the purpose of enhancing image retrieval. The first solution proposes a semantic framework for image representation and retrieval that can efficiently handle complex human-wise queries. The second solution proposes a semantic evaluation for auto-generated image annotations based on similarity measurement.

In the proposed image representation and retrieval framework, semantics are integrated by employing external knowledge sources and query expansion in the retrieval process. A set of developed and off the shelf parsing tools are used to obtain a full semantic understanding for relating the natural language queries and image annotations. The user query is parsed and next fused with the external knowledge sources in a query expansion process to infer supplemental knowledge about the terms of the query and hence increasing the

searchability of the query over image captions while increasing the average Recall. For providing relevance, a relation similarity metric is proposed to rank retrieved images based on their similarity scores.

For the purpose of evaluation of auto-generated image annotations, a semantic evaluation metric that measures the similarities of the generated annotations towards a set of human-written reference annotations is used.

Experiments are conducted on the Flickr datasets using a large set of natural language queries. The proposed solutions have been compared versus existing related systems for different evaluation measures: Number of Retrievals (Recall), Retrieval Accuracy, expansion rule hits, Pearson's, and Kendall's correlations.

The results show that the proposed image representation and retrieval framework that integrates external sources and inference for a semantic rule-based query expansion outperforms related systems with 40% increase in the system Recall at approximately 100% retrieval accuracy. This is due to the high usage of the expansion rules which are 88% and 61% for event and entity rules respectively.

Regarding the proposed semantic evaluation for auto-generated annotations, the results show that the proposed annotation evaluation approach outperforms existing automatic annotation evaluation metrics by achieving a Pearson's ρ correlation of 0.73 and Kendall's τ rank correlation of 0.49 for system and caption-level correlations , respectively.

Table of Contents

List of Figures

List of Tables