Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 2, February 2014, pg.704 – 708

RESEARCH ARTICLE

MULTIMEDIA QA GENERATION BY USING SEARCH DIVERSIFICATION

Nandhini.N[1], Ramya.K[2], Sandeepa.P[3]

[1,2,3] V.S.B Engineering College, Affiliated to Anna University, Karur

[1] nandhinikamalini@gmail.com; [2] meramyak@gmail.com; [3] sandeepaperumalsamy@gmail.com

Abstract- Community Question answering (cQA) services have gained popularity for the past few years. It supports community users to post and answer questions and also it enables general users to acquire information from a set of answered questions. Though existing cQA forum provides textual answers alone, it is not much informative for many questions. In order to enhance textual answers in cQA with suitable media data MMQA method has been introduced. This method consists of three parts, Answer medium selection, Query generation for multimedia search and multimedia data selection and presentation. This method automatically finds out which type of multimedia information should be added to get an elaborated textual answer and also it automatically gathers data from the web to enhance the answer. By processing a collection of question answer pairs and adding them to a dataset it can set up a novel multimedia question answering (MMQA) method as users can find multimedia answers by comparing questions with those in the dataset. This MMQA method not only provides image and video for direct question answers but also give answers for more complex questions. The multimedia search diversification method is used here to collect the relevant answers based on questions. The result shows that it provides more satisfactory answers to the users and also it is more effective.

Keywords— Question answering; multimedia search; reranking; search diversification; query generation

I. INTRODUCTION

The amount of information on the web has grown exponentially throughout the years, and simultaneously multimedia content has also been increased in vast amount which covers almost all topics. As a result, when looking for information, users are often confused by the vast quantity of results from search engines. To reduce this problem Question Answering (QA) technique has been introduced which automatically answers a question in natural text language. This QA technique helps the users in extracting the most relevant information where the users are often bewildered in browsing vast quantity of information by the search engines like google and yahoo[1]. The community Questions Answering (CQA) service like yahoo! Answers, allows the people to ask question and it either answers automatically or allows other users to provide some textual answers[2]. But a textual answer does not provide sufficient information for certain questions. Let us consider an example "which country won the cricket world cub 2011?" and "How to install linux" it provides longest textual descriptions alone. In these cases, we are providing textual answers along with media information, which helps users easy to understand.

In order to enhance the community contributed textual answers with appropriate multimedia information the following phases are used[2].

1. Answer medium selection- Based on the QA pairs the answers will be selected that is which type of medium data should be added. In these medium selection has been classified into four types: Text, Text + image, Text+ video, Text+ image+ video. This method will automatically gather images, videos and other combination of information to improve the quality of textual answers.

2. Query generation for multimedia search- To select multimedia information, informative queries has to be generated. The informative queries can be generated by using three-class classification model. The QA pair will be given and queries are generated from the questions, answers and from the QA pairs respectively.

3. Multimedia data selection and presentation- From the generated queries, the images and videos are collected vertically by using multimedia search engines. To enrich the textual answers reranking and duplicate removal is performed to get an accurate videos and images.

These three types of methods are used to collect the multimedia based information for QA, which is called as MultiMedia Question Answering (MMQA)[1][8]. This MMQA method does not answers the questions directly; alternatively it uses multimedia content to enhance the community-contributed answers. Section II describes the related work with existing system. In section III we discuss about the architectural model for multimedia search. Section IV describes about the query generation and multimedia search diversification method and section V concludes the paper and points out future work.

II. **RELATED WORK**

In 1960's QA system has been started to investigate and it mainly focused on expert system in particular domains. The text-based QA has attained popularity in the year 1990's. Depending upon the type of questions and expected answers, QA can be summarized into following classes: open domain-Based QA, Restricted-domain QA, Definitional QA and List QA[1]. The automatic QA still faces some difficulties in answering complex questions and cQA is an alternate approach. The existing cQA like Yahoo! Answers, Wiki Answers, and metafilter supports pure text-based answers alone which is not sufficient for users. To overcome this problem multimedia search has been introduced which adds images and videos along with text. Multimedia QA system relies on video optical character recognition (VOCR) and automatic speech recognition (ASR) [4].

The amount of digital information stored on the web has been grown in vast; hence extracting the desired information is an important task. To overcome this problem multimedia search has been introduced which is classified into two categories: text-based search and content-based search [5]. Text-based search is based on the text queries and term based specification, where it matches the text with media in the web to retrieve the media data, and in order to improve the performance some machine learning methods are used which automatically annotate the entities for gathering information. Various social websites like Flicker, Face book uses the manually annotated media entities along with text-based search method which faces some technical issues. To overcome this problem content-based search is used which filters the information by analysing the content in the media instead of metadata. The content-based retrieval has some limitations such as high computational cost, difficulty in finding the visual queries and large gap between the visual description and user's semantic expansion.

Multimedia search reranking algorithm is used for improving the search relevance by extracting the visual information of images and videos [4]. The Reranking algorithm has been categorized into two techniques: pseudo relevance feedback and graph-based reranking [6]. Pseudo relevance feedback is used to collect the relevant samples and that samples are assumed to be irrelevant. A classification or ranking model is used for ranking the samples and provide feedback by labelling the results as relevant or irrelevant. Graph classification method is based on the two premises. First one is variance between the initial ranking list and processed ranking list should be small. Second one is visually similar samples should be ranked very nearer. This algorithm constructs a graph where the images or videos represent the vertices and pair wise similarities represent the edges.

These two approaches rely on the visual similarities between two media entities. It should measure the resemblance like color, texture, shape and so on; Almost query adaptive technique is used for estimating the similarities. Let us consider an example as finding out the person. Here we need to identify the similarities of facial characteristics. Based on this identification, queries are classified into two classes namely person related query or non-person related query [12].

III. **ARCHITECTURE MODEL**

The Multimedia QA technique has the following phases: Answer medium selection, Query generation for multimedia search and multimedia data selection.

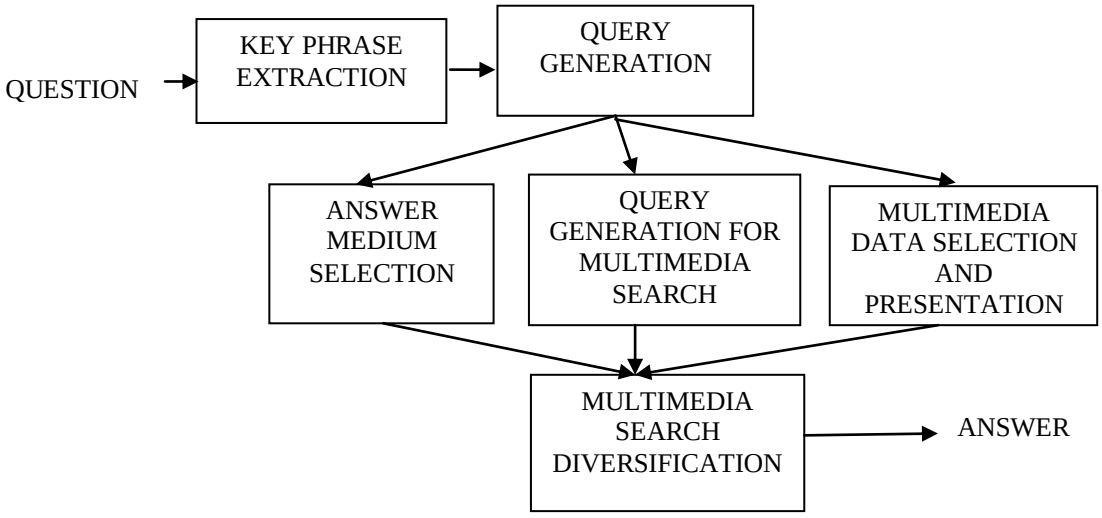

Fig. 1 The overall framework system

A. Answer medium selection

Answer medium selection determines what type of medium to be selected in order to enhance the textual answers. For example, "when india got republic day?" in this type of questions textual answer alone is sufficient. Whereas "Who is the father of india?" In this case textual answer together with image provides sufficient information to the user. Here images will enrich the textual answers. Let us consider another example as "How to install Skype in pc?" In these type of questions, textual answers along with the video information is essential. So it will opt for text+video. In some situation questions may also need all things together i.e. text+image+video which include both videos and image along with text. The Answer medium selection is classified into three methods as follows [12]: 1. Question based classification- In this approach the questions are classified based on two steps. First one is to identify the interrogative words (yes/no) and then it will directly answer the question. Second one is to classify the interrogative words using the naïve bayes classifier. 2. Answer based classification- In this approach verb and bigram words are extracted from the Question. The verbs will help to identify how to enrich the textual answers i.e. either by means of image or video. 3. Media resource analysis- In this method the information from web search engines are accumulated and converted to various forms for answering the questions [8].The limitation in this answer medium selection is, it does not support the text+image+audio+video information.

B. Query generation for multimedia search

This method is used for generating the queries before performing the search in multimedia search engine. Queries will helps to retrieve the most relevant images and videos from the web. First step is query extraction i.e to extract informative keywords from questions and answers [6]. Second step is query selection i.e either from the questions or the answers or the combination of both the question and answers. This query selection method includes the features of POS histogram and search performance prediction [1]. This Query generation for multimedia search supports only the 42-dimensional search prediction for each QA pairs.

C. Multimedia data selection and presentation

This method uses the queries which are generated earlier to perform the searching of image and video content from the web search engine [11]. Here search engine uses the graph-based reranking method to identify whether the query belongs to either person related query or non person related query. The face detection method is used to identify the most returned images and from that we can decide the person related query. The person related queries extract the 256-D Local binary pattern features [12] from the largest faces of images or video frames. The non-person related query requires extracting the 428-dimensional visual features and then it is to be raranked.

IV. MULTIMEDIA SEARCH DIVERSIFICATION METHOD

In the literature several failures has been observed. For instance, if the generated queries are tedious and complex then the system may fail to generate relevant multimedia answers. The video with textual answers scheme are good, but it is not informative to the users and at the same time the videos are providing unwanted information. Another issue is the lack of diversity for generating media data. To overcome this problem, the multimedia search diversification method has been introduced which provides more relevant, accurate and better answer to the users and hence it improves the performance of MMQA. This search diversification method provides the various answers for similar queries. The result of diversification method considers the extra information related to the data or distance among result set of elements [7]. Diversification search compromises between finding relevant elements to the query and achieving the variety in the result set. It considers the following three methods for answering the QA in effective manner.(i) content (ii) novelty (iii) semantic coverage.

A. Content based diversification method

It retrieve the answers based on user query like text, image and video and it differentiate the results to each other. Content based diversification method is based on the query, which selects k answers and selects the high relevance answers from the set k[10]. This method is used to avoid the un necessary answers in the cQA. K-similar diversification search defined as:-

$$Y= X' \subseteq X, |X'| = k \; F(q,X')$$

B. Novelty based diversification method

Novelty based diversification is the new result which is obtained by the users when compared to the previous one. This method is used for answering the similar queries. This diversification method is used to provide dissimilar answer for the same query and also it helps to improve the quality of answers. Novelty based diversification method provides informative answers to the users and also it provides good quality of videos along with textual answers.

C. Semantic coverage based diversification

In some cases the queries submitted to the retrieval system are ambiguous. For users to gather the relevant answers it is suitable to diversify the search results and hence top-ranked answers are displayed. Based on these method user requirements are satisfied and the lack of diversity is reduced.

V. CONCLUSION AND FUTURE WORK

In this literature, we describe the motivation and evolution of MMQA system which uses the media data by leveraging textual answers in cQA. This approach first predicts which type of medium is suitable for enriching the text based QA pair and then it automatically generates the query based on the QA and performs the multimedia search. Eventually query-adaptive reranking and duplicate removal has been performed to obtain more relevant and accurate media information which will be provided along with the textual answers. In Existing system MMQA approach has been used which faces some issues and in order to overcome this, the multimedia search diversification method has been proposed. This method improves the efficiency of cQA service and simultaneously it achieves better performance. The future work is focused on reducing the traffic between the cQA, system and users and also efficient retrieval of multimedia data.

REFERNCES

[1] Liqiang Nie, Meng Wang, Member, IEEE, Yue Gao, Zheng-Jun Zha, Member, IEEE,and Tat-Seng Chua, Senior Member, IEEE "Beyond Text QA: Multimedia Answer Generation by Harvesting Web Information"- IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 15, NO. 2, FEBRUARY 2013.

[2] T.-S. Chua et al., ''From Text Question-Answering to Multimedia QA on Web-Scale Media Resources,'' Proc. 1st ACM Workshop on Large-Scale Multimedia Retrieval and Mining, ACM Press, 2009.

[3] G. Li et al., ''Video Reference: Question Answering on YouTube,'' Proc. 17th ACM Int'l Conf. Multimedia, ACM Press, 2009, pp. 773-776.

[4] E. Parzen and F. Hoti, On Estimation of a Probability Density Function and Mode,'' Annals of Mathematical Statistics, vol. 33, no. 3, 1962.

[5] F. Schroff, ''Harvesting Image Databases from the Web,'' Proc. 11th Int'l Conf. Computer Vision, IEEE Press, 2007.

[6] D. Liu et al., ''Tag Ranking,'' Proc. 18th Int'l Conf.World Wide Web, ACM Press, 2009, pp. 351-360.

[7] K. Wang, Z. Ming, and T.-S. Chua, ‘‘A Syntactic Tree Matching Approach to Finding Similar Questions in Community-Based QA Services,’’ Proc.32nd Int’l ACM Sigir Conf. Research and Development in Information Retrieval, ACM Press, 2009.

[8] J.L. Song, ‘‘Scable Image Retrieval Based on Feature Forest,’’ Proc. Asian Conf. Computer Vision, Springer Press, 2009.

[9] H. Feng, A. Chandrashekhara, and T.-S. Chua, ‘‘Atmra: An Automatic Temporal Multi-Resolution Analysis Framework for Shot Boundary Detection,’’ Proc. Int’l Conf. Multimedia Modeling (MMM), ACM Press, 2003.

[10] H. Bay et al., ‘‘Speeded-Up Robust Features (SURF),’’ Computer Vision and Image Understanding, vol. 110, no. 3, 2008, pp. 346-359.

[11] S. Lazebnik, C. Schmid, and J. Ponce, ‘‘Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories,’’ Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition (CVPR), IEEE CS Press, 2006.

[12] D.R. Radev et al., ‘‘Evaluating Web-based Question Answering Systems,’’ Proc. Int’l Conf. Language Resources and Evaluation, 2002

