Bilingual Text-Based Image Retrieval Using PDA's (BTBIRUP)

Amer N. Abuali, Loai F. Al-Zoua'bi, Haifa Y. Abu-Addose, Shaima Y. Hamam

Abstract – The aim of this paper project is to provide a powerful combination of information retrieval approaches and the new technology advances (mobile technology). We examined the capabilities of mobile devices and PDA's in retrieving images based on Text-based image retrieval approach (TBIR) by using their caption which written in Arabic and English languages. The results were promising in helping Arab community in Image Retrieval field. Copyright © 2010 Praise Worthy Prize S.r.l. - All rights reserved.

Keywords: CBIR, TBIR, Bilingual Dictionary-Based Approach, CLIR

I. Introduction

Many IR systems have been proposed and implemented using different approaches and technologies. Content-based image retrieval (CBIR) for example retrieves images based on the content characteristics such as color, shapes, texture, etc.; however, another promising and simpler approach to image retrieval is the text-based (TBIR) approach providing access to images based on captions describing the contents of images such as the Google image search engine.

TBIR allows for combining one of the most aspects and approaches to information retrieval; that is, to submit a query in one language (Source language) and the results are in another language (Target language) [1]. This bilingual behavior expands the users' knowledge and exploits the opportunity to take advantage of the available information on the internet. There are different techniques for implementing such a bilingual system: the translation-based and the statistical-based techniques. In this paper we focused on one of the implementations of the translation-based technique, which is the Dictionary-based approach.

Most IR systems have been developed for desktop computers; usually retrieving information from the web, but the advances in technology has brought new ideas and approaches such as mobility, which can be employed for IR systems. The increasing number of mobile devices and the capabilities of such devices have enabled many opportunities to develop mobile based applications. Although mobile device are surrounded by some limitations, but the emergent need for a mobile based IR system has led to the development of our paper.

This proposed paper will enable users to search the web for the desired images by entering Arabic queries (Source language) then the system translates the query using the dictionary to English language (Target language) and retrieve the required images seamlessly from the images database.

II. Image Retrieval

Searching the Web, which is full of information of all kinds and subjects, is not a simple task. Users generally find information using search engines - websites that match the users input with existing Web pages using a search algorithm. The input to these search engines often consist of keywords or other written text like a question. But the Web contains not only text, but also information in other modalities such as images. These images have to be retrieved using one of the image retrieval systems. There are two approaches for image retrieval; Content-based image retrieval, and Text-based image retrieval.

II.1. Content-Based Image Retrieval (CBIR)

Content-based image retrieval (CBIR) has become popular in the last few years. CBIR systems try to return those images that are visually most similar to an example image; similarity is based on a set of low-level image features. Features that can be used to index images are color, texture, shape and spatial layout [2]. However; the empirical studies have shown that using image features to find similar images is usually insufficient [3]. First, it is difficult for user to specify visual queries with low-level visual features (color, shape, texture...). Second low level image features cannot precisely describe user information needs. There is a gap between low-level visual descriptions and user’s semantic expectation

II.2. Text-Based Image Retrieval (TBIR)

Text-based image retrieval approach (TBIR) depends on retrieving images via caption or text associated within each image. Many image repositories have this information (e.g. Web image collections, stock photographic databases, medical image repositories and historic photographic collections) enabling simple text-based access. Indeed, because of the success of text-based methods techniques to automatically assign
keywords or categories to images is also an intense research area [4].

II.3. Cross Language Image Retrieval (CLIR)

One area of recent interest in the field of information retrieval (IR) is cross-language retrieval (CLIR) that provides multilingual access to document collections [4]. Given that images are “language independent”, ideally the language used to express the associated texts or textual queries should not affect the success of retrieval, i.e. an image with a caption written in English should be searchable in languages other than English. Providing multilingual access to image repositories offers benefits to both the owners and end-users alike: enabling wider access to images which might otherwise be unsearchable. The first task of (CLIR) system is translation. User query and caption must be in the same language for retrieval process to be carried out. One approach is to translate the captions of all images in the collection into the user’s query, but the problem for this approach is the large number of captions to be translated, this takes a large amount of time. Another approach is to translate the user’s query into the language of image Caption. In this approach minimum translation is needed. Several approach used to query translation [5], Machine translation, statistical model, and using Dictionary.

II.4. Arabic/English Text-Retrieval Using Bilingual Dictionary

Bilingual Dictionary [5] is a method used to translate the user’s query (source language) into the (target language) stored in the dictionary (Fig.1).

In Arabic-English retrieval the Arabic query is translated into English using the bilingual dictionary, then the query is matched against the English caption collection.

Although Bilingual dictionary has suffer from several problem such as missing words, lack of proper noun, etc, “In Arabic machine translation systems, there is an increasing need for robust methods for the extraction of Arabic Multiword Terms (MWTs) to reduce the expected ambiguity inherent in word-to-word matching”, [11] but the simplicity of using this dictionary for translation make its good choices for cross-language retrieval.

III. Image Retrieval Using Mobile Devices

Traditional and classical image retrieval systems are conducted on the desktop computer; usually retrieving information from the web, but the advances in technology has brought new ideas and approaches such as mobility, which can be employed for IR systems.

The increasing number of mobile devices and the capabilities of such devices have enabled many opportunities to develop mobile based applications. Although mobile device are surrounded by some limitations, but the emergent need for a mobile based IR system has led to the development of our project.

In addition, the mobile devices industry is going through phenomenal change over the past few years with significant advances in areas of communications and multimedia.

IV. Related Work

In the context of this paper, there are different works for image retrieval using mobile devices. M-MUVIS [6] is a Content-Based System for image retrieval using mobile devices. This system has been tested with an image database. They have obtained quite encouraging results based on content-based retrieval. And they are facing limiting factors such as hardware/software features in mobile devices and unpredictable network responses.

In another system, Hirotaka Sonobe [7] proposed image retrieval system of fishes using mobile computing. In his system, the name and attributes of fishes are retrieved using a mobile device. In Order to retrieve information of fishes, the system uses the result of the retrieval is displayed on a web page constructed on the characteristics of shape and color extracted from a fish image. In the previous systems the approaches that uses is CBIR rather than TBIR. In our system we will use the TBIR approach.

V. Text-Based Image Retrieval Using PDA System

V.1. System Architecture

Our system is client-server system architecture. A mobile devices or PDA is used as s client side, and the server machine is used as a server side.

The reason why the system is client-server is the limitations of computational power and the limited size of storage for mobile devices (client side).

In our system the users can input their queries in Arabic language on PDA and press search then this request will pass from PDA (client side) to the server machine (server side), and the results will displayed on the PDA screen. These results consist of the images and their captions. (Fig. 2) describes the architecture of our proposed system.
Fig. 2. Basic Model of our proposed System

V.2. Interface Consideration for PDA in IR System

User interface design for mobile devices has been discussed for many years now, with the consensus that a mobile interface should require a different interaction style from that of the GUI desktop interface, and that attempts to cram all the functionality of desktop system into a mobile device are a mistake [8].

In addition to the processing and memory issues, the designer has to consider the screen size of the mobile devices. Therefore, using mobile devices in IR system must lead to create interface that have the simplicity and friendly for user to fit with his expectations.

The interface of our system will illustrate in (Fig. 3).

Fig. 3. User Interface of Our System

VI. Methods & Techniques

In our system there are number of methods and techniques are applied. Some of these methods are used in most of traditional information retrieval (IR) systems. But others are used in advanced information retrieval systems such as cross language image retrieval. In this paper we combined these methods to obtain new methodology for Image retrieval systems. The main purpose of our paper is retrieving the relevant images depends on their caption using mobile devices. To achieve this purpose we should following sequence of methods and techniques. Our methods for developing this system begin by collecting the required data for testing and evaluating the system result. As mentioned above this collection is relatively small and consist of the images itself and their captions (Image name, Image number, and Image description). This collection is stored in database to perform the required processing on it. Pre-processing operations will performed on image captions such as indexing technique by building an inverted file, and removing of stopwords, etc. these operation will enhanced both the search and retrieval in order to obtain better results in reasonable time. Because our system is bilingual text-based image retrieval, we used the bilingual dictionary which is a method used to translate the user’s query (source language) into the target language stored in the dictionary. After applying the previous methods on this system the final step is displaying the results (Images) on PDA screen. At this step, we take into consideration the limitation of mobile screen, such as small size and other factors; in our paper we put into account these limitations and we try to overcome these factors.

VII. Our Collection

To test any IR system it should have a collection of test data to ensure that the system behave well and satisfy user’s expectations. This collection should be choosing carefully to reflect the system functionalities. Building or choosing collection is not easy work and need to spend a lot of time and money. For image retrieval systems there are many standard collections used to testing these systems, for example St Andrews Collection (SAC). “The St. Andrews collection consists of 28,133 images, all of which have associated textual captions written in British English (the target language). The captions consist of 8 fields including title, photographer, location, date and one or more pre-defined categories (all manually assigned by domain experts)” [9]. In our system we built the small collection manually from the scratch, the reason for this is the inability to access one of the standard collections that is used. Our collection consisted of 161 images all of which have a text caption written in both languages Arabic and English. And each caption consisted of 3-fields (Image number, Image name, and image Description) as in the (Fig.4). In this system we used the Image Description field to retrieve images and discarded the two other fields. The image description in our collection consisted of 2544 terms and 3572 word occurrences; the average caption length was 18 words, with maximum length of 35 words.
VIII. The Experiments

The usability experiment was set up as a preliminary exploration of how capable users were at cross language image searching. It was decided to set a series of known item searching tasks [1]. The implementation of our system can perform by preparing group of queries in Arabic languages in order to tests the system results. These queries are typed and prepared depends on the system’s collection. The results from these queries will be pre-known by the users; for example in our collection there are 5 images relevance to the (عمليات التفاضل) query. Once the system run the query will be translated into target language (English language, "World Miracles"), then the results will be obtained into both languages. Finally, the query will be evaluated and examined in order to measure the system effectiveness. To illustrate the experiments of this system consider the following query:

Query1: 
عمليات التفاضل
Number of relevant images in the collection in both languages: 11
The retrieved image is: 10
The result of this query after running the system is illustrated in (Fig. 5).
In our experiment we prepared 4 queries to test the system results and to help in evaluating the effectiveness. These queries are presented in the next section (system evaluation).

**IX. System Evaluation**

In general, every system should be evaluated by applying some of measurement techniques in order to assess the system's performance and functionality. In term of Information Retrieval systems, the evaluation can be conducted to assess the effectiveness of these systems. We means by “effectiveness” that the ability of the system to retrieve the relevant documents (images) while at the same time holding back non-relevant one, and it can be measured by using several measures such as Precision, Recall, Novelty, Fallout, F-Measure etc.

In our system we used Precision and Recall measures in order to evaluate the system effectiveness. The formula for each measure as following:

\[
\text{Precision} = \frac{\text{number of relevant doc retrieved}}{\text{Total numbers of doc retrieve}} \quad \text{Recall} = \frac{\text{number of relevant doc retrieved}}{\text{total number of relevant docs}}
\]

For this evaluation we prepared 4 queries in Arabic languages (Q1, Q2... Q4). For each query we calculated the precision and recall and the results are presented in (Table 1).

<table>
<thead>
<tr>
<th>Query number</th>
<th>Recall</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>14/23=61%</td>
<td>14/20=70%</td>
</tr>
<tr>
<td>Q2</td>
<td>16/17=94%</td>
<td>16/20=80%</td>
</tr>
<tr>
<td>Q3</td>
<td>8/11=72%</td>
<td>8/10=80%</td>
</tr>
<tr>
<td>Q4</td>
<td>12/12=100%</td>
<td>12/19=63%</td>
</tr>
</tbody>
</table>

After we calculated the Recall/Precision for each query, we computed the interpolated recall-precision average were used as input for graphical presentations of the runs. Precision average at 11 standard recall level (0, 0.1, 0.2... 0.9, 1.0) are used to compare system performance. The recall-precision average is computed by summing the interpolated precision of whole run at the specified standard recall cut of values [10]. The interpolated precision was measured at the 11 recall levels of (0.0, 0.1, 0.2... 1.0). The measured interpolated precision was then averaged over the set of our queries.

<table>
<thead>
<tr>
<th>Recall Level</th>
<th>Interpolated precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100%</td>
</tr>
<tr>
<td>10%</td>
<td>100%</td>
</tr>
<tr>
<td>20%</td>
<td>100%</td>
</tr>
<tr>
<td>30%</td>
<td>94.25%</td>
</tr>
<tr>
<td>40%</td>
<td>95%</td>
</tr>
<tr>
<td>50%</td>
<td>94.5%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Recall Level</th>
<th>Interpolated precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>60%</td>
<td>89%</td>
</tr>
<tr>
<td>70%</td>
<td>67.2%</td>
</tr>
<tr>
<td>80%</td>
<td>67.7%</td>
</tr>
<tr>
<td>90%</td>
<td>44.8%</td>
</tr>
<tr>
<td>100%</td>
<td>35.75%</td>
</tr>
</tbody>
</table>

Once they were averaged with other queries, a composite precision-recall curve showing 11 points can be graphed as illustrated in Fig. 6.

**X. Conclusion**

This system was developed to combine the information retrieval systems and the advance technology (mobile technology). There were some limitations that Bilingual dictionary had suffered from several problem such as missing words, lack of proper noun, etc. In another hand, mobile device are surrounded by some limitations such as small screen size. Another major limitation is the inability to access one of the standard collections that is used, forcing us to build manually from scratch a small collection which resulted in facing problems when evaluating the system performance. But
the emergent needs to system that can help the Arab community in Image Retrieval field have led to develop our paper. In conclusion, we can evaluate the system behavior more precisely when we have a realistic collection of images to conduct a realistic evaluation.

References


Authors’ information

Amer N. Abuali: drabuali@yahoo.com
Loai F. Al-Zoua'bi: Loosi.Foad@gmail.com
Haifa Y. Abu-Addose: haddose@yahoo.com
Shaima Y. Hamam: Shaimayah@yahoo.com