Multi-Model Search Engine for Text and Image Retrieval

Diksha Chawla [1] 
PDM College of Engineering, Bahadurgadh

Vineet Shrivastava [2] 
Babu Banarsi Das Institute of Technology, Ghaziabad

Abstract: With the popularity of the network and development of multimedia technology images on web are increasing. Due to these large growing amounts of visual and multimedia data underline the need to create access methods that offer more than simple text based queries or based on matching exact database field. We propose a refining search algorithm to narrow down the search further from retrieved images. We propose a system that can allow user to generate Multi-Model queries (Text +Image) across language boundaries and provide subjectively more useful results than keyword and image search engine.

Keywords: Text, Image, Multi Model, Retrieval

Introduction

There has been a rapid increase in the size of digital image collection. Every day, many images have been generated by many groups including journalist, Engineers, historians, designers, teachers, artist and advertising agencies. The image needs and usages vary among these users. They may require access to images based on primitive features such as color, texture or shape or associated text [16, 17]. With this ever increasing large collection of data user may have to face many problems like searching for and retrieving relevant content. One approach which uses text as a query which uses Meta tags or text available in the context of image [12]. The success of this approach limited by the quality of contextual material. An alternative approach like video google enables image search using image as a query by finding similar region of image in database. This method exploits only visual content which may not necessarily be sufficient [12]. With the increasing usage of networking users’ needs online help on many issues, they can learn wide variety of task such as [13] how to set up home network, backup files or change the speed of mouse cursor. Such type of resources is often available online for user help. Software developers always provide up to date information and provide guidance. These resources are created by third party experts, sites offering tutorial and online help (E.g EHow.com) [13].

Recently content Based image retrieval [4, 18] becomes an active research area. In this different techniques are provided which is used to retrieve and index images from database based on visual content. To build such type of system we have to deal with many challenges including different type of textual data present in images. Different type of images exist on web and before answering queries and retrieving relevant image we have to deal with different type of text data present in such images. For example, Fig 1 a) Shows document images that contain text and few graphics content. Fig 1 b) shows scene text which occurs naturally as a part of the scene and contain important semantic information.

Figure 1(a): Document Images

Figure 1(b): Scene Text
2. Applications
The proposed method has been used in wide range of applications such as:
1. GUI based computer applications – Online GUI can assist users to know how to perform wide variety of task.
2. Cross lingual search – It helps user in searching relevant document without any language restrictions.
3. Document analysis – Different type of documents that contain text embedded in images for filling semantic gap.
4. Detection of vehicle license plate – Text extraction can be used to supervise traffic in real time. Vehicle license can be recognized easily from traffic accidents, which can improve the efficiency of transportation system.
5. Content based image filtering – In this image spam can be detected and pornography, reactionary and fraud words can easily be filtered.

3. Proposed Solution
In this paper, the work has been carried out to provide user Multi-Model Search engine which can extract, recognize text and extract images from web. It also enables text information extraction from any type of images including both document and real scene images.

Metrits
1. It enables text queries under cold start conditions even when resources have not been assigned any tags.
2. It offers fast and intuitive method to describe an interface by capturing single screen shot rather than thinking of many keywords.
3. It offers users Multi-Model search for online computer help.
4. It also offers Multi-lingual search so that language would not be a barrier for searching required document.
5. The benefits of combining text extraction and content based retrieval allow the user to narrow down the range of results found so that a more precise result of image is found.

Demerits
1. Huge range of vocabulary is required and to be updated frequently to extract meaningful information.
2. Some text and images are not detected as maximally stable extraction regions(MSER’s) due to the instability of colors caused by the disturb of the illumination conditions, hence many resources are waste on them.
3. Most of the maximally stable extraction regions are repeating with each other.
4. Characters that are above or below baseline segmented into two components which makes recognition difficult.

4. Identification steps
Our technique for searching image is based on text present in an image and pattern matching concept.

Work Flow
1. Image acquisition and preprocessing
The first stage of any vision system is image acquisition stage. After the image has been obtained preprocessing can be applied. Preprocessing in an improvement of the image data that suppresses unwilling distortions or enhances some image features important for further processing.

2. Text extraction
After preprocessing next step is to detect text regions which is printed against shaded or texture background or is embedded in and images examples photographs, maps, advertisement etc.

3. Multi-Lingual Character Recognition
Multilingual information applications which are driven by modeling lexical resources[12]to high degree-the quality of lexicon present is the main bottleneck for quality of performance. As a first step document image is segmented into various zones, regions and word tokens using suitable geometric properties. Within each zone the words are compared to dictionaries associated with various candidate languages and the language that exhibits the highest confidence factor is identified as the language of the zone. Subsequently zone is split further into regions. The language for each word is determined.

4. Vocabulary and indexing
Vocabulary recognition is one of the crucial step in searching image content. Vocabulary[3] which is present in image usually consists of large sample of features from a representative corpus of images. Image vocabulary extraction enables indexing of images. Documents of text contain distribution of words, and thus can compactly be summarized by their word counts. The occurrence of words tend to be sparse across different documents, an index maps word to the files in which they can occur can take keyword query and immediately produce relevant document.

Conclusion
The need of Text and image retrieval is to provide user convenient and faster way to access GUI applications and with less human intervention. The proposed work on different limitation of already existing methods and generate a prototype which help user to generate online web content without specifying much detail. The method considers various features of Image like shape, color, texture, character, etc and extract relevant document.

References
[1] Satish Kumar, Sunil Kumar, Dr.S.Gopinath” Text extraction from images”, International Journal of advanced research in computer Engineering and Technology, volume 1, June 2012.


