

Content Based Image Retrieval for Space Weather Application

Esraa Zeki Mohammed

State Company for Internet Services, Ministry of Communication, Kirkuk, Iraq.

(Received November 29, 2015, accepted March 13, 2016)

Abstract. The research presented in this paper was aimed to develop an CBIR system to aid in automatically detect and retrieve solar flares images based on detected edges and color features. In general, the work consists of two phases: (1) enrollment phase, which consist of feature extraction based on Conventional Color Histograms (CCH) and edge detection for sunspots and active regions on the sun using Sobel filter, (2) retrieving phase, which use the similarity measurement to retrieve most similar images from the database for a given query sample. The conducted tests were carried on 250 images and give good precision and recall rates.

Keywords: color histogram, image retrieval, similarity measure, sobel filter, solar flares, sunspots.

1. Introduction

Digital imagery is getting more popular in many perspectives. Private photo collections, medical imaging, solar phenomenon and geographical information systems are only some to mention. As the computation power is growing and the cost of storage media is decreasing, the size of digital image collections is increasing rapidly. There is a need for techniques that enables us to access and retrieve the huge amount of information embedded in these image collections. Now, simple manual browsing mechanism is getting cumbersome even with private collections. So, automatic image retrieval is inevitable [1].

The field of content-based image retrieval (CBIR) focuses on the analysis of image content and the development of tools to represent the visual content in a way that can be efficiently searched and compared. CBIR techniques extract visual features and perform the indexing, querying and retrieval based on such features. General CBIR approaches are intended for browsing large archives of general (arbitrary) content [2]. More recently, CBIR systems are starting to emerge in different applications, and the solar flares phenomenon studying domain is currently cited as one of the principle application domains for content-access technologies.

The term "space weather" refers to adverse conditions on the Sun that may affect space-borne or ground-based technological systems and can endanger human health or life. The importance of space weather is increasing day after day because of the way solar activities affect life on Earth and it will continue to increase as we rely more and more on different communication and power systems [3]. The established effects of space weather activities on our daily lives and can be summarized as follows:

Ground based systems: induced electric field and currents can disrupt the normal operation of high voltage power transmission grids. Pipeline, telecommunications cables, metallic oil and gas pipelines and railway signaling [4].

Communication systems: wireless communication systems suffer from interruption of service like frequency jamming and dropped communications due to radio bursts caused by solar microwave emission. Solar activity can be produce X-ray that disrupt poin-to-point high frequency radio communications and radio noise that interferes with communication and radar systems [5,6].

Space based systems: adverse space weather conditions can be cause anomalies and system failures and increased drag on movement of satellites and spacecraft leadind to slow-down, change in orbits and shorter life-times of missions. Other radiation hazards include direct collision damage and/or electrical effects, caused by charged particles [7].

Astronomical observations have a long history, dating back to ancient times. However the technological progress in recent years opened new areas of research and exploration of previously unknown astronomical phenomena and dependencies among them. Since most of the astronomical data come in the form of images that lead to make image processing techniques natural means of data analysis. Some of these techniques methods have been used in astronomy for many years and led to revolutionary discoveries and observations. One of the areas relatively undeveloped in terms of image processing is solar image feature extraction [8].

With a substantial increase of solar image databases an automated detection and verification of various features of interest is becoming increasingly important for a reliable forecast of the solar activity and space weather [9].

The visual contents of the images are extracted and described by feature vectors that form a feature database. During the retrieval stage, the users provide the retrieval system with a sample image (query image). The system then changes this query into feature vectors. In the matching stage, the system calculates the similarities between the feature vectors of the query sample and those of the images in the database, and then performs the retrieval accordingly [10].

Many studies have applied the concepts of feature extraction and retrieval to analyze and retrieve images such as:

Csillaghy, et al [11], presented a method that summarizes the image information content by partitioning the image in regions with same texture. They call this process texture summarization. Second, indexing features are generated by examining the distribution of parameters describing image regions. Indexing features can be associated with global or local image characteristics. Both kinds of indexing features are evaluated on the retrieval system of the Zurich archive of solar radio spectrograms. The evaluation shows that generating local indexing features using self-organizing maps yields the best effectiveness of all tested methods.

Seo, et al [12] presented a ROI (Region-Of-Interest) based medical image retrieval system that is considered a combination of feature descriptors and initial weights for similarity matching. For semantic ROI segmentation, they create attention window to remove the meaningless regions included in the image such as background and propose a quad-tree based ROI segmentation method.

Banda, et al [13], presented an extensive comparative analysis of several different domain-specific datasets in order to provide some guidance for the solar physics community on the well-researched field of medical image analysis allowing them to transfer knowledge from one applied field to their own. They report on the transfer of image parameters that produce good results for medical images to the domain of solar image analysis. Using the first solar domain-specific benchmark dataset that contains multiple types of solar phenomena they discovered during their work for constructing a content-based image retrieval (CBIR) system for NASA's Solar Dynamics Observatory (SDO) mission that they could take advantage of the research on the analysis of images in the medical field.

Jyothi et al [14] proposed a retrieval system using region based image retrieval system, finding region in the pictures using a new image segmentation method by improved mountain clustering (IMC) technique and features are extracted using a set of orthogonal set of moment functions for describing images. The achieved average precision rates were in the range of 68% to 76% considering a database covers brain tumor, tooth decay, lung cancer, and tuberculosis images.

2. Concepts and Methods

2.1. CBIR

In 1992, the National Science Foundation of the United States organized a workshop on visual information management systems in order to identify new directions in image database management systems. At this conference, Kato [10] introduced the term Content-Based Image Retrieval (CBIR) to describe automatic retrieval of images from a database. He emphasized the use of color and shape as the criteria for the automatic image retrieval process. Since then, the term CBIR has been adopted to describe an image-retrieving process that is used for large collections of images; which is based on features that can be automatically extracted from the images themselves [15].

Content Based Image Retrieval is a set of techniques for retrieving semantically-relevant images from an image database based on automatically-derived image features. The main goal of CBIR is the retrieval efficiency during image indexing and retrieval, thereby reducing the need for human intervention in the indexing process. The computer must be able to retrieve images from a database without any human assumption on specific domain.

One of the main tasks for CBIR systems is the similarity comparison. It is involved with the extraction of feature signatures of every image depending on its pixel values and, then, defining the rules for comparing images. The extracted features are considered as the image representation structure which used for measuring