i-Search Software Requirement Specification

Anagha Anand

Anirudha M. S.

Rakshitha Vibhu

Yashwin B. P.

TABLE OF CONTENTS

|  |  |
| --- | --- |
| Introduction | |
| Problem Statement | |
|  | Problem Explanation | | |
|  | Problem Solution | | |
| Features of The Project | | |
| General Description | |
|  | Users and Characteristics |
|  | Product Perspective |
|  | Overview of Functional Requirements |
|  | Overview of data Requirements |
|  | General Constraints, Assumption Dependencies, Guidelines |
|  | User View of Product Use  Hardware and Software Requirements |

System Design Description

|  |  |
| --- | --- |
| Preliminary Design | |
|  | Layout of Content Based Image Retrieval System |
|  | About Technology |
|  | Design Considerations |

**Introduction**

Image retrieval is concerned with techniques for storing and retrieving images both efficiently and effectively.

Robust Content-based image retrieval (referred to as RCBIR), which is based on automatically extracted primitive features such as color, shape, texture, and even the spatial relationships among objects, has been employed since the 1990’s . In the last ten years, a great deal of research work on image retrieval has concentrated on CBIR technology.

Image databases and collections can be enormous in size, containing hundreds, thousands or even millions of images. The conventional method of image retrieval is searching for a keyword that would match the descriptive keyword assigned to the image by a human categorizer. Currently under development, even though several systems exist, is the retrieval of images based on their content, called Content Based Image Retrieval, CBIR. While computationally expensive, the results are far more accurate than conventional image indexing. Hence, there exists a tradeoff between accuracy.

The need for Content- Based image retrieval is to retrieve images that are more appropriate, along with multiple features for better retrieval accuracy. Usually in search process using any search engine, which is through text retrieval, which won’t be so accurate. So, we go for Content- Based image retrieval. Content- Based Image Retrieval also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR) .

The similarity measurements and the representation of the visual features are two important issues in Content-Based Image Retrieval (CBIR).

Given a query image, with single / multiple object present in it; mission of this work is to retrieve similar kind of images from the database based on the features extracted from the query image. In this we use features like

* RGB Color
* GCH (Global Color Histogram)
* LCH (Local Color Histogram).

Further to Image retrieval we can able to compare the several pairs of Images and find out whether it is exactly matching or not.

**Problem Statement**

The problem involves entering an image as a query into a software application that is designed to employ CBIR techniques in extracting visual properties, and matching them. This is done to retrieve images in the database that are visually similar to the query image.It’s a resolution and format dependent.

Image databases and collections can be enormous in size, containing hundreds, thousands or even millions of images. The conventional method of image retrieval is searching for a keyword that would match the descriptive keyword assigned to the image by a human categorizer. Currently under development, even though several systems exist, is the retrieval of images based on their content, called Content Based Image Retrieval, CBIR.

While computationally expensive, the results are far more accurate than conventional image indexing. Hence, there exists a tradeoff between accuracy and computational cost. This tradeoff decreases as more efficient algorithms are utilized and increased computational power becomes inexpensive.

Here we can use only jpeg format images as a query image due to which it becomes format dependent and other formats are considered error. We cannot use higher resolution images as input image because of this restrictions it get difficult to use different images as query image.

The limitations of CBIR are

* Query process is difficult as a user has to specify query image by selecting from exiting systems.
* Impartial for very large databases.
* Image retrieval based on the shape is very difficult to automate.
* The 100 percent of accuracy can not be attained.

|  |  |  |
| --- | --- | --- |
| **Problem Solution**  The solution initially proposed was to extract the primitive features of a query image and compare them to those of database images. The image features under consideration were color and texture. Thus, using matching and comparison algorithms, the color, the shape and texture features of one image are compared and matched to the corresponding features of another image. This comparison is performed using color, and texture distance metrics. In the end, these metrics are performed one after another, so as to retrieve database images that are similar to the query.  The similarity between features was to be calculated using algorithms used by well known CBIR systems such as QBIC. For each specific feature there was a specific algorithm for extraction and matching.  **Features of the Project**  The main feature of this thesis work is to retrieve images that are similar to query image from a large database. We use content- based search, for high accuracy multiple features like color, texture and shape is incorporated. Color feature extraction is done through “Global Color Histogram (GCH)” and “Local Color Histogram”, and Average RGB Color Value.  CBIR’s advantages   * Retrieve images that are more appropriate. * Uses multiple features for better retrieval accuracy. * Makes use of the contents of image themselves, rather than relying on human-inputted metadata such as captions or keywords. * Retrieve similar kind of images from the database based on the features extracted from the query image. * Use features like RGB Color, GCH (Global Color Histogram) and LCH (Local Color Histogram). * Compare the several pairs of Images and find out whether it is exactly matching or not.   CBIR’s disadvantages   * Sometimes it gives approximate output not the accurate. * Has to be enchanced further in future.   **General Description**  Under this topic we get the description of the users who uses the system along with their characteristics and it specifies the product. It also describes briefly the functional and data requirements of the project.  **Users and their Characteristics**  The users of this product are military agents crime detectors and architectural and engineering designer in software systems, interior designers, medical diagonisors etc.   * User will give the query image .This image should be a real world image. * User also has to mention the features on which the related images should be retrieved like global color histogram, local color histogram and average RGB etc. * Later the user will get the related images for the query image as the output.   **Product Perspective**  Our product is developed to get the related images for the query image from the database. It is helpful in web designing who wants to search the whole database images for the specific type of images for designing purpose. The product meets the 90% of the accuracy level.   * Crime prevention * The military * Architectural and engineering design * Fashion and interior design * Journalism and advertising * Medical department |  |  |

**Overview of Functional Requirements**

**Purpose:** To retrieve the related images from database based on similarity features.

Retrieve similar kind of images from the database based on the features extracted from the query image and makes use of the contents of image themselves, rather than relying on human-inputted metadata such as captions or keywords.

**Inputs**: Query images from the real world entity and compare the several pairs of inputs Images and find out whether it is exactly matching or not.

**Outputs:** List of related images is displayed along with the total time required.

Use features like RGB Color, GCH (Global Color Histogram) and LCH (Local Color Histogram), and produces related output images.

**Overview of Non-Functional Requirements**

**Security:** Ensure that this particular program is secured and safe for the users, and does not cause any hinderance when the execution is taking place.

**Performance:** Performance of the software will be evaluated at the end, and will be compared, and the list of images would be displayed.

**Overview of Data Requirements**

The product requires the database to store the images to be compared with query image. Images in the database and the query images features have to be extracted which is required in similarity measurements. The features extracted will be based on Average RGB, Local color histogram and Global color histogram. So the images of different world entities are stored in database.

**General Constraints, Assumptions Dependencies, Guidelines**

Our product is entirely based on image processing and requires JPEG format of the image to process with. Users must have an enough memory to hold feature extracted so that all the similarities result can be stored temporally. The system should also specify that the object is an image file but not a text file.

**User View of Product Use**

User can use this product for the following:

* Web designing purpose**.**
* Security systems
* Criminal detection
* Image processing

**Hardware and Software Requirements**

The hardware and software requirements of the system which are required for the implementation of the project in a system.

SOFTWARE REQUIREMENTS

Technologies : Java, AWT, Swings

JDK : Version 8

DataBase : MySQL Client and Server

Operating system : Microsoft Windows 7

HARDWARE REQUIREMENTS

Processor: Pentium core i5 Processor

RAM :4 GB(minimum)

**SYSTEM DESIGN DESCRIPTION**

The design description consists of preliminary design and detail design. Preliminary design specifies the high level design and detail design specifies the low level design.

Image databases and collections can be enormous in size, containing hundreds, thousands or even millions of images. The conventional method of image retrieval is searching for a keyword that would match the descriptive keyword assigned to the image by a human categorizer. Currently under development, even though several systems exist, is the retrieval of images based on their content, called Robust Content Based Image Retrievalusing multiple features extraction, RCBIR. While computationally expensive, the results are far more accurate than conventional image indexing.

Hence, there exists a tradeoff between accuracy and computational cost. This tradeoff decreases as more efficient algorithms are utilized and increased computational power becomes inexpensive.

**Preliminary Design**

The purpose of the design is to plan the solution of a problem specified by the system requirements. The design of the system is perhaps the most critical factor affecting the quality of the software and has a major impact on the later phases, particularly testing and maintenance.

System design aims to identify the modules that should be in the system, the specifications of these modules and to interact with each other to produce the desired results. At the end of the system design all the major data structures, file formats, output formats as well as major modules in the system and their specifications are decided.

**Layout of Content Based Image Retrieval System**

* The administrator has to insert the images one at a time or folder wise into the database before starting with the query image.
* The user interface gives the query image whose related image from database has to be extracted. The images should be from real world entity.
* Features of the query image is extracted and stored temporally for comparison.
* In the same way the features of images in the database is extracted.
* User specifies the feature on which the related database image has to be extracted as output. Features considered are color histogram like local and global color histogram, average RGB.
* After the comparison if the image of database is related to the query image then that image will be added in the list of output images.

**About Technology**

The technology used in implementing CBIR is divided into three layers. It is as shown in the table below. The basic functionalities are developed in Java class member functions. Here the Java Swing procedures and functions residing in the front-end call Java class member functions.

As with securing a network, securing a database by looking at the various layers that are involved is an effective approach. Security of databases can be defined as preventing unauthorized or accidental disclosure, alteration, or destruction of data .

In addition, the confidentiality of data that exists in the database must be considered, as should the availability of that data. The following section will discuss a secure database design; while not all-inclusive, it should provide a good, basic starting point.

Access to information contained in the tables must be properly regulated. This can be done with control over direct access to the tables, and also through views. Views and privileges assigned to the views can be created to limit users to only see specified portions of data contained within a table .

Through the use of the selects, projections and joins, existing relations between tables in a relational database, as well as a single table, can be created. Control over the read, insert, update and delete commands must also be assigned appropriately within those views.

Role-based authentication should be considered when adding access to any database. Typical roles for access include administrator, user, programmer and operator. For the first three roles, it is fairly obvious what access should be granted; it is the operator role that can be a sticking point. Operators are expected to play an essential part in the production operation of a system, yet they are often restricted in what type of access they are granted.

Segregation of duties should be considered in the operator role, instead of just granting one operator control over an entire process. Operators’ roles do need to be carefully defined and kept within the realm of production support as much as possible. Furthermore, all roles should have logging enabled to keep track of what occurs .

It is important to remember that changes made to the database, whether structural or to the data itself, must be tracked and regulated by interested parties. Whether formal or informal, the process must be defined and followed by all roles defined in the database structure.

**Design Considerations**

This section describes many of the issues, which need to be addressed or resolved before attempting to devise a complete design solution**.**

**Related Hardware and Software:**

Hardware Requirements: The hardware requirements for running the project are as follows:

* 4GB RAM(Minimum)
* Webcam (optional)

Software Requirements: The software requirements for running the project are as follows:

* Operating system: Windows 7

Languages used: Java

**Operating System:**

The operating system used is any Windows operating system. The user can access the application through software’s (Java) in Windows operating system. Windows is the most popular and user-friendly operating system. It sufficiently caters to the needs of the application on the client side. It’s good enough to support all sorts of activities required by user. Because user may not be technical graduates it’s easier for them to use the application in the windows operating system.