CONTENT BASED IMAGE RETRIEVAL

Sahil Jain¹, Kiranmai Pulaparthi², Chetan Fulara³

¹,²,³Student, Maharaja Agrasen Institute of Technology, GGSIP University, New Delhi, India

Abstract: The purpose of this Paper is to describe our research on different feature extraction and matching techniques in designing a Content Based Image Retrieval/Processing system. Due to the enormous increase in image database sizes, as well as its vast deployment in various applications, the need for CBIR development arose. This paper outlines a description of primitive feature extraction techniques like: texture, color, and shape. Once these features are extracted and used as the basis for a similarity check between images, the various matching techniques are discussed.

This paper proposes novel system architecture for CBIR system which combines techniques including content based image and color analysis, as well as data mining techniques. This is a study to propose segmentation module to build the CBIR system. It also includes concept of neighborhood color analysis module which also recognizes the side of every grids of image. The study also includes the implementation of 4 algorithms namely: - K-MEANS, SIFT, SURF and BRIEF algorithm.

Keywords: Content based images retrieval; K-Means Clustering; Feature extraction; Image retrieval; SIFT; SURF; BRIEF.

INTRODUCTION

As processors become increasingly powerful, and memories become increasingly cheaper, the deployment of large image databases for a variety of applications have now become realizable. Databases of art works, satellite and medical imagery have been attracting more and more users in various professional fields — for example, geography, medicine, architecture, advertising, design, fashion, and publishing. Effectively and efficiently accessing desired images from large and varied image databases is now a necessity. CBIR or Content Based Image Retrieval is the retrieval of images based on visual features such as color, texture and shape. Reasons for its development are that in many large image databases, traditional methods of image indexing have proven to be insufficient, laborious, and extremely time consuming. These old methods of image indexing, ranging from storing an image in the database and associating it with a keyword or number, to associating it with a categorized description, have become obsolete. This is not CBIR. In CBIR, each image that is stored in the database has its features extracted and compared to the features of the query image. It involves two steps:-

1) Feature Extraction: The first step in the process is extracting image features to a distinguishable extent.

2) Matching: The second step involves matching these features to yield a result that is visually similar.[2]

Image matching is a fundamental aspect of many problems in computer vision, including object or scene recognition, solving for 3D structure from multiple images, stereo correspondence, and motion tracking, etc.
Image processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. Main stages in Image processing and matching are as follows:

- Detection of local features in images present in the stored database.
- Descriptor Extractor- Describing the features to obtain a large dictionary to be used for matching the test images.
- Training the machine and making it learn about the stored images by passing a feature vector.
- Detecting and describing features of test image and matching it with the feature vector.
- Returning the output.

**Examples of CBIR applications are:-**

- **Security Check:** Finger print or retina scanning for access privileges.
- **Intellectual Property:** Trademark image registration, where a new candidate mark is compared with existing marks to ensure no risk of confusing property ownership.
- **Medical Diagnosis:** Using CBIR in a medical database of medical images to aid diagnosis by identifying similar past cases.
- **Crime prevention:** Automatic face recognition systems, used by police forces.

**RELATED WORK DONE**

We do not yet have a universally acceptable algorithmic means of characterizing human vision, more specifically in the context of image understanding. Hence it is not surprising to see continuing efforts towards it, either building up on prior work or exploring novel directions. The matching measure, termed integrated region matching (IRM), has been constructed for faster retrieval using region feature clustering and the most similar highest priority (MSHP) principle[5].

Color has been an active area of research in image retrieval, more than in any other branch of computer vision. Color makes the image take values in a color vector space. The interest in color may be ascribed to the superior discriminating potentiality of a 3-dimensional domain compare to the single dimension domain of gray-level images.

![Data flow in essential query by content](image)

**Fig 1.** Data flow in essential query by content

Understanding the nature and scope of image data plays a key role in the complexity of image search system design. Factors such as the diversity of user-base and expected user traffic for a search system also largely influence the design. Along this dimension, we classify search data into the following categories: Personal Collection, Domain-Specific Collection, Enterprise Collection, Archives and Web. [9]

An important parameter to measure user-system interaction level is the complexity of queries supported by the system. From a user perspective, this translates to the different modalities she can use to query a system.
There are various querying modalities such as: Keywords, Free-text, Image, Graphics, Text-based, Content-based queries etc.

There are various kinds of feature descriptors which are used for mapping an image with the other image. Some of them are: corners, edges, ridges, colors, textures, intensity, etc., locality, pose invariance, distinctiveness, repeatability. [6]

An image retrieval system called Wavelet-Based Color Histogram Retrieval (WBCHIR) was proposed. It is based on the combination of color and texture features. The color histogram for color feature and wavelet representation for texture and location information of an image. This reduces the processing time for retrieval of an image with more promising representatives. [11]

In December 2012, a system was proposed in which the CBIR system was divided into two parts: learning and querying. Learning step tells about the training process in which a huge amount of sample images are input to the system in the first step, then the feature extraction process takes place and then clustering is performed. K-means cluster algorithm is selected to cluster the training data. [4]

SIFT algorithm was introduced for the detection and description of keypoints. It proved out to be a boon for image processing as it was based on the extraction of keypoints which was independent of scale variance and rotation variance. It comprised of mainly five steps: Scale Space Extrema detection, Keypoint Localization, Orientation Assignment, Feature description, Feature matching.

A more efficient algorithm which was developed for image retrieval was SURF (Speeded Up Robust Features). It is a local feature detector and descriptor that is used for object recognition, 3-D reconstruction, classification or registration. It is several times faster than SIFT and is claimed to be robust against different image transformations than SIFT. [13]

Another algorithm which was discovered to enhance the speed of processing was BRIEF (Binary Robust Independent Elementary Features). It is a high speed algorithm which provides smoother image extraction and is also independent of any scales or rotation in the input or dataset. For each detected keypoint (FAST), Sample all intensity pairs ($I_1, I_2$), typically 256 in total) according to pattern around the keypoint. If $I_1 < I_2$, add a 1 to the descriptor, otherwise 0. [8]

Chin-Chin Lai have proposed an interactive genetic algorithm (IGA) to reduce the gap between the retrieval results and the users’ expectation. They have used color attributes like the mean value, standard deviation, and image bitmap. They have also used texture features like the entropy based on the gray level co-occurrence matrix and the edge histogram. They compared these methods with other approaches and achieved better results. [1]

Color is one of the most reliable visual features that are also easier to apply in image retrieval systems. Color is independent of image size and orientation, because, it is robust to background complication. First a color space is used to represent color images. Typically, RGB space where the gray level intensity is represented as the sum of red, green and blue gray level intensities. Swain and Ballard proposed histogram intersection, an L1 metric as the similarity measure for color histogram. Color histogram is the most common method for extracting the color features of colored images. Color histograms are widely used for CBIR systems in the image retrieval area. [3]
Image Retrieval Using Color Histogram: Histogram represents the distribution of intensity other color in the image. The image retrieval consists of the following stages: Query image is given from the user, Histogram of the Color image is calculated, Color Histogram of the database images are calculated, Euclidean Distance is calculated, sorted the distance in ascending. [12]

**INTRODUCTION TO ALGORITHMS**

1) **K-Means Cluster**

K-means clustering algorithm first defined the size of K clusters. Based on the features extracted from the images themselves, K-means allocates those into the nearest cluster. The algorithm calculates and allocates until there is little variation in the movement of feature points in each cluster.[10]

There are four modules in the main system architecture they are:

- Segmentation and Grid module
- K – Means clustering module
- Feature Extraction Module
- Neighborhood Concept Module

**Fig 2.** Flow chart of the system architecture.

**Fig 3.** Sample of Segmentation and grid module and Color feature extraction
2) **SIFT - Scale Invariant Feature Transform**

The SIFT approach, for image feature generation, takes an image and transforms it into a "large collection of local feature vectors". Each of these feature vectors is invariant to any scaling, rotation or translation of the image.

It involves 5 major steps:-

- Scale Space Extrema Detection
- Keypoint Localisation
- Orientation Assignment
- Feature Description
- Feature Matching
3) SURF- Speed Up Robust Feature

SURF creates a “stack” without 2:1 down sampling for higher levels in the pyramid resulting in images of the same resolution. In keypoint matching step, the nearest neighbor is defined as the keypoint with minimum Euclidean distance for the invariant descriptor vector. Lower used a more effective measurement that obtained by comparing the distance of the closest neighbor to that second-closest neighbor.

4) BRIEF Algorithm - Binary Robust Independent Elementary Feature

BRIEF is a general-purpose feature point descriptor that can be combined with arbitrary detectors. It is robust to typical classes of photometric and geometric image transformations. BRIEF is targeting real-time applications leaving them with a large portion of the available CPU power for subsequent tasks but also allows running feature point matching algorithms on computationally weak devices such as mobile phones. It provides a shortcut to find the binary strings directly without finding descriptors. [7]

The first important thing is Detection; in this step you want to detect points-of-interest or keypoints, and what that means is that you want to choose local points (basically small patches) that you think are interesting in the image, there are many ways to do that; this paper doesn't contribute in this area. However, it seems that they use SURF feature detector and SURF keypoints

After detection is done, Feature Description follows. You know the interesting points in the image and now you want to describe them (basically you want to describe the points/patch around the interesting points). SIFT is one popular feature descriptor.
COMPARISON AMONG VARIOUS ALGORITHMS

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Features</th>
<th>Advantages</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-Means</td>
<td>1) Cluster the training data. 2) Outputs the clustering result as a learning code book.</td>
<td>1) Accurate and efficient. 2) Less computation time. 3) Easy to implement. 4) Well developed over 50 years.</td>
<td>1) Difficult to predict ‘k’ with fixed number of clusters. 2) Doesn’t provide good results for images with high feature similarities. 3) Doesn’t work well with non-globular cluster.</td>
</tr>
<tr>
<td>SIFT (Scale Invariant Feature Transform)</td>
<td>1) Scale Space Extrema Detection 2) Keypoint Localization 3) Orientation Assignment 4) Feature Description 5) Feature Matching</td>
<td>1) Locality 2) Distinctiveness 3) Quantity 4) Efficiency 5) Extensibility</td>
<td>1) Being a 128-vector, is relatively slow to compute and match. 2) Generally doesn’t work well with lighting changes and if blur.</td>
</tr>
<tr>
<td>SURF (Speeded Up Robust Feature)</td>
<td>1) Exploit the integral image. 2) Creates a grid around the keypoint. 3) Divides each grid cell into subgrids.</td>
<td>1) Reuse the calculations 2) Maintain the robustness to rotation, scale illumination change. 3) 2X faster than Difference of Gaussian. 4) Improves on SIFT by using a box filter approximation.</td>
<td>1) Invariance reduces the ability to discriminate. 2) Difficult to track edges robustly. 3) View dependent. 4) More fragile</td>
</tr>
<tr>
<td>BRIEF (Binary Robust Independent Elementary Feature)</td>
<td>1) Feature Detection. 2) Feature Description. 3) Matching.</td>
<td>1) Highly discriminative. 2) Almost 4 times faster both to build and match. 3) High recognition rate.</td>
<td>1) Can only tolerate small amount of rotation. 2) Absence of any method to find the feature.</td>
</tr>
</tbody>
</table>

REFERENCES


