Content based Image Retrieval using Machine Learning Technique

Charulata Leuva
Department of Computer Science & Engineering
GTU, Ahmedabad, India

Abstract

The fast growth of computer technologies and the coming of the World Wide Web have increased the amount and the complex difficulty of combining video, sound, words and pictures together information. A Content Based Image Retrieval (CBIR) system has been developed as an efficient image retrieval tool where by the user can provide their question to the system to allow it to retrieve the user’s desired image from the image collection. However the usual relevance responses to something or helpful returned information method to support the user question based on the representative image selection and weight ranking of the images retrieved. The Support Vector Machine(SVM) has been used to support the learning process to reduce the semantic gap between the user and the CBIR system. SVM can classify the data into relevance training set and Gabor Filtering will extract the feature from the given image dataset. It can also improve the performance of CBIR. Also solve the imbalance training set. 

Keywords: Image Retrieval, Content Based Image Retrieval, Semantic Gap, Relevance Feedback, Support Vector Machine, Gabor Filtering

I. INTRODUCTION

A. Image Retrieval

Image Retrieval is a method to retrieve the images from the given database. Image retrieval gives the images as per user request. But traditional Image Retrieval technique does not give the relevant images as per user requirement. Therefore Content-based image retrieval is introduced to retrieve visual contents of an image such as color, shape, texture, and spatial layout. Text-based image retrieval uses traditional database which give the images as per user input the text. Which manage images [1]. Through Text description images can arrange in hierarchical way and it will browse the images as per the text description. However, since automatically generating descriptive texts that type of images is not feasible, even text-based image retrieval systems require manual annotation of images. Obviously, manual annotation of image is a crucial and expensive task for large image databases, and is often subjective, context-sensitive and incomplete. As a result, traditional image retrieval is difficult for large database system [1].

Fig. 1: Diagram for content-based image retrieval system

B. Image Retrieval Methods

There are basically two types of image retrieval methods:
- Image Meta Search(Text Based)
- Content Based Image Retrieval
1) Image Meta Search (Text Based):

Image Meta Search is also known as Text Based Image Retrieval. It is retrieve images based on keyword, label or metadata. And it is not give relevant result as per user requirement. It is traditional retrieval method [2].

2) Content Based Image Retrieval:

The requirement for development of CBIR is enhanced due to large volume of images as well as the widespread uses in multiple fields. Texture, color, shape and spatial layout are the underlying traits to represent and index the images [2].

C. Content Based Image Retrieval

In [1] nowadays Numerous CBIR systems have been proposed, yet at the same time now the numerous issue of recovering pictures on the premise of their pixel content remains to a great extent unsolved. Basically four types of techniques are used for Content based image retrieval:

1) Query Techniques

CBIR has Different implementations query techniques. In Query by example user give the example technique and CBIR provide resultant technique as per search upon. Query technique removes the difficult which can arise when trying to describe the images with words also. Query Techniques will give better result than traditional technique [1].

2) Semantic Retrieval

Semantic Retrieval is method which begins the inquiry with the assistance of user making a request like "find pictures of Busses". This sort of open-ended errand is extremely troublesome for PCs to perform – Busses may not generally be confronting the camera or in the same posture. Numerous CBIR frameworks thusly by and large make utilization of lower-level elements like composition, shading, and shape. These components are either utilized as a part of mix with interfaces that permit less demanding data of the criteria or with databases that have as of now been prepared to match feature [1].

3) Relevance Feedback (Human Interaction)

Relevance Feedback derives the images in relevant and irrelevant with the user intent. User will separate the images with their voting to the images. In this techniques user interaction take part for retrieving relevant images [1].

4) Machine Learning

Machine Learning techniques are more useful for Content Based Image Retrieval. It is most suitable for complex semantics but difficult to implement. There are two types Supervised Learning and Unsupervised Learning which both are associated low level features with query concepts. There are many transform based techniques which are useful to retrieve the texture based feature extraction [1].

![Flow Chart of Content Based Image Retrieval](image)

In fig 2 present block diagram of content based image retrieval which represent how to extract the feature of images and how to matching the images with different similarity matrices techniques

D. Color Histogram

In [17] A histogram is the distribution of image into the number of pixels. The quantity of components in a histogram relies on upon the quantity of bits in every pixel of image .. For example, if we consider a pixel depth of n bit, the pixel Values will be in between 0 and 2n, and the histogram will have 2n elements. For very large data sets, color, moments can be computed based on color histogram as follows.

\[
\text{mean} = \frac{\sum_{i=0}^{255} i \cdot h(i)}{\sum_{i=0}^{255} h(i)} \quad (3)
\]

\[
\text{variance} = \frac{\sum_{i=0}^{255} h(i) \cdot (i - \text{mean})^2}{\sum_{i=0}^{255} h(i)} \quad (4)
\]

Where \( h \) is the histogram of the image.
E. **Color Coherent Vector (CCV)**

In [14] Color histogram does not give the spatial data of pixels. This may bring about comparative color conveyance for distinctive pictures. Color coherent vector addresses this issue. In CCV every histogram container is parcelled into two sorts: Coherent and Incoherent. Pixel worth has a place with a huge casually shaded district falls into lucid sort. Else it falls into incoherent. A colorlucid oherent speaks to this grouping for every color in the image.

F. **Color Correlogram**

In [15] the color correlogram was purposed to describe the color dispersion of pixels, as well as the spatial relationship of sets of hues. The main and the second measurement of the 3-D histogram are the colors of any pixel pair and the third measurement is their spatial separation. Color correlogram is a table listed by shading sets, where k-section for (i,j) indicates likelihood of discovering a pixel of color "j" at separation "k" from pixel "i" in the picture. Give i a chance to speak to the whole arrangement of picture pixels and Ic(j) speak to the arrangement of pixels.

\[
y_{i,j}(k) = Pr_{p1 \in I_{i}, p2 \in I_{j}} |p1 - p2| = k]
\]

Where \( i, j \in \{1, 2, ..., N\}, k \in \{1, 2, ..., d\}, \) and \( |p1 - p2| \) is the distance between pixels p1 and p2.

G. **Similarity Measurements**

In [15] Similarity measurement is coordinating the component with image database. In this first concentrate the components of images and contrast and image database. When elements of images in the database are separated and the user question is framed, the query items are acquired by measuring the comparability between the preextracted components of the picture database and the investigated user's inquiry. Rather than definite coordinating, substance based image retrieval ascents visual similarities between a query image and images in a database. Numerous comparability measures have been produced for image recovery in view of exact evaluations of the circulation of elements as of late. We mean \( D(I, J) \) as the separation measure between the question picture I and the picture J in the database; and \( f(I) \) as the quantity of pixels in container I of 1.

H. **Semantic Templates**

Semantic Template is a set of general characteristics that are calculated from the number of images stored in the database. It traces a link between the characteristics of high and low-level features. It usually defined as a function of a concept representative calculated from a group of sample images.

<table>
<thead>
<tr>
<th>Comparison between various Semantic Reduction Gap Techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td>Techniques</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td>Object Ontology</td>
</tr>
<tr>
<td>Machine Learning</td>
</tr>
<tr>
<td>Relevance Feedback</td>
</tr>
<tr>
<td>Semantic Templates</td>
</tr>
<tr>
<td>Web Image Retrieval</td>
</tr>
</tbody>
</table>

I. **Wavelet Transform**

In [11] Wavelet transforms are taking into diminutive waves, called wavelets, of varying frequency & constrained length of time. Discrete wavelet transform remodel the image in four unique parts high frequency part (HH), high low frequency part (HL), Low high frequency part(LH), lower part (LL) vertical parts is 1-level picture deteriorations then process snippets of all recurrence part than store and utilization it as highlight to acquire the pictures. Surface entropy and contrast, clumsiness are the for the most part utilized properties. Factual components of dark levels were one of the productive strategies to classifying texture.

The Gray Level Co-occurrence Matrix (GLCM) is utilized to concentrate second request measurements from a image. GLCMs have been utilized gainfully for composition estimations. From Gray Level Co-event Matrix all the components are pondered and put away into the database. The utilization of Gray Level Co-occurrence Matrix gives great result however it is in spatial space so
it is more error pron. CCH (Contrast Context Histogram) to figure out the component of the question picture and different pictures put away in the database. CCH is in spatial space and it introduces worldwide appropriation. The MPEG Descriptors has been utilized like Edge Histogram Descriptor for surface. The Edge histogram separates edges as indicated by their course

J. Gabor Wavelet Transform:

In [11] The Gabor transform is demonstrated suitable for the element extraction, on the grounds that it minimizes the instability of the two dimensional verbalizations in the space and recurrence. Those 2D channels can be customizable in the introduction and scale. Hence, indicators of lines and borders, and measurable computations of miniaturized scale highlights in a particular region are made ordinarily to portray the data of the surface. The system of Gabor has been utilized as a part of a few uses of images preparing, for instance, for the image representation by the meaning of wavelets set that supply a complete representation of the same ones and the recuperation of images in view of the composition. There are uses of Gabor algorithm likewise for the reorganization.

K. Support Vector Machine (SVM)

SVM is a machine learning technique which supervised the learning process. The principle motivation behind SVM is to construct ideal isolating hyper planes. It recognize information and recognizes designs which are utilized for order and regression analysis. It takes an arrangement of information and produces a derived capacity called classifier or regression. The fundamental point is to draw hyper arrange as wide as would be prudent for a decent detachment that implies biggest separation to closest preparing information of pixel qualities. The separation between two hyper planes is the edge of the hyper planes regarding the example. The reason for SVMs is to boost this separation. On the off chance that separation of pixels to hyper arrangement is extensive than speculation error of classifier is low [10].

Support Vector Machines were presented as a machine learning strategy by Cortes and Vapnik in 1995. The fundamental instinct behind them is that given a two-class training set they anticipate its information focuses in a higher dimensional space and attempt to indicate a most extreme edge isolating hyper plane between the information purposes of two classes. This hyper plane is ideal as in it sums up well to unseen information. A more in the depth of SVMs takes after, in view of Burges and Vapnik in 1998. The training data of SVMs comprises of information focuses that are vectors of genuine esteemed numbers. The dataset is then anticipated to higher dimensional feature space, utilizing a capacity that fulfills Mercer's condition, the kernel function.

L. Algorithm for Proposed System Methodology

1) Step 1: Load Dataset and Extract Image Names
2) Step 2: Extract Image Name from query ImageFeature Vector
3) Step 3: Construct Labels
4) Step 4: Split training/testing sets
5) Step 5: Classify the dataset
6) Step 6: Create Confusion Matrix as per Precision and Recall
II. EXPERIMENTAL RESULTS

Precision = No. of relevant images retrieved / Total No of images retrieved
Recall = No. of relevant images retrieved / Total No of relevant images in the database

A. Similarity Matrices (Mahanthan Distance) L1 using SVM Function

<table>
<thead>
<tr>
<th>Total no. of images in the database</th>
<th>No of images give as input</th>
<th>Accuracy Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>499</td>
<td>05</td>
<td>82.60%</td>
</tr>
<tr>
<td>499</td>
<td>10</td>
<td>81.20%</td>
</tr>
<tr>
<td>499</td>
<td>20</td>
<td>79.60%</td>
</tr>
</tbody>
</table>
III. CONCLUSION

During the reviewed and analysis of various papers concentrates only on retrieving of image files. The main CBIR components including image feature representation, indexing and system design. Extracting feature is very efficient for retrieving the images. Also analyzed various CBIR techniques as well as gather the knowledge about different Machine Learning Approaches

In future this work can be enhanced to retrieve the audio and video file by using these features or modifying them. Also here we have done only binary classification using SVM; we can extend it further using Multi-class SVM and classify more features at a time.

REFERENCES

[1] Dr. Fuhui Long, D. H. FUNDAMENTALS OF CONTENT-BASED IMAGE RETRIEVAL.