International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 06 | June 2020 www.irjet.net p-ISSN: 2395-0072

Content based Video Retrieval

Mrs. Madhura Prakash[1]

Assistant Professor, Dept. of ISE
BNM Institute of Technology
Bangalore, India

Apoorva S[2], Arshiya Firdouse[3], Bhumika V Raj[4], Charul Kandoi[5]

Students, Dept. of ISE
BNM Institute of Technology
Bangalore, India

Abstract - *In today's world, huge amount of video data is being generated on a daily basis. The video information on any topic is much easier for the user to understand than textual information and hence most of the users prefer videos over textual information. Video databases are growing ever fast. In this case, searching for a particular video in a huge and varied collection of videos becomes challenging. Also, a user might have an image and may want to retrieve videos containing similar information. Then in such cases, searching a huge database becomes difficult and highly time consuming. The proposed system reduces the manual task of searching similar videos thereby saving the user's time and effort. The video frames can be annotated to perform search for retrieving videos containing similar content. But this technique will not take into account the context information of the video. The proposed system retrieves video of similar content by taking into account texture, edge features as the context is also considered. The proposed system extracts key frames from each of the video in the database. A key frame which is a single image that represents the video. For all the frames extracted, multiple features like the color and texture features are obtained. Subsequently, a comparison between the query or the test image and the database videos is done by measuring Euclidean distance.*

Key Words: ***Key Frame Extraction, Texture Feature, GLCM, Color Histogram, Edge Detection algorithm Euclidean Distance***

I. INTRODUCTION

Content Based Video Retrieval (CBVR) is used to retrieve desired videos from a large collection of videos on the basis of features that are extracted from the videos. A video has features like color, texture. These feature helps to describe the contextual information contained in the videos. Large volumes of videos are acquired and stored on computers due to frequent use of digital video devices in various areas. Video retrieval system is used for searching, browsing and restoring videos from excessive database of digital videos. The ability to manipulate and access stored videos is widely used by the users of different fields in various ways. The users find it difficult to locate the desired video in varied and huge collection of videos. Videos containing useful information are under-utilized unless CBVR systems are capable of retrieving desired videos by selecting relevant videos and filtering out the undesired videos. Search for solutions regarding the problems related to video retrieval has become the wide area for research and development.

Content based video retrieval and analysis is the one of the most important and recent research areas in the Image processing domain. Content based video retrieval can be used for multiuser systems for video search and browsing which are useful in web applications. Video retrieval system has acquired much importance because of the richness of the visual information contained in the video, thus Text based video retrieval methods have become inefficient methods. Content based video retrieval methods are applied in order to improve the effectiveness of the content-based methods. And these methods can play essential rules which are used in media collection and enhance retrieval accuracy. Therefore, content-based video retrieval (CBVR) plays a vital role instead of text-based retrieval in multimedia systems.

With few advances' multimedia technologies, digital TV and information highways, more and more video data have been captured, produced and stored. However, without appropriate techniques that can make the video content more accessible, all these data are hardly usable. So, the research on the management of video data is now a new field. But the differences between multimedia and textual data in continuity

make traditional database technology unavailable for access to handle the multimedia information. Consequently, the content-based access and retrieval become a proper solution.

Hence, Content-based video retrieval systems have established as powerful tools for finding specific content in ever-growing large-scale video collections. Video retrieval tools are typically built around a retrieval engine that returns a ranked video list according to various features.

II. METHODOLOGY

The proposed method consists of a database of video files from which key frames are extracted. From the key frames that are obtained from the videos, features like color, texture, edge are extracted from each of the image.

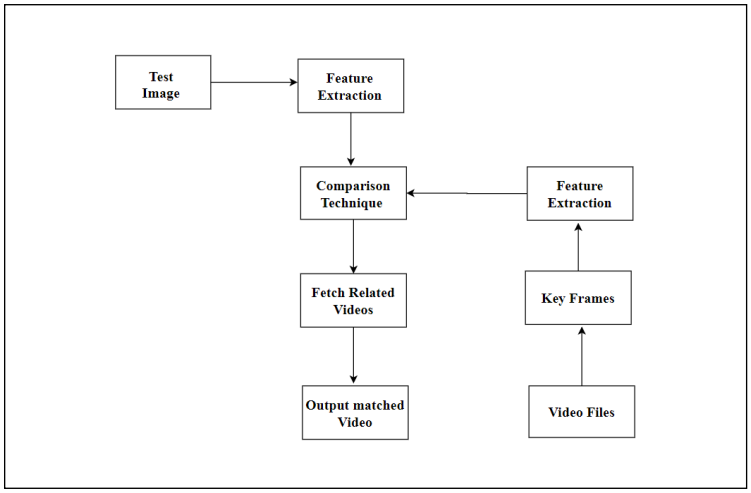

Fig. 1 Overview of the proposed CBVR System

Further, when a user inputs a test image or a query image for the video the user wants to retrieve, the color and texture features are extracted from the query image in a manner similar to the feature extraction from the database videos.

The general idea of our CBVR system is:
1)For the Videos stored in the database, Key frames are extracted using Key Extraction Method by implementing katna.

2)Feature Extraction is done using two methods:
i)Texture Feature: It is extracted using GLCM for the key frames and input image.
ii)Color Feature: It is extracted using Color Histogram for the key Frames and input image.
iii)Edge Feature: It is extracted using Canny Edge Detection.
3)Similarity Measure which is done using Euclidean Distance which compares the input image and key frames extracted.

A. Key Frame Extraction

Since there is lot of redundancy in the video for simplicity, we can select key frames from the video so that the key frame will represent the whole of the video.
A key frame is a representative image for a video. Key frame extraction plays an important role in the analysis of large amount of video files. Key Frame extraction reduces the useless information of the video. The key frame extraction is done using katna. The number of key frames to be extracted from a video can be limited. There are two modules: Video Module and Image Module.

Video Module: This module extracts the key frames from the video, the frames that provide the most accurate information about video content. The module extracts key frames based on Brightness score filtering and entropy/contrast score filtering. It employs the K-Means Clustering algorithm using image histogram to select the best frame from the clusters based on image blur detection.

Image Module: This module handles the smart cropping of the image. Smart cropping identifies the part or the area of the image where most of the focus is required to interpret the information contained in the image. The selection criteria for cropping the image would be Edge, Saliency or Face detection features that are detected in the image.

B. Feature Extraction

The process of feature extraction is useful when there is a need to reduce the number of resources required in processing without losing important or relevant

information. The Feature Extraction can also reduce the amount of redundant data for a given analysis.

Feature Extraction is basically the process of transforming the input data into a set of features, where the features are extracted carefully to perform a desired task.

Features like the Texture Feature and Edge Detection features are extracted from the key frames (image) obtained in the previous step.

I. Texture Feature Extraction:

- The Texture feature plays an important role in the analysis of any given image. The texture feature of an image are complex visual patterns that have characteristic such as brightness, color, shape and size.
- The image texture gives information about the color arrangement or intensities in an image or selected region of an image.
- Here, the texture feature extraction is done using the Gray-Level Co-Occurrence Matrix (GLCM).
- The GLCM is a second order statistical approach that considers the relationship between groups of two pixels in the original image.
- The GLCM is a widely and most commonly used approach for texture feature extraction because a few lower order methods only tell the probability of finding a Gray level in a particular pixel and do not consider neighbour pixel relationship. The other high order methods that consider relationship among three or more pixels are theoretically possible but not commonly implemented due to calculation time and interpretation difficulty.
- The GLCM is a tabulation of how often different combination of Gray levels co-occur in an image or a section of an image.
- GLCM is also called as Co-Occurrence matrix or Co-Occurrence distribution i.e. how the pixel values co-occur at a given offset and how they are distributed throughout the image.

II. Color Feature Extraction:

- Color content of an image does not change much when there is an object motion. Hence it provides a significant cue for retrieving objects which have significant color information. Similar images can also be identified using color information.
- Moreover, Human Vision system is more sensitive to color information than Gray levels and hence color feature is an important candidate for feature extraction.
- The Color Features of an image are extracted using the Color Histogram Technique.
- The Color Histogram represents the image from a different perspective. It represents the frequency distribution of color bins in an image.
- The Color Histograms are defined as a set of bins where each bin denotes the probability of pixels in the image of a particular color. Typically, each pixel in an image will be assigned to a bin of color histogram of that image, so for the color histogram of an image, the value of each bin is the number of pixels that has the same corresponding color.

III. Edge Feature Extraction:

- Edge Detection is a technique for finding the boundaries of objectss within images.
- It works by detecting discontinuities in brightness. The points at which brightness chnages sharply are typically organized into a set of curved line segments termed as edges.
- The Edge Detection Algorithm used in this paper is Canny Edge Detection Algortihm.
- The Color Histogram Method is dependent of the color of the object, ignoring its shape and texture.
- Also color histograms have high senstivity to noisy interference such as lighting intensity, which is one of the resons of using Canny Edge Detection over Color Histograms. The Canny Edge Detection Algorithm is a multi-step algorithm that can detect edges with noise supressed at the same time.

- Canny Edge is considered to be a better edge detection:
  i. Non Maximum Suppression: Edges candidates which are not dominant in their neighbourhood aren't considered to be edges.
  ii. Hysteresis Process: While moving along the candidates, given a candidate which is in the neighbourhood of an edge the threshold is lower.

These two steps reduce the number of "False" edges.

- The Canny Edge Detection also smooths the image using Gausian Filter to reduce the noise in the image.
- The Steps of Canny Edge Detection can be broken down into 5 steps:
  1. Apply Gaussian filer to smooth the image in order to remove the noise.
  2. Find the intensity gradients of the image.
  3. Apply non-maximum suppression to get rid of spurious response to edge detection.
  4. Apply double threshold to determine potential edges.
  5. Track edge by hysteresis: Finalize the detection of edges by suppressing all the other edges that are weak and not connected to strong edges.

- The Steps of Canny Edge Detection algorithm is as shown in figure 2.

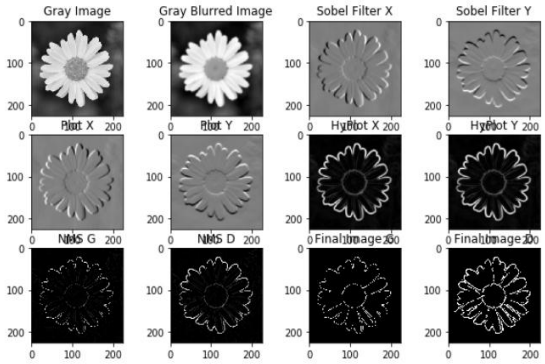

Fig. 2 Steps of Canny Edge Detection

C. SIMILARITY MEASURE:

- Similarity Measure plays a vital role in the proposed system to measure the similarities between input query image features and the video database features.
- The Euclidean Distance is the method employed to measure the similarity between input query image features and the video database features.
- Euclidean distance is the square root of the sum of squared differences between the corresponding element.

III. RESULTS

The experiment is performed on a PC with Intel core, 500GB hard disk, 8GB RAM Memory and run on a 64-bit windows 10 Operating System. The video database consists of 30 videos with 3 categories: wildlife, traffic and flower as shown in the fig. 3. The wildlife consists of 15 videos, traffic consists of 6 videos and flower consists of 9 videos. From the 30 videos, 150 frames are extracted as shown in the fig. 5. There are 75 frames extracted from wildlife, 30 frames from traffic and 45 frames from flower.

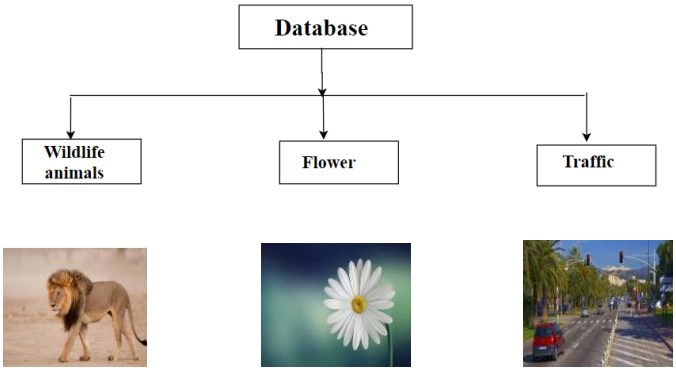

Fig. 3 Video Categories

Fig. 4 Video Database

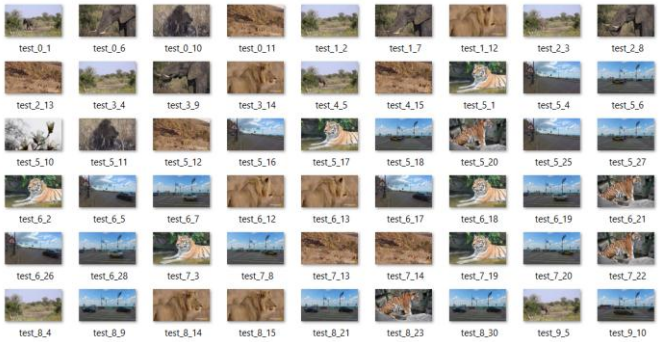

Fig. 5 Extracted Key Frames

The proposed system uses multiple features like edge detection features and texture features. The Canny Edge Detection is used for edge detection from the image and the GLCM texture feature is used to extract texture feature from the image.

The work done considering a combination of GLCM Texture Feature and Color Histogram showed an accuracy of only about 50-55%, because the color histogram is a low-level feature, sensitive to noise. Whereas, retrieving videos using a combination of edge detection and the GLCM texture feature showed a much higher and better combination, it showed an accuracy of 65-70%. Hence, in the proposed work, a combination of Edge Detection and GLCM Feature Extraction has been used to retrieve relevant videos.

The same process of feature extraction is repeated for the query image input by the user. Then the Euclidean distance is calculated to measure the similarity between the query image and the frames and hence retrieve the desired videos.

IV. CONCLUSIONS

Searching for a particular video in a huge and enormous set of videos require a lot of time and effort. The proposed system reduces the delay by automating the task of video retrieval by saving the users time, effort and energy. The proposed extracts multiple features from an image to aid in the process of video retrieval. To retrieve the desired videos correctly, only one feature like the color feature may not be enough. Hence, the system also considers the texture feature. Multiple feature extraction from an image would provide better accuracy compared to single feature. The accuracy of the system is about 70%. To achieve a better accuracy, the system could be trained with a larger and varied set of videos with higher configuration system and high-speed processing systems.

REFERENCES

[1] Aasif Ansari and Muzammil H Mohammed "Content based Video Retrieval Systems-Methods, Techniques and Challenges", International Journal of Computer Applications (0975 – 8887) Volume 112 – No. 7, February 2015.

[2] K.S.Thakre, A.M.Rajurkar and R.R.Manthalkar "Video partitioning and secured key frame extraction of MPEG Video", International Conference on Information Security & Privacy (ICISP2015), 11-12 December 2015, Nagpur, INDIA.

[3] Janarthanan.Y, Balajee J.M and Srinivasa Raghavan.S "Content based Video Retrieval and Analysis using image processing", International Journal of Pharmacy and Technology, December 2016.

[4] B. M¨unzer, M.J. Primus, M. Hudelist, C. Beecks, W. H¨urst, K. Schoeffmann "Content based Video Retrieval and Human Computation unite", 978-1-5090-6067-2 © 2017 IEEE.

[5] Manisha Verma, Balasubramanian Raman "A Hierarchical Shot Boundary Detection Algorithm Using Global and Local Features", December 2017 DOI: 10.1007/978-981-10-2107-7_35.

[6] Tejeswi Potluri, Gnaneshwara O Nitta "Content based Video Retrieval using Dominant color of the truncated blocks of Frame", Journal of Theoretical and Applied Information Technology 2016.

[7] Priyanka Ramesh Malode, S. V. Gumaste "Feature Extraction from Video Using Ordered Dither Block Truncation Coding (ODBTC)And RGB Color Space Model ", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 5, Issue 6, June 2016.