Content Based Video Retrieval

Ankita Muzalda (B14CS005)
Mentor : Dr. Chiranjoy Chattopadhyay
Department of Computer Science and Engineering
Indian Institute of Technology, Jodhpur

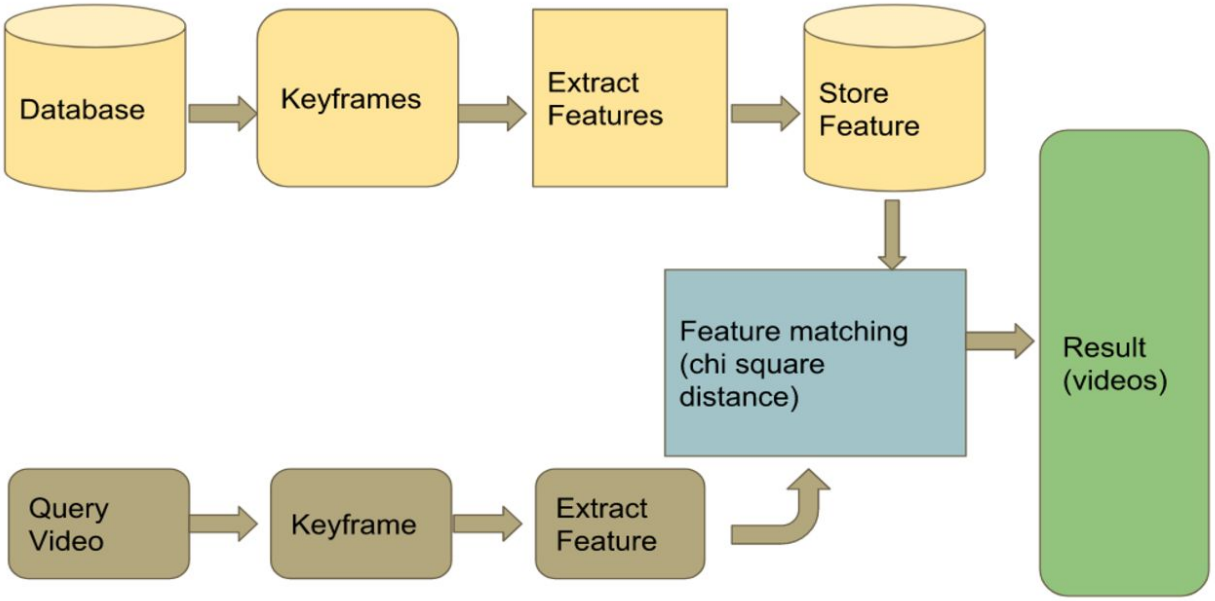
Motivation:

With increase in digital data it is hard to scan all the data for desired output and hard to give file name uniquely.So content of the data gives as accurate information about the file and with the help of content analysis we can easily access the desired file within less time,space and with less number of calculations.

Objective:

To develop a video Retrieval system first we need to find a representative frames of video we can call it keyframe which contains all the essential information of the video and stores it .we know that video files are large takes more times to process it, so keyframe are used to minimize the process time.By using the keyframe we can apply content based image retrieval method to retrieve similar video. So for the given query video, extract keyframe from it than extract feature of the keyframe match it with previously stored features of the keyframes of the database videos.

Methodology:



Algorithm :

keyframe Extraction Method :

Method 1 :

1. Extract all frames of the video.
2. Convert each frame to grayscale image
3. Find average intensity of the frame
4. Find consecutive difference of average intensity of frame
5. Draw a graph of the differences
6. Find the frame with highest and lowest difference in graph as key frame

Limitations :

It gives good results in case of still camera and single object video. For more than one object motion two keyframes are not sufficient .To increase the keyframes we can apply threshold to the difference in average intensity.

Method 2:

a single visual descriptor cannot capture all the pictorial details needed to estimate the changes in the visual content of frames and the visual complexity of a video shot.we must take into account both color properties and structural properties, such as texture.In this algorithm we have selected colour histograms, edge direction histogram and Multiresolution wavelet analysis.

Color histograms are frequently used to compare images because they are simple to compute and tend to be robust regarding small changes in camera viewpoint.

Two Sobel filters are applied to obtain the gradient of the horizontal and edges of the luminance frame image.

Multiresolution wavelet analysis provides representations of image data in which both spatial and frequency information are present

Distance between two colour histogram

$$d_H(H_t , H_{t+1}) = 1 - \sum[\min(H_t(j), H_{t+1}(j)]$$

Where H_t and H_{t+1} is colour histogram of F(t) and F(t+1)

The difference between two edge direction histograms (d_D)

$d_D(D_t , D_{t+1}) = \sqrt{(\sum(D_t(j) - D_{t+1}(j))^2}$

The difference between two edge direction wavelet statistics (d_w)

$d_D(W_t , W_{t+1}) = \sqrt{(\sum(W_t(j) - W_{t+1}(j))^2}$

The three resulting values (to simplify the notation we have indicated them as d_H , d_W and d_D only) are mapped into the range [0, 1] and then combined to form the final frame difference measure (d_{HWD}) as follows:

$$d_{HWD} = d_H d_W + d_W d_D + d_D d_H$$

For Keyframe Selection dynamically selects the representative frames by analyzing the complexity of the events depicted in the shot in terms of pictorial changes. Draw a graph of the cumulative frame difference by frame difference previously obtained.Sharp slopes indicate significant changes in the visual content due to a moving object, camera motion or the registration of a highly dynamic event.Find high curvature point.
Select a frame as midpoint between two consecutive high curvature point as keyframe for the video.

Features Extracted :

1. Colour Feature:
    - convert the image to the HSV color space
    - grab the dimensions and compute the center of the image
    - divide the image into four rectangles/segments (top-left,top-right, bottom-right, bottom-left)
    - construct an elliptical mask representing the center of the image
    - loop over the segments
        - construct a mask for each corner of the image, subtracting the elliptical center from it
        - extract a color histogram from the image, then update the feature vector.

- extract a color histogram from the elliptical region and update the feature vector
- return the feature vector

2. ORB (Oriented FAST and Rotated BRIEF):
   ORB if Fast keypoint detector and BRIEF descriptor . It takes the descriptor of one feature in first set and is matched with all other features in second set using some distance calculation. And the closest one is returned.

Similarity Measure :

To find similarity between two images we have to find feature vector and measured the similarity between the features

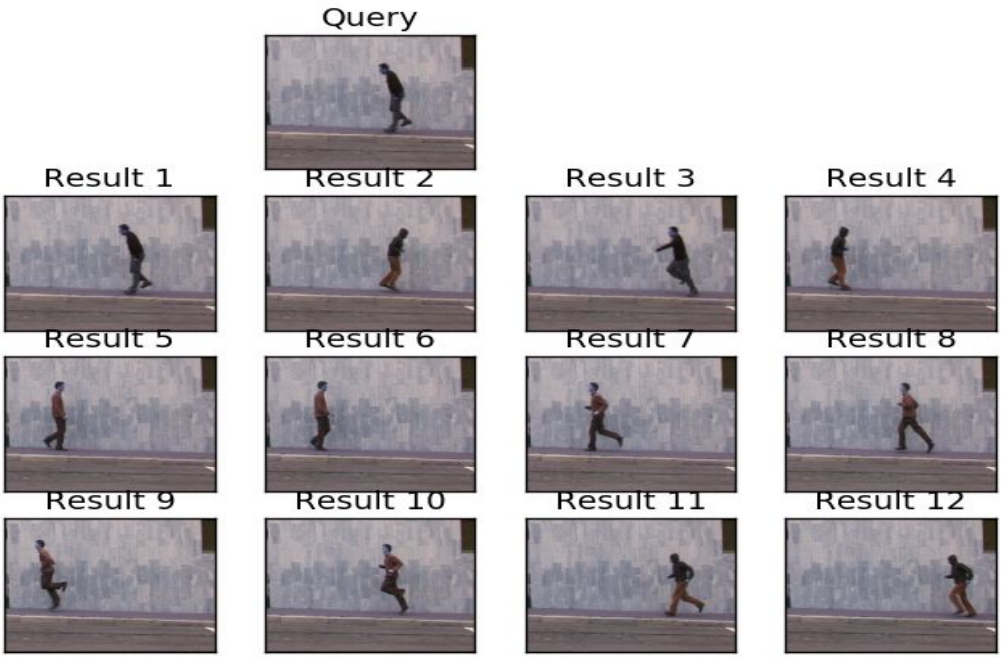
1. Chi-squared
   d(x,y) = sum((xi-yi)^2 / (xi+yi)) / 2

Images that have a chi-squared similarity of 0 will be deemed to be ***identical*** to each other. As the chi-squared similarity value increases, the images are considered to be ***less similar*** to each other.

2. cv2.BFMatcher() : Brute-Force matcher ,It takes the descriptor of one feature in first set and is matched with all other features in second set using some distance calculation. And the closest one is returned.

Results:

Keyframe similar to the query video.



The result is improved using new keyframe method as compare to tradition keyframe extraction method . By using this method we can overcome the limitation of the previous method.Clearly new approach is better but can be improved by some motion detection or Object movement.ORB feature is more efficient than simple colour feature.With enough number of input data we can apply machine learning algorithm to get more accurate result.

References:

1. *G. Ciocca and R. Schettini, "An innovative algorithm for keyframe extraction in video summarization," Journal of Real-Time Image Processing (Springer), vol. 1, no. 1, pp. 69–88, 2006*.

2. C. Sujatha and U. Mudenagudi, "A Study on Keyframe Extraction Methods for Video Summary," *2011 International Conference on Computational Intelligence and Communication Networks*, Gwalior, 2011, pp. 73-77

3.Y. Yang, B. C. Lovell and F. Dadgostar, "Content-Based Video Retrieval (CBVR) System for CCTV Surveillance Videos," 2009 Digital Image Computing: Techniques and Applications, Melbourne, VIC, 2009, pp. 183-187.

4. *W. Wolf, "Key frame selection by motion analysis," 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings, Atlanta, GA, 1996, pp. 1228-1231 vol. 2.*
doi: 10.1109/ICASSP.1996.543588

5. Idris, F., Panchanathan, S.: Storage and retrieval of compressed images using wavelet vector quantization. J. Vis.Lang. Comput. 8, 289–301 (1997)