

Discovering Thematic Object in a Video

¹Shalini N, ²Sharada K A

Abstract: In this paper we are identifying the frequently appearing object in a video. This frequently appearing object is called thematic object. If an object appears many times in a video such objects are called thematic object. Identifying this frequently appearing object in a video is helpful for object search and tagging of that object. To identify thematic pattern in the video we must give an object as an input and then we try to find corner points of that object by Harris corner detection algorithm. Later we can find the similarity between the reference image and test frame by extracting the descriptors around the corner point. We are mining the video to identify the common patterns that appears in that video. The proposed approach will help to identify the object even when there is partial occlusion and variation in the viewpoint.

Index Terms: Common Patterns, Corner points, Key object, Video data mining.

I. Introduction

We are given with a collection of videos. In that we need to identify the object which appears frequently. This frequently appearing object is called thematic object [1] or key object. Identifying this key object is helpful for visual object search and detection [2]. By keeping this key object we can be able to retrieve all the video frames which contain this key object.

Discovering objects in video is a challenging task [3]. By discovering, we mean that the object can be of any kind. Without having any prior knowledge about the object type or its position, we would like to identify an object from a video that occurs over a period of time. This is particularly challenging when the image sequence has low resolution and consists of highly cluttered background.

Object identification can be defined as [4] the process of segmenting an object of interest from a video scene and keeping track of its motion, orientation, occlusion etc. in order to extract useful information.

For identifying this key object, we characterize each video as a sequence of frames. We need to check whether this key object appears in the rest of the frames or not. It is not easy to

Identify whether that object is present in each frame or not. Because the shape of the object may vary from one frame to another frame. The view point of the object may vary i.e. an object can be viewed from front view or from side view. Sometimes there can be partial occlusion of the object. In all such cases we need to identify the key object in each and every frame of the video.

The video is nothing but the collection of frames. Each frames of the video contain some images. All the frames of the video may or may not contain the thematic object. So we need to

Identify which frames of the video contain this thematic object. To identify the thematic object, local feature points are calculated for every image. Local feature point is nothing but the corner points. The corner points can be identified by Harris corner detection method. After identifying the corner point, a window is generated around each corner point. A window is nothing but a descriptor which is collection of pixels around the corner point. Window is also called a patch around the corner point. Like this descriptors are extracted for the given image and the frames of the video. These patch descriptors of the given image and frame of video are compared each other to find similarity. If any similar object is present in the frame of the video, such frames are identified and saved so that it can be review in the later stage. The similarity between the given image and the frame of the video is done to identify the common object. Like this thematic object is identified in the given video. Identifying thematic object in a video is helpful for object search, tagging and video summarization

To identify this frequently appearing object we first convert the image into set of features. This feature vector is used for matching to find whether the object of interest [7] is present in the rest of the video.

II. Literature Survey

To identify common visual patterns in the image, some previous work identify an image as a graph consisting of visual primitives such as corners, interest points.

Shalini N, Computer Science and Engineering, East West Institute of Technology, Bangalore, India, +91-8971810862., (e-mail: shalinitn@gmail.com).
Sharada K.A, Assistant Professor, Computer Science and Engineering, East West Institute of Technology, Bangalore, India, +91-9448855473., (e-mail: sharadaa1234@gmail.com)

Yang and Cohen [5] use affine transformation for object recognition. An object is recognized by affine invariants to establish the correspondence between the vertices of a test image with a reference image. The algorithm used in this will recognize an object that must be represented as polygonal outliners and also as a set of scattered feature points. It uses point matching approach for recognition. However, if the objects have different shape then they have identical convex hulls.

Tan and Ngo [10] propose an approach to discover common patterns in a set of images by region matching. Here image is segmented into regions. Histogram is drawn for each segmented region. Histogram is the graphical representation of the given image. By the histograms of different segments, the presence of the common object is identified. If histograms are similar then there is common object is present in it. This is called region matching

Ying Shan and Harpreet [6] use a histogram to solve sequence to sequence matching problem. It will identify an object in the presence of large appearance and pose variations and also background clutter. However this approach uses intensity profile feature. This method is not invariant to illumination changes

Liu and Chen [2] propose an approach for video object discovery, which extracts unknown object form video. This approach uses video data mining and object oriented nature for video content analysis. But this provides a rough position for the object of interest.

III. SYSTEM MODEL

To find the object of interest in a given video and to find its reoccurrences in the video we use different modules. First is the Feature extraction which converts the image into a set of features or feature vectors which is used for further analysis. Next is the Feature matching, where the features extracted from the input image are matched with the stored template or reference model and a recognition decision is made.

All thematic recognition systems have to serve two distinguished phases. The first one is referred to as enrolment or training phase and the second one is referred to as the operational or testing phase.

To identify thematic object in a video, we characterize video as a collection of video frames, each frame as an image sequence. Each image is characterized by collection of local features. A local feature is a property of an image located on a single point or small region. It is a single piece of information describing the distinctive property of the object. Examples for the local features of an object are color or gray value of a pixel. For object recognition task, the local feature must be invariant to illumination changes, noise, scale changes and changes in view direction. These local features are also called visual primitives. We match the visual primitives to identify the presence thematic object in a video. In the initialization step, we discard uncommon visual primitives that find few matches among the rest of the image because they will not belong to thematic object. For the remained visual primitive, we make a larger visual group and check the commonness for the extent of thematic object in a video.

Various videos and images are collected and stored as a dataset. First we need to select an image. We need to select a video to identify whether the same query image is present in it or not. The feature points are identified by Harris corner point detection. A patch window of 21*21 is identified for every feature point by extract descriptor module. The match function is applied to the patch descriptor to find there is any match between the given image and the frames of the video. Like this every frame of the video is compared to find whether the thematic object is present in a video or not.

IV. BLOCK DIAGRAM

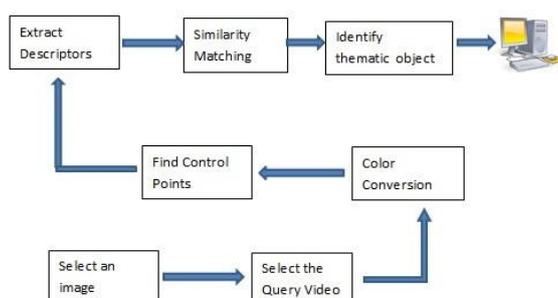


Figure 1

Above shows the block diagram for identifying thematic object in a video. First we are selecting an image and we are selecting a video, we are doing color conversion for both the reference image and the test video. For this we are finding the control points by using harris corner detection method. We are extracting the descriptors

around the corner point to do similarity matching. If similar object is found then it is displayed stating it as a thematic object.

V. SYSTEM MODULES

Here we are showing the different modules with their description.

A. Video acquisition module:

We can read the input video by uigetfile command which allow us to browse the video in which we are going to given as the input to the system. So we can get the video by selecting it from database.

B. Video preprocessing module:

The size of the video is verified and it is resized to 320x240 of frame size. These videos are sampled at 2 frames per second.

So, in video preprocessing step frames are extracted from the video and they are resized if they are not of proper size.

C. ROI Selection:

ROI is nothing but region of interest. Careful selection of video is essential for ROI analysis. For this we select an image and it is kept as a reference frame. This is used to compare the frames of the video to identify the thematic object is present in the video.

D. Color conversion:

Color is the brains reaction to a specific visual stimulus. We can precisely describe color by measuring its spectral power distribution which leads to a large degree of redundancy.

There are three basic quantities they are

Radiance is the energy that flows from the light source and it is measured in watts.

Luminance is a measure of energy i.e. what an observer perceives from a light source and it is measured in lumens. Brightness is a subjective descriptor which difficult to measure.

E. Edge Detection:

The purpose of edge detection [8] in general is to significantly reduce the amount of data in an image while preserving the structural properties to be used for further image processing.

The Canny Edge Detection algorithm has five steps:

1. Smoothing:

Blurring of image to remove noise.

2. Finding Gradients:

The edges should be marked where the gradients of the image has large magnitudes.

3. Non-maximum suppression:

Only local maxima should be marked as edges.

4. Double Thresholding:

Potential edges are determined by Thresholding.

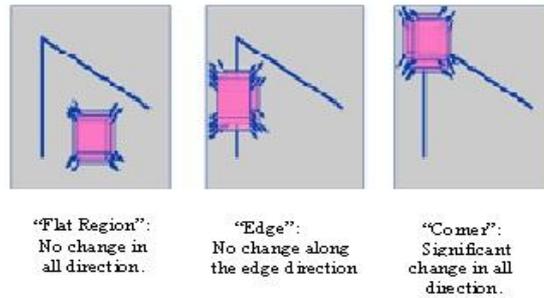
5. Edge tracking by hysteresis:

Final edges are determined by suppressing all the edges that are not connected to a strong edge.

F. Harris Corner Point Identification:

Harris Corner point detector [9] is a popular interest point detector. It has strong invariance to rotation, scale, illumination variation and noise. The Harris Corner detector is based on the local auto-correlation function. The local auto-correlation function measures the local changes with the patches shifted by a small amount in different directions.

Harris Corner Detector Basic Idea:



G. Similarity matching:

The match function will read two images. The local feature points and descriptors are found out for the two images. Distance ratio value is initialized. It is the value where the similarity between the two images has to be computed. A match is found out if the ratio of vector angles from the nearest to second nearest neighbor is less than distance ratio. For each descriptor in the first image, its match to second image is selected. For this, first it is required to transpose the descriptor matrix. Dot product of the matrix and its transpose is done. Inverse cosine is applied for the dot product matrix and it is sorted. When sorted, if any of the value is less than the distance ratio value, there is a match between the two images. If value is more than the distance ratio, there will be no match between the two images. The similarity between the two images is identified.

The pseudo code for finding the corner points is shown below. The different steps such as reading an image as an input, the preprocessing steps are shown in below. The detection of edge is done by using built in cranny’s edge detection algorithm. The number of corner points required can be specified by user. Finally the corner points can be detected by Harris corner detection algorithm.

```
%im = im2double(rgb2gray(imrgb));
g1 = fspecial('gaussian', 9,1);
g2 = fspecial('gaussian', 11,1.5);
G1 = g1;
G2 = g2;
img1 = conv2(im, g1, 'same');
Ix = conv2(img1, [-1 0 1], 'same');
Iy = conv2(img1, [-1;0;1], 'same');
Ix2 = conv2(Ix.*Ix, g2, 'same');
Iy2 = conv2(Iy.*Iy, g2, 'same');
IxIy = conv2(Ix.*Iy, g2, 'same');
R = (Ix2.*Iy2 - IxIy.*IxIy)
./ (Ix2 + Iy2 + eps);
R([1:20, end-21:end], :) = 0;
R(:, [1:20, end-21:end]) = 0;
nonmax = inline('max(x)');
Rmax = colfilt(R, [3 3], 'sliding', nonmax);
Rnm = R.*(R == Rmax);
[y, x, strength] = find(Rnm);
[yy ii] = sort(-strength);
nn = min(topn, length(ii));
y = y(ii(1:nn));
x = x(ii(1:nn));
```

The below figures shows the sample results to identify corner points. Fig 3.1 is the selection of image for identification of corner points. Fig 3.2 shows the identification of corner points by applying Harris method.

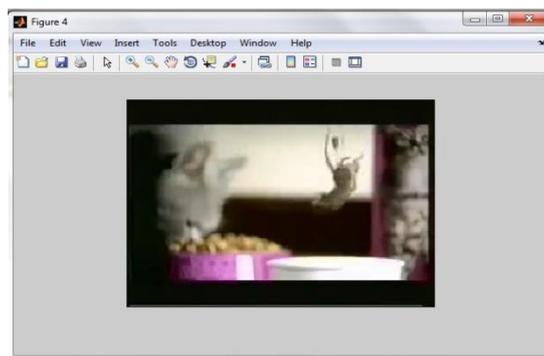


Fig 3.1

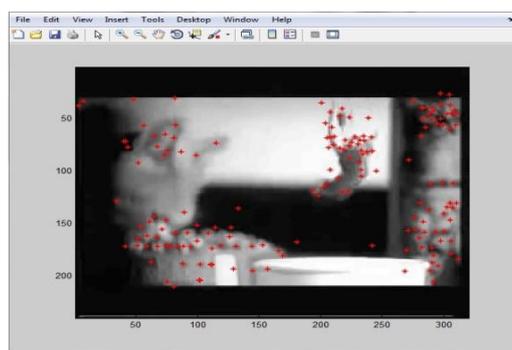


Fig 3.2

VI. CONCLUSION AND FUTURE WORK

In this paper, we are identifying the frequently appearing object in a video. This frequently appearing object is also called thematic object. Identifying this thematic object is helpful for object search and their tagging. It is also helpful for video summarization. Our approach will identify this frequently appearing object even in the presence of background clutters, partial occlusion. We are identifying the thematic object by considering the region of interest in the selected object. Then we are applying the Harris algorithm to select the corner points in the image. Here we are identifying only one thematic object at a time. Our future work is to identify multiple thematic object in a single frame.

ACKNOWLEDGMENT

Special thanks to East West Institute of Technology for helping us in our work and supporting the research. Visveshvaraya Technological research centre are also helping us in this work.

REFERENCES

- [1] Junsong Yuan, Gangqiang Zhao, Yun Fu, Zhu Li, Aggelos K. Katsaggelos and Ying Wu, "Discovering Thematic Objects in Image Collections and Videos" IEEE Transactions on Image Processing vol.21, No. 4, April 2012.
- [2] Davide Liu, Tshun Chen: "DISCOV :A Framework for Discovering Objects in Video", IEEE Transactions on Multimedia, Vol. 10,No.2, February 2008.
- [3] David Liu and Tshun Chen: "A Topic-Motion Model for Unsupervised Video Object Discovery".
- [4] Alok K. Watve: "Object tracking in video scenes".
- [5] Zhengwei Yang and Fernand S. Cohen, "Image Registration and Object Recognition Using Affine Invariants and Convex Hulls", IEEE Transactions on Image Processing, Vol 8, No. 7, July 1999.
- [6] Ying Shan, Harpreet S. Sawhney, Art Pope: "Measuring the Similarity Of Two Image Sequences".
- [7] David Liu, Gang Hua, Tsuhan Chen: " A Hierarchical Visual Model for Video Object Summerization".IEEE Transactions on Pattern analysis and Machine Intelligence vol 32 December 2010.
- [8] John Canny: "A Computational approach to Edge Detection". Pattern Analysis and Machine Intelligence, IEEE Transactions on PAMI Nov 1986.
- [9] Harris, C., Stephens : "A Combined Corner and Edge Detector", Proceedings of Alvey Vision Conference, 1988.
- [10] Hung Khoon Tan and chongwah Ngo " Localized matching using earth mover's distance towards discovery of common patterns from small image samples: