ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

Moving Object Extraction Based on Background Reconstruction

Prajakta A Patil[1], Prachi A Deshpande[2]

Lecturer, Dept. of E&TC., JSPM's BSOITR, Pune, Maharastra, India[1]

Lecturer, Dept. of E&TC., JSPM's BSOITR, Pune, Maharastra, India[2]

ABSTRACT:- Object extraction from video is one of the most important areas of video processing in which objects from video sequences are extracted and used for many applications such as surveillance systems. This thesis describes the theoretical bases, development and testing of moving object detection framework. Many systems uses motion and color information to detect the moving object from the video sequences, but these methods cannot produce robust segmentation results. We propose an algorithm for automatically detecting and segmenting a moving object from single concept videos, i.e. videos which have only one object category of interest but may have multiple object instances with pose, scale, etc. variations. Given such a video, we first reconstruct the background image from the input frames. From that background image we can construct motion cues based on motion information and colour cues based on color information present in frames. Then combine that both motion and color cues in MRF framework to get objet which is separated from the background.

KEYWORDS : color cues, motion cues, MRF Framework.

I. INTRODUCTION

Object extraction from video is one of the most widely researched areas. Under this topic we present an unsupervised algorithm for automatically detecting and segmenting a moving object from a monocular video. Detecting and segmenting a moving object from a video with limited object motion is challenging. Since existing automatic algorithms rely on motion to detect the moving object, they cannot work well when the object motion is sparse and insufficient [1]. In our current system we present an unsupervised algorithm to learn object color and locality cues from the sparse motion information. We first detect key frames with reliable motion cues and then estimate moving sub-objects based on these motion cues using a Markov Random Field (MRF) framework. From these sub-objects, we learn an appearance model as a color Gaussian Mixture Model. To avoid the false classification of background pixels with similar color to the moving objects, the locations of these sub objects are propagated to neighbouring frames as motion cues. Finally, robust moving object segmentation is achieved by combining these learned color and motion cues with in a MRF framework. We first create a empty mask of frame by using background reconstruction method. And estimate it's color cues and motion cues on the base of object color and motion information. Finally we combined that color mask with motion mask we get the object extracted from the background.

In order to decrease the cost of systems and make straightforward the algorithms, much explore effort has been dedicated to extract significant features from time-series inertial signals. To name a few, Lim *et al.* computed relationship coefficients of the absolute value of acceleration and the absolute value of the first and second derivatives of acceleration to form feature vectors.

II. REVIEW ON LITERATURE

The search for the relevant information in the large space of image and video databases has become more challenging. An image can be represented as a set of low level visual features such as color, texture and shape features. While several image extraction systems rely on only one feature for the extraction of relevant images, it has been shown that an appropriate combination of relevant features can yield better retrieval performance. Nevatia et al. and Davis et al. both decomposed an object shape model in a hierarchical way to train object part detectors [8]. These detectors are used to describe all possible configurations of the object of interest (e.g. pedestrians). Gorelick and Basri collected a set of object silhouette exemplars. To extract the object of interest, the authors over-segmented the input image and determined the segments which best matched the associated templates. To deal with multiple human instances with large pose deformations, Niebles et al. applied a human body detector on each frame, and their detection results were refined by pose density estimation function and probability diffusion between adjacent frames.

- Video Object Segmentation Using Background Subtraction

The [10] idea of background subtraction is to subtract the current image from the still background, which is acquired before the objects move in. After subtraction, only non-stationary or new objects are left. This method is especially suitable for video conferencing and surveillance applications, where the backgrounds remain still during the conference or the monitoring time. Nevertheless, there are still many annoying factors such as similar color appearing in both foreground and background areas, changing of lighting condition, and noise which have prevented us from using a simple difference and threshold method to automatically segment the video objects.

III. PROPOSED SYSTEM

In this work an approach to motion segmentation based on optimization on a Background reconstruction is proposed. This approach integrates motion estimation, the fusion of multiple visual cues and segmentation in one single optimization process. For the general purpose of video processing, the back ground is usually considered as the scene without the presence of objects of interest, such as human objects or moving vehicles. Background is usually composed of non-living objects that remain passively in the scene. In a video about a general environment, the background can consist of both stationary and moving objects. Our proposed system depends on the object motion as well as color information. We extact motion cues and color cues of the object from the video sequences. And finally combine that both the cues in the MRF model to produce the robust segmentation result. In our system we overcame all the above mention problems.

- Video Segmentation

Videos are actually sequences of images, each of which called a frame, displayed in fast enough frequency so that human eyes can percept the continuity of its content. It is obvious that all image processing techniques can be applied to individual frames. Besides, the contents of two consecutive frames are usually closely related. Video segmentation refers to the identification of regions in a frame of video that are homogeneous in some sense [11].

- Object Detection and Tracking

Object detection in videos involves verifying the presence of an object in image sequences and possibly locating it precisely for recognition. Object tracking is to monitor object spatial and temporal changes during a video sequence, including its presence, position, size, shape, etc. This is done by solving the temporal correspondence problem, the problem of matching the target region in successive frames of a sequence of images taken at closely-spaced time intervals. These two processes are closely related because tracking usually starts with detecting objects, while detecting an object repeatedly in subsequent image sequence is often necessary to help and verify tracking.

- System Block diagram

The system diagram of our object detection method is shown in Figure1. Our method depends on a six stage process to extract objects with their features in video imagery. Here we consider moving objects in a video as some compact regions with different apparent motion from the background. To differentiate between the background and moving object, If a pixel/region has significant different apparent motion from the background, it mostly likely belongs to a moving object, otherwise it will be a pixel from background. We can define the motion cues as the difference between the pixel motion and the background motion.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

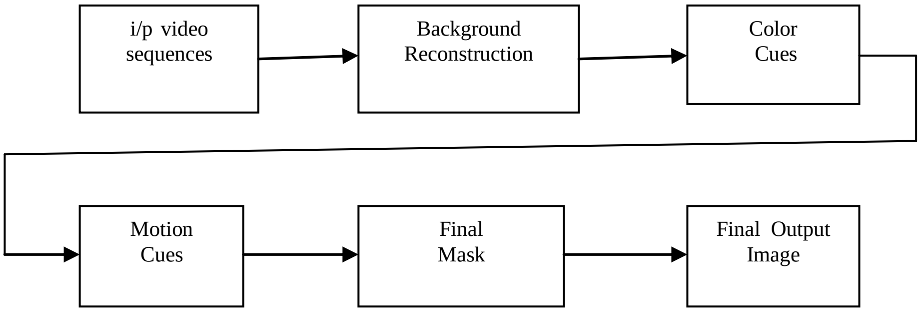

Figure1: Video Object Segmentation System

We focus on VOE in single concept videos captured by a monocular camera in static or arbitrary types of background. Instead of assuming that the background motion is consistently dominant and different from that of the foreground we relax this assumption and allow foreground objects to be present in scenes which have marginal but complex background motion (e.g., motion induced by sea waves, swaying trees, etc.). We also ignore the video frames with significant motion variations due to shot changes or abrupt camera movements[11]. To make our method robust and not require any user interaction, we start from multiple local motion cues, and integrate the induced shape and color models to get final output image.

- I/P Video Sequences

A sample video with a stationery background and a single object moving is considered. This video is converted into frames using software. And the operation is performed on each frame separately. These frames are provided in sequence as input to the program.

- Background Reconstruction

We adopted the assumption that background is most often visible in the image sequence, in other words, background pixels appear in the image sequence with the maximum frequency [5]. Assume that there is no background in the video. Background will appear with the concept of foreground object. With this assumption we first construct the background in video sequences. Because by using this method we will get robust object segmentation result. For ex. if there is small motion in object, after subtracting current image form next or previous image some part of the moving object is also subtracted and we cannot get whole moving object.

Each frame in the video sequences formed by RGB color component. Separate out the entire RGB component from each frame. Create a separate plane of all R, G, B component for every image. Then we will find out the median value of the pixel component in from all R plane, similarly from all G plane and all B plane [10]. Consider there are ‘n’ numbers of R plane, and then we have to construct the median R plane, for the we will find out the median value of all pixel on the same position in each frame. Consider the position A, then the median value of pixel at position A from all R plane can be calculated as-

$$\text{Median (R)} = \frac{\sum_{i=i}^{n} A(1,i)}{n}$$

Similarly, median for G plane and B plane can be calculated as-

$$\text{Median (G)} = \frac{\sum_{i=i}^{n} A(1,i)}{n}$$

$$\text{Median (B)} = \frac{\sum_{i=i}^{n} A(1,i)}{n}$$

Above equations help to construct median of R, G, B planes. By combing these median planes we will get background image reconstructed.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

- Color Cues

Color is low level cue for object detection that can be implemented in computationally fast and effective way for locating object. It also offers robustness against geometrical changes under a stable and uniform illumination field. Some cases it clearly discriminates object from background. Background reconstruction is the first part of this thesis after that we have to prepare the color mask. To get the color mask the current RGB frame can be converted into the grey scale [16]. After converting the RGB image into gray levels we consider the pixel values in the form 0 and 1. And calculate the mean value of the current frame. Consider there are 'n' numbers of pixel in the frame and 'x' represents any pixel, then mean value of 'x' can be calculated as-

$$\text{Mean (x)} = \frac{\sum_{i=1}^{n} xi}{n}$$

Mean value of pixel component in current frame can be calculated by using above formula. Here, we choose some threshold value 'T' for intensity variation. Based on this threshold value there is possibility of variance to be occurred. We will select that variable value up to 2 only. Based on these variable values of pixels, standard deviation can be calculated. Then we will find out the the value by subtracting every pixel value from the mean value. From this difference value we can create an empty mask, we can called it as color mask. There should be possibility of different values of different pixels. If the pixel value is >2* standard deviation value, we will keep that value as it is, otherwise we will set that value to zero. Now our final color mask is ready to use. And we can apply that mask on each image to get the color cues.

- Motion Cues

For the motion mask to calculated three consecutive frames takes into account i.e. current frame, previous frame and the next frame. Then for motion mask difference between each to frame could be estimated. First, the current image and the background will take into consideration and estimate the difference between these two frames we can called it as MM1. Then the current image and the past image will take into consideration and estimate the difference between these two frames we can called it as MM2. Then the current image and the next image will take into consideration and estimate the difference between these two frames we can called it as MM3. Then motion mask 1 is the mask estimated from the ANDing operation between the current frame and background frame mask (MM1) with the current frame and the past frame mask (MM2).

M1=and(MM1,MM2)

Then motion mask2 is the mask estimated from the ANDing operation between the current frame and background frame mask (MM1) with the current frame and the next frame mask (MM2).

M2=and(MM1,MM3)

By using and operation all zero or negative value can be neglected from the equation. Now the mask contains only positive values. Final motion mask can be estimated by performing ORing operation with the above two masks i.e. mask M1 and Mask M2.

MMask=or(M1,M2)

Discussion uptil now we got the two masks color mask as well as motion mask.

- Final Mask

Final mask can be obtained by adding above two masks. So we can say that our final mask can be estimated by ORing motion mask and color mask.

Mask=or(CCmask,MMask)

In the above expression CCmask Indicates the color mask while the MMmask Indicates the motion mask. And at this case we get our object which are extracted from the video sequences as well as background image.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

IV. SYSTEM PERFORMANCE

Consider a home-made video of a moving pen held. In this video pen is moving slowly but the background image is constant. The position of pen between two consecutive frames differs by their position. Here the length of this video is 2 sec and the size 804×566 is considered, because if we take large length of video, the number of frames also increases and it will take more time for it's execution. So the result of our method for this video is shown below.

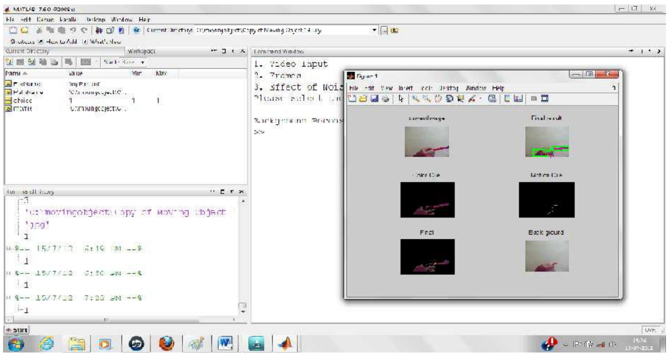

Figure 2: Matlab window appear on the screen

After executing program and giving video input the following type of matlab window will appear on the screen of the computer mention in figure 2.

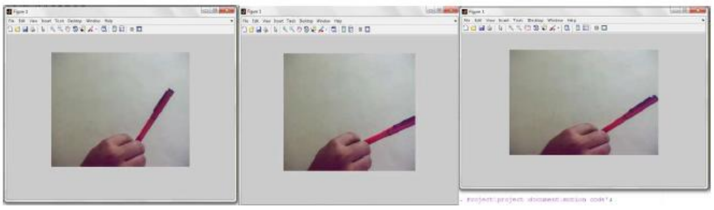

Figure 3: Frame no.4, frame no.16 frame no. 20.

The input video sequence is first divided into number of frames out of that total number of frame no.4,16and 20 indicated in figure 3.

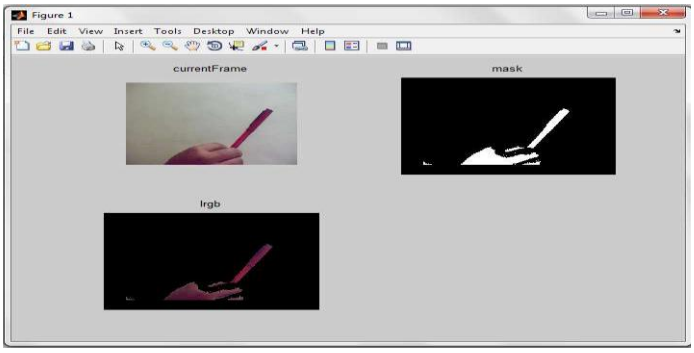

Figure 4: Extracted mask

For calculating the motion cue, first the motion mask is estimated and the RGB component of the image separate out shown in the figure 4.

.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

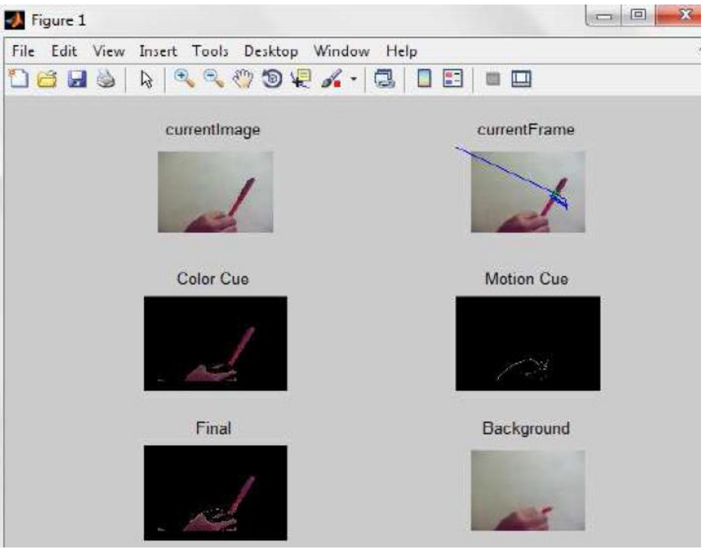

Figure 5: Final output of video of moving pen

Final extracted output with motion and color mask and reconstructed background is shown in following figure 5. Also it tracks the motion of the object. Here the moving object can be extracted from its background.

V. CONCLUSION

Our algorithm learns the moving object cues from key frames, which are extracted based on motion cues. When some parts of the object never move, the object cues may not capture their characteristics. In practice, our algorithm can identify most of these parts as moving objects because it incorporates the interaction between neighbouring pixels with MRF Prior. But it is still possible our algorithm will miss some of these parts if they are far from other identified moving regions both in color and space. This is a fundamental problem of automatic object segmentation methods (including ours) that only rely on the video itself to infer the object of interest. Incorporating external information can help solving this problem.

REFERENCES

1. Bo Wu, Ram Nevatia, and Yuan Li, “Segmentation of Multiple, Partially Occluded Objects by Grouping, Merging, Assigning Part Detection Responses” University of Southern California Institute for Robotics and Intelligent Systems Los Angeles, CA 90089-0273.
2. D. Zhong and S.-F. Chang,” Video Object Model and Segmentation for Content-Based Video Indexing”, To appear in IEEE Intern. Conf. on Circuits and Systems, June, 1997, Hong Kong.
3. Ediz S_aykoly Ugur Gudukbay and Ozgur Ulusoy, “A Semi_Automatic Object Extraction Tool for Querying in Multimedia Databases” Department of Computer Engineering Bilkent University Bilkent Ankara Turkey.
4. Eric Galmar ,”Representation and Analysis of Video Content for Automatic Object Extraction” EURECOM Department of MultiMedia Communications 2229, route des Crêtes B.P. 193 06904 Sophia-Antipolis FRANCE.
5. Fardin Akhlaghian Tab, Golshah Naghdy and Alfred Mertin, “WELL DEFINED VIDEO OBJECT EXTRACTION SUITABLE FOR SCALABLE WAVELET BASED OBJECT CODING”.
6. Jinhui Pan, Shipeng Li, Ya-Qin Zhang, “AUTOMATIC EXTRACTION OF MOVING OBJECTS USING MULTIPLE FEATURES AND MULTIPLE FRAMES”, Tsinghua University Microsoft Research, China.
7. Kuo-Chin Lien and Yu-Chiang Frank Wang, “AUTOMATIC OBJECT EXTRACTION IN SINGLE-CONCEPT VIDEOS”. National Science Council of Taiwan via NSC 99-2221-E-001- 020 and NSC 100-2631-H-001-013.
8. Liang Yu1, Ya guang kong, “An Algorithm of Background Reconstruction Based on Morphology and PIC” Proceedings of the 2009 International Workshop on Information Security and Application (IWISA 2009).
9. Ninad Thakoor ,Jean Gao, “ Automatic Extraction and Localization of Multiple Moving Objects with Stereo Camera in Motion”.
10. Nir Friedman, Stuart Russell, “Image Segmentation in Video Sequences: A Probabilistic Approach”, Computer Science Division University of California, Berkeley, CA 94720-1776.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

11. Pei Yin, Student Member, IEEE, Antonio Criminisi, John Winn, Member, IEEE, and Irfan Essa, Senior Member, IEEE, "Bilayer Segmentation of Webcam Videos Using Treebased Classifiers" IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VO0. NO. 0, JANUARY 0000.
12. Petrus Johannes Withagen," Object detection and segmentation for visual surveillance".
13. Ross Kindermann, J. Laurie Snell, Contemporary Mathematics , "Markov Random Field and Their Applications".
14. Surya Sumpeno, Hariadi Mochamad and Takafumi Aoki, "A Standard Performance Evaluation Tool for Video Object Extraction", Graduate School of Information Sciences, Tohoku University Aoba-yama 05, Sendai 980-8579 Japan.
15. Thomas Meier, Member, IEEE, and King N. Ngan, Senior Member, IEEE , "Video Segmentation for Content-Based Coding" IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 9, NO. 8, DECEMBER 1999.
16. Thomas Meier and King N. Ngan, Senior Member, IEEE," Automatic Segmentation of Moving Objects for Video Object Plane Generation", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 8, NO. 5, SEPTEMBER1998.