

Color Detection and Tracking from Live Stream – An Extensive Survey

Sagar Pandey

Student, Department of Information Technology
Institute of Engineering & Management,
Kolkata, West Bengal, India

Subhabrata Sengupta

Faculty, Department of Information Technology
Institute of Engineering & Management,
Kolkata, West Bengal, India

ABSTRACT

Segmentation refers to a technique of picking out a particular part of any object. If we talk about image segmentation it is basically a method of picking out a particular area of that image. As we know a video is nothing but a sequence of frames in an orderly manner, hence video segmentation could also be achieved in this manner. Video segmentation is in great demand now a day with the advancement of AI. Few important applications of video segmentation could be a face, colour detection system, an object tracking system etc. Detection and tracking of any moving object is a very important part of video surveillance and hence a lot of emphasis is done on it. Lots of researches had been taking place in the domain of object detection and tracking from live stream and various algorithms were suggested to obtain the most effective results. In this survey, we would cover all important researches and works done on object detection and tracking especially the color as an object part and will also discuss the benefits of doing this. In short we will make out how video segmentation is going to open the gate for a great and smart future ahead.

Keywords

Image segmentation, video segmentation, skin detection, edge & corner detection. Color detection and tracking.

1. INTRODUCTION

With the advancements of artificial intelligence, there has been a lot of focus on video segmentation and tracking as it opens the door for future. A lot of research has already been done to accomplish segmentation from any live stream and also many are still in progress. Color segmentation is another important part of object segmentation process. This is a process through which colour could be detected from a live stream.

The scope of detecting and tracking object especially moving object is quite large in the field of computer vision. Various purposes like video surveillance, live conferencing etc have a lot to do with object detection and tracking. High resolution frames at high frame rates are the major challenges faced in order to track object

There are a number of color models which help us in understanding the object in a better way, some of them are explained below:

1.1 The RGB Color Model

The RGB (Red, Green, and Blue) color model uses a Cartesian coordinate system and form a unit cube shown in Fig. 1.

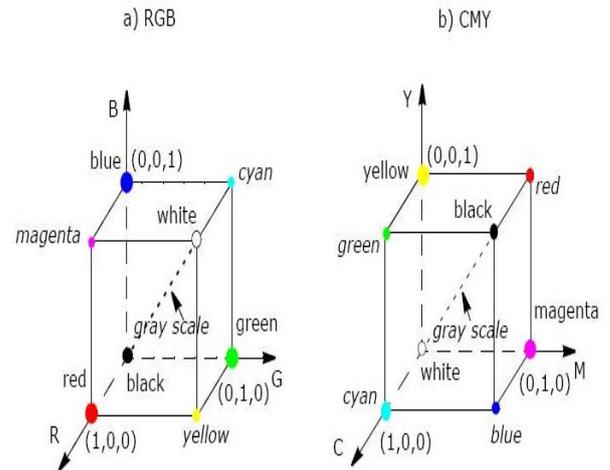


Fig. 1: The RGB Color Model

The RGB color model consists of the three primary colors – red, green and blue. The main idea behind this model is that combining these three primary colours in various proportions, various other colours could be produced. As the RGB color model is hardware oriented it is used in many image capturing, processing and rendering devices.

1.2 The HSV Color Model

The HSV in the HSV color model stands for Hue, Saturation, and Value. It was introduced by A. Smith and is basically a user governed color model. It is based on the spontaneous request of the user's tint, shade, and tone. The region within which the HSV color model is defined is a hex cone. Figure 2 shows how a hex cone looks like:

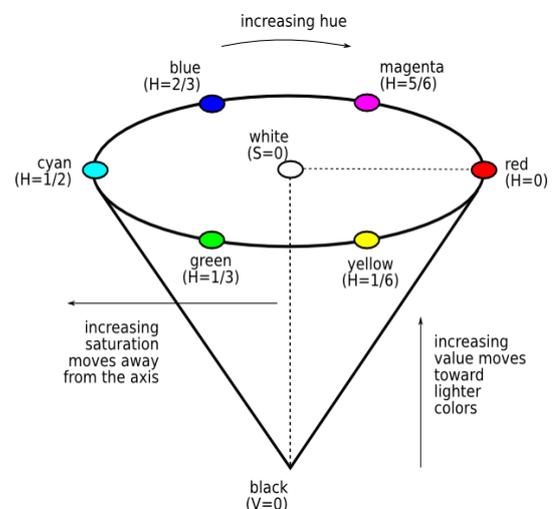


Fig. 2: The HSV Color Model

1.3 The HLS Color Model

The HLS in the HLS color model stands for Hue, Lightness, and Saturation. This color model was introduced by W. Ostwald.

The definition of Hue is the same as in the HSV model. In the HSL color model black occupies the bottom tip of hex cone while white occupies the top tip of hex cone as shown in the figure. This is actually the measure of lightness. The value of lightness is 0 in case of black and 1 in case of white. In fact, HLS is very much similar to HSV with some deformities, where white is placed on top of the top cone. Saturation is radical measurement from the vertical gray axis, ranging from 0 on the vertical gray axis to 1 on the surface of the hex cone. We can find the maximum saturated hues at $S = 1$ and $L = 0.5$.

Hence with the help of various color models we can easily get details of any image like the colors used, hue level, saturation level etc. just like the images the details from a live stream video can also be detected as videos are nothing but sequences of images known as frames. Once the details of an image are obtained the segmentation could be achieved and we can detect and track color from a given live stream.

It appears like a double hex cone region as if the HSV model with its own image above it as shown in Fig. 3.

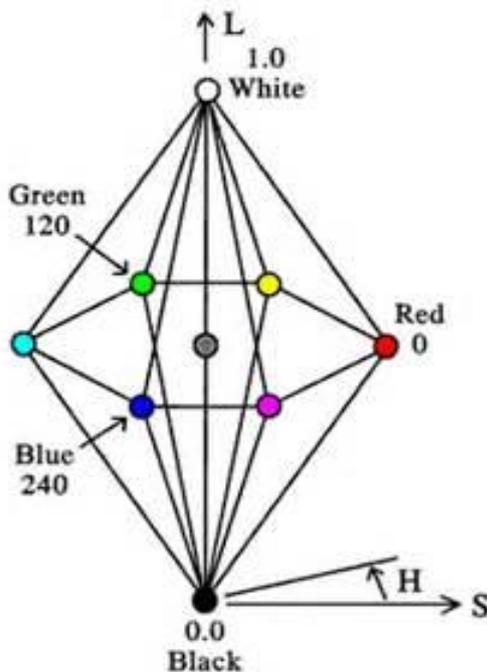


Fig. 3: The HLS Color Model

2. RELATED WORK

Kalisa Wilson et al in her paper on “Real- Time Tracking for Multiple Objects Based on Implementation of RGB Color Space in Video” proposed an algorithm to track multiple color objects in real time. [1] The framework suggested by her was basically based on using the steps of image processing like the thresholding. Using her algorithm in the real-time video first pre-processing was done and then the color discrimination part was achieved and finally segmentation and tracking part.

Ssu-Wei Chen et al in his paper on “Moving Object tracking Based on Background Subtraction Combined Temporal Difference” proposed a new tracking method that uses Three Temporal Difference(TTD) and the Gaussian Mixture

Model(GMM) approach for object tracking. [2] The three temporal differences i.e. TTD method suggested by them make use of continuous image subtraction.

Vandana S. Bhatia et al in his paper on “Face detection system using HSV color model and morphing operations” proposed an algorithm which uses the fact that the human faces have unique color features. [3] Using skin color to detect face regions is a much faster approach. First the image was read in the RGB format and then converted to the HSV model. Then the image gets converted to the binary format and finally segmentation was done and face gets detected.

Ravikant Gupta et al in his paper on” Human Face Detection By YCbCrHs Technique” proposed a feature based algorithm to detect faces. [4] The suggested algorithm used spatial filters to detect feature points from the image and then does the grouping part using various constraints at geometric and gray level. A probabilistic framework was then used to identify if the features extracted matches to a particular face or not.

Prasad Kalane et al in his paper on ” Target Tracking Using Kalman Filter” used kalman filter technique to detect any target. [5] A target can be anything, maybe a human or an umbrella or anything. He added that there are two key steps in target tracking:

1. Detecting the moving objects on which we are interested
2. Keeping the track of the interested objects from frame to frame.

Isaac Cohen et al in his paper on ” Detecting and Tracking Moving Objects for Video Surveillance” proposed to choose an affine model which approximates well the general perspective projection, while having a low numerical complexity. [6]

R E Kalman et al states that objects can be represented in three ways: Point tracking, Kernel tracking and Silhouette Tracking. Point tracking can be achieved using various Deterministic and statistical methods available. Kernel tracking makes the use of various Templates, multi-view appearance models and density based appearance models. Silhouette Tracking uses methods like matching the shapes method and contour evolution methods.

Table1. Various Tracking Methods

| Type of tracking | Methodology | Advantages | Limitations |
|------------------|-------------|---|---|
| Point tracking | MGE tracker | Can handle entries | Costly |
| Point tracking | GOA tracker | Handles misdetection | Assume no object entries and exists |
| Kernel Tracking | Mean shift | 1. Mean shift is an application independent tool suitable for real data analysis. 2. Does not assume any predefined | 1. Non trivial selection of window size. 2. “Shallow” modes can arise due to inappropriate window |

| | | | |
|---------------------|--------------------|---|--|
| | | shape on data clusters. | sizes. |
| Kernel method | Layering method | Suited for multiple object detection | High cost |
| Kernel method | KLT Method | It is faster than traditional techniques for examining far fewer potential matches between the images | Efficient but don't satisfy today AI applications |
| Silhouette tracking | State space models | Can handle complex models for rigid and non-rigid objects | Entry and exit of objects are difficult to handle, gestures recognition is not so accurate |

Andres Alarcon Ramirez et .al in his paper on “A New Algorithm for Tracking Objects in Videos of Cluttered Scenes” proposed an algorithm that constituted of two steps: the first step is estimating the direction of the object. [7] Once the direction of the interested object is estimated, we can know where the object is going to lie in the next frame and the process continues. Hence with this approach, we can detect the frames and thereby the motion and changing directions of the interested object and hence our interested object could be detected and tracked.

Chris Harris et. al in his paper on “A COMBINED CORNER AND EDGE DETECTOR” proposed a technique to detect edges and corners in an image, junction would then be considered as points where various edges meets. [8] To pursue the above approach he started from the **Moravec's corner detector**

Moravec's corner detector considers a local window in the image, and then determined the average changes in the intensity of the image that resulted due to shifting of the window by a little amount in various directions. Three cases needed to be considered:

1. If the image that was windowed has almost constant intensity, then all shifts will result in only a small change;
2. If the window stands or sit an edge, then any shift along the edge will result in a small change, whereas any shift perpendicularly to the edge will result in a large change;
3. If the windowed patch is either a corner or an isolated point, then all shifts will result in a large change. A corner can thus be detected easily by finding when the shift is large even for any minimum change.

Katja Nummiaro et. al proposed a colour-based object tracking in multi-camera environments [9] . In this the object was placed in an environment having multiple cameras. To

support multiple cameras they used more than one histogram for a target. The cameras were placed at different angles in order to extract the features of the object from various angles. Since all the cameras keeps a track of the object independently there must be a correspondence between the histograms

Marcus Thaler et .al proposed a real time person detection and tracking in panoramic video method. [10] This is a useful technique for detecting any player in particular sports. As we know in any live event of a sport there are a number of cameras trying to take the best scene of the play, this technique can be used to detect any player.

Wen-cheng Wang et .al proposed a face detection method to be used in color images. [11] He analysed four kinds of color space: RGB, HIS, YUV, and YCbCr. The main aim was to study the clustering features of the skin color in different color spaces and establish a practical and reliable color mixing model, and finally, use YCbCr color space for skin color detection, screen the candidate regions by integrating the geometric characteristics of the face, and then remove the similar skin region and other regions of body exposed, and continue the process to narrow search scope.

Hani k. Mohair et .al proposed a method to detect human skin color. [12] In this method first the original image was transferred to color space and then skin detection was done. Once the skin areas were detected, after dilation and erosion, face was detected.

Monika Deswal et. al in her paper on” A Fast HSV Image Color and Texture Detection and Image Conversion Algorithm” proposed an algorithm to detect the texture of an image. [13] First the RGB image was converted to the HSV image by removing noise and a corresponding database of HSV values was prepared. Then the level of H, S, and V were found and was normalized and finally the texture was calculated.

Douglas Chai et.al in his paper on” Face Segmentation Using Skin-Color Map in Videophone Applications” proposed a method of face segmentation using the skin-color map. [14] The approach of his algorithm is as follow: first the head and shoulder image was taken as input and then color segmentation followed by density regularization, luminaries' regularization, geometric correction, contour extraction was done and finally segmented facial region was obtained.

Umapathy Eaganathan et.al explained about different color models such as RGB, YCbCr, HSV, HSI, CIELab in medical image processing and it additionally explored how these color models can be derived from the basic color model RGB. He did an analysis of various colour models in medical image processing.

S. Chitra et.al did a comparative study for two color spaces HSCBCr and YCbCr in skin color detection. [15] The paper concentrated on how the input images get converted to the HSV and YCbCr model and how the values of H, S, Cb, and Cr are collected and then to verify if these values satisfies the threshold values. If the pixels lie in the threshold range then that pixels will be considered as skin region otherwise it is a non skin region. Thus using this skin and non skin area could be detected easily. The algorithm concluded that HSCbCr gave better results than YCbCr model.

Intaek Kim et.al proposed a multi-target tracking using color information proposed a method that tracks multiple objects in

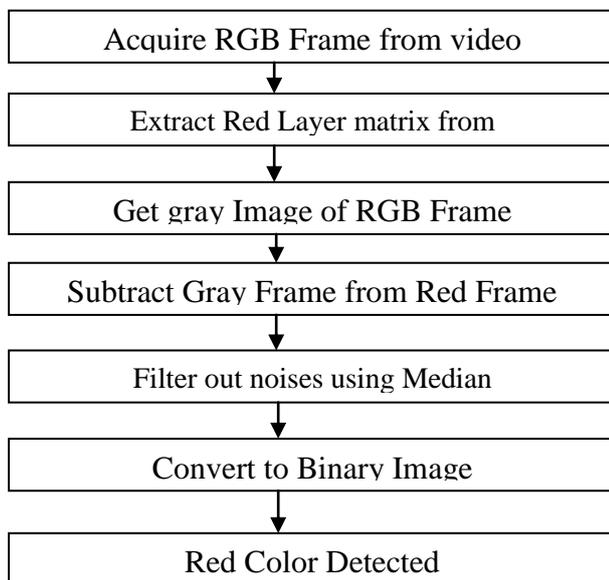
a video applying extended Kalman filter and color information. [16]

Akriti Kaushik et.al proposed a RGB color sensing technique. [17] They used sensor to detect RGB color .color sensing is of great use in the field of medical science. The main idea behind this approach was that when the color object is placed before the sensor the color gets detected and the LED of the same color get on; it uses 8 colour LEDs that include three primary colors- red, green, blue and also include magenta, yellow and cyan as well as black and white

Hasan Fleyeh et.al proposed a road and traffic detection sign color detection and segmentation technique [18]. The color segmentation algorithm is carried out by placing a RGB camera on the moving car. The camera captures images from the road side and then converts it to the HSV color model as the HSV color model deals with the hue condition and hue is invariant under any type of light condition be it dark or bright.

3. PROPOSED ALGORITHM

The following Flow chart shows how the Color detection and tracking algorithm is processed:



Algorithmic Steps:

Step 1: To acquire RGB framework from a live stream, once done the red color matrix gets extracted from the RGB framework.

Step 2: After this the gray image of RGB framework is obtained and the gray Frame is subtracted from the Red Frame.

Step 3: All the noises are filtered out using the median filter.

Step 4: The diff Frame is then converted to corresponding binary image with roper threshold value.

Step 5: Finally the red color is detected and outline is made in order to identify it from the live stream

4. CONCLUSION & FUTURE WORK

With the steps shown in the above section we can easily detect and track color using tools like MATLAB. The above algorithm is simple and easy to implement with high percentage of accuracy

As of now we can easily conclude that color detection and tracking is one of the key steps towards the future that will be dominated by artificial intelligence. Color detection and tracking has many roles in various fields which also include the medical science.

A lot of researches have already been conducted and many are still going on in the field of object tracking which is a key in the field of artificial intelligence. Furthermore there is a lot of work still left to be done so that the scope of object tracking and its uses could be broadened.

5. ACKNOWLEDGEMENT

We are indebted to the experts who have contributed towards development of the template. The authors would like to acknowledge to Institute of Engineering & Management, Kolkata, West Bengal, India.

6. REFERENCES

- [1] Kalisa Wilson, Real- Time Tracking for Multiple Objects Based on Implementation of RGB Color Space in Video, International Journal of Signal Processing, Image Processing and Pattern Recognition Vol.9, No.4, (2016), pp.331-338
- [2] Ssu-Wei Chen, Luke K. Wang, Jen-Hong Lan, Moving Object tracking Based on Background Subtraction Combined Temporal Difference, International Conference on Emerging Trends in Computer and Image Processing (ICETCIP'2011) Bangkok Dec., 2011
- [3] Vandana S. Bhata* and Jagadeesh D. Pujaria, Face detection system using HSV color model and morphing operations, International Journal of Current Engineering and Technology
- [4] Ravikant Gupta, Satyaprakash Pandey Yogesh Tayal, Pramod Kumar Pandey, D.V.B Singh, Human Face Detection By YCbCrHs Technique, International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS)
- [5] Prasad Kalane, Target Tracking Using Kalman Filter, International Journal of Science & Technology, Vol. 2 Issue 2, April 2012d
- [6] Isaac Cohen G'erard Medioni, Detecting and Tracking Moving Objects for Video Surveillance, IEEE Proc. Computer Vision and Pattern Recognition Jun. 23-25, 1999. Fort Collins CO
- [7] Andres Alarcon Ramirez and Mohamed Chouikha, A New Algorithm for Tracking Objects in Videos of Cluttered Scenes, International Journal of Information Technology, Modeling and Computing (IJITMC) Vol.1, No.2, May 2013
- [8] Chris Harris & Mike Stephens, A COMBINED CORNER AND EDGE DETECTOR, Plessey Research Roke Manor, United Kingdom © The Plessey Company pic. 1988
- [9] Katja Nummiaro, Esther Koller-Meier, Tom' a' s Svoboda, Daniel Roth, and Luc Van Gool, Color-Based Object Tracking in Multi-Camera Environments, In Proceedings of the DAGM'03, Springer LNCS 2781, pp. 591-599, Sep 2003
- [10] Marcus Thaler, Werner Bailer, Real-time Person Detection and Tracking in Panoramic Video, JOANNEUM RESEARCH, DIGITAL – Institute for

Information and Communication Technologies
Steyrergasse 17, 8010 Graz, Austria

- [11] Wen-cheng Wang, A Face Detection Method Used for Color Images, *International Journal of Signal Processing, Image Processing and Pattern Recognition* Vol. 8, No. 2 (2015), pp. 257-266
- [12] Hani K. Al-Mohair, Junita Mohamad-Saleh and Shahrel Azmin Suandi, Human skin color detection :A review on neural network perspective, *International Journal of Innovative Computing, Information of Control* Volume 8, Number 12, December 2012.
- [13] Monika Deswal, Neetu Sharma, A Fast HSV Image Color and Texture Detection and Image Conversion Algorithm, *International Journal of Science and Research (IJSR)*
- [14] Douglas Chai, Face Segmentation Using Skin-Color Map in Videophone Applications, *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY*, VOL. 9, NO. 4, JUNE 1999
- [15] S. Chitra, Comparative Study for Two Color Spaces HSCbCr and YCbCr in Skin Color Detection, *Applied Mathematical Sciences*, Vol. 6, 2012, no. 85, 4229 - 4238
- [16] Intaek Kim, Malik Muhammad Khan, Tayyab Wahab Awan, and Youngsung Soh, Multi-Target Tracking Using Color Information, *International Journal of Computer and Communication Engineering*, Vol. 3, No. 1, January 2014
- [17] Akriti Kaushik1, Aastha Sharama, RGB COLOR SENSING TECHNIQUE, *International Journal of Advance Research in Science and Engineering*
- [18] Hasan Fleyeh, Road and Traffic Sign Color Detection and Segmentation A Fuzzy Approach, *MVA2005 IAPR Conference on Machine Vision Applications*, May 16-18, 2005 Tsukuba Science City, Japan