EFFECTIVE VIDEO SALIENCY DETECTION FOR BACK GROUD AND FOREGROUND SEPERATION OF IMAGES

Hema Rajini N
Department of Computer Science and Engineering, Alagappa Chettiar Government College of Engineering and Technology, Karaikudi – 630003, Tamilnadu, India

Abstract: **-The issue of enumerating the dependability of computational saliency in a video, which can be utilized to propel the processing algorithms used for saliency-based video calculation and unwavering quality of the procedure are tended to in this module. Weight based spatio-temporal approach is used to enumerate such reliability in twofold: 1) investigating the spatial connections in both the saliency graph and the eye-fixation map.2) taking in the weight-average based spatiotemporal relationships, which characterize a dependable saliency outline. First examination on spatio-temporal eye fixation information from the general public CRCNS dataset and examination of a typical element in human visual consideration is done, which ascertains the relationship in saliency inside a pixel and its immediate neighbors.This algorithm estimates a pixel wise uncertainty map that replicates the assurance in the associated computational saliency map by relating it to the pixel's saliency of its direct neighbor.The difference of a pixel from its nearby neighborhood is estimated in the saliency graph to gauge such vulnerabilities. Then we add a weight value to each pixel, thus producing more reliable output. The objective of this module is to produce more reliable saliency map. The simulation tools used for this research modules is MATLAB.**

Keyword: saliency map, spatio-temporal features, superpixel, adaptive histogram equalization

1. INTRODUCTION

Saliency detection is usuallytermed as the automatic estimation of significant (salient) objects and regions of images without any prior guess or knowledge. Saliency usually is defined as the difference between a pixel and its neighborhood.

In [3] direction based methodology is utilized to recognize remarkable locales in recordings by prevailing camera movement evacuation. This strategy catches long haul protest movements to produce video saliency maps and build up a brought together calculation that can take a shot at recordings taken by both stationary and moving cameras with no earlier data [7-9]. Be that as it may, the districts with complex movement presents commotions in the video that decreases the exactness of saliency graph.

In [6] originator has built two models dependent on mid and object level highlights, separately. It thinks about their exhibitions against those dependent on low-level highlights. The real disadvantage is that it needs consideration in unsupervised and regulated conduct, individually to enhance execution. The paper [7] has proposed a algorithm using both spatial and temporal features. But it generates a less reliable saliency map.Larger part of existing examination endeavors center around computational saliency models [1-3], even-though, less consideration has been given to evaluating the unwavering quality of the produced saliency maps [4-5].

Despite the fact that picture saliency models can be utilized for saliency discovery freely in every video frame, their saliency identification execution is for the most part lower than spatiotemporal saliency models, which likewise considers the temporal data within the video frames.

Rest of this paper is sorted as pursues. In Section 2, the issues identified with the current strategies are recognized. Segment 3 offers definite portrayal of proposed calculations for distinguishing the saliency in a video. Investigation results and dialogs are marked in Section 4. Ending the paper with area 5 which incorporates the output and further upgrade.

2. PROBLEM IDENDIFICATION

The model of SP spatiotemporal saliency, which comprises of five segments checked utilizing the dash-spotted lines, i.e., include extraction, superpixel-level global saliency, superpixel-level spatial saliency, pixel-level temporal/spatial saliency induction, and pixel-level spatiotemporal saliency production [10]. In this paper, we manufacture the proposed model on the superpixel portrayal of input video frame, in light of the fact that the highlights removed at superpixels, which are over-portioned districts and fit well to the limits between the background regions and the salient objects, are more significant than block dimension and pixel-level highlights.

Addition to it, the superpixel portrayal encourages to protect the very much characterized object limits[11].Using extracted motion and color histograms, we measure the temporal saliency and spatial saliency for each superpixel by considering the contrast between its superpixel-level and the casing level histogram and also its likenesses with different superpixels in both temporal and spatial area[12-15].It assesses the movement peculiarity of superpixels and adventure a plan of transient saliency expectation and acclimation to quantify the superpixel-level saliency and mutually assess the global difference and spatial sparsity of superpixels to gauge the superpixel-level spatial saliency.

To produce the saliency graph the pixel-level precision, a pixel-level saliency deduction strategy is utilized to change the superpixel-level measures both in spatial and temporal means into the pixel-level temporal/spatial saliency outline [16-17]. At last, an adaptive fusion strategy is used to incorporate the temporal saliency outline and the spatial saliency outline of the pixel-level for producing the spatiotemporal saliency graph.

3. SOLUTION-PROPOSED METHOD

3.1 Solution-The proposed method:

Weight average based spatiotemporal correlations

1. Eye-fixation information for the spatio-temporal estimation is considered from the public CRCNS data-set and validate that normally there is a high-relationship in-between the saliency of a pixel and its immediate neighbors.
2. Then, a map for the pixel-wise uncertainty is assessed that mirrors our trust in the computational saliency map relating the value of the pixel to the estimations of its immediate neighbors in a computationally productive manner.
3. Then average weight values for each pixels is calculated separately and applied, thus more reliable saliency map is derived.

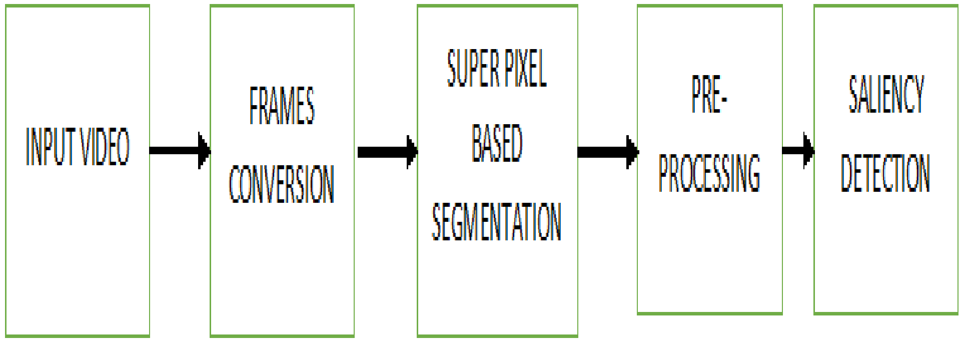


Fig 3.1 Block diagram

3.2.1Input video:

The video in which the salincy is to be detetcted is given as a input. The input video may be in any format such as mp4, avi etc… the inputed video is converted to frames in the next stage.

3.2.2Frame conversion:

The next process is that the inputed video is converted into frames. Since the processing is done by comparing the frames, to identify the saliency in the video. The charecteristics of the video is determined.we calculate the,

1. No of frames.
2. Frame rate.
3. Width of the frame.
4. Height of the frame.

3.2.3Superpixel based segmentation:

- SPS isolates a picture to a progression of unpredictable squares as indicated by surface likeness, brilliance closeness, and shape coherence.
- The hubs of the diagram are the substances that we need to seperate, and here, they are the pixels; the edges between two hubs relate to the comparability with which these two hubs have a place with one gathering.
- Calculate the global threshold esteem Tg on the component picture to discover an esteem which can expand the difference among background and object components by understanding,

  $V(T) = PCPS(\mu C - \mu S)$

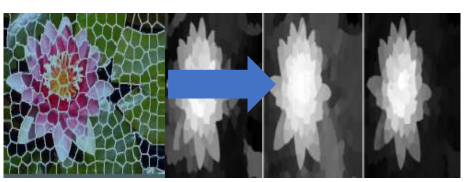
  where PC and PS are the event probabilities of the background and object components, individually. μC and μS are the mean of background and object individually. At last, Tg is set to be max{V (T)}(1 < T < 255).Calculate the maximum value Wmax and minimum value Wmin for the whole feature image.
- Calculate the maximum value Lmax and minimum value Lmin for each superpixel.
- Determine the local threshold for each superpixel by the following criteria:

  Tl = if Lmax < Tg

  Tl = Lmin, if Lmin < Tg

  Tl = ½ Sl + ½ Tg, otherwise

where Sl refers to the threshold of the superpixel



3.2.4Preprocessing:

The frame is first converted into grayscale. Then adaptive histogram equalisation technique is applied on it and outputed.

1. Gray scale conversion:Gray scale images are the measure of the intensity of light at each pixel according to the particular weighted combination of frequencies.
2. Adaptive histogram equalisation:This procedure is utilized to enhance the difference in the contrast of the images.It is reasonable for enhancing the neighborhood differentiation and enchancing the edges in every block of a picture.

3.3 Weight based spatio-temporal cue extraction:

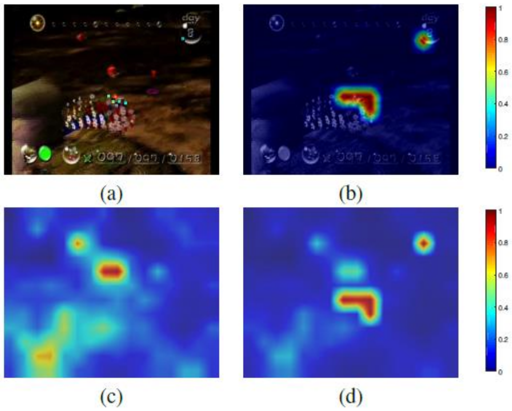


Fig 3.2 True uncertainty data.

(a) Eye fixation superimposed format of the original video;
(b) Modified eye fixation map;
(c) Results for the detection of the saliency;
(d) Identified uncertainty.

Simple linear iterative clustering (SLIC) calculation is utilized to segment the frames into various superpixels.The super pixel level spatial and temporal saliency values are calculated. Then a weight value is applied to each pixels and then the saliency map is generated.

3.4Performance analysis:

The Rand index specifically in data clustering, is a proportion of the comparability between two information clustering's. The Rand list has an incentive somewhere in the range of 0 and 1, with 0 showing that the two information clustering's don't concur on any pair of focuses and 1 demonstrating that the information clustering's are actually the equivalent.

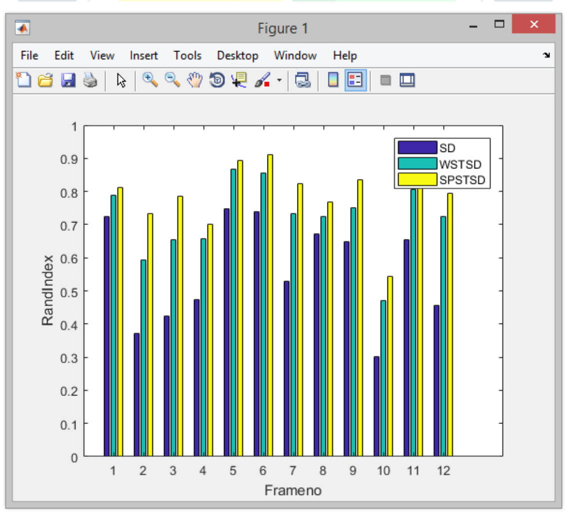


Fig 3.3 frame no VS randindex

The division exactness is estimated mulling over the global consistency error. The global consistency error (GCE) expect that one of the divisions must be a refinement of the other, and forces every single neighborhood refinement to be a similar way.

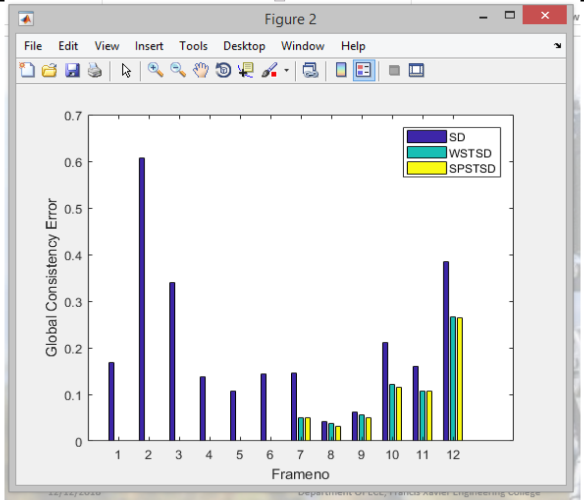


Fig 3.4 frame no VS global consistency error

Frame no	Rand index			Global consistency error			Boundary displacement error			Variation of information		
	SD	WSTSD	SPSTSD	SD	WSTSD	SPSTSD	SD	WSTSD	SPSTSD	SD	WSTSD	SP S0.13 TSD
1	0.72	0.78	0.82	0.18	-	-	0.3	0.02	0.02	0.2	0.14	0.13
2	0.35	0.59	0.73	0.6	-	-	0.28	0.02	0.02	0.56	0.3	0.3
3	0.43	0.65	0.78	0.36	-	-	0.28	0.1	0.2	0.44	0.16	0.15
4	0.49	0.65	0.7	0.15	-	-	0.3	0.04	0.03	0.28	0.24	0.22
5	0.72	0.88	0.9	0.1	-	-	0.34	0.04	0.03	0.14	0.12	0.1
6	0.72	0.86	0.92	0.16	-	-	0.25	0.01	0.01	0.16	0.12	0.1
7	0.52	0.72	0.82	0.16	0.05	0.05	0.68	0.01	0.01	0.22	0.20	0.18
8	0.68	0.72	0.78	0.03	0.03	0.02	0.32	0.08	0.08	0.18	0.15	0.4
9	0.62	0.74	0.82	0.06	0.05	0.04	0.32	0.28	0.22	0.2	0.15	0.13
10	0.28	0.48	0.55	0.2	0.12	0.11	0.28	0.03	0.02	0.44	0.43	0.35
11	0.62	0.82	0.84	0.18	0.10	0.10	0.38	0.02	0.02	0.2	0.19	0.18
12	0.44	0.72	0.8	0.38	0.28	0.08	0.1	0.01	0.01	0.48	0.48	0.35

The boundary displacement error (BDE) measures the normal relocation blunder of limit pixels between two sectioned pictures. Especially, it characterizes the error of one limit pixel as the separation between the pixel and, the nearest pixel in the other boundary picture.

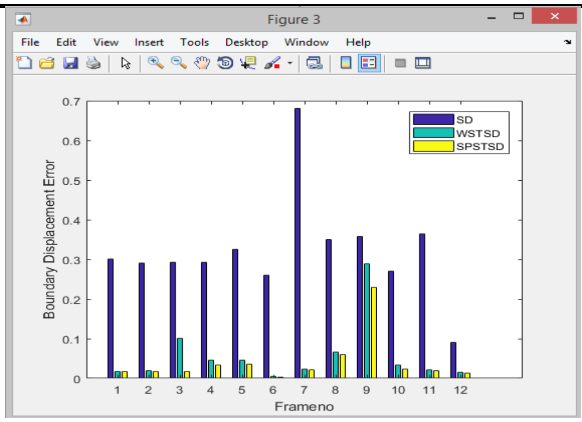


Fig 3.5 frame no VS boundary displacement error

The Information Variation (VoI) metric defines the distance between two segmentations as the average conditi onal entropy of one segmentation compared to the other, and thus approximatelymeasures the amount of ran domness in one segmentation that can not be explained by the other.

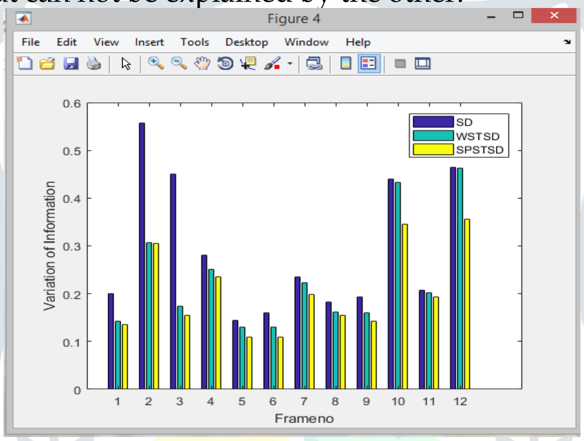


Fig 3.6 frame no VS variation of information

4. RESULT AND DISCUSSION

The video in which the salincy is to be detetcted is given as a input. The input video may be in any format such as mp4, avi etc… The inputed video is converted to frames in the next stage. Then each frames are further processed.

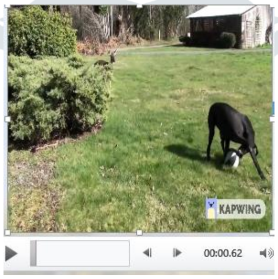


Fig 4.1: Input video

Simple linear iterative clustering (SLIC) calculation s utilized to segment the frames into various superpixels.The input video is segmented based on super pixel and changed over into frames as shown in the fig 4.2. Then the frames are further processed using various algorithms.

Fig 4.2 Video to frame conversion

The input frames are converted into gray scale so as the color images can be further processed. Grayscale pictures are different from one piece bi-tonal dark and white pictures. It is the proportion of the power of light at every pixel as indicated by a specific weighted blend of frequencies. Thus based on the intensity of light the image is converted into gray scale image format as shown in the figure 4.3.

Fig 4.3 Gray scale conversion

Then the gray scale converted images are further processed using adaptive histogram equalization technique.It varies from the common histogram technique in regard that the versatile technique registers a few histograms, each comparing to a particular area of a picture, and uses them to redistribute the softness estimations of the picture. it is consequently appropriate for enhancing the neighborhood and upgrading meanings of edges in every part of a picture. It gives an output as shown in the figure 4.4.

Fig 4.4 Adaptive histogram equalization output

The output of the adaptive histogram equalization is further processed and a saliency map, which is the global measure of conspicuity, is generated as shown in the figure 4.5.It is the combined output of both the saliency and temporal values.

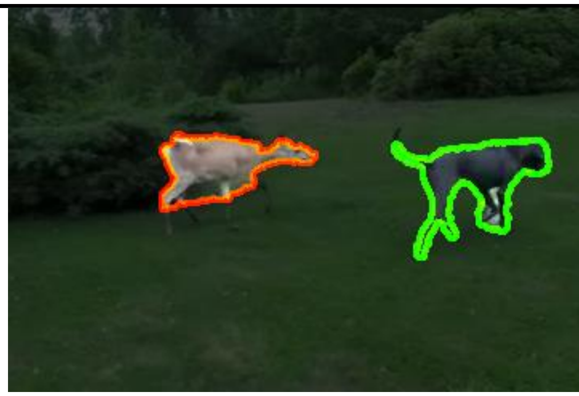


Fig 4.5 Saliency detection

5. CONCLUSION & FUTURE ENHANCEMENT

In this research module, the difficulty in detecting the video saliency by quantifying the uncertainty is studied. A procedureis expressedto determine theuncertainty by pixel-wise and it is plotted on the saliency map that relies on the common feature of human fixation. The proposed algorithm is formulated according tosuper pixel segmentation and the spatiotemporal neighborhoods. Finally a weight value is added to each pixel to get more reliable saliency map.

Moreover, we have demonstrated that the proper size of nearby neighborhood is basically controlled by the video substance and has a noteworthy effect on the calculation execution. This calculation is unsupervised and computationally efficient. The performance of the calculation is additionally improved by utilizing a weight-average value of maps from various scales relying upon the video content.

This can be additionally upgraded by expanding the unwavering quality in finding the saliency of a video. The proposed technique can be amazingly useful, either as a free target evaluation methodology for saliency area figuring's, or as a convincing strategies for quality control for saliency-based video dealing with the applications.

REFERENCES

1) W. Wang, J. Shen, and L. Shao, "Consistent video saliency using local.gradient flow optimization and global refinement," IEEE Transactions on Image Processing, vol. 24, no. 11, pp. 4185–4196, Nov 2015.
2) H. Kim, Y. Kim, J. Y. Sim, and C. S. Kim, "Spatiotemporal saliency detection for video sequences based on random walk with restart," IEEE Transactions on Image Processing, vol. 24, no. 8, pp. 2552–2564, Aug 2015.
3) Ms. S. JebaPon Elizabeth, Mrs.R.VedhapriyaVadhana, "Comparing PSNR by using Integer and Haar Wavelet Transform in Reversible Data Hiding Technique," International Journal of Communication and Computer Technologies,Volume 02 – No.10 Issue: 01 Mar 2014,ISSN NUMBER : 2278-9723.
4) C. R. Huang, Y. J. Chang, Z. X. Yang, and Y. Y. Lin, "Video saliency map detection by dominant camera motion removal," IEEE Transactions on Circuits and Systems for Video Technology, vol. 24, no. 8, pp. 1336–1349, Aug 2014.
5) R.VedhapriyaVadhana, E.Maheswari, "Parametric Blur Estimation Using Modified Radon Transform For Natural Images Restoration," International Journal of Advanced Research in Biology, Ecology, Science and Technology (IJARBEST), Vol. 1, Issue 1, April 2015.
6) T. Disilva, R. VedhapriyaVadhana and RubaSoundar, "Reliable Stereo Matching with Depth Measurement in Videos," I J C T A, 9(4), 2016, pp. 35-42 © International Science Press 1,2,3.
7) Y. Fang, Z. Wang, W. Lin, and Z. Fang, "Video saliency incorporating spatiotemporal cues and uncertainty weighting," IEEE Transactions on Image Processing, vol. 23, no. 9, pp. 3910–3921, Sept 2014.
8) T. Alshawi, Z. Long, and G. AlRegib, "Unsupervised uncertainty estimation in saliency detection for videos using temporal cues," in IEEE Global Conf. on Signal and Information Processing (GlobalSIP), Orlando, Florida, Dec. 14-16. IEEE, 2015.
9) R. VedhaPriyaVadhana, and K. RubaSoundar, "A Novel Video Inpainting Technique Based on Digital Notch Filtering Method", Copyright © 2017 American Scientific Publishers ,Vol. 14, 1239–1244, 2017.
10) Liang and X. Hu, 'Predicting eye fixations with higher-level visual features', IEEE Transactions on Image Processing, Vol. 24, No. 3, pp. 1178–1189, March 2015.

11) D. Mahapatra, S. O. Gilani, and M. K. Saini, 'Coherency based spatio-temporal saliency detection for video object segmentation', IEEE Journal of Selected Topics in Signal Processing, Vol. 8, No. 3, pp. 454–462, June 20.
12) R.Nikitha, C.K.Sakavi, H.Mehnaz, M.Rajalakshmi, Christo Ananth, 'High Resolution Image Reconstruction with smart camera network'.IJARBEST, Volume I, issue IV, July 2015.
13) R. VedhaPriyaVadhana, and K. RubaSoundar ," Effective Motion Tracking Of Moving Persons/Objects Using Mcmc Sampling Method", IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 1, Issue 5, Oct-Nov, 2013, ISSN: 2320–8791.