VISUAL SALIENCY FOR SALIENT OBJECT DETECTION BASED ON TOPDOWN METHOD

G.P.Krishna moorthy
PG scholar
Department of ECE,
Excel Engineering college,
Komarapalayam-600 025
Namakkal (Dist)
Tamil Nadu, India.
krishnaapalanisamy@gmail.com

K.Meignanamoorthy
PG scholar
Department of ECE,
Excel Engineering college,
Komarapalayam-600 025
Namakkal (Dist)
Tamil Nadu, India.
k.meignanamoorthy@gmail.com

S.Koutheesh
Assistant Professor,
Department of ECE,
Excel Engineering college,
Komarapalayam-600 025
Namakkal (Dist)
Tamil Nadu,India.
kowtheesh@gmail.com

Abstract-**For many applications in graphics, design and human computer interaction, it is essential to estimate the visual saliency of images. This project based on visual saliency detection method that combines the respective merits of top down saliency method and bottom up saliency method to achieve more accurate saliency maps. An algorithm of coarse saliency region is first obtained using convex hull of interest points and also analyzes the saliency information with mid level visual information using super pixels. Laplacian sparse subspace clustering method used to group super pixels with local features and analyze the results with respect to the coarse saliency region to compute the prior saliency map. Using low level visual information based on the convex hull to compute saliency at each pixel.**

Index terms – **saliency, visual attention, interest point detector, convex hull, clustering, super pixel, saliency map.**

I. INTRODUCTION

In the domain of computer vision, efforts have been made in modeling the mechanism of human attention, especially the bottom-up attention mechanism Such a process is also called visual saliency detection. Human vision system doesn't process whole parts of the image. Visual saliency is the distinct subjective perceptual quality which makes some items in the image stand out from their neighbors and immediately grab our attention. Visual attention may be a solution to the inability to fully process all locations in parallel. It is then attracted towards salient visual locations. This location is sufficiently different from its surroundings to be worthy of your attention. This bottom-up deployment of attention towards salient locations can be strongly modulated Thus a lone red object in a green field will be salient and will attract attention in a bottom-up manner. Using bottom-up and top down method for the direction of selective attention. The attention focus only depends on the visual features from the image.

The saliency map combines information from each of the feature maps into a global measure where points corresponding to one location in a feature map project to single units in the saliency map. Saliency at a given location is determined by the degree of difference between that location and its surround. An integrated technology of top-down and bottom-up visual attention used to the solution of the feature selection and saliency detection problems in object extraction and categorization of natural scene images.

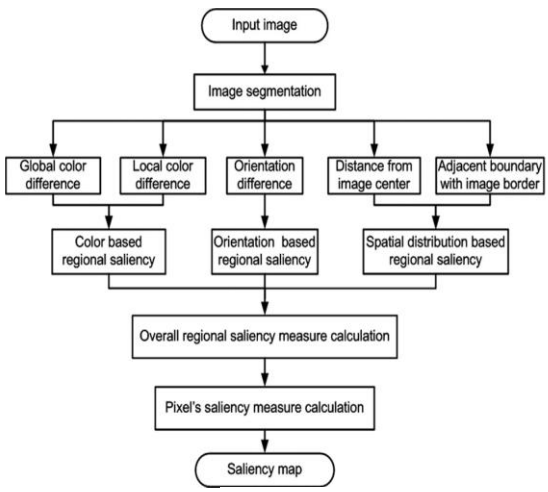

Fig 1 Flow of operation

While existing bottom-up saliency models have presented good results but it has several problems. One of the main problems in bottom-up method is object surrounding also highlighted and inability to highlight the particular salient object. Finally, some bottom-up algorithms are computationally expensive as they often entail exhaustive search at multiple scales and neighborhoods to compute saliency measure at each pixel.

The Coarse saliency region is obtained using convex hull of interest point detection and saliency information of mid level visual information retrieved from super pixels. Laplacian sparse subspace clustering method to group super pixels and analyze the results with coarse saliency region to compute the prior saliency map. Using the low level visual information based on the convex hull to compute the observation probability, thereby facilitating conclusion of Bayesian saliency at each pixel.There are many computational models based on the bottom-up mechanisms of the visual attention, which are designed to determine the saliency of the image from the visual features of the scenes. The models that generate topographical maps of the visual saliency from the images are called Saliency Maps.

In this paper, we address these problems with a Bayesian saliency model by exploiting low and mid level information to generate saliency maps based on the center-surround principle. First, we compute the convex hull of interest points for a coarse region estimation of the salient object. The prior map of the proposed Bayesian saliency model is computed based on the coarse saliency region and our Laplacian sparse subspace clustering method. In contrast to segmenting salient object which measures saliency with contrast between scanned windows and surrounding regions, our formulation is more efficient and effective by exploiting mid level information and better metrics. Our clustering method extends the sparse subspace clustering by introducing a regularization term with a Laplacian matrix at the superpixel level. This regularization term further enforces similar superpixels to be clustered into the same group.

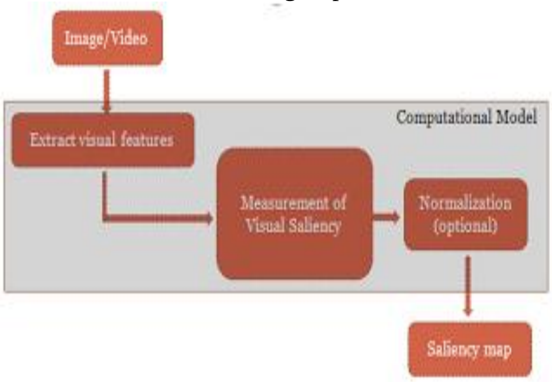

Fig 2 Computation model

II. PROPOSED ALGORITHM

In this section we give brief description about our saliency model using low and mid level visual information. we present the coarse estimation of saliency regions based on the convex hull that includes all the detected interest points from low level information. The prior distribution is estimated based on a novel clustering algorithm with mid level information represented by super pixels and coarse saliency regions. Saliency computed based on the center-surround principle with coarse saliency regions and thus the Bayesian visual saliency map for an image. An interest point detector is used to locate salient points of the object. Next, the convex hull enclosing the salient points is computed to estimate the approximate salient region. Based on the estimated region, we formulate the saliency detection as a Bayesian inference problem for estimating the posterior probability at each pixel *v* of the image

$$P(sal|v) = \frac{P(sal)P(v|sal)}{P(sal)P(v|sal)+P(bk)P(v|bk)}$$

$$P(bk) = 1 - P(sal)$$

where *p(sal|v)* is a shorthand for probability of predicting a pixel being salient $p(sal = 1|v)$, *p(sal)* is the prior probability of being salient at pixel *v* and *p(bk)* is the prior probability of a pixel belonging to the background. In addition, *p(v|sal)* and *p(v|bk)* are the likelihood of observations (shorthands for $p(v|sal = 1)$ and $p(v|bk = 1)$). The goal is to produce a saliency map where the probability of each pixel being salient is estimated.

A. Interest point detector

Existing methods based on the bottom-up method and center-surround principle usually operate with low level information. That is pixel standouts from the neighborhood of a pixel is computed as its saliency value.Compute for all pixels the horizontal and vertical derivatives *fx* and *fy* .Compute for each pixel the matrix:

$$M = \begin{bmatrix} \sum_{(x_k,y_k)\in W} f_x^2(x_k,y_k) & \sum_{(x_k,y_k)\in W} f_x(x_k,y_k)f_y(x_k,y_k) \\ \sum_{(x_k,y_k)\in W} f_x(x_k,y_k)f_y(x_k,y_k) & \sum_{(x_k,y_k)\in W} f_y^2(x_k,y_k) \end{bmatrix}$$

W: Neighborhood of the pixel. The summation may be weighted by a Gaussian function. Compute for each pixel scalar interest measure. Find local maxima above a certain threshold and report these as detected feature point locations. Salient points to compute the approximate location of the salient object, and then exploit the properties of this region to compute the final saliency map. This approach not only helps alleviate high computational cost but also makes the proposed saliency map more discriminative. Most interest point detectors do not exploit color information and may be sensitive to the background noise.

The information of a color image derivative is computed based on its frequency or probability, and a transformation function is developed to boost color saliency. The color Harris detector is then operated on the image to detect feature points. Compared with the intensity-based feature detectors, the boosted color saliency points are shown to be more stable and informative. In this paper, we use the color boosted Harris point operator to detect corners or contour points of salient regions in a color image. The saliency points provide useful spatial information regarding the interesting object in the scene. We eliminate those near the image boundary, and compute a convex hull to enclose all the remaining salient points.

An interest point is a point in the image which in general can be characterized as follows:

- It has a clear, preferably mathematically well-founded, definition,
- It has a well-defined *position* in image space,
- The local image structure around the interest point is rich in terms of local *information contents* (e.g.: significant 2D texture), such that the use of interest points simplify further processing in the vision system.

The use of interest points to the notion of regions of interest, which have been used to signal the presence of objects, often formulated in terms of the output of a detection step. Most of the interest point detectors do not exploit color information and may be sensitive to the background noise to exploit gradient information of both grayscale and color channels to measure saliency. The color boosted Harris point operator to detect corners or contour points of salient regions in a color image. The saliency points provide useful spatial information regarding the interesting object in the scene. Eliminating noise in image near the image boundary, and compute a convex hull to enclose all the remaining salient points. As the detected interest points usually surround the salient object, this method provides coarse region estimation.Convex hull set of points given contains only one point, that point is its own convex hull. The convex hull of two unique points is the line segment between the two points.

Algorithm: Convex Hull (P)

Input: A set P of points in the plane.

1. Sort the points by x-coordinate, resulting in a sequence $P_1 \ldots . P_n$.

2. Put the points p1 and p2 in a list L_{upper}, with P1 as the first point.

3. for i 3 to n

4. do Append pi to L_{upper}.

5. while L_{upper} contains more than two points and the last three points in L_{upper} do not make a right turn

6. do Delete the middle of the last three points from L_{upper}.

Output: A list containing the vertices of CH(P) in clockwise order.

B.Sparse Subspace Clustering

To cluster data points that included in a union of low-dimensional subspaces. Among infinitely many possible representations of a data point in terms of other points, a sparse representation corresponds to selecting a few points from the same subspace. It encourages solving a sparse optimization program whose solution is used in a spectral clustering framework to deduce the clustering of the data into subspaces. For solving the sparse optimization program is in general NP-hard, we examine a convex relaxation and under appropriate conditions on the arrangement of the subspaces and distribution of the data, the proposed minimization program conquer in recovering the thirst sparse representations. The proposed algorithm is dynamic and can handle data points near the intersections of subspaces. It can contract directly with data nuisances, such as noise, sparse outlying entries, and missing entries, by consolidate the model of the data into the sparse optimization program.

Fig 3(a) Input Image (b) After Spectral Clustering

SCC can deal with noisy data but requires knowing the number and dimensions of subspaces and assumes that subspaces have the same dimension. Complexity of building the multiway similarity grows exponentially with the dimensions of the subspaces. a sampling strategy is employed to reduce the computational cost. Using advances in sparse and low-rank, recovery algorithms, the Sparse Subspace Clustering (SSC), Low-Rank Recovery (LRR) and Low-Rank Subspace Clustering (LRSC) algorithms pose the clustering problem as one of finding a sparse or low-rank representation of the data in the dictionary of the data itself. The solution of the corresponding global optimization algorithm is then used to build a similarity graph from which the segmentation of the data is obtained. We obtain a sparse representation for each data point whose nonzero elements ideally correspond to points from the same subspace. The next step of the algorithm is to infer the segmentation of the data into different subspaces using the sparse coefficients.

Algorithm: SPARSE SUBSPACE CLUSTERING

Input: A set of data points $Y \in R^{m \times n}$, and the number of desired clusters k.

1. Solve the l^1- minimization problem (1) to get the collection of c_i.
2. Form an affinity matrix $A = |C|^T + |C|$, where $C = [c_1, c_2, \ldots\ldots c_n]$.
3. Construct a Laplacian matrix $L = I - D^{-1/2} A D^{-1/2}$ using A, where $D = diag\{d_i\}$ with $d_i = \sum_{j=1}^{n} A_{ij}$.
4. Obtain eigenvector matrix $V \in R^{n \times k}$ which consists of the first k normalized eigenvectors of L corresponding to its k smallest eigenvalues.
5. Get the segmentations of the data by performing k-means on the row of V

Output: The cluster assignments of Y

TABLE 1
CLUSTERING ERROR (%) OF DIFFERENT ALGORITHMS ON THE HOPKINS 155 DATASET WITH THE 2F-DIMENSIONAL DATA POINTS

Algorithms	LSA	SSC	LRR	LRR-H	LRSC	SSC
2 Motions						
Mean	4.23	2.89	4.10	2.13	3.69	1.52
Median	0.56	0.00	0.22	0.00	0.29	0.00
3 Motions						
Mean	7.02	8.25	9.89	4.03	7.69	4.40
Motions	1.45	0.24	6.22	1.43	3.80	0.56
All						
Mean	4.86	4.10	5.41	2.56	4.59	2.18
Motions	0.89	0.00	0.53	0.00	0.60	0.00

Applying subspace clustering algorithms to the data set while using the original 2F-dimensional Feature trajectories and transfer the data into a 4ndimensional subspace (n is the number of subspaces) using PCA are shown in Tables 1 and 2, respectively.

TABLE 2

CLUSTERING ERROR (%) OF DIFFERENT ALGORITHMS ON THE HOPKINS 155 DATASET WITH THE 4N-DIMENSIONAL DATA POINTS OBTAINED BY APPLYING PCA

Algorithms	LSA	SSC	LRR	LRR-H	LRSC	SSC
2 Motions						
Mean	3.61	3.04	4.83	3.41	3.87	1.83
Median	0.51	0.00	0.26	0.00	0.26	0.00
3 Motions						
Mean	7.65	7.91	9.89	4.86	7.72	4.40
Motions	1.27	1.14	6.22	1.47	3.80	0.56
All						
Mean	4.52	4.14	5.98	3.74	4.74	2.41
Motions	0.57	0.00	0.59	0.00	0.58	0.00

In both cases, SSC obtains a small clustering error, outperforming the other algorithms. This suggests that the separation of different motion subspaces in terms of their principal angles and the distribution of the feature trajectories in each motion subspace are sufficient for the success of the sparse optimization program, hence clustering. Global spectral clustering-based approaches try to resolve these issues by building better similarities between data points using global information.

C. Laplacian Sparse Subspace Clustering

The SSC method performs well in clustering, it uses only low level visual information. In LSSC, handles the mid level Visual information using super pixels for clustering. Compute a constraint matrix to encode the relationship between each pair of super pixels A Laplacian regularization term based on the constraint matrix is introduced in to enforce that similar super pixels should have similar sparse coefficients. Denote the constraint (affinity) matrix as *W*, the resulting optimization problem is:

$$\min \| \hat{U}_i c_i - u_i \|_2 + \lambda \| c_i \|_1 + \alpha/2 \sum_{ij} \| c_i - c_j \|^2 W_{ij} =$$

$$\min \| \hat{U}_i c_i - u_i \|_2 + \lambda \| c_i \|_1 + \alpha \, tr(CLC^T)$$

$$\text{Subject to } C_i^T 1 = 1$$

Where *L* is the Laplacian matrix defined as *L* = *H* − *W*, and *H* is the diagonal matrix with row sums, $H_{ii} = \sum_j W_{ij}$. The parameter α is the weight that balances the effect of the constrained regularization term. As a super pixel contains similar pixels and preserves the structural information of a salient object, this representation is adopted for clustering due to its stability and efficiency.

LSSC method to group super pixels one of the main issues with spectral clustering to construct an effective adjacency matrix that describes the similarity between each pair of pixels accurately. It is usually redundant and inefficient to consider a fully connected graph, and existing methods compute the matrix using the local neighborhood information.

Fig 4 (a) Input Image (b) SSC (c) LSSC

D: Evaluation of Convex Hulls and Clustering Algorithms

Convex hull is the smallest convex set that contains that subset. The convex hull of a set *q* of points is the smallest convex polygon *p* for which each point *q* is either on the boundary of *p* or in its interior. A set of points is defined to be convex, if it contains the line segments connecting each pair of its points. The convex hull of a given set *X* may be defined as

1. The (unique) minimal convex set containing *X*.
2. The intersection of all convex sets containing *X*.
3. The set of all convex combination of points in *X*.

Each image extracts the point and removes those near the image boundary. The average number of detected points and the standard deviation based on the images used this concept. To measure the overlap of the convex hull of detected interest points and the region containing salient region with labeled ground truth. The average ratio is more than 60% and thus the convex hulls provide good estimates of salient objects. To utilize the clustering results to compute the prior saliency map. Evaluating the results using other clustering methods. To plot the precision-recall curves, use fixed thresholds to determine whether a pixel is salient or not.

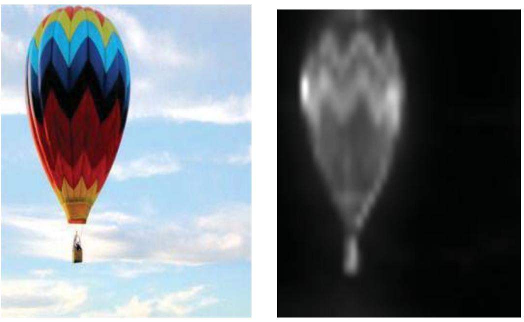

Fig 5(a) Input Image (b) Detection of Saliency Object

E: Evaluation of Saliency Map Models

The saliency map combines information from each of the feature maps into a global measure where points corresponding to one location in a feature map project to single units in the saliency map. Saliency at a given location is determined by the degree of difference between that location and its surround. some sample results where brighter pixels indicate higher saliency probabilities. In the existing method estimates saliency at each pixel by its color contrast to the average of the entire image, and it does not work well when the salient region and the background have similar color or the object of interest is relatively large.

This model is able to better estimate saliency maps at pixel level within and on the contour of the objects in cluttered backgrounds. When multiple objects are present in an image, the salient point detector determines the boundary of the largest salient region containing them all.

Fig 6 (a) Input Image(b) Bottom up Saliency Method (c) Top down Saliency Method

Let N_I be the number of pixels in *I*, and N_I *(f (v))* $f \in \{l, a, b\}$ be the count that region *I* contains *f (v)*. To represent color histograms of pixels in *O* with N_O, N_O *(f (v))* $F \in \{l, a, b\}$.the saliency value at pixel v is computed as

$$p(v|sal) = \prod_{f\in\{l,a,b\}} \frac{NI(f(v))}{NI}$$

$$p(v|bk) = \prod_{f\in\{l,a,b\}} \frac{No(f(v))}{No}$$

Segmentation salient objects based on these saliency maps and evaluate the segmented results. The threshold from 0 to 255 to obtain different segmentations and compute the precision-recall curves. As shown in Fig 3.5(a), our saliency model achieves the best performance up to a precision rate above 0.9.

III.CONCLUSION

Bottom-up saliency model with in the Bayesian framework using low and mid level visual information. Based on the informative saliency points, an algorithm that does not entail exhaustive scan of windows or choosing proper neighborhood scale. A new image clustering method at the super pixel level using sparse representation and apply it to compute the prior distribution of saliency. The prior probability map within the Bayesian inference framework and evaluate our method on a data set of 1,000 images with labeled ground truth. Experimental results demonstrate of effectiveness of our clustering method and saliency map model. This method generates more discriminative saliency maps with higher precision and recall than state-of-the-art algorithms.

REFERENCES

[1] Chenlei Guo,and Liming Zhang, "A Novel Multiresolution Spatiotemporal Saliency Detection Model and Its Applications in Image and Video Compression" vol. 19, no. 1, january 2010.

[2] Christian Siagian, Laurent Itti,"Rapid Biologically-Inspired Scene Classification Using Features Shared with Visual Attention" vol. 29, no. 2, February 2007.

[3] Claudio M. Privitera, Lawrence W. Stark, "Algorithms for Defining Visual Regions-of-Interest: Comparison with Eye Fixations" vol. 22, no. 9, September 2000.

[4] Dashan Gao, Sunhyoung Han, Nuno Vasconcelos,"Discriminant Saliency, the Detection of Suspicious Coincidences, and Applications to Visual Recognition" vol. 31, no. 6, June 2009.

[5] E. Rahtu, J. Kannala, M. Salo, and J. Heikkilä, "Segmenting salient objects from images and videos," in *Proc. Eur. Conf. Comput. Vis.*, 2010, pp. 366–379.

[6] E. Elhamifar and R. Vidal, "Sparse subspace clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Mar. 2009, pp. 2790–2797.

[7] Jen-Tzung Chien, Meng-Sung Wu, "Adaptive Bayesian Latent Semantic Analysis" IEEE Transactions on audio, speech, and language processing, vol. 16, no. 1, January 2008.

[8] Junwei Han, King N. Ngan, Mingjing Li, and Hong-Jiang Zhang,"Unsupervised Extraction of Visual Attention Objects in Color Images" IEEE Transactions on circuits and systems for video technology, vol. 16, no. 1, January 2006.

[9] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254–1259, Nov. 1998.

[10] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Mar. 2010, pp. 2376–2383.

[11] Song Wang, Jeffrey Mark Siskind,"Image Segmentation with Ratio Cut" IEEE Transactions on Pattern Analysis and machine intelligence, vol. 25, no. 6, June 2003.

[12] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H.-Y. Shum, "Learning to detect a salient object," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 2, pp. 353–367, Feb. 2011.

[13] W. Förstner and B. Moonen, "A metric for covariance matrices," Dept. Geodesy & Geoinformatics, Stuttgart Univ., Stuttgart, Germany, Tech. Rep., 1999.

[14] Yulin Xie, Huchuan Lu, Ming-Hsuan Yang," Bayesian Saliency via Low and Mid Level Information" vol. 22, no. 5, May 2013.

[15] Z. Tang, Z. Miao, Y. Wan, and J. Li, "Automatic foreground extraction for images and videos," in *Proc. IEEE Int. Conf. Image Process.*, Mar. 2010, pp. 2993–2996.

