Contents lists available at www.innovativejournal.in

Asian Journal of Computer Science and Information Technology

Journal homepage:http://www.innovativejournal.in/index.php/ajcsit

ISSN 2249 - 5126

A REVIEW : SALIENT FEATURE EXTRACTION USING K-MEDIODS CLUSTERING TECHNIQUE

Jyoti Verma*, Vineet Richhariya

CSE Department LNCT , Bhopal, India.

Contents lists available at www.innovativejournal.in

INNOVATIVE JOURNAL

Asian Journal of Computer Science and Information Technology

Journal homepage:http://www.innovativejournal.in/index.php/ajcsit

AJCSIT

ISSN 2249 - 5126

A REVIEW : SALIENT FEATURE EXTRACTION USING K-MEDIODS CLUSTERING TECHNIQUE

Jyoti Verma*, Vineet Richhariya

CSE Department LNCT , Bhopal, India.

ARTICLE INFO

Corresponding Author:
Jyoti Verma
CSE Department LNCT , Bhopal, India.

KeyWords: Clustering, Feature Extraction, K-mediods, Saliency

ABSTRACT

Human eye is perceptually more sensitive to certain colors and intensities and objects with such features are considered more salient. Detection of salient image regions is useful for applications like image segmentation, adaptive compression, and region-based image retrieval. In this paper we propose a method to first determine salient regions in images using low-level features of luminance and color and then extract the features. The method is fast, easy to implement and generates high quality saliency maps of the same size and resolution as the input image. We identify salient regions as those regions of an image that are visually more conspicuous by virtue of their contrast with respect to surrounding regions.

INTRODUCTION

Salient objects have the quality to visually stand out from their surroundings and are likely to attract human attention. A key property that makes an object salient is the visual difference to the background. A polar bear is salient on dark rocks, but almost invisible in snow. The detection of visual saliency is of high interest in many computer vision applications, ranging from general object detection in web images, over image thumb nailing, to computing a joint focus of attention in human robot interaction [1-3].

Visual saliency and, more general, visual attention have been widely investigated in neurobiology and psychophysics [4] and many computational models [5-7] have been built based on such finding. Recently, several saliency approaches came up that are based on computational and mathematical ideas and usually less biologically motivated. These approaches range from the computation of entropy [8,9], over determining features that best discriminate between a target and a null hypothesis [10], to learning the optimal feature combination with machine learning techniques [11].

Identifying visually salient regions is useful in applications such as object based image retrieval, adaptive content delivery, adaptive region-of-interest based image compression, and smart image resizing [12-15]. We identify salient regions as those regions of an image that are visually more conspicuous by virtue of their contrast with respect to surrounding regions.

RELATED WORK

Ma and Zhang [13] propose a local contrast-based method for generating saliency maps that operates at a single scale and is not based on any biological model. The input to this local contrast-based map is a resized and color quan-tized CIELuv image, sub-divided into pixel blocks. The saliency map is obtained from summing up differences of image pixels with their respective surrounding pixels in a small neighborhood. This framework extracts the points and regions of attention. A fuzzy-growing method then segments salient regions from the saliency map.

Hu et al. [16] create saliency maps by thresholding the color, intensity, and orientation maps using histogram entropy thresholding analysis instead of a scale space approach. They then use a spatial compactness measure, computed as the area of the convex hull encompassing the salient region, and saliency density, which is a function of the magnitudes of saliency values in the saliency feature maps, to weigh the individual saliency maps before combining them.

Itti et al. [6] have built a computational model of saliency-based spatial attention derived from a biologically plausible architecture. They compute saliency maps for features of luminance, color, and orientation at different scales that aggregate and combine information about each location in an image and feed into a combined saliency map in a bottom-up manner. The saliency maps produced by Itti's approach have been used by other researchers for applications like adapting images on small devices [17] and unsupervised object segmentation [18, 19].

Segmentation using Itti's saliency maps (a 480x320 pixel image generates a saliency map of size 30x20 pixels) or any other sub-sampled saliency map from a different method requires complex approaches. For instance, a Markov random field model is used to integrate the seed values from the saliency map along with low-level features of color, texture, and edges to grow the salient object regions [18].

Ko and Nam [19], on the other hand, use a Support Vector Machine trained on the features of image segments

to select the salient regions of interest from the image, which are then clustered to extract the salient objects.
Recently, Frintrop et al. [20] used integral images [21] in VOCUS (Visual Object Detection with a Computational Attention System) to speed up computation of center-surround differences for finding salient regions using separate feature maps of color, intensity, and orientation.

SALIENCY DETECTION AND SALIENCY MAP

This section presents details of saliency determination and its use in segmenting whole objects. Using the saliency calculation method described later, saliency maps are created at different scales. These maps are added pixel-wise to get the final saliency maps. The input image is then over-segmented and the segments whose average saliency exceeds a certain threshold are chosen.

Saliency calculation:

Saliency is determined as the local contrast of an image region with respect to its neighborhood at various scales. This is evaluated as the distance between the average feature vector of the pixels of an image sub-region with the average feature vector of the pixels of its neighborhood. This allows obtaining a combined feature map at a given scale by using feature vectors for each pixel, instead of combining separate saliency maps for scalar values of each feature.

Saliency Detection Algorithm:

At a given scale, the contrast based saliency value ci,j for a pixel at position (i; j) in the image is determined as the distance D between the average vectors of pixel features of the inner region R1 and that of the outer region R2 as

$$Ci, j = D\left[\left(\frac{1}{N1}\sum_{p=1}^{N1} vp\right),\left(\frac{1}{N2}\sum_{q=1}^{N2} vq\right)\right]$$

where N1 and N2 are the number of pixels in R1 and R2 respectively, and v is the vector of feature elements corresponding to a pixel. The distance D is an Euclidean distance if v is a vector of uncorrelated feature elements, and it is a Mahalanobis distance (or any other suitable distance measure) if the elements of the vector are correlated. In this work, we use the CIELab color space, assuming sRGB images, to generate feature vectors for color and luminance. Since perceptual differences in CIELab color space are approximately Euclidian, D in Equation given below is:

$$ci;j = ||v1-v2||$$

where v1 = [L1; a1; b1]T and v2 = [L2; a2; b2]T are the average vectors for regions R1 and R2, respectively. Since only average feature vector values of R1 and R2 need to be found, we use the integral image approach as used in [21] for computational efficiency. A change in scale is affected by scaling the region R2 instead of scaling the image. Scaling the filter instead of the image allows the generation of saliency maps of the same size and resolution as the input image. For an image of width w pixels and height h pixels, the width of region R2, namely wR2 is varied as:

$$w/2 \geq wR2 \geq w/8$$

assuming w to be smaller than h (else we choose h to decide the dimensions of R2). This is based on the observation that the largest size of R2 and the smaller ones (smaller than w=8) are of less use in finding salient regions . The former might highlight non-salient regions as salient, while the latter are basically edge detectors. So for each image, filtering is performed at three
different scales and the final saliency map is determined as a sum of saliency values across the scales S:

$$mi;j = \sum ci;j$$

for all i € [1;w]; for all j € [1; h] where mi;j is an element of the combined saliency map
M obtained by point-wise summation of saliency values across the scales.

PROPOSED ALGORITHM

Clustering is the process of grouping a set of objects into classes or clusters so that objects within a cluster have similarity in comparison to one another, but are dissimilar to objects in other clusters (Han et al 2001). Unfortunately, K-means clustering is sensitive to the outliers and a set of objects closest to a centroid may be empty, in which case centroids cannot be updated. For this reason, K-medoids clustering are sometimes used, where representative objects called medoids are considered instead of centroids. Because it uses the most centrally located object in a cluster, it is less sensitive to outliers compared with the K-means clustering. Among many algorithms for K-medoids clustering, Partitioning Around Medoids (PAM) proposed by Kaufman and Rousseeuw (1990) is known to be most powerful. However, PAM also has a drawback that it works inefficiently for large data sets due to its complexity (Han et al, 2001). This is main motivation of this paper. The algorithm can be summarized as follows.

1. Initialize: randomly select *k* of the *n* data points as the medoids
2. Associate each data point to the closest medoid. ("closest" here is defined using any valid distance metric, most commonly Euclidean distance, Manhattan distance or Minkowski distance)
3. For each medoid *m*
   3.1 For each non-medoid data point *o*
   3.1.1.Swap *m* and *o* and compute the total cost of the configuration
4. Select the configuration with the lowest cost.
5. Repeat steps 2 to 5 until there is no change in the medoid.

CONCLUSION

We presented a novel method of finding salient regions in images, using low level features of color and luminance, which is easy to implement, noise tolerant, and fast enough to be useful for real time applications. It generates saliency maps at the same resolution as the input image.

REFERENCES

[1] Dominik A. Klein and Simone Frintrop : Center-surround Divergence of Feature Statistics for Salient Object Detection
[2] L. Marchesotti, C. Cifarelli, and G. Csurka. A framework for visual saliency detection with applications to image thumbnailing. In Proc. of ICCV, 2009
[3] B. Schauerte and G. A. Fink. Focusing computational visual attention in multi-modal human-robot interaction. In Proc. of Int'l Conf. on Multimodal Interfaces (ICMI), 2010.
[4] H. Pashler. The Psychology of Attention. MIT Press, Cambridge, MA, 1997.
[5] J. K. Tsotsos. Towards a computational model of visual attention. In Early Vision and Beyond. MIT Press, 1995.

[6] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. IEEE Trans. On PAMI, 20(11), 1998
[7] S. Frintrop. VOCUS: A Visual Attention System for Object Detection and Goal-directed Search, volume 3899 of Lecture Notes in Artificial Intelligence (LNAI). Springer, 2006.
[8] S. Frintrop, E. Rome, and H. I. Christensen. Computational visual attention systems and their cognitive foundation: A survey. ACM Trans. on Applied Perception, 7(1), 2010
[9] T. Kadir, A. Zisserman, and M. Brady. An affine invariant salient region detector. In Proc. of ECCV, 2004.
[10] D. Gao and N. Vasconcelos. Bottom-up saliency is a discriminant process. In Proc. of ICCV, 2007
[11] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H.-Y. Shum. Learning to detect a salient object. IEEE Trans. on PAMI, 2009.
[12] Radhakrishna Achanta, Francisco Estrada, Patricia Wils, and Sabine SÄusstrunk : Salient Region Detection and Segmentation
[13]. Y.-F. Ma and H.-J. Zhang. Contrast-based image attention analysis by using fuzzy growing. In Proceedings of the Eleventh ACM International Conference on Multimedia, pages 374{381, November 2003.
[14]. V. Setlur, S. Takagi, R. Raskar, M. Gleicher, and B. Gooch. Automatic image retargeting. In Proceedings of the 4th International Conference on Mobile and Ubiquitous Multimedia (MUM'05), pages 59{68, October 2005.
[15]. S. Avidan and A. Shamir. Seam carving for content-aware image resizing. ACM Transactions on Graphics, 26(3):10, July 2007.
[16]. Y. Hu, X. Xie, W.-Y. Ma, L.-T. Chia, and D. Rajan. Salient region detection using weighted feature maps based on the human visual attention model.Springer Lecture Notes in Computer Science, 3332(2):993{1000, October 2004.
[17]. L. Chen, X. Xie, X. Fan, W.-Y. Ma, H.-J. Zhang, and H. Zhou. A visual attention model for adapting images on small displays.ACM Transactions on Multimedia Systems, 9:353{364, November 2003.
[18]. 5. J. Han, K. Ngan, M. Li, and H. Zhang. Unsupervised extraction of visual attention objects in color images. IEEE Transactions on Circuits and Systems for Video Technology, 16(1):141{145, January 2006.
[19]. B. C. Ko and J.-Y. Nam. Object-of-interest image segmentation based on human attention and semantic region clustering. Journal of Optical Society of America A, 23(10):2462{2470, October 2006
[20].S. Frintrop, M. Klodt, and E. Rome. A real-time visual attention system using inte-gral images. InInternational Conference on Computer Vision Systems (ICVS'07), March 2007.
[21]. P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'01), 1:511{518, December 2001.

AUTHORS PROFILE

Jyoti Verma
Completed Bachelor of engineering in Information Technology from RKDF Institute of science and technology , Rajiv Gandhi University, Bhopal (M.P.), India in 2009. Final Year Student Master of technology in Software Engineering from Lakshmi Naraian college of Technology, Rajiv Gandhi University, Bhopal(M.P.), India

Vineet Richariya
Vineet Richariya is Head Of Department of Computer Science And Engineering, Lakhmi Narain College Of Technology, Bhopal, India. He did his M.Tech(Computer Science & Engineering) from BITS Pillali in year 2001. He did his B.E (Computer Science & Engineering) from Jiwaji University, Gwalior, India in year 1990.