

# Weighted Principal Component Analysis based Visual Saliency Detection

M. Shivashankari (AP),  
Department of ECE  
K. Ramakrishnan College of Technology  
Samayapuram, Trichy, Tamilnadu

G. Dheepika, M. Hemalatha, R. Nallammai  
Department of ECE  
K. Ramakrishnan College of Technology  
Samayapuram, Trichy, Tamilnadu

**Abstract-**In image processing technique, we are implementing weighted principal component analysis based visual saliency detection. This algorithm is capable of reducing the dimension of the image. It consists of color texture and luminescence which helps us to identify the depth of an image. User can measure the saliency by integrating three elements as follows: the dissimilarities between image patches, which were evaluated in the reduced dimensional space, the spatial distance between image patches and the central bias. A mechanism of pixel reduction can be done by, indicating a bias for human fixations to the center of the image. We extracted the weighted principal components (WPCs) by sampling the patches from the current image. Then the patches are represented by the coefficients of weighted principal components using WPCA analysis. Then the patches are represented by the coefficients of principal components using WPCA analysis. Based on the compact representation of patches, two types of distinctiveness are introduced 1) center-surround contrast, 2) global rarity. Experimental results show that our method outperforms the current algorithm based on predicting human fixation. It can also be widely used in gesture based and medical applications.

## I. INTRODUCTION

Visual attention has been extensively studied recently due to its wide applications in image adaptation, quality assessment, contrast enhancement, and video compression. Numerous saliency models have been proposed by psychologists, neurophysiologists and computer scientists to imitate human visual attention. According to the motivation of inferring visual attention, saliency models can be categorized into top-down models and bottom-up models. Top-down models are task-driven while bottom-up models are data-driven and more related to the nature of Human Visual System (HVS). Visual salience (or visual saliency) is the distinct subjective perceptual quality which makes some items in the world stand out from their neighbours and immediately grab our attention. Visual attention may be a solution to the inability to fully process all locations in parallel. However, this solution produces a problem. Visual salience helps your brain achieves reasonably efficient selection. Early stages of visual processing give rise to a distinct subjective perceptual quality which makes some stimuli stand out from among other items or locations. Our brain has evolved to rapidly compute salience in an automatic manner and in real-time over the entire visual field. Visual attention is then attracted towards salient visual locations. The core of visual salience is a bottom-up, stimulus-driven signal that announces "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 or even sometimes overridden by top-down, user-driven factors. Thus, a lone red object in a green field will be salient and will attract attention in a bottom-up.

In addition, if you are looking through a child's toy bin for a red plastic dragon, amidst plastic objects of many vivid colors, no one color may be especially salient until your top-down desire to find the red object renders all red objects, whether dragons or not, more salient. Visual salience is sometimes carelessly described as a physical property of a visual stimulus. It is important to remember that salience is the consequence of an interaction of a stimulus with other stimuli, as well as with a visual system (biological or artificial). As a straight-forward example, consider that a color-blind person will have a dramatically different experience of visual salience than a person with normal color vision, even when both look at exactly the same physical scene.

The basic principle behind computing salience is the detection of locations whose local visual attributes significantly differ from the surrounding image attributes, along some dimension or combination of dimensions. This significant difference could be in a number of simple visual feature dimensions which are believed to be represented in the early stages of cortical visual processing: color, edge orientation, luminance, or motion direction. Wolfe and Horowitz provide a very nice review of which elementary visual features may strongly contribute to visual salience and guide visual

search. A simple framework to think about how salience may be computed in biological brains has been developed over the past three decades. According to the framework, incoming visual information is first analyzed by early visual neurons, which are sensitive to the various elementary visual features of the stimulus. This analysis, operated in parallel over the entire visual field and at multiple spatial and temporal scales, gives rise to a number of cortical feature maps, where each map represents the amount of a given visual feature at any location in the visual field. Within each of the feature maps, locations which significantly differ from their neighbors are highlighted, as further discussed below. Finally, all highlighted locations from all feature maps combine into a single saliency map which represents a pure salience signal that is independent of visual.

According to several models, the relative contributions of different feature maps to the final saliency map are dependent upon the current behavioral goals and subjective state of the observer. In the absence of any particular task, such as, for example, during casual viewing, attention is drawn towards the most salient locations in the saliency map, as detected, for example, via a winner-take-all mechanism. This, in turns, triggers motor actions which direct the eyes and the head towards salient visual locations.

The essence of salience lies in enhancing the neural and perceptual representation of locations whose local visual statistics significantly differ from the broadly surrounding image statistics in some behaviorally relevant manner. This basic principle is intuitively motivated as follows. Imagine a simple search array as depicted below, where one bar pops-out because of its unique orientation. Now imagine examining a feature map which is tuned to stimulus intensity (luminance) contrast: because there are many white bars on a black background, early visual neurons sensitive to local intensity contrast will respond vigorously to each of the bars.

Based on the pattern of activity in this map, in which essentially every bar elicits a strong peak of activity, one would be hard pressed to pick one location as being clearly more interesting and worthy of attention than all the others. Intuitively, hence, one might want to apply some normalization operator  $N(\cdot)$  which would give a very low overall weight to this map's contribution to the final saliency map. The situation is quite different when examining a feature map where neurons are tuned to local vertically oriented edges. In this map, one location (where the single roughly vertical bar is) would strongly excite the neural feature detectors, while all other locations would elicit much weaker responses. Hence, one location clearly stands out and hence becomes an obvious target for attention. It would be desirable in this situation that the normalization operator  $N(\cdot)$  gives a high weight to this map's contribution to the final saliency map.

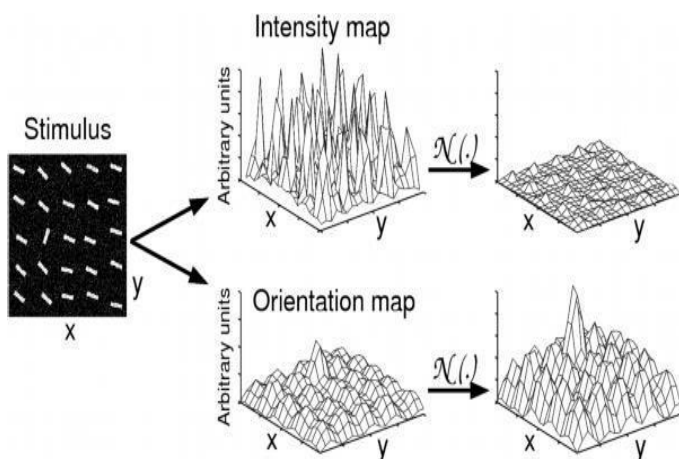


Fig 1: Saliency depends on context and on how unique of a response is elicited by a given item

As for comparatively low-complexity images, they usually have definitely separated foreground objects and background regions. Therefore we are able to search

for their saliency maps through first down sampling the input image signal properly and then conducting local comparisons. Conversely, it is not easy to clearly distinguish foreground and background layers in images of comparatively high complexity. In this regard, it was found that using global-based strategy is more suitable for detecting salient regions.

The CWS technique is established based on this inspiration, namely computing image complexity feature to weight and combine local and global features. Given an input visual signal, we first introduce how to estimate image complexity. The free energy based brain theory is proved to have a very close relation to human sensation of visual saliency and quality. As a consequence, the free energy principle is used for measuring the image complexity. To be more concretely, the free energy principle can be regarded as a unified brain theory combining popular brain theories in biological and physical sciences.

Essentially, this principle makes a basic premise that the human cognitive process is controlled by a so-called internal generative model in the brain. Using this internal generative model, the brain is capable of analyzing and predicting the given scene in a constructive way. In other ways, we can also consider this way to be a probabilistic model consisting of a likelihood term and a prior term.

## II.EXISTING WORK

As a powerful statistical procedure in data analysis, PCA (principal component analysis) is fully exploited to convert color space and produce compact patch representation. Images are first converted to linearly uncorrelated channels and divided into non overlapped patches. Then the patches are represented by the coefficients of principal components using PCA analysis. Additional experiments on salient object detection and image retargeting shows that the proposed model can achieve better performance than traditional models. Visual attention has been extensively studied recently due to its wide applications in image adaptation, quality assessment, contrast enhancement, and video compression. Numerous saliency models have been proposed by psychologists, neurophysiologists and computer scientists to imitate human visual attention. According to the motivation of inferring visual attention, saliency models can be categorized into top-down models and bottom-up models. Top-down models are task-driven while bottom-up models are data-driven and more related to the nature of Human Visual System (HVS). We focus on bottom-up model in this project.

## III.PROPOSED WORK

We used the Weighted Principal component analysis, because it reduces the weight of the image therefore time consumption will be reduced. The WPCA which is implemented using discrete wavelet transform. It provides high resolution of images when compared to the PCA. It can reduce the more dimensions of images and provide a vector PSNR ratio. This system is proposed based on

WPCA analysis in this paper. By fully exploiting the power of WPCA, the color channels are de-correlated and a compact representation of patch is obtained. And then visual distinctiveness are detected both locally and globally based on this representation. Local distinctiveness measures center-surround contrast while global distinctiveness evaluates the rarity compared to the entire image. Experimental results demonstrate that the WPCA-based color space conversion and patch representation can improve the accuracy of human fixations prediction. And the proposed algorithm achieves superior accuracy against the mainstream algorithms on predicting human fixations.

IV. ALGORITHM

Weighted Principal component analysis (WPCA) Weighted principal component analysis (WPCA) is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called

BLOCK DIAGRAM



principal components. The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. The resulting vectors are an uncorrelated orthogonal basis set. WPCA is sensitive to the relative scaling of the original variables.

WPCA is sensitive to the scaling of the variables. If we have just two variables and they have the same sample variance and are positively correlated, then the WPCA will entail a rotation by 45° and the "loadings" for the two variables with respect to the principal component will be equal. But if we multiply all values of the first variable by 100, then the first principal component will be almost the same as that variable, with a small contribution from the other variable, whereas the second component will be almost aligned with the second original variable. This means that whenever the different variables have different units (like temperature and mass), WPCA is a somewhat arbitrary method of analysis. (Different results would be obtained if one used Fahrenheit rather than Celsius for example.) Note that Pearson's original paper

was entitled "On Lines and Planes of Closest Fit to Systems of Points in Space" – "in space" implies physical Euclidean space where such concerns do not arise. One way of making the WPCA less arbitrary is to use variables scaled so as to have unit variance, by standardizing the data and hence use the autocorrelation matrix instead of the auto covariance matrix as a basis for WPCA. However, this compresses (or expands) the fluctuations in all dimensions of the signal space to unit variance. Mean subtraction is necessary for performing WPCA to ensure that the first principal component describes the direction of maximum variance. If mean subtraction is not performed, the first principal component might instead correspond more or less to the mean of the data. A mean of zero is needed for finding a basis that minimizes the mean square error of the approximation of the data. Mean-centering is unnecessary if performing a principal components analysis on a correlation matrix, as the data is already centered after calculating correlations. Correlations are derived from the cross-product of two standard scores (Z-scores) or statistical moments (hence the name: Pearson Product-Moment Correlation). Also see the article by Kromrey and Foster-Johnson on "Mean-centering in Moderated Regression: Much Ado About Nothing".

An auto encoder neural network with a linear hidden layer is similar to WPCA. Upon convergence, the weight vectors of the K neurons in the hidden layer will form a basis for the space spanned by the first K principal components. Unlike WPCA, this technique will not necessarily produce orthogonal vectors. WPCA is a popular primary technique in pattern recognition. It is not, however, optimized for class separability. An alternative is the linear discriminant analysis, which does take this into account.

In this proposed system WPCA analysis is used to facilitate the balanced operation on each color channel and avoid the over-emphasis on redundant information caused by redundant color channels, the image is converted to linearly uncorrelated channels. In order to distinguish salient patches from non salient ones, a robust and compact patch is represented by WPCA. The K most principle components are extracted to capture important features while the less principle components are discarded to eliminate the impacts of noises.

Given the compact patch resenatation, we introduce two schemes to evaluate patch distinctiveness both locally and globally. Local distinctiveness focuses on center-surround contrast while global distinctiveness deals with rarity of WPCA coefficients. The combination of two schemes provides a comprehensive measurement of saliency. The proposed model achieves better performance than state-of-the-art algorithms.

PCA Use for Image Compression

Data volume reduction is a common task in image processing. There is a huge amount of algorithms based on various principles leading to the image compression. Algorithms based on the image color reduction are mostly loss but their results are still acceptable for some applications. The image transformation from color to the

gray-level (intensity) image  $I$  belongs to the most common algorithms. Its implementation is usually based on the weighted sum of three color components R, G, B according to relation

$$I = w_1R + w_2G + w_3B$$

The R, G and B matrices contain image color components, the weights we were determined with regards to the possibilities of human perception. The PCA method provides an alternative way to this method.

The idea is based on Equation

$$x = ATy + mx$$

Where the matrix  $A$  is replaced by matrix  $A_l$  in which only  $l$  largest (instead of  $n$ ) Eigen values are used for its forming. The vector  $\hat{x}$  of reconstructed variables is then given by relation

$$\hat{x} = A_l^T y + mx$$

True-color images of size  $M \times N$  are usually saved in the three-dimensional matrix  $P$  with size  $M \times N \times 3$  which means that the information about intensity of color components is stored in the 3 given planes.

## V.CONCLUSION

A patch-wise saliency detection algorithm is proposed based on WPCA analysis in this paper. By exploiting the power of WPCA, the color channels are completely decorrelated. This algorithm is capable of reducing the dimension of the image. It consists of color texture and luminescence which helps us to identify the depth of an image. User can measure the saliency by integrating three elements as follows: the dissimilarities between image patches, which were evaluated in the reduced dimensional space, the spatial distance between image patches and the central bias. A mechanism of pixel reduction can be done by, indicating a bias for human fixations to the center of the image. And then, a compact WPCA-based representation of patch is proposed, upon which, the visual distinctiveness is defined both locally and globally. Local distinctiveness detects patches with high center-surround contrast while global distinctiveness highlights patches with high rarity compared to the entire image. Experimental results demonstrate that the PCA-based color space conversion and patch representation largely improve the accuracy of human fixations prediction. The proposed algorithm achieves superior accuracy against the state-of-the-art algorithms and additional experiments on salient object detection and image retargeting further demonstrate the superiority of the proposed PCA-based model over traditional models. The future works of this paper may include incorporating top-down factors and extending the proposed model into spatial-temporal domain for video saliency detection.

## REFERENCES

- [1] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, "Visual saliency detection with free energy theory," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp.1552– 1555, Oct. 2015.
- [2] J. Liu, X. Yang, G. Zhai, and C. Chen, "Visual saliency model based on minimum description length," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Montreal, QC, Canada, May 2016, pp. 990– 993.
- [3] A. Borji, M.-M. Cheng, H. Jiang, and J. Li, "Salient object detection: A survey," *arXiv preprint arXiv:1411.5878*, 2014.
- [4] J. Zhang et al., "Minimum barrier salient object detection at 80 FPS," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit. (CVPR)*, 2015, pp. 1404– 1412.
- [5] J. Kim, D. Han, Y.-W. Tai, and J. Kim, "Salient region detection via high-dimensional color transform," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit. (CVPR)*, Columbus, OH, USA, 2014, pp. 883–890.
- [6] Z. Liu, W. Zou, and O. Le Meur, "Saliency tree: A novel saliency detection framework," *IEEE Trans. Image Process.*, vol. 23, no. 5, pp. 1937–1952, May 2014.
- [7] M.-M. Cheng, N. J. Mitra, X. Huang, P. H. S. Torr, and S.-M. Hu, "Global contrast based salient region detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 569–582, Mar. 2015.
- [8] X. Zhou, Z. Liu, G. Sun, L. Ye, and X. Wang, "Improving saliency detection via multiple kernel boosting and adaptive fusion," *IEEE Signal Process. Lett.*, vol.23, no. 4, pp. 517–521, Apr. 2016.
- [9] W. Zhu, S. Liang, Y. Wei, and J. Sun, "Saliency optimization from robust background detection," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.(CVPR)*, Columbus, OH, USA, 2014, pp. 2814–2821.
- [10] J. Sun, J. Xie, J. Liu, and T. Sikora, "Image adaptation and dynamic browsing based on two-layer saliency combination," *IEEE Trans. Broadcast.*, vol. 59, no. 4, pp. 602–613, Dec. 2013.
- [11] H.-C. Shih, "A novel attention-based key-frame determination method," *IEEE Trans. Broadcast.*, vol. 59, no. 3, pp. 556–562, Sep. 2013.
- [12] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang, "Saliency detection via-based manifold ranking," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit. (CVPR)*, Portland, OR, USA, 2013, pp.3166–3173.