Saliency Based Object Detection and Enhancements in Static Images

Rehan Mehmood Yousaf[1(✉)], Saad Rehman[2], Hassan Dawood[1], Guo Ping[3], Zahid Mehmood[1], Shoaib Azam[4], and Abdullah Aman Khan[5]

[1] Software Engineering Department, UET, Taxila, Pakistan
Rehanmehmoodyousaf@gmail.com,
{hassan.dawood,zahid.mehmood}@uettaxila.edu.pk
[2] College of Electrical and Mechanical Engineering, NUST, Islamabad, Pakistan
saadrehman@ceme.nust.edu.pk
[3] Image Processing and Pattern Recognition Laboratory, Beijing Normal University, Beijing, China
pguo@bnu.edu.cn
[4] Gwangju Institute of Science and Technology, Gwangju, Korea
shoaibazam@gist.ac.kr
[5] HITEC University, Taxila, Pakistan
abdkhan@hotmail.com

Abstract. Human visual system always focuses on the salient region of an image. From that region the salient features are obtained and can be collected by generating the saliency map. Natural statistics measures are used to measure the saliency from data collection of natural images. ICA filters are used to generate the saliency map that can blur the image. We have improved it by using different techniques like edge detection and morphological operations. By applying these algorithms we have successfully reduced the blur in images. That makes the salient objects more prominent by sharpening the edges. Proposed method is also compared with the state-of-the-art method like Achanta model.

Keywords: Saliency · Edge detection · Morphological image processing · AUC score

1 Introduction

The human visual system is enriched by modern technology and lot of work has been done to improve the saliency map. Improvement in saliency map can increase the object detection and tracking. Different techniques are being used by the latest researchers such as particle filters, log maps, background subtraction, feature extraction and feature description. Feature extraction and description are used for image matching and recognition [30–33]. Saliency is basically making the most prominent features salient so that the machine visual system can recognize the important information in an image. Saliency can be done through different techniques that includes difference of Gaussian, independent component analysis filters [2], spectral residual and spatial-temporal [26].

K. Kim and N. Joukov (eds.), *Information Science and Applications 2017*,
Lecture Notes in Electrical Engineering 424, DOI 10.1007/978-981-10-4154-9_14

Recently many models have tried to explain the above mentioned problem for instance Hou and Zhang proposed a method which deals with spectral residual method by using the Fourier transform [26]. Then Achanta *et al*. proposed a method based on colour and luminance of the image and obtained the well-defined regions of the image [25]. Natural statistics also plays an important role to obtain the saliency map. The independent component analysis filter is used to obtain linear features, used in the saliency algorithm, obtained by applying on the natural images. By using the edge detection and morphological operations we have improved the performance of the system. The rest of the paper is organised as: Sect. 2 briefly explains the ICA. In Sect. 3 we explained our proposed methodology. Sections 4 and 5 are about edge detection and the morphological operations. In Sect. 6, we have discussed the results and compared with baseline method.

2 Independent Component Analysis

Independent component analysis [2] is used to recover independent signals from the measured signals. The measured signals are a linear combination of independent signals. Therefore an equal number of independent and measured signals are obtained.

Independent component analysis is defined as

$$Xi = a_1 s_1 + a_2 s_2 + \ldots + a_n s_n \tag{1}$$

Or in matrix form

$$X = AS \tag{2}$$

Where X_i belongs to every measured signal, S is independent signals and A is n x n matrix called the mixing matrix. An alternative form of (2) can be obtained if and only if matrix A is invertible that is

$$W = A^{-1} \tag{3}$$

So,

$$S = WX \tag{4}$$

This means that each independent signal S_i can be expressed as a linear combination of measured signals. So, by estimating the W, independent signal S can be obtained. We assume that each signal is a random variable. The Central Limit Theorem states that if the sum of several independent random variables, such as those in S, tends towards a Gaussian distribution [3]. So $x_i = a_1 s_1 + a_2 s_2$ is more Gaussian than either s_1 or s_2. The Central Limit Theorem also implies that if the combinations of the measured signals in X with minimal Gaussian properties are obtained, then that signal will be one of the independent signals. To achieve this we have to measure the nongaussianity of WX [3, 22–24]. To measure the nongaussianity the Negentropy approximation is used.

We have applied the FASTICA [2] algorithm as used by Lingyun Zhang and Tim K. Marks [1] and enhanced the images by using different operations; hence improved the salient features in saliency map to a noticeable result.

The following are results and 3D plots of ICA filters applied on three different images.

3 Proposed Methodology

Saliency map is generated by ICA has some drawbacks like edges are blur so salient features are not easy to obtain. We have improved the saliency map generated by ICA. Sobel operator is applied on the saliency map that was generated by ICA. Morphological operator (dilation) is then used to improve the edges by which salient features can be obtained efficiently. The visual results clearly show that proposed methodology improves the performance of the system.

Flow chart of proposed method is as follows (Fig. 2).

4 Edge Detection

Edge detection technique enhances the edges of an image by sharpening the image edges, Sobel operator is found to be good edge detector [10]. The Sobel operator is a discrete function computing the gradient of the intensity in an image. The operator uses two 3×3 kernels which are convolved with the original image to calculate approximations of the derivatives-one for horizontal changes, and one for vertical [10]. Image is represented as B, and H_x and H_y are two filters to compute the horizontal and vertical derivative approximations, the computations are as follows:

$$\mathbf{H_x} = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} * \mathbf{B} \quad \mathbf{H_Y} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \mathbf{B}$$

Where * is the 2-dimensional convolution operation.

Since the Sobel kernels can be decomposed as the products of an averaging and a differentiation kernel, they compute the gradient with smoothing. For example, H_x can be written as

$$\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}$$

The x-coordinate is defined here as increasing in the “right”-direction, and the y-coordinate is defined as increasing in the “down”-direction. At each point in the image, the resulting gradient approximations can be combined to give the gradient magnitude, using

$$\mathbf{H} = \sqrt{\mathbf{H}_x^2 + \mathbf{H}_y^2}$$

Gradient direction can be calculated as:

$$\boldsymbol{\alpha} = \mathbf{atan2}\,(\mathbf{H}_y,\ \mathbf{H}_x)$$

Where, for example, α is 0 for a vertical edge that is darker on the right side.

5 Morphological Operation

Morphological operations are used for extracting meaningful components from the images. There are different operations like dilation, erosion, opening, closing etc.

Dilation operator is used in our experiments [11]. Dilation [11] is used to thicken or grow objects in an image. The dilation process takes two pieces of data as inputs. The first is the original image and the second one is structuring element (also known as kernel). Structuring element is the one through which thickening process is controlled in dilation operation. The line structuring element is applied on the image pixels from the start till end. The SE is applied in the form of the line in the image every time and a change in the pixel values according to the SE i.e. the change appears in places where the line affects the pixel and finally thicken the points we wanted and making the edges thicker and clear.

Dilation function is defined in term of set operation. The dilation of C and D is defined as

$$C \oplus D = \left\{ z | (D')_z \cap C \neq \phi \right\}$$

Where ϕ is the empty set, D is structuring element and C is the binary image. In other words, dilation of C and D is the set consisting of all elements of D' such that its origin remain in C.

Morphological operation (Dilation) is applied to edge detection images to enlarge the boundaries of the regions of the salient features.

6 Results

To verify the proposed algorithm and with base line algorithm the image dataset is used as in [21]. For experiments 1000 images randomly selected from 10000 images and computed the results. Some images are shown in Fig. 3 after applying our proposed method.

6.1 Results by Applying Algorithm on the Dataset

7 Comparison Between ICA, Edge Detection Technique and Morphological Operation

Saliency map of original image is generated by applying ICA filter image. These saliency maps are shown in Fig. 1, with their 3D plots that clearly show the most prominent and salient features in the images. From figure it can be shown that the edges are not clear so sobel and dilation operators are used to improve the image quality. The results clearly show that the edges are sharper in whole image. This gets us to the point nearer to the object detection rather just projector out the prominent features. The histogram of the sobel operator shows us the peaks referring to the edges of the object and filter out purely the object. Results of the sobel operator is improved by applying another operation i.e. dilation; that provide the perfect results for object detection as shown in Fig. 3. The uncompleted edges have been thickened and completed by dilation. The histogram of dilation images gives the clear peaks and also those peaks that are blurred before. So, it is concluded after comparison between three of the techniques that when applied in hybrid these three techniques gives us the clear detected object (Fig. 4).

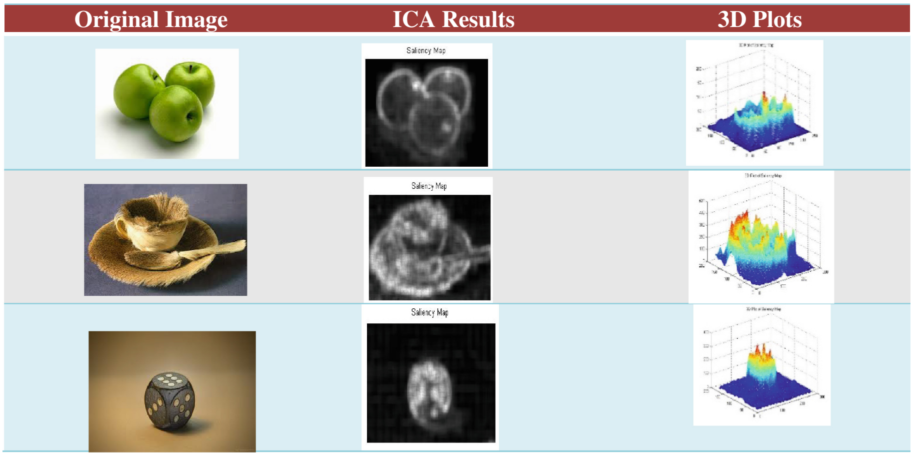


Fig. 1. Saliency map using ICA filters.

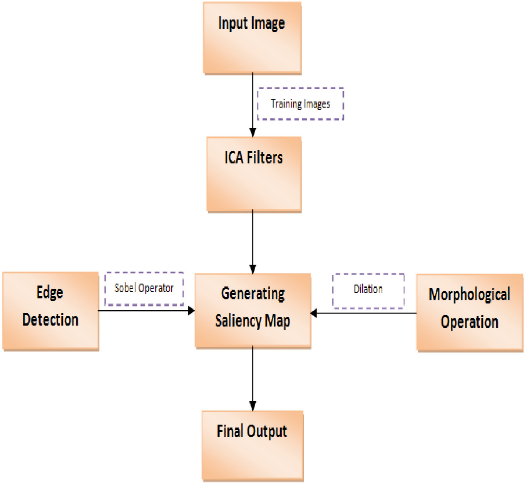
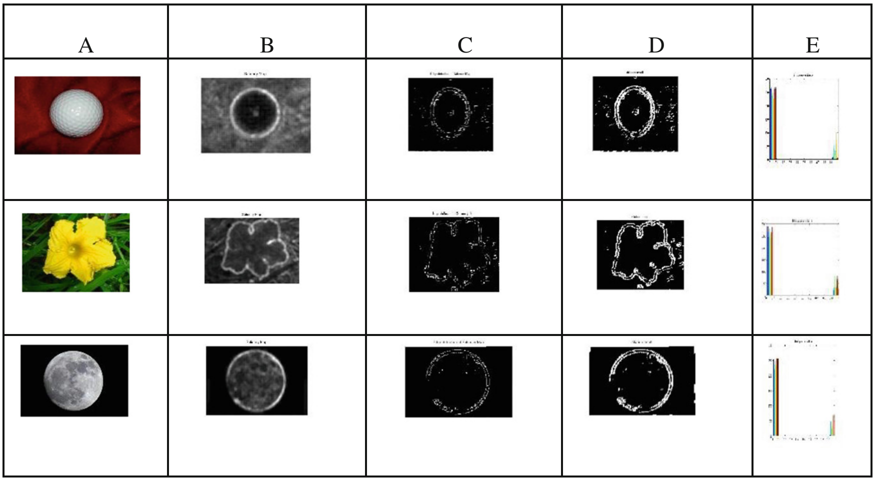


Fig. 2. Flow Diagram

Fig. 3.

The results can be visualized as under:



(a) Original Images (b) Saliency maps (c) Edge detection (d) Morphological operation (e) Histogram

Fig. 4.

8 Comparison of Two Computational Models and the AUC Score

The comparison between two models, which are the Achanta et al model and the ICA based saliency model, are based on the testing of area under the ROC curve score.

The area under the ROC curve needs a dataset that has been tested on human visual system. The dataset are used for calculating an AUC consists of random number of observers in a free viewing scenario on 135 different images [28]. The scenario is kept as to obtain the best possible results. The data for these random numbers of observations are viewed on is viewing 135 different images and their point of focus is computed. These values are used to compute the eye fixation map. Then the Achanta [28] model is used to compute the saliency maps of the same 135 images. These saliency maps are then compared with the eye fixation results and a score for each image is computed. Similarly the ICA based saliency maps are also compared to the eye fixation results and score is computed again. The score limits from 0 to 1. If the comparison results in score near to 1 the similarity is maximum thus the result is good and if it's near 0 then the similarity is minimum. We have only computed the comparison for first 8 images

The results of comparison between the AUC score is as under (Fig. 5):

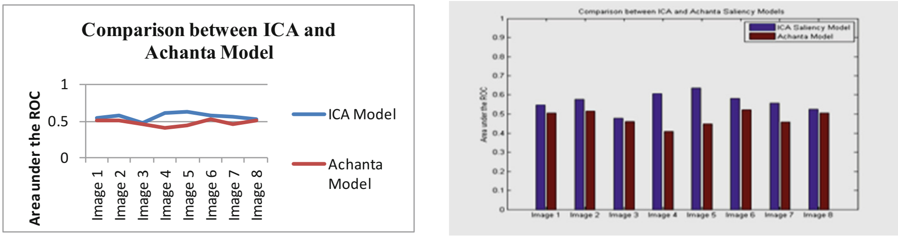


Fig. 5.

9 Conclusions

To generate the saliency map, ICA filter is used that gives the linear features used in saliency algorithm. To enhance the salient region edge detection technique with sobel operator is used which shows the accurate region of the saliency map and distinguishes from background. Using the same edge detection results, morphological operation (dilation) is applied that has brighten the edges, improves the lines and curves of edge detection result. Dilation is basically used to complete the incomplete boundaries of the region and thickens it contains. After applying all these operations the results clearly shows the effects, improving the saliency map and the edge detection has enhanced the results of saliency map. The morphological technique highlights the final result of the feature detection. The AUC ROC score also proves the models results better than many models as baseline model.

References

1. Zhang, L., Marks, T.K., Tong, M.H., Shan, H., Cottrell, G.W.: SUN: a bayesian framework for saliency using natural statistics. J. Vis. **8**(7), 1–20 (2008)
2. Hyvarinen, A., Oja, E.: A fast fixed-point algorithm for independent component analysis. Neural Comput. **9**(7), 1483–1492 (1997)
3. Bell, A.J., Sejnowski, T.J.: An information-maximization approach to blind separation and blind deconvolution. Neural Comp. **7**, 1129–1159 (1995)
4. Borji, A., Itti, L.: State-of-the-art in visual attention modeling. IEEE Trans. Pattern Anal. Mach. Intell. **35**, 30–43 (2010)
5. Itti, L., Koch, C., Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. IEEE Trans. Pattern Anal. Mach. Intell. **20**(11), 1254–1259 (1998)
6. Judd, T., Ehinger, K., Durand, F., Torralba, A.: Learning to predict where humans look. In: IEEE 12th International Conference on Computer Vision (2009)
7. Bruce, N., Tsotsos, J.: Saliency based on information maximization. In: Weiss, Y., Schölkopf, B., Platt, J. (eds.) Advances in Neural Information Processing Systems, vol. 18, pp. 155–162. MIT Press, Cambridge (2006)
8. Umbaugh, S.E.: Digital Image Processing and Analysis: Human and Computer Vision Applications with CVIPtools, 2nd edn. CRC Press, Boca Raton (2010). ISBN 9-7814-3980-2052

9. Jähne, B., Scharr, H., Körkel, S.: Principles of filter design. In: Handbook of Computer Vision and Applications. Academic Press, Cambridge (1999)
10. Sobel, I.: History and Definition of the Sobel Operator (2014)
11. Dougherty, E.R.: An Introduction to Morphological Image Processing (1992). ISBN 0-8194-0845-X
12. Efford, N.: Digital Image Processing: A Practical Introduction Using JavaTM. Pearson Education, Upper Saddle River (2000)
13. Gonzalez, R.C., Woods, R.E.: Digital Image Processing, 3rd edn. Prentice Hall, Upper Saddle River (2008). ISBN 0-13-168728-8
14. Haralick, R., Shapiro, L.: Computer and Robot Vision, vol. 1, Chap. 5. Addison-Wesley Publishing Company, Boston (1992)
15. Jain, A.: Fundamentals of Digital Image Processing. Prentice-Hall, Upper Saddle River (1986)
16. Itti, L., Koch, C.: Computational modeling of visual attention. Nat. Rev. Neurosci. **2**(3), 194–203 (2001)
17. Avraham, T., Lindenbaum, M.: Esaliency: Meaningful attention using stochastic image modeling. IEEE Trans. Pattern Anal. Mach. Intell. **99**(1) (2009)
18. Bruce, N.D.B., Tsotsos, J.K.: Saliency, attention, andvisual search: an information theoretic approach. J. Vis. **9**(3), 1–24 (2009)
19. Walther, D., Itti, L., Riesenhuber, M., Poggio, T., Koch, C.: Attentional selection for object recognition — a gentle way. In: Bülthoff, Heinrich, H., Wallraven, C., Lee, S.-W., Poggio, Tomaso, A. (eds.) BMCV 2002. LNCS, vol. 2525, pp. 472–479. Springer, Heidelberg (2002). doi:10.1007/3-540-36181-2_47
20. http://www.ece.rice.edu/~dhj/courses/elec531/notes.pdf
21. Cheng, M.-M., Mitra, N.J., Huang, X., Hu, S.-M.: SalientShape: group saliency in image collections. Visual Comput. (2013)
22. Hyvarinen, A., Oja, E.: Independent component analysis: algorithms and applications. Neural Netw. **13**, 411–430 (2000)
23. Hyvarinen, A., Karhunen, J., Oja, E.: Independent Component Analysis, 481 pages. John Wiley & Sons, Toronto (2001)
24. McKeown, M.J., Makeig, S., Brown, G.G., Jung, T., Kindermann, S.S., Bell, A.J., Sejnowski, T.J.: Analysis of fMRI data by blind separation into independent spatial components. Hum. Brain Mapp. **6**, 160–188 (1998)
25. Achanta, R., Hemami, S., Estrada, F., Susstrunk, S.: Frequency-tuned salient region detection. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2009, pp. 1597–1604, June 2009
26. Hou, X., Zhang, L.: Saliency detection: a spectral residual approach, Department of Computer Science, Shanghai Jiao Tong University
27. Achanta, R., Estrada, F., Wils, P., Süsstrunk, S.: Salient region detection and segmentation. In: Gasteratos, A., Vincze, M., Tsotsos, John, K. (eds.) ICVS 2008. LNCS, vol. 5008, pp. 66–75. Springer, Heidelberg (2008). doi:10.1007/978-3-540-79547-6_7
28. Mancas, M., Le Meur, O.: Memorability of natural scenes: the role of attention. In: Proceedings of the International Conference on Image Processing (IEEE ICIP 2013), Melbourne, Australia, September 15–18 (2013)
29. Dawood, H., Dawood, H., Guo, P.: A hybrid image feature descriptor for classification. In: 2015 11th International Conference on Computational Intelligence and Security, (CIS 2015) (EI), pp. 58–61 (2015)
30. Hassan, D., Dawood, H., Guo, P.: Removal of random-valued impulse noise by Khalimsky grid. In: Asia Pacific Conference on Multimedia and Broadcasting (APMediaCast 2015) (EI), vol. 1 (2015)

31. Wang, Y., Dawood, H., Guo, P., Yin, Q.: A comparative study of different feature mapping methods for image annotation. In: The Seventh International Conference on Advanced Computational Intelligence (ICACI 2015). (EI) (2015)
32. Dawood, H., Dawood, H., Guo, P.: Texture image classification with improved weber local descriptor. In: Rutkowski, L., Korytkowski, M., Scherer, R., Tadeusiewicz, R., Zadeh, Lotfi, A., Zurada, Jacek, M. (eds.) ICAISC 2014. LNCS (LNAI), vol. 8467, pp. 684–692. Springer, Heidelberg (2014). doi:10.1007/978-3-319-07173-2_58
33. Dawood, H., Dawood, H., Guo, P.: Global matching to enhance the strength of local intensity order pattern feature descriptor. In: Guo, C., Hou, Z.-G., Zeng, Z. (eds.) ISNN 2013. LNCS, vol. 7951, pp. 497–504. Springer, Heidelberg (2013). doi:10.1007/978-3-642-39065-4_60