Saliency based Object Detection and Enhancements in Static Images

By

REHAN MEHMOOD YOUSAF

2011-NUST-MS PHD-ComE-23

Supervisor

Dr. SAAD REHMAN

Department Of Computer Engineering

A thesis submitted in the partial fulfillment of the requirements for the degree of Masters of Computer Engineering (MSCE)

College of Electrical and Mechanical Engineering (E&ME)
National University of Sciences and Technology (NUST)
March 2014

IN THE NAME OF ALLAH
THE MOST GRACIOUS
THE MOST MERCIFUL

College of Electrical and Mechanical Engineering

A center of excellence for quality education and research

Certificate

Certified that the contents of thesis document titled **"(Saliency based Object Detection and Enhancements in Static Images)"** submitted by Rehan Mehmood Yousaf has been found satisfactory for the requirement of degree.

Advisor: Dr.SaadRehman

Committee Member1: Dr. Umar Munir

Committee Member2: Dr. Ali Hassan

Committee Member3: Dr. Farhan Riaz

DEDICATION

This thesis is dedicated to my parents, my wife and most of all my friends who have always supported me and last but not the least my advisor who have given his precious time in guiding me to the best.

CERTIFICATE OF ORIGINALITY

I hereby declare that the research paper titled **"Saliency based Object Detection and Enhancements in Static Images"** is my own hard Work and to the best of my knowledge. It holds no materials which was awhile ago distributed or was composed by someone else, nor material which to a certain degree has been acknowledged for the honor of any degree or certificate at E&ME or whatever possible instruction organization, with the exception of where due affirmation, is made in the theory. Any viable commitment made to the examination by others, with whom I have worked at E&ME or somewhere else, is expressly recognized in the postulation.

Author Name: Rehan Yousaf

Signature: _______________

Acknowledgement

First of all I would like to thank almighty Allah on bestowing his blessings on me and making me able to do what I wanted to achieve in my work.

I would specially like to thank my supervisor Dr. Saad Rehman was consistent efforts in guiding me to right path. Many times during my thesis work I use to get off the track but my supervisor guided and helped me to his best. Most of the time he took time out from his busy schedule and helped me and so many times I got stuck in places from where I could not have moved on if Dr. Saad wouldn't be there. Thanks to you sir!!!!

I would like to mention here my parents praying for me constantly and supporting me regularly motivating me without them I won't be standing where I am standing today

My friends helped me technically and morally throughout my work and been my replacement at many occasions.

My committee members also guided me in different aspects and future work throughout the whole procedure and the motivation from Dr. Ali Hassan gathered me together and made me what I m today.

Lastly I would like to thank my wife. She was a true supporter and understood me in all times. She took my responsibilities and spared me for my thesis work. Without her I think it was impossible for me manage my thesis.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ICA	Independent Component analysis
DIP	Digital Image Processing
S.E	Structuring Element
AUC	Area Under The Curve
ROC	Receiver Operating Characteristics
GBVS	Graphical Based Visual Saliency
FT	Fourier Transform
PR	pattern Recognition
CV	Computer Vision
EEG	Electroencephalogram.

LIST OF FIGURES

ABSTRACT

Whenever we see an image the visual system focuses on the prominent features present in an image. These prominent features are known as the salient features which we are obtaining by generating the saliency map. As done before we are also using the natural statistics that is measuring the saliency from data collection of natural images.

We have used the ICA filters to generate the saliency map. Our proposed work is to improve the blur result given by the ICA filter by using different techniques. The results of our work are very clear. The prominent feature is improved by applying the techniques.

The bottom up saliency plays a vital role for the basic saliency map but the difference majorly occurs, when we improve the saliency map by applying different filters and edge detection techniques thus resulting in object detection and improving the performance of our saliency map.

Chapter 1 Introduction

This thesis focuses on partial object detection which includes three phases that is implementing ICA filter [2] for computing saliency map which is already done by Zhang [1]. The second phase is to improve the results of saliency map and the third phase is very important which is the comparison between saliency map [10] results and the proposed methodology.

1.1 Motivation

When I came across searching my thesis topic I already had interest in Digital Image Processing. Two more fields that are pattern recognition and computer vision had been also attractive for me when I was studying my courses so I decided to pick a topic from a field which I always admired.

Object detection nowadays has been much seen to be researched by different people and is very important as far as the application point of view is concerned. Many modern researches have been working on object detection problems in static images as well as in videos. I came across many techniques used for object detection such as particle filters [3] [4]. While doing my literature review I came across a technique called saliency map. This technique is being used in so many recent applications in the modern time.

Consequently discovering objects of interest has immediate advantages in various requisitions, particularly those that oblige content mindfulness. Feature layering is one such region. In situations like web-visits, system data transfer capacity reserve funds might be accomplished by adaptively doling out additional bits to confronts, which are more significant instead of foundation. Household robots and self-sufficient vehicles can distinguish protests in their environment for route and impact shirking. An alternate region is that of programmed picture re-focusing on, which is the undertaking of changing perspective proportions of pictures to suit diverse, showcase gadgets without contorting the paramount substance of the picture. Reconnaissance and security provisions can profit from programmed gatecrasher location. Programmed number plate discovery in vehicle picture databases can defend the security of managers or help computerized stopping frameworks. Prescription and science can immensely profit from programmed location of tumors or cell structures. Picture furthermore feature database and indexing frameworks can benefit tremendously from learning of picture content. Clearly, just about all parts of our current lives that include advanced pictures can exploit consequently discovering objects of interest.

1.2 Problem Statement

Computing the saliency map itself is a big task. The saliency map gives us the salient features or in other words we can say the salient object in an image. But the results according to me are vague. My point of view is that the saliency map results can be made more clear and obvious by applying different algorithms and reach to a point where we can include the results in detection of the object.

1.3 Proposed Solution

The proposed solution for improvement in saliency map results can be done by applying different image processing techniques and some noise removal algorithms which can give better results than the saliency map.

This will lead to three phase procedure which will include improvement of results using edge detection techniques [5] [6] and morphological operations [7] [8] in the first phase. The second phase will include their histograms [9] and the third phase will include the comparison between saliency results and the proposed solution.

1.4 Outline of the Report

The next parts of the report are organized in following manner: Chapter 2 gives the literature review of the saliency map generation. The 3rd chapter will explain the implementation of proposed methodology. In chapter 4, results and discussion. Chapter 5 will explain the future work and conclusion of the topic.

Chapter 2 Literature Review

2.1 ICA FILTERS

ICA filter [2][11]concept starts from a scenario which says that if you imagine a closed place where two people are speaking simultaneously and there are two microphones placed in that room. What will happen is that both the speakers' voices will be merged together and both the microphones will record the mixture of both the speakers. Both the microphones are placed at a certain distance at certain places. Now the point of focus is to know the speech signals separately. For this we have to estimate the sources s1 and s2. The certain parameters a_{11}, a_{12},a_{21} and a_{22} that depend on the distance of the microphones placed.

This can be illustrated by a simple diagram:

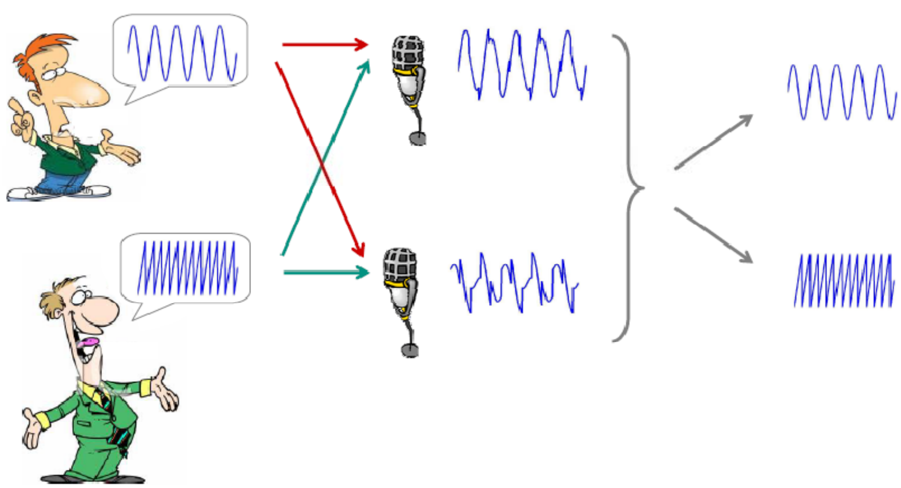


Fig 2.1

Elaborating more we can define the Independent Component analysis by the following diagram defining the sources and the recorded signals and the parameters:

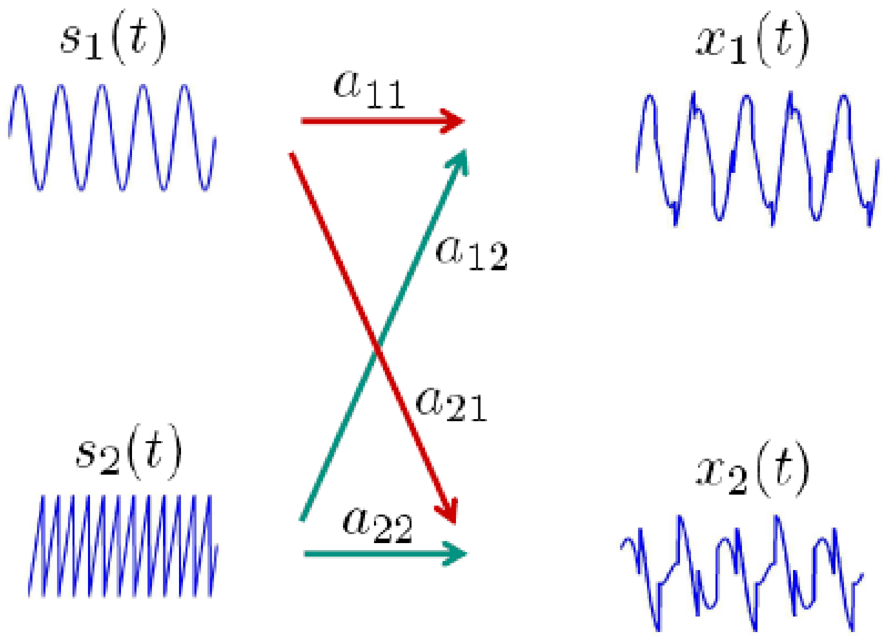


Fig 2.2

Assumptions

linearity
no delay
statistically independent sources

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} s_1(t) \\ s_2(t) \end{pmatrix}$$

Fig 2.3

Consider the waveforms in fig 2.4 and fig 2.5

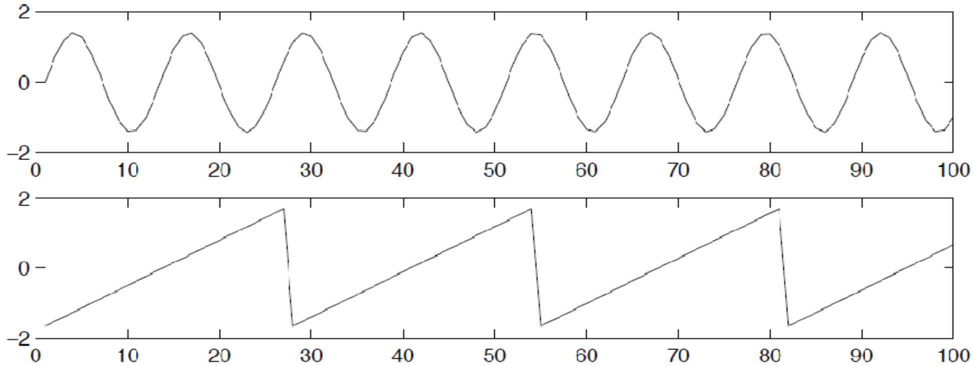


Figure 1: The original signals.

Fig 2.4

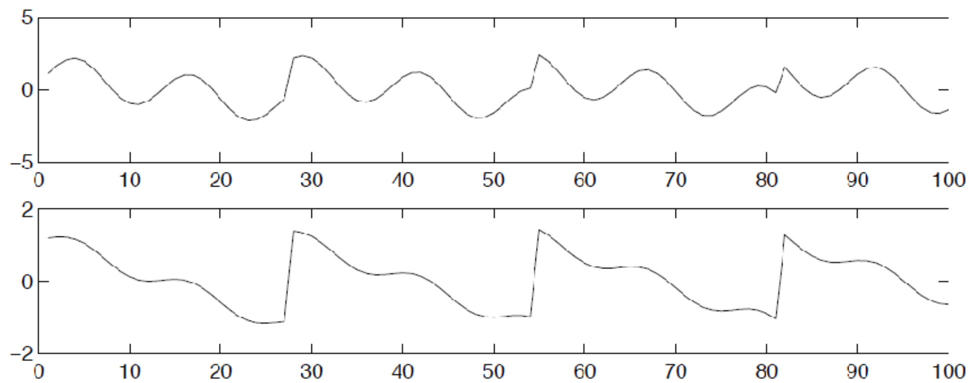


Figure 2: The observed mixtures of the source signals in Fig. 1.

Fig 2.5

Obviously the original speech signal would no look like in fig2.4 and also that in fig 2.5 for the mixed speech signal. But the task is to obtain the data in fig 2.5 using the data of fig 2.4.

The key point in this whole scenario is the a_{ij} factor. If we know what the a_{ij} actual value the problem becomes much easier. But if a_{ij} is unknown as in this case the problem becomes very much difficult.

Now simple solutions to find out the original signals we have to take an assumption that source s1 and source s2 are independent. If they are independent we can estimate a_{ij} by the procedure of independent component analysis [2].

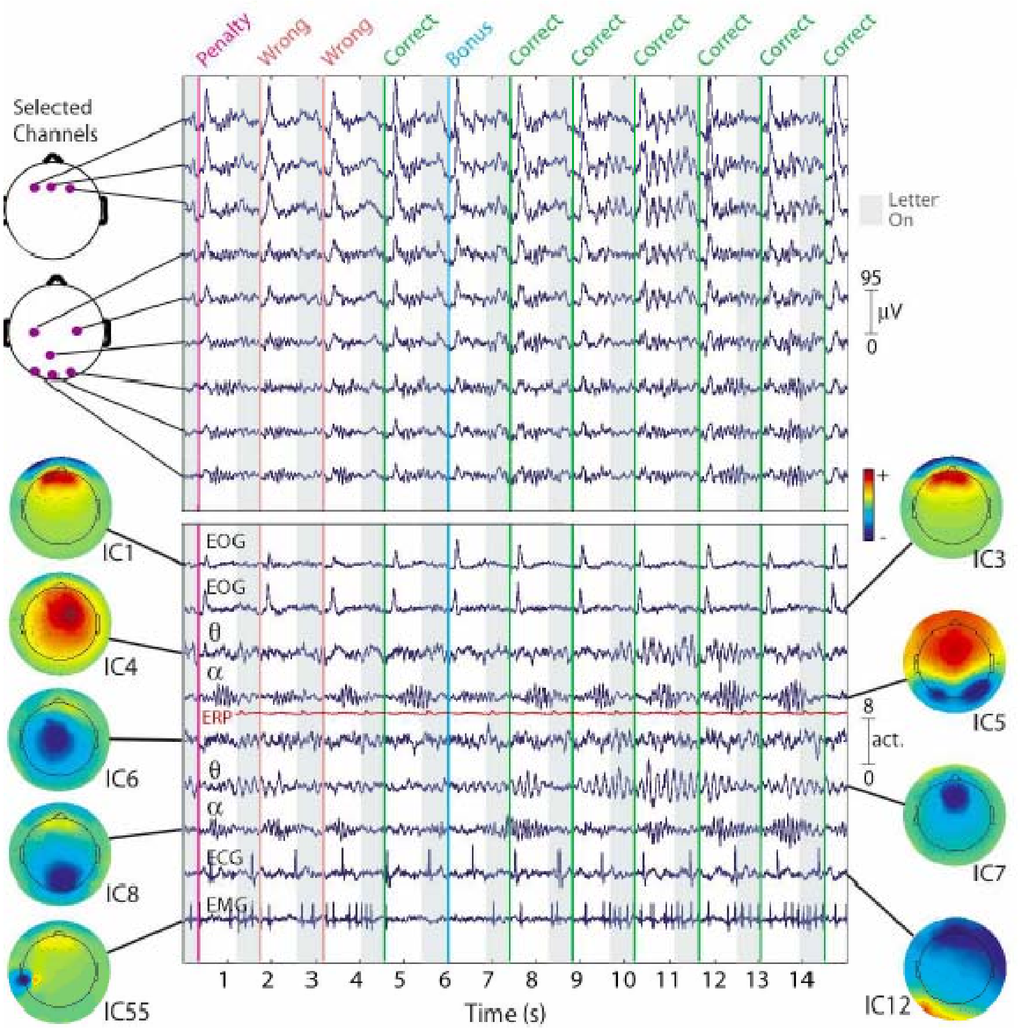


Fig 2.6

The above diagram is an ICA of EEG (electroencephalogram). The EEG of a brain gives a mixture of electrical potentials of the brain. But the issue is we want to get the independent component signals of the different part of the brain. In this matter ICA can play a vital role in separating out different signals of different parts of the brain from the mixture of signal we get.

Now let us generalize the terms in a simpler form to estimate the factors. Let us use vector matrix form for x_i (t),a_{ij} and s_i (t) [2] [11]. We call the first one in a general form X, A and S respectively.

So the final equation becomes

$$\mathbf{x} = \mathbf{As} \qquad (1)$$

This model of equation is finally called the independent component analysis.

As the bold lower case symbols denote vectors and bold upper case symbols denote matrices.

So x and s are vectors and A is a matrix. Which is also called the MIXING MATRIX [12] [13].

We have only information about x and both the other components are unknown. We have to estimate both the parameters A and s from the information of x.

Now starts the assumptions which are used to estimate the parameters. The first one is that the s_i components are statistically independent. We also have to assume that the mixing matrix A is a square matrix in some case we can ignore this. Now after estimating A we can compute its inverse matrix defined as W.

$$\mathbf{x} = \mathbf{Ws} \qquad (2)$$

Noise in this model is also assumed as nil although in many estimation models it is considered as included.

2.1.1 INDEPENDENCE:

Random variables R and S are independent if the conditional probability of any with respect to R is simply the probability of any. In different words, learning a benefit of S tells all of us nothing about R. This may be expressed since the equation:

$P(R|S) = P(R)$ Because $P(R|S) = P(R, S)/P(S)$, in which $P(R, S)$ would be the joint thickness function of any and S.

$$P(R, S) = P(R)*P(S) \qquad (3)$$

Another important feature of record independence is usually

mean (g1(R) g2(S)) = mean (g1(R)) mean (g2(S)) for just about any functions g1 as well as g2 as well as $R \neq S$. We could also consider the covariance among R as well as S.

$$Cov(R, S) = mean(R*S) - mean(R)*mean(S) \qquad (4)$$

For a random varying R, cov(R, R) is equal to the variance of any. If R and S are independent, then mean(R*S) =mean(R)*mean(S) along with the covariance will be zero. The covariance associated with two statistically independent variables is actually zero. The converse is just not always correct. Just because the covariance is usually zero may not necessarily mean R as well as S usually is independent. Even so, in the actual special event of Gaussian variables,

zero covariance may imply self-reliance. This feature of Gaussian variables is needed to come across columns associated with W in WX=S. Previously we all stated that every measured transmission in X can be a linear mix off the independent signals in S. The mixing matrix A new is invertible in a way that A-1=W. Each of the dependent elements in S may also be expressed as being a linear mix off the scored signals in X (S=WX).

The Central Limit Theorem claims that the sum of the several independent random variables, such since those in S, appears towards a new Gaussian supply.

So xi = a1s1 + a2s2 is usually more Gaussian as compared to either s1 or maybe s2. One example is, the sum of a pair of dice approximates a new Gaussian distribution using a mean associated with seven.

The Central Limit Theorem ensures that if we can find combining the scored signals in X with minimal Gaussian houses, then of which signal will be one of several independent indicators. Once W is established this is a simple make any difference to change it to identify A. In order to locate this transmission, some approach to measure the actual nongaussianity associated with wX is necessary. There usually are several ways to achieve this.

2.1.2 HOW TO ESTIMATE INDEPENDENCE:

KURTOSIS:

Kurtosis is the basic classical method used for the measurement of non guassianity of a random variable supposes K. when it is assumed that K is of unit variance then kurtosis is simply the normalized model of fourth model [2][11].

$$\text{Kurt}(k) = E\{k^4\} - 3(E\{k^2\})^2 \qquad (5)$$

NEGENTROPY:

The second basic method used for the measurement of nongaussianity is the negentropy. It is based on the quantity of differential entropy [14][15].

BASIC DATASET:

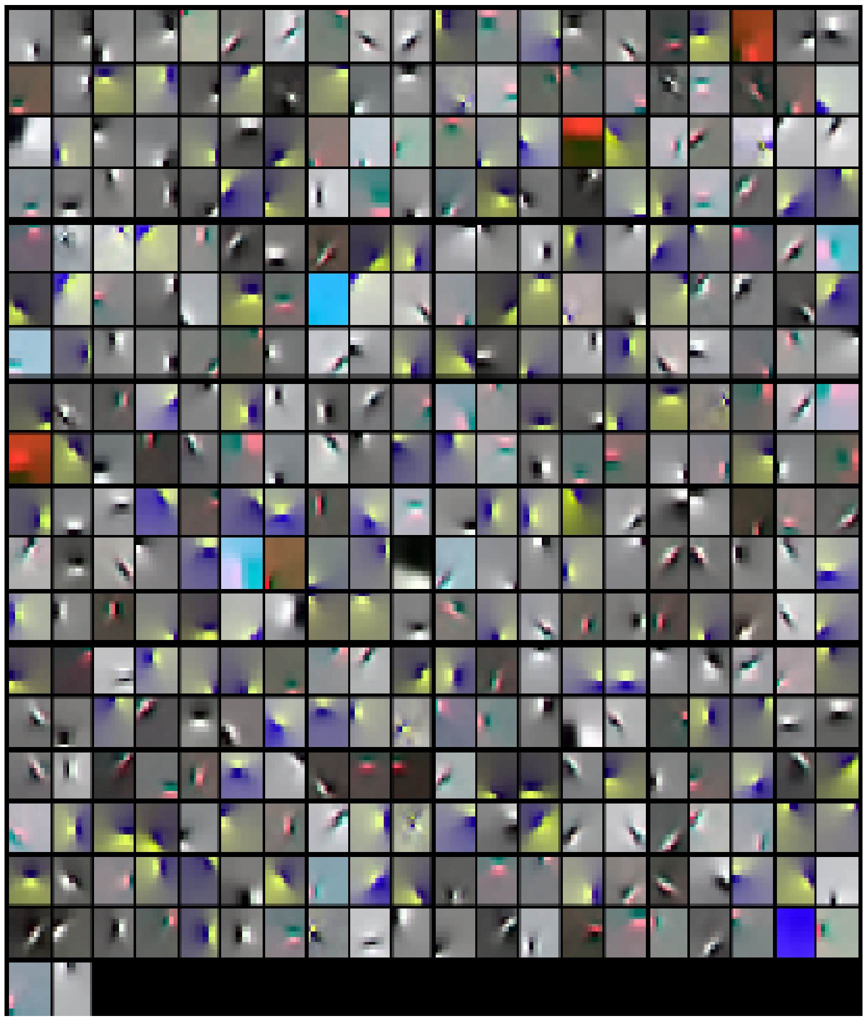


Figure 2: The 362 linear features learned by applying a complete independent component analysis (ICA) algorithm to 11×11 patches of color natural images from the Kyoto dataset.

Fig 2.7

2.2 SALIENCY MAP:

Saliency map [16] [17] is nothing but the first thing in an image that is the eye catcher. Whenever we see an image or scenery the object of focus in that scenery is basically the salient feature which needs to be identified and projected from that image. This also includes some part of object detection.

Now different people have produced different models to compute the saliency map. Each technique is different from the other and focuses on a certain parameter in the image and computes the saliency map.

The general structure of computational models is as follows:

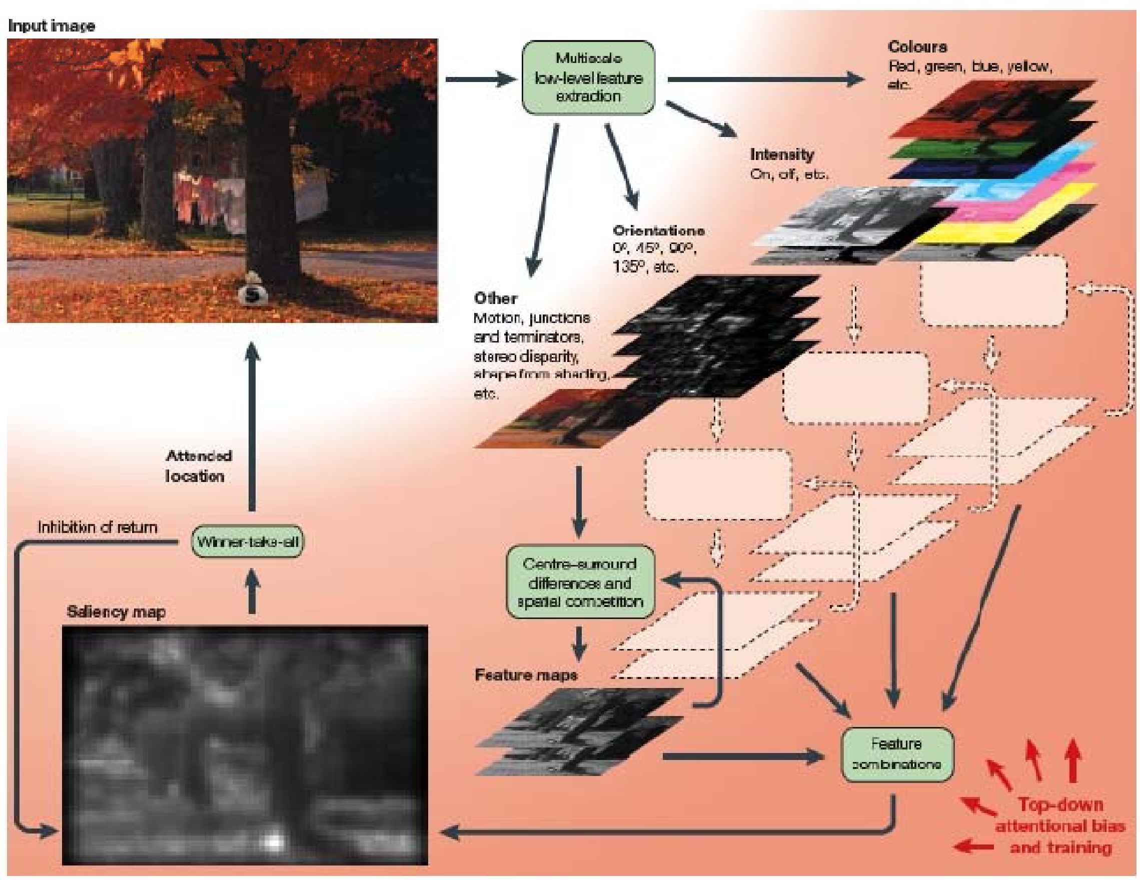


Fig 2.8

2.2.1 MODELS USED FOR COMPUTING SALINCY MAP:

There are several models which have computed saliency map by different techniques from which some of them are mentioned below.

ACHANTA:

Diagnosis of creatively salient image regions is useful for apps like target segmentation, adaptive compression setting, and target recognition [18] [19] [20]. On this paper, we introduce a for salient region detection that will outputs total resolution saliency maps with well-defined restrictions of salient objects. These restrictions are stored by retaining substantially additional frequency content from the original image than additional existing approaches. Our technique exploits top features of color as well as luminance, is straightforward to apply, and is actually computationally productive. We review our method to all 5 state-of-the-art salient region detection methods that has a frequency area analysis, ground truth, and also a salient target segmentation app. Our technique outperforms your five algorithms both on the floor truth evaluate and around the segmentation process by obtaining both higher precision as well as better recollect.

CONTEXT AWARE:

This model draws attentions to upon uncovering an important area of the picture which has the primary articles regarding facts although not only salient target [21] [22].

Gofer man et 'al. suggested a context-aware saliency diagnosis model. Salient picture locations tend to be discovered structured upon some ideas regarding man awareness: 1) local low-level

Concerns such as color and also distinction, 2) Global concerns which often suppress frequently transpiring characteristics though preserving characteristics which deviate on the norm, 3) Graphic organization principles which often declare that graphic kinds might have One or various centers regarding the law of gravity in relation to that the kind is actually Sorted, and also 4) High-level aspects, such as man faces. They applied their particular saliency approach to two software: retargeting and also summarization.

RANDOM SURROUND:

This specific model uses a straightforward course of action to analyze the image saliencies which often contains processing regional saliencies around some arbitrary parts inside the square condition in line with the undertaking awareness [23]. Then your Gaussian filtering will be used on remove the noise on the images. Shade space will be produced in addition to broken down in to L*, a* in addition to b* programmes. Saliency maps are usually produced with these types of programmes. Apart from n arbitrary subscription house windows are usually developed around all L*, a* in addition to b* sales channel. A final saliency road will be produced by simply fusing the Euclidean convention regarding saliency calculated with 3 over programmes.

WAVELET BASED SALIENCY:

This specific type focuses on removing the lower degree features as a result of wavelet transform next the idea creates the function chart which often definitely points out the features for instance edges as well as texture [24].

GRAPH BASED VISUAL SALIENCY (GBVS):

In this technique it will extract characteristic roadmaps on several spatial skin scales [25]. Any scale-space pyramid will be very first produced from photograph capabilities: Intensity, color, as well as alignment. Then, a new fully-connected graph over all grid places associated with each and every characteristic guide is made. Weights among a pair of nodes are usually assigned proportional towards likeness associated with characteristic ideals as well as their spatial long distance.

PHASE BASE FOURIER TRANSFORM:

In this particular model the method works by using the particular cycle spectrum to help determine the particular saliency map [26]. It really is identical within process for the amplitude spectrum employed by spectral residual model yet differs in the component of concentrate that is phase.

RESULTS OF ZHANG COMPUTING SALIENCY MAP:

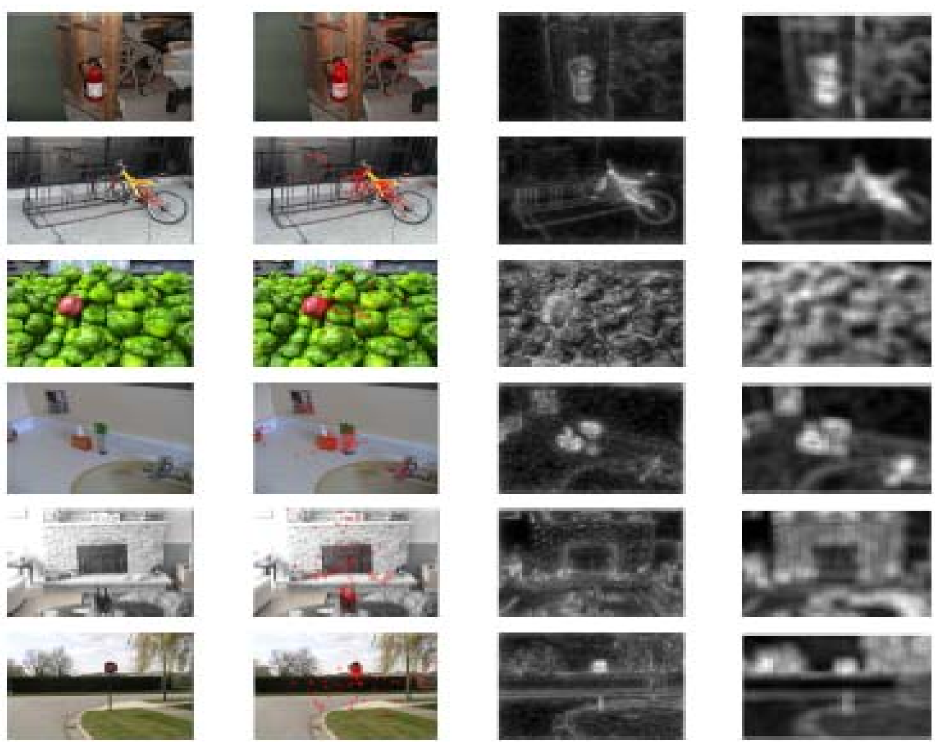


Figure 3: Examples of saliency maps for qualitative comparison. Each row contains, from left to right: An original test image; the same image with human fixations (from Bruce & Tsotsos, 2006) shown as red crosses; the saliency map produced by our SUN algorithm with DoG filters (Method 1); and the saliency map produced by SUN with ICA features (Method 2).

Fig 2.9

2.3 EDGE DETECTION:

In an image there are many frequencies which tend to change at every pixel value [27] [28]. The points where frequencies change rapidly are the points which can be detected by the help of a technique called edge detection in image processing. This field has been old enough and many researchers have worked on this technique and produced wonders.

Edge detection is also used in medical images to detect the minor portions which are not visible and also helps in connecting the broken lines.

2.3.1 SOBEL OPERATOR:

Basically the sobel operator function is a process of convolving the image with a filter having some integer values both in vertical and horizontal direction [29] [30].

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. If we define B as the source image, and H_x and H_y are two images which at each point contain the horizontal and vertical derivative approximations, the computations are as follows:

$$Hx = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} * B \qquad (6)$$

$$Hy = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * B \qquad (7)$$

where * here denotes 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, Hx can be written as

=

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:

$$H = \sqrt{Hx2 + Hy2} \qquad (8)$$

Using this information, we can also calculate the gradient's direction:

$$\alpha = \text{atan2}\,(Hy, Hx) \qquad (9)$$

where, for example, α is 0 for a vertical edge which is darker on the right side.

The following results and histograms are obtained by edge detection (sobel operator) technique on the saliency map of the above three images respectively.

2.4 MORPHOLOGICAL OPERATOIONS:

Morphological operations are used for extracting meaningful components from the images. There are different operations like dilation, erosion, opening, closing etc [31] [32].

In our case we have used the dilation operation. Dilation is used to thicken or grow objects in 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. There are various types of structuring element available; in this case we have used the Line structuring element. The line structuring element is applied on the image pixels from the start till end. The SE applies in the form of the line in the image every time and changes the pixel values according to the SE i.e. the change appear in places where the line affects the pixel. By this technique we are able to finally thicken the points we wanted and making the edges thicker and clear.

Mathematically the dilation is defined in term of set operation. The dilation of C and D is defined as

$$C \oplus D = \{z \mid (D')_z \cap C \neq \phi \} \quad (10)$$

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.

Chapter 3 Proposed Methodology

3.1 METHODOLOGY

we have applied different algorithms on the saliency map to improve the featured object which was little blur before and also noise reduction was our target.

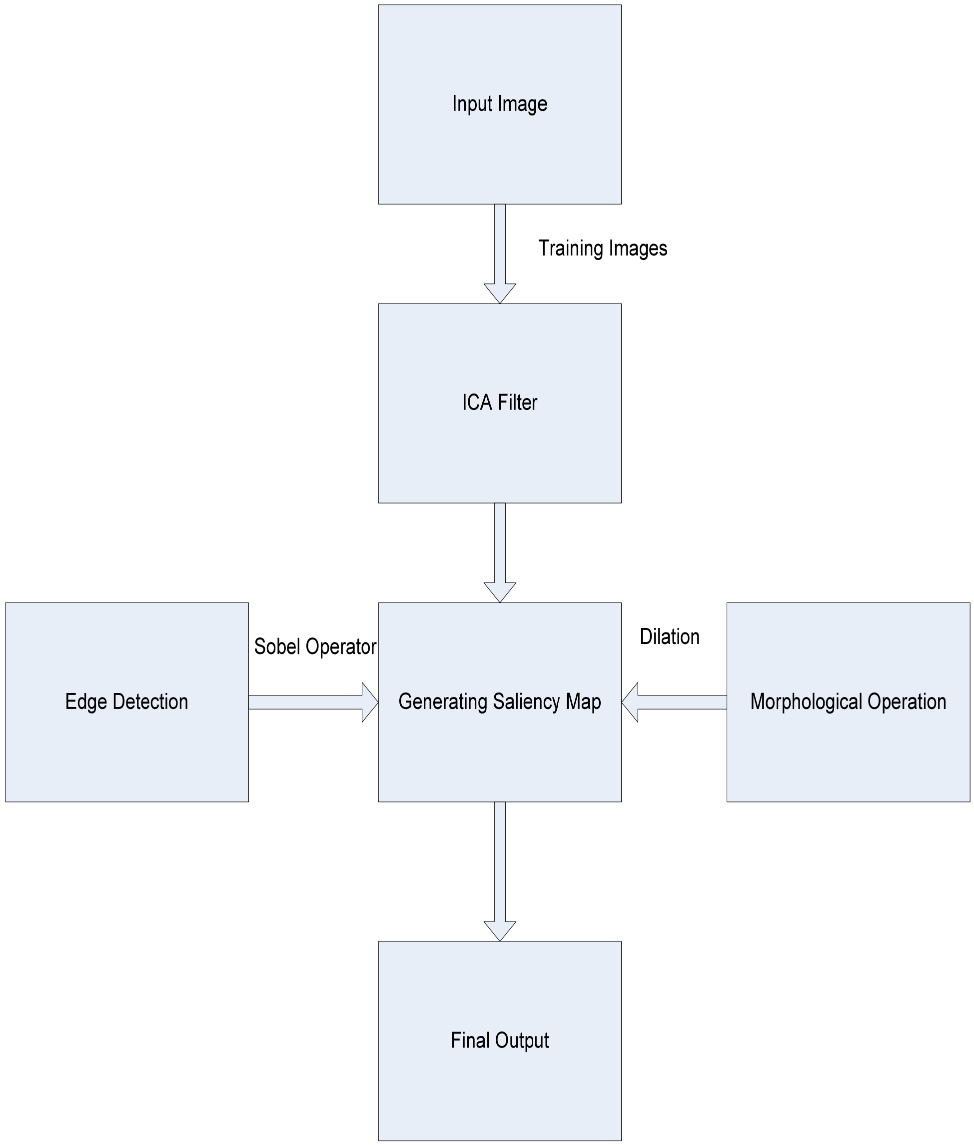
The basic flow chart of our proposed methodology is as follows:

Nowadays visual systems are being improved and the machine visual system is somehow tried to get close to that of human visual system. Still we cannot reach completely to efficiency of the human visual system.

Now the saliency maps generated by Zhang et al [1] has been used and tried to improve the results.Zhang et al [1] has used the Independent Component Analysis (ICA) computed by Hyvarinen et al [2] and computed the saliency maps.

Our proposed methodology is to improve the results of saliency map generated by Zhang et al [1].

3.2 FLOWCHART



Flow Diagram

Fig 3.1

Applying edge detection on the saliency map converted the saliency map in the form of edges of the salient feature and making the rest of the image black. In edge detection we have applied particularly sobel operator which is good for feature edge detection.

While applying sobel operator the saliency map results were made noise free as the rest of the image became black only focusing on the salient feature. Now the salient feature we get after applying sobel operator gives us a image just focusing on the salient feature but the broken edges makes it difficult to just the exact shape of the featured salient object.

To fill in the broken edges and to make the salient feature more visible we applied the morphological operator in particular the dilation process.

The structuring element is chosen very wisely in order to keep the shape of the salient object.

The dilation then finally gives us the final result which is a pure salient feature which we wanted in the saliency map results.

The histograms clearly show the distribution high in one of its parts as all the data is filtered in only one frequency and lies in that frequency.

3.3 RIGOROUS TESTING:

In the last part we have taken a dataset of 10000 images [33]. These ten thousand images are random and don't match each other. I have applied the algorithm on 1000 random images taken from the dataset and confirmed the results of the algorithm. From these 1000 images I have added some of them to my research paper and some of them are added to result chapter.

3.4 AREA UNDER THE ROC CURVE:

The area under the ROC curve is basically the test by which we can accomplish the result authenticity. We have used a dataset of eye fixation map of 15 users in free view scenario on 12 videos each containing 300 frames [34].

This procedure includes the computation of fixation map from the data given by the dataset. Then the saliency maps computed by ICA filter technique are compared to these fixation map results.

The mean score of the area under the ROC curve tends to be 0.79 which is extremely good. Score of the area under the ROC curve ranges from 0 to 1. The 0 indicates the minimum side and the 1 indicates the maximum side. As 0.79 is near to 1 it indicates that the saliency map generated are much near to that of the real human visual system and explains the authenticity of the work.

Chapter 4 Result and Discussion

A. RESULTS OF ORIGINAL IMAGE AND SALIENCY MAP WITH THEIR 3D PLOTS:

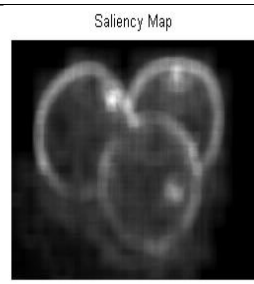
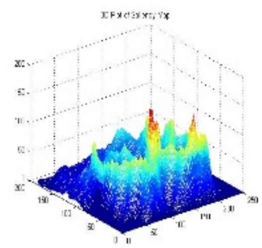
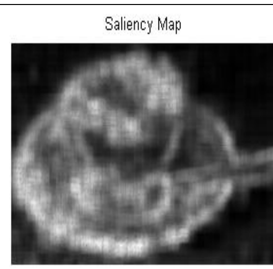
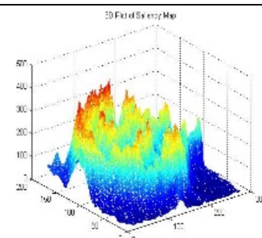
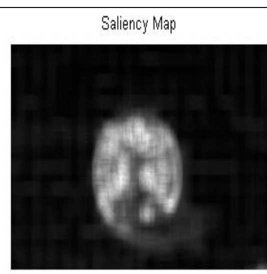
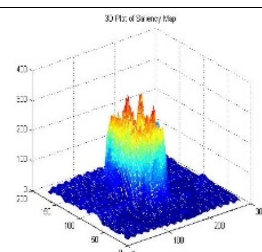
Original Image	ICA Results	3D Plots
		
		
		

Fig 4.1

The above results show the original image on which the algorithm is applied to compute the saliency map and its 3d plots are generated as showing the main focus on the salient object which covers most the centre part where high frequency components contain the data .

B. RESULTS OF SALIENCY MAP AND EDGE DETECTION (SOBEL OPERATOR) AND THEIR HISTOGRAMS:

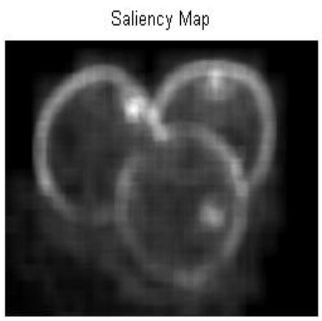
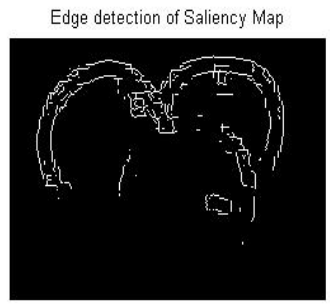
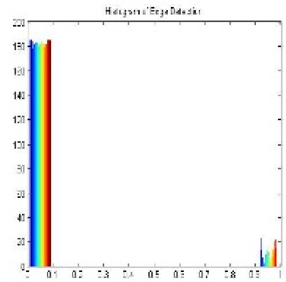
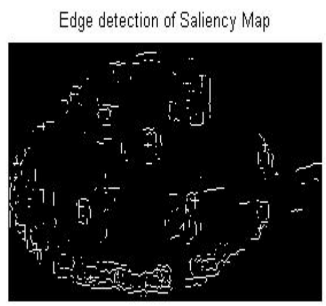
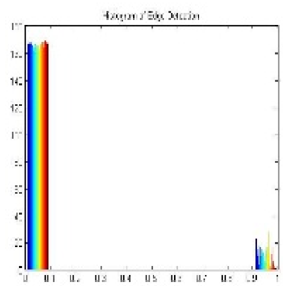
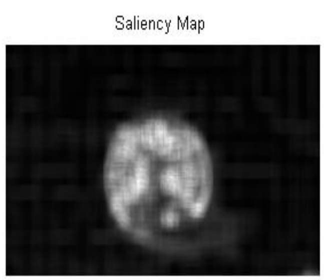
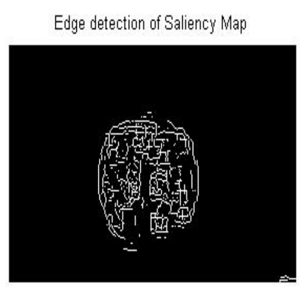
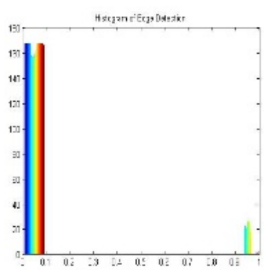
Saliency Map	Edge Detection (Sobel)	Histograms
		
		
		

Fig 4.2

In fig 4.2 the results of the previous fig 4.1 are under gone the edge detection method with a sobel operator highlighting the main content of the salient part. The histogram is computed to show the salient parts and the pixel repetition in this regard. The sobel operator gives clear edges of the image and all the information is highlighted.

C. RESULTS OF EDGE DETECTION AND DILATION AND THEIR HISTOGRAMS:

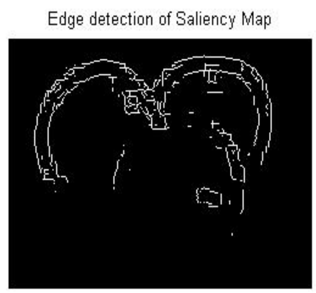
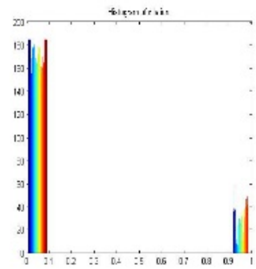
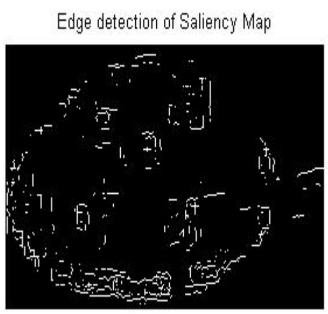
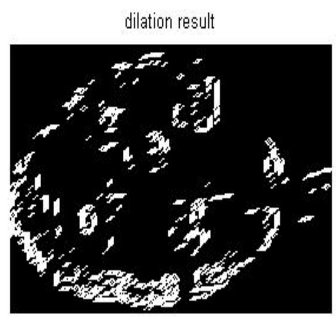
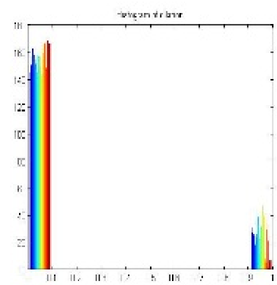
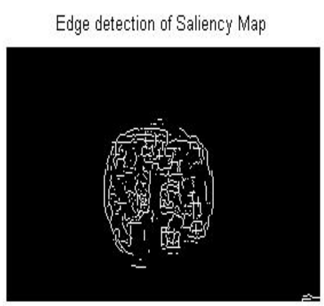
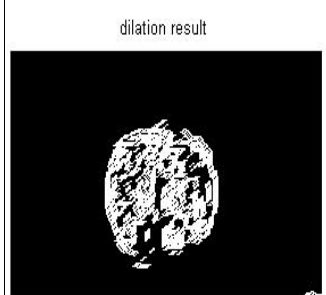
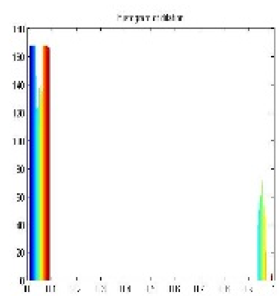
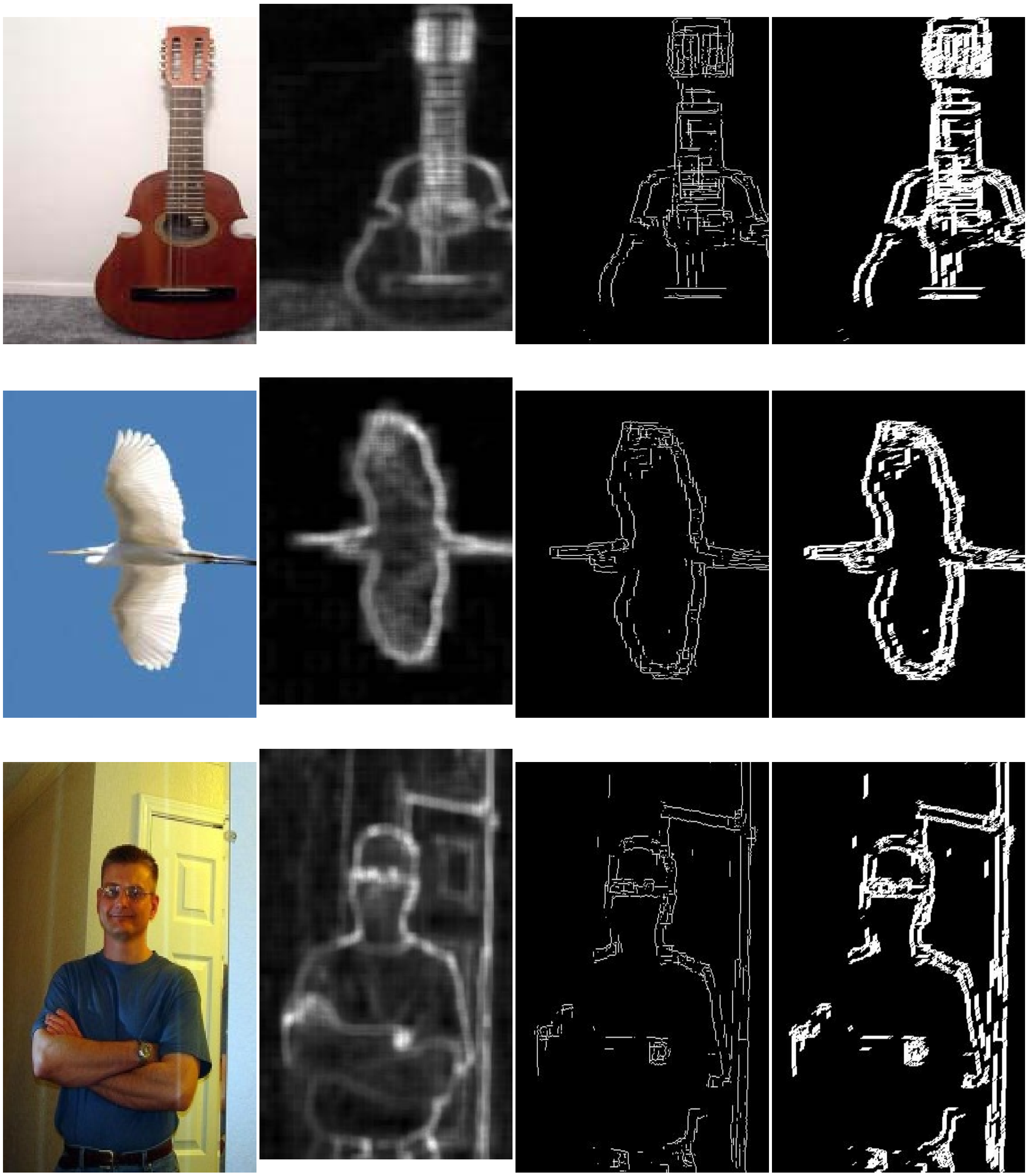
Edge Detection (Sobel)	Dilation Result	Histograms
		
		
		

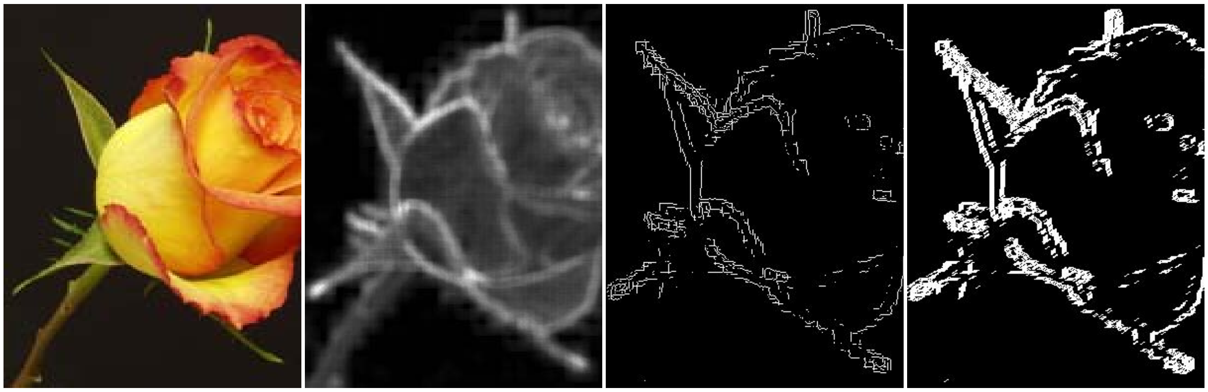
Fig 4.3

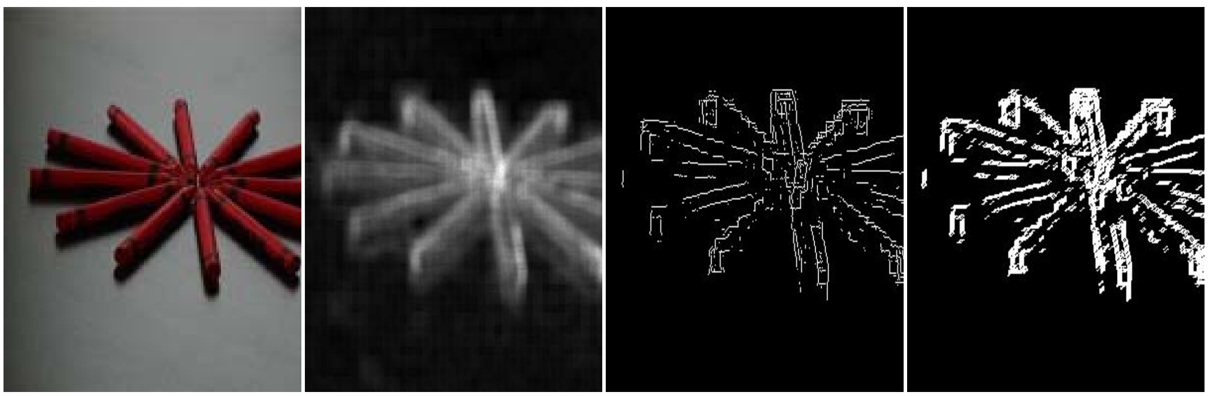
In fig 4.3 the edge detections results are used and morphological operations are applied on it specifically dilation with a certain structuring element which completes the edges which were broken in the previous result and hence giving the clear salient features and completing the salient part. The histogram shows an improvement and the repeating pixels are increased and hence making the objective clearer.

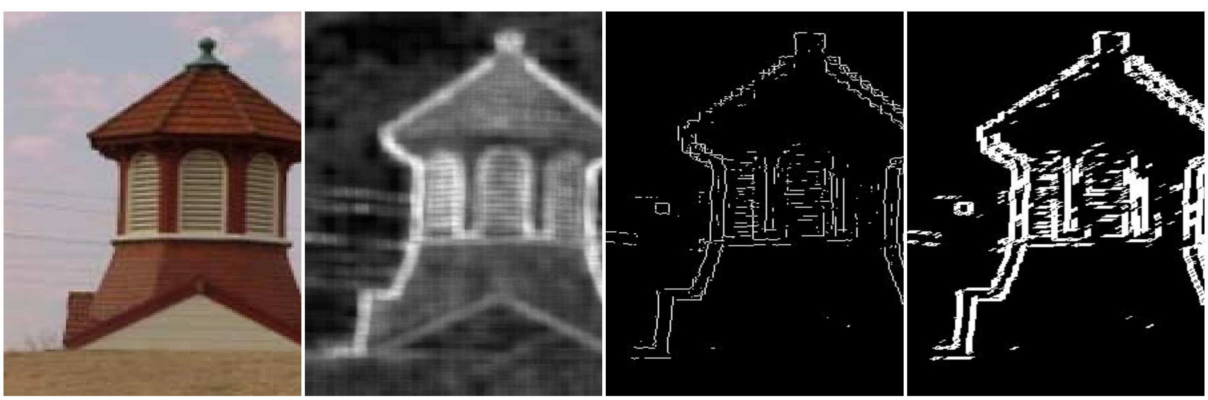
D. RESULTS OF ALGORITHM ON A DATASET:

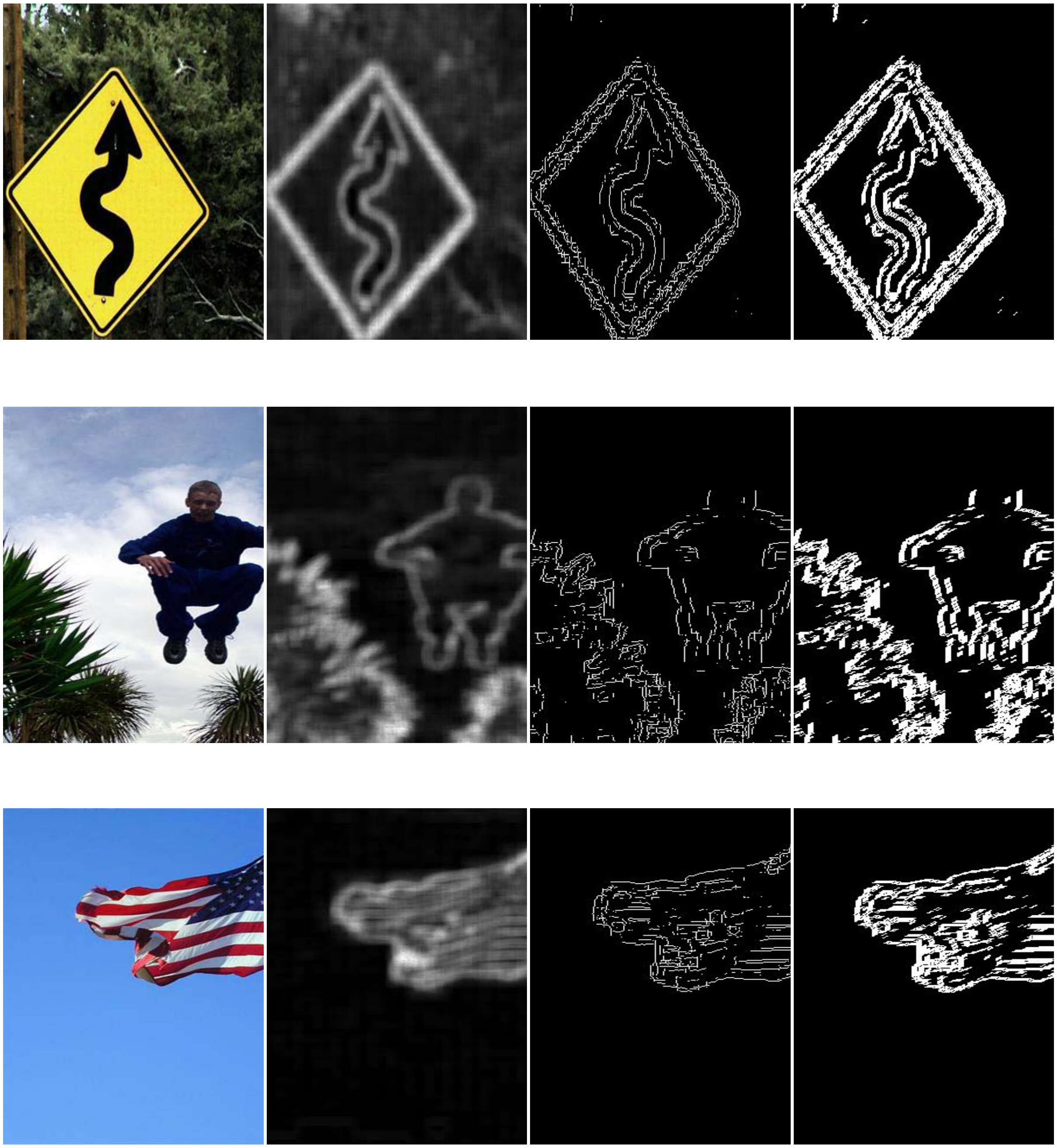
Original image	Saliency map	Edge Detection (Sobel)	Dilation Result











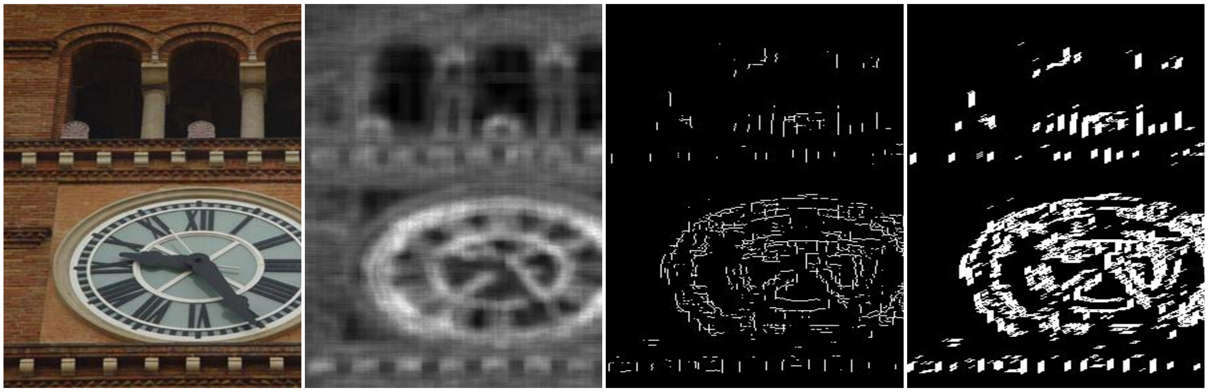


Fig 4.4

For testing the algorithm on different images for verification of the algorithm accuracy we have used a dataset. The dataset is taken from THUS 10000 MSRA DATASET. We have applied the algorithm to 1000 images randomly in 10000 images and computed the results. From these 1000 images above are some of the results of the images. The results show their saliency maps sobel results and the final results of dilation.

E. GENERALIZED RESULTS:

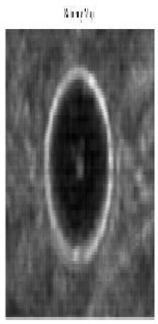
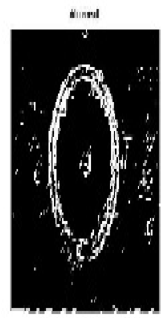
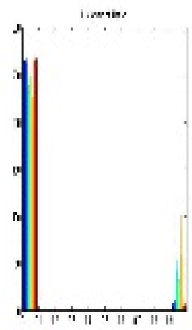
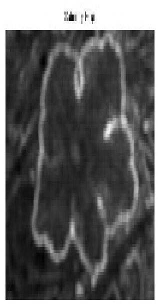
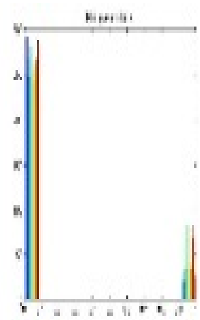
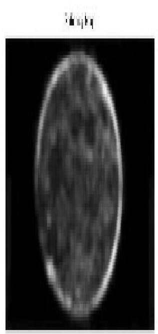
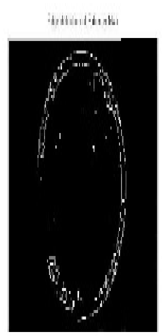
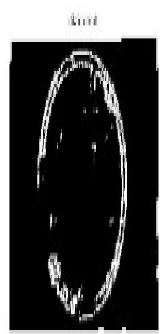
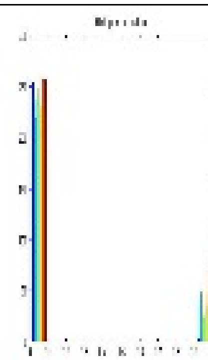
A	b	C	D	E
				
				
				

Fig 4.5

4.1 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 which has been tested in human visual system. The dataset used for calculating an AUC consists of random number of observers in a free viewing scenario on 135 different images [29]. The scenario is kept as to obtain the best possible results. The data for these random number of observers is viewing 135 different images and their point of focus is computed. We used these values 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.

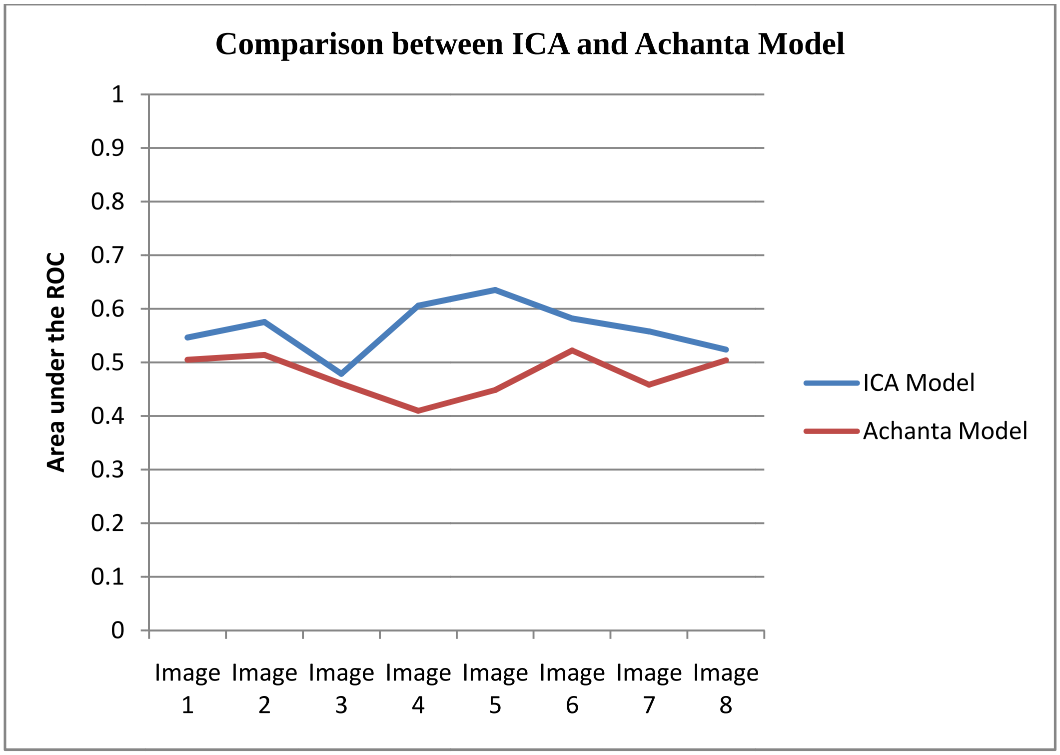

Comparison between ICA and Achanta Model
Area under the ROC
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Image 1
Image 2
Image 3
Image 4
Image 5
Image 6
Image 7
Image 8
ICA Model
Achanta Model

Fig 4.6

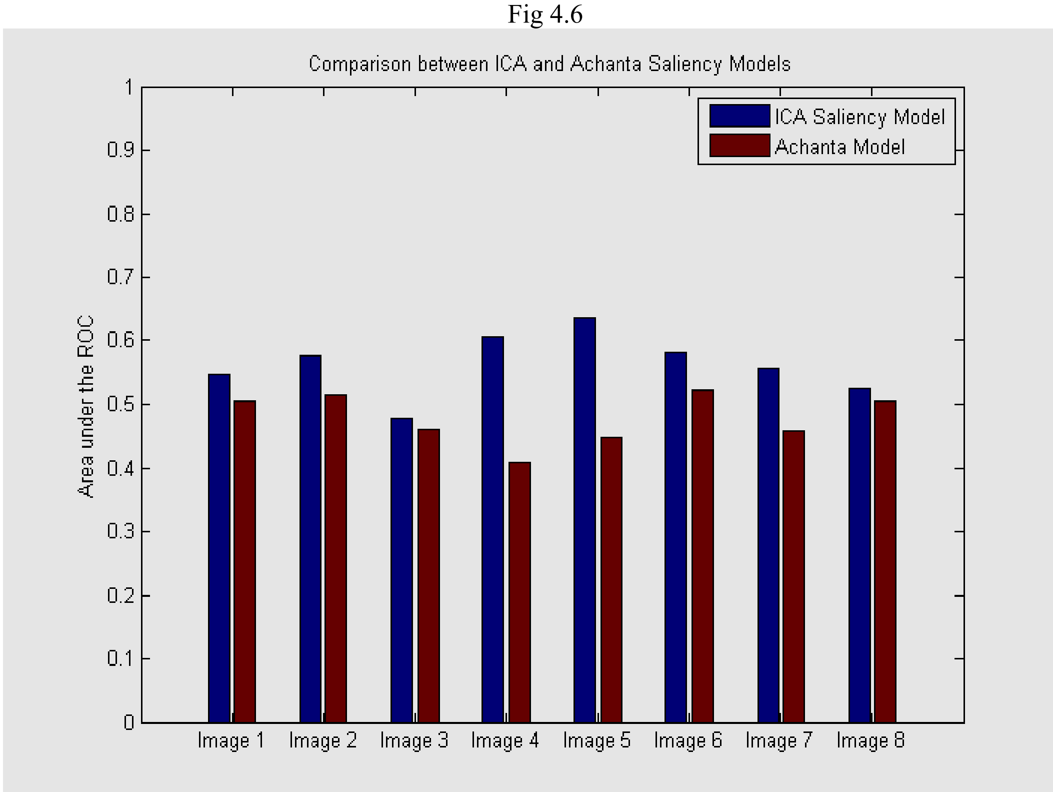

Comparison between ICA and Achanta Saliency Models
ICA Saliency Model
Achanta Model
Area under the ROC
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Image 1
Image 2
Image 3
Image 4
Image 5
Image 6
Image 7
Image 8

Fig 4.7

4.2 DISCUSSION:

The saliency map generated has been improved in a way that salient feature which is be detected is highlighted.

Both the algorithms edge detection and dilation has immensely enhanced the results that is the salient feature is projected out from the image and dilation helped it to make the result more clear and also defining boundaries of salient object.

As the area under the ROC curve is also good enough to mark the saliency map as a good result and the sobel operator and the dilation operator enhances it as far as the visual analysis is concerned

CHAPTER 5 FUTURE WORK AND CONCLUSION

5.1 Future work:

In this chapter we will discuss the future work that can be done in the time to come and also the conclusion of the thesis.

The future work which can be done as far as the saliency map and the improved version of saliency is concerned is that this can be extended to the videos.
We can take one video and break it into frames and then apply the algorithm and see the results. We can also take numerous video dataset and also apply the saliency map on these videos in the form of frames and once we get the saliency maps we can join the video frames into a complete video.

Else we can also apply improved technique as the visual analysis is concerned and also apply on the videos and get the results.
This algorithm can be modified for purely detection of objects and different features in a still image also in videos. This can be similar to the surveillance of objects in a scenario.

Also one more addition to the work can be done. Eye fixation maps can be used to compute the area under the ROC curve score to verify the efficiency of the technique and can be comparable to the human visual system both in still images and also in videos.

5.2 CONCLUSION:

To generate the saliency map, ICA filter is used which gives us 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 which brightens the edge detection results and improves the lines and curves of edge detection result. Dilation is basically used to complete the incomplete boundaries of the region and thickens it. After applying all these operations we have produce results clearly showing the effects, improving the saliency map and the edge detection has enhanced the results of saliency map and finally the morphological technique highlights the final result of the feature detection.

REFERENCES:

1.Lingyun Zhang, Matthew H. Tong, Tim K. Marks, Honghao Shan & GarrisonW.Cottrell (2008). SUN: A Bayesian framework for saliency using natural statistics. Journal of Vision, 8(7):32, 1-20,

2.Hyvarinen, A., & Oja, E. (1997). A fast fixed-point algorithm for independent component analysis.Neural Computation, 9(7), 1483-1492.

3.Arulampalam, M.S., Maskell, S., Gordon, N. & Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-gaussian Bayesian tracking. IEEE Transactions on Signal Processing. 50(2):174-188.

4.Lin Z, Q. L. Tracking in Clutter based on Mean Shift Embedded Particle
Filter. ICCET, Vol6,p6331-6335,2010

5.D. Ziou and S. Tabbone (1998) "Edge detection techniques: An overview", International Journal of Pattern Recognition and Image Analysis, 8(4):537–559, 1998

6.J. M. Park and Y. Lu (2008) "Edge detection in grayscale, color, and range images", in B. W. Wah (editor) Encyclopedia of Computer Science and Engineering, doi 10.1002/9780470050118.ecse603

7.An Introduction to Morphological Image Processing by Edward R. Dougherty, ISBN 0-8194-0845-X (1992)

8.Morphological Image Analysis; Principles and Applications by Pierre Soille, ISBN 3-540-65671-5 (1999), 2nd edition (2003).

9. Scott, David W. (1979). "On optimal and data-based histograms". Biometrika 66 (3): 605–610. doi:10.1093/biomet/66.3.605

10.L. Itti, ”Quantifying the Contribution of Low-Level Saliency to Human Eye Movements in Dynamic Scenes,“ Visual Cognition, vol. 12, no. 6, pp. 1093-1123, 2005.

11.A. Hyvärinen and E. Oja. Independent Component Analysis: Algorithms and Applications. Neural Networks, 13(4-5):411-430, 2000.

12.A.D.Back and A.S.Weigend.A first application of independent component analysis to extracting structure from stock return. Int.J.on Neural Systems.

13.P.Comon. Independent component analysis a new concept of Signal Processing.

14.T.M.Cover and J.A.Thomas.Elements of Information Theory.John Wiley Sons 1991.

15.A.Papoulis.ProbabilityRandom Variablesand Stochastic Processes.McGraw-Hill 3rd edition 1991.

16. L. Itti, C. Koch, E. Niebur, A Model of Saliency-Based Visual Attention for Rapid Scene Analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 11, pp. 1254-1259, Nov 1998.

17.Itti, L., & Koch, C. (2001). Computational modeling of visual attention. Nature Reviews Neuroscience, 2(3), 194-203.

18.[Achanta et al., 2008] Radhakrishna Achanta, Francisco Estrada, Patricia Wils, and Sabine Ssstrunk. Salient region detection and segmentation. In Antonios Gasteratos, Markus Vincze, and John Tsotsos, editors, Computer Vision Systems, volume 5008 of Lecture Notes in Computer Science, pages 66{75. Springer Berlin / Heidelberg, 2008.

19.[Achanta et al., 2009] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk. Frequency-tuned salient region detection. In Computer Vision and Pattern Recog-nition, 2009. CVPR 2009. IEEE Conference on, pages 1597 {1604, june 2009.

20.[Avraham and Lindenbaum, 2006] Tamar Avraham and Michael Lindenbaum. Attention-based dynamic visual search using inner-scene similarity: Algorithms and bounds. IEEE Transactions on Pattern Analysis and Machine Intelligence, 28:251{264, 2006.

21.[Goferman et al., 2010] Stas Goferman, Lihi Zelnik-Manor, and Ayellet Tal. Context aware saliency detection. In CVPR'10, pages 2376{2383, 2010.

22.[Wolfe et al., 1989] Jeremy M. Wolfe, Kyle R. Cave, and Susan L. Franzel. Guided search: An alternative to the feature integration model for visual search. Journal of Experimental Psychology: Human Perception and Performance, Vol 15(3):419{433, 1989.

23.Tadmeri Narayan V. Random center- surround approaches for modeling visual saliency. Bielefeld: Universitätsbibliothek; 2013.

24.Wavelet-Based Salient Region Extraction Dong-Woei Lin and Shih-Hsuan Yang Department of Computer Science and Information Engineering National Taipei University of Technology Sec. 3, Chung-Hsiao E. Rd., Taipei, Taiwan.

25.[Harel et al., 2007] Jonathan Harel, Christof Koch, and Pietro Perona. Graph-based visual saliency. In Advances in Neural Information Processing Systems 19, pages 545{552. MIT Press, 2007.

26.Spatio-temporal Saliency detection using phase spectrum of quaternion fourier transform Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference by Chenlei Guo,Qi Ma ; Liming Zhang.

27. Lindeberg, Tony (2001), "Edge detection", in Hazewinkel, Michiel, Encyclopedia of Mathematics, Springer, ISBN 978-1-55608-010-4

28.BARGHOUT, Lauren. "SYSTEM AND METHOD FOR EDGE DETECTION IN IMAGE PROCESSING AND RECOGNITION." WIPO Patent No. 2007044828. 20 Apr. 2007.

29. Irwin Sobel, 2014, History and Definition of the Sobel Operator

30.Feature Detectors - Sobel Edge Detector

31.mage Analysis and Mathematical Morphology, Volume 2: Theoretical Advances by

Jean Serra, ISBN 0-12-637241-1 (1988)

32.An Introduction to Morphological Image Processing by Edward R. Dougherty, ISBN 0-8194-0845-X (1992)

33.SalientShape: Group Saliency in Image Collections. Ming-Ming Cheng, Niloy J. Mitra, Xiaolei Huang, Shi-Min Hu. The Visual Computer, 2013.

34. M. MANCAS, O. LE MEUR, 2013, "Memorability of Natural Scenes: the Role of Attention", *Proceedings of the International Conference on Image Processing (IEEE ICIP 2013), Melbourne, Australia, September 15-18.*