AUTOMATED IMAGE RETARGETTING AND CROPPING USING SALIENCY DETECTION AUTOMATIC CONTEXT AWARE CROPPING

Hakimuddin Asghar[1], Adeel Ahmad[2], Umapathy Eaganathan[3]

[1,2] BSc (Hons) in Intelligent Systems, Asia Pacific University, Malaysia

[3]Faculty in Computing, Asia Pacific University, Malaysia

ABSTRACT

Image processing is an important field which is enhanced since few decades. One of its essential use in image resizing has received a lot of attention because it is being used in every field of science. Previously, traditional scaling was used to crop the image and resize it but it was unable to preserve the salient region and remove the unwanted data. Hence, in this study, an automatic resizing approach is used to crop the image based on salient object detection and remove the extra background which is not of equal importance as the salient items. A GUI based system is created to facilitate the user to input images and to specify the height and width of the cropping size. The user defined size is then applied automatically to the image based on the salient point. Finally, after getting the results, these results are tested based on different sizes. Analysis is then performed on the final outcomes and the method used is evaluated to find out the efficiency of the method used.

Keywords—*Saliency; Context; Threashold; GBVS; Cropping Saliency_map*

I. INTODUCTION

Image processing aspects has become an interesting topic in various applications or field to make use of image processing and analyzation techniques (Bovik, 2009). Every branch of science which uses different devices to collect image data. This sort of data is in the form of objects which is often multi-dimensional and can be arranged to make it appropriate for human viewing. However, sometimes data is not clearly depict and small in size and contains undesired data which should be removed and resized.

Image targeting or resizing is a technique in image processing which important to display an images without losing its contents. It is necessary to resize an image because different electronic devices such as mobile phone, tablets and computers required adjustment of an image (Kumar & Babu, 2010). The quality of an image should be preserved and contents can be kept by removing an undesired regions and obtaining better color balance. An automatic cropping image technique is used to improve the visual quality and unwanted area can be removed. There are various images and videos are created which are not informative because the specific region are not

apparent to human observers (Jin, et, al., 20120. Moreover, there is an overwhelming amount of extracting background or image display is very small.

II. PROBLEM STATEMENT

Image resizing is an important aspect of image processing which is being utilized in every field of science such as medical, astronomy, Microscopy, and Radiology. Moreover, there are many devices which allocates different size adjustment to an image and sometime, due to size of an image, the important or salient part of the image is cropped unexpectedly. In last few decades, there were many techniques and methods used to perform image retargeting without preserving the salient object (Yan, et al., 2009). An important techniques used was traditional scaling but it does not keep the salient object and unable to preserve the important contents in an image.

Losing contents of an image will destroy the quality of an image and desired region will be lost. If the image retargeting is used in medical fields where it loses the important contents, then it is useless to make use of image processing. Furthermore, there were two cropping techniques used which are attention-based and region-based to solve this issue and determine the salient region.

III. AIMS

"To crop an image without losing salient object and resizing to make use of it for different electronic devices".

IV. OBJECTIVES

The objective of an image processing content preservation and cropping is as follows,

- To crop an image automatically and preserve the salient object.
- To make use of automated cropping for different devices size adjustment.
- To remove unwanted data or region which effects the quality of image.

V. DESCRIPTION AND DESCUSSION OF PROPOSED ALGORITHM

There are various methods and algorithms have been proposed and practiced in the past for an image cropping. The main purpose of these algorithms was to enhance and resize the images while preserving the important image. Moreover, attention-based and region-based algorithm focus on the main region in the image. Salient object is detected and computed over the image. It is computed based on region based entropy size and distance from the image centre. However, the cropping box is set on the on the salient object. The system works manually by setting the cropping box or bounding box on the image. The image may lose its important region because the bounding box size is fixed and works manually. The accuracy of resulting output is not satisfied in image processing.

To improve the efficiency and effectiveness, the developer propose appropriate and better algorithm to improve the image resizing. The accuracy of the results obtained is higher than other algorithms. The diversity of image introduces issues in cropping which must be resized for an optimal display or to use in different applications. Image is resized based on aspect ratio and size to satisfy the requirements (Lin, et al., 2015). To overcome this problem, a context-aware cropping is an optimal solution which take the image contents into consideration to preserver the important region and distortion is minimized. Previously, researchers have proposed many algorithms to solve this issue but the topic is still a subject of ongoing investigation. The detection of an important or salient part in region of an image is crucial part of computer vision (Vaquero, et al., 2014).

Furthermore, an important role of image resizing is to estimate where is the important region which needed to be preserved. To perform saliency detection, the values of [0, 1] are assigned to every pixel and find out the higher value. The higher values region has higher importance. Saliency map draw the important area which represents the human attention area. Furthermore, there are two approaches in saliency detection map, first one is bottom up method and last one is top-down method. The bottom-up method for low level features such as color, intensities and edge detection. Top-up approach is used for semantic information such as important region detection (Avidan & Shamir , 2013).
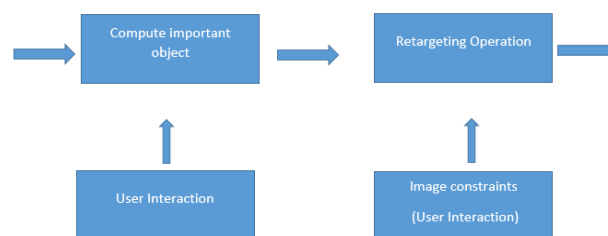

Fig. 1. Context Aware Cropping Approach

Now, making effective use of Salient Objects detected along with threshold surroundings is very important in this study for cropping the image. Once the Salient object(s) is detected the developer can now implement an algorithm to crop the images. This requires studying the imcrop function in MATLAB at its initial level to begin with the cropping process. Here, the most important piece of code to understand is the:

$$I2 = imcrop (I, [76\ 68\ 130\ 112]);$$

This function accepts several parameters. Among these parameter, the one used in this study is the one shown above. This accepts 2 parameters. (1), the input image and (2) the cropping rectangle. This rectangle has four parts. Horizontal and vertical coordinates of the upper left corner (x1, y1) and the width (w) and height (h) of the cropping size.

$$[x\quad y\quad width\quad height]$$

The width and height are specified by the user in the system and x and y values are calculated based on the most salient object detection. The width and height are provided by the user as a percentage and is easily calculated by multiplying it to the original image. x and y need to be found however.

The following steps are followed by the developers to calculate these coordinates and to receive the output:
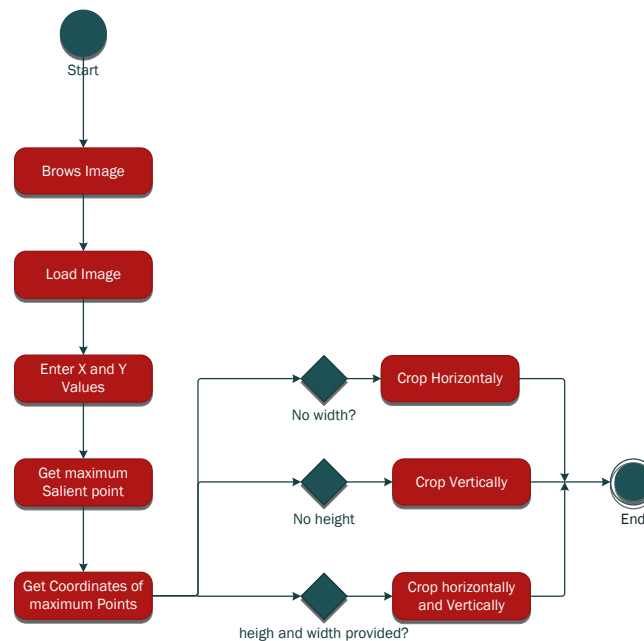
I. SYSTEM DESCRIPTION & ANALYSIS

A. Description and Discussion of the system

Automated Image Resizing system was designed and implemented to crop images without losing the important content. The software and hardware specification are required to develop the system which is as follows:

Software Specifications:

Software: MATLAB 2014

Operating System: Windows 10

Plug-in: GBVS – for Salient object detection

B. System GUI and Description

1) Main Window

Graphical User Interface was designed to get user inputs. When system is executed, the main window will open which is required user to brows for an image, provide the horizontal and vertical cropping values in the form of percentage as shown in the figure below.
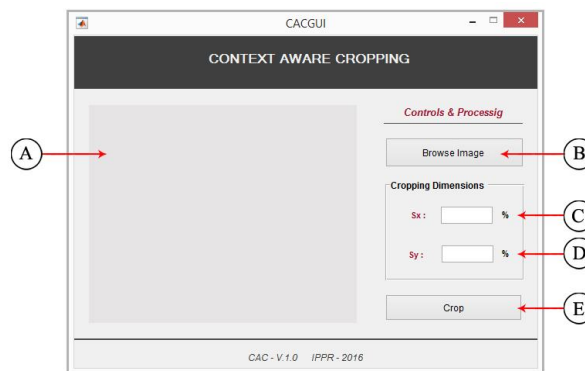
Fig. 2.   Main Window –System

Button **(B)** is the first step of execution, which brows the image and load the image to section **(A)**. The text fields **(C)** and **(D)** are used to get inputs from user and last step if performed which is cropping the image based on user input values.

2) Browse an Image

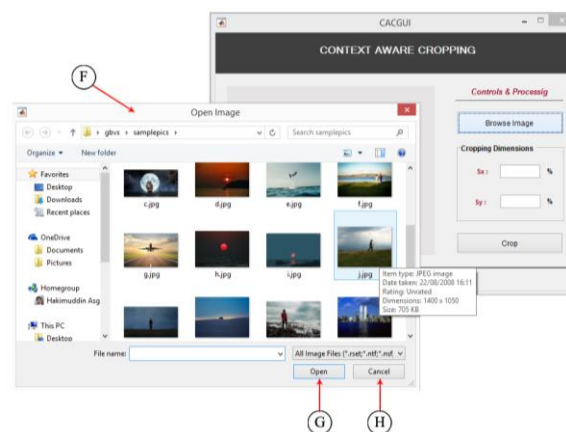Button **(B)** is pressed and brows an image window is shown in the figure below.

Fig. 3.   Image Browsing

The figure shows the image browsing step to be performed. The user is required to choose the path **(F)** for an image selection and press the Open button **(G)** to crop the image or cancel to stop the execution process **(H).**

3) Load Image

After selecting path and image to crop, user is required to press the Crop button which is shown in the figure below.
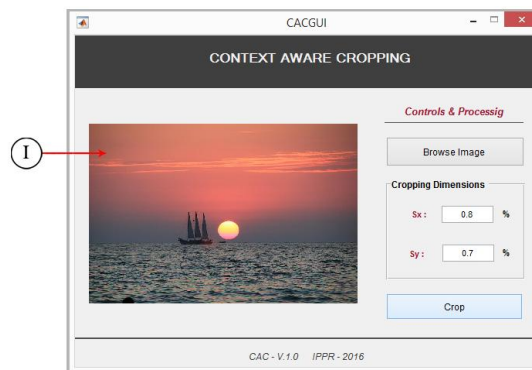
Fig. 4. Load Image

Figure show that, input image has been loaded in the window and values are provided to crop the image. The salient object in the image is small and it will be preserved when cropping the image. The output results will be shown the results obtained section.

VI. RESULTS AND ANALYSIS

In order to produce results, photos are selected from a span of several categories. These include humans, architecture, tools and objects, planets & moons, nature, vehicles and other objects. Hence, giving a wide variety of images and categories for testing and further analysis. Each image is cropped several times based on the inputs specified by the user in Sx (Height) and Sy (Weight). As mentioned earlier, each cropping window requires four parameters: horizontal and vertical coordinates of the upper left corner (x1, y1) and the width (w) and height (h) of the cropping size. Here the upper left coordinates of the window are automatically detected so they does not need to be produced in the results. However, the developer(s) will analyze the images based on the height and width of cropping window to check if it fits the coordinates and secures the salient parts of the image. Especially the most subjective and important part of the image. Several cropping outputs are clustered together for each respective image to provide a profound understanding for the user and to provide the developer(s) with a strong basis for analysis and comparison

In the following, the developer(s) describe these results obtained, discuss their computation and analyze different cropping windows on each respective image:

C. Saliency Obtained

Original images (a) from figure [8] to figure [17] are reduced horizontally, vertically and on both axes. Among these figures, figures [8] to [12] are used to test random samples on an accepted minimum cropping percentage. While figures [13] to [17] are used to test extreme cases and are used to analyze the limitations. After getting the original image, saliency map is produced by implementing the GBVS function. This is represented in each subpart (b) of the image. This figure contains the maximum saliency point in grayscale. Further produced is another subpart of each figure that shows threshold values and most salient parts. This is represented as a percentage value and covers all important points within the percentage defined. It is 75% by default and shall be

maintained throughout during this project. An overlay is shown for each image in its respective figure. This overlay is simply a combination of saliency map and original image to make the user understand the important points in the image.

All images from [8] to [12] contain subplot images from (e) to (l) that define different cropping sizes. Further images are tested and a similar combined result is produced for each image. This is attached in the Appendix to provide more results to the user for extensive study in this subject. The results tested in this section along with the extreme and limitations case are enough to produce results in this section.

A total of 10 x 8 = 80 samples are analyzed in this section. Among the samples, 5 x 8 = 40 are produced for random images and 5 x 8 = 40 images are produced and analyzed for extreme and limitation cases. For the appendix, 7 x 8 = 56 samples are provided for further study. Hence, a total of 17 x 8 = 136 samples are produced in this documentation.

Lastly, after understanding the saliency, now cropping is performed and results are produced as follows:
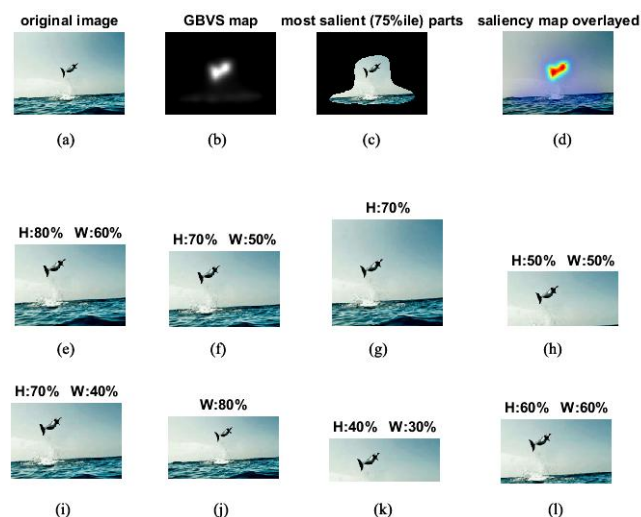
D. Results and Analysis – Performed

Fig. 5. Photo cropping. (a)Original image, with (b) Saliency Map, (c) Threshold and (d) Overlay from GBVS Function. (e)(l)Computed Cropping Regions based on user's Sx and Sy values. (e) Sx: 0.8, Sy: 0.6 (f) Sx: 0.7, Sy: 0.5 (g) Sx: 0.7, Sy: 1 (h) Sx: 0.5, Sy: 0.5 (i) Sx: 0.7, Sy: 0.4 (j) Sx: 0.8, Sy: 1 (k) Sx: 0.4, Sy: 0.3 (l) Sx: 0.6, Sy: 0.6.

The original image in [8] (a) is cropped and is shown in sub-figures (e) to (l). In figure (e), when height is 80% and width is 60%, the important element i.e. the dolphin is secured along with surrounding necessary points effectively. In figure (f), when height is 70% and width is 50%, the width doesn't have a greater impact on the image and still helps secure the dolphin and the water burst. However, the change in height causes some surrounding sea area to be cropped. In figure (g), when height is 70%, the dolphin is centered properly in the image and the sky and the sea, including the water burst is preserved. In figure (h), when height is 50% and width is 50%, the sky and sea is cropped out but the dolphin is secured and centered. The sky is still clear but the water burst is lost. In figure (i), when height is 70% and width is 40%, the dolphin and water burst are both preserved and centered in the image. Sea is effected on the left side of the image but the important points are

preserved. In figure (j), when width is 80%, the dolphin and water burst are preserved in the center. However, some dark edges from the sea are gone. In figure (k), when height is 40% and width is 30%, the figure is tested on extreme values. Even when the width and height is so much reduced, the most salient part is preserved but the other threshold items are lost. In figure (l), when height is 60% and width is 60%, the salient objects are secured but the edge of the water burst falls on the cropping line and hence some data might be lost.
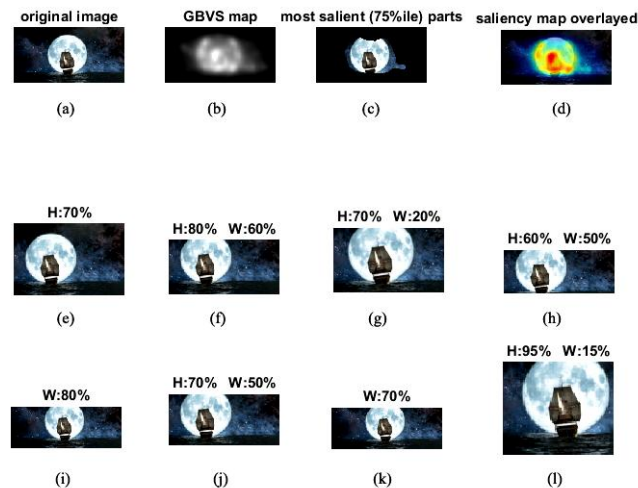

Fig. 6. Photo cropping. (a)Original image, with (b) Saliency Map, (c) Threshold and (d) Overlay from GBVS Function. (e)(l) Computed Cropping Regions based on user's Sx and Sy values. (e) Sx: 0.7, Sy: 1 (f) Sx: 0.8, Sy: 0.6 (g) Sx: 0.7, Sy: 0.2 (h) Sx: 0.6, Sy: 0.5 (i) Sx: 1, Sy: 0.8 (j) Sx: 0.7, Sy: 0.5 (k) Sx: 0.7, Sy: 1 (l) Sx: 0.95, Sy: 0.15.

The original image in [9] (a) is cropped and is shown in sub-figures (e) to (l). In figure (e), when height is 70%, the important element i.e. the moon and the boat are secured along with surrounding necessary points effectively. In figure (f), when height is 80% and width is 60%, the width doesn't have a greater impact on the image and still helps secure the points. However, the change in height causes some data loss in the moon. This is the result of the salient object being too large to cover the whole height. In figure (g), when height is 70% and width is 20%, some portion of the moon is lost but the width cropping is performed accurately. And the sea surroundings and the moon boundaries are preserved. In figure (h), when height is 60% and width is 50%, the sky and sea are cropped along with the bottom portion of the boat. Again this is due to the salient object being too big. The algorithm applied tries to minimize the risk and attain the central highest salient object however. In figure (i), when width is 80%, the side portions of the image are effectively cropped but this implies an impact on stretching that causes the enlarging effect that in turn compromises on the height i.e. the moon loses some upper portion data. In figure (j), when height is 70% and width is 50%, the width is cropped effectively however again due to the suspecting objects size, some upper and lower portions are lost. Figure (k), has the same effect as figure (i) i.e. when width is 70%, some upper boundaries of the moon are lost. This however is less than that the previous 80%. Lastly, in figure (l), when height is 95% and width is 15%, cropping is tested on extreme values. Even when the width and height is so much reduced, the most salient parts are centered but due to the enlarged size of the subject, much information is lost i.e. all other threshold items are lost
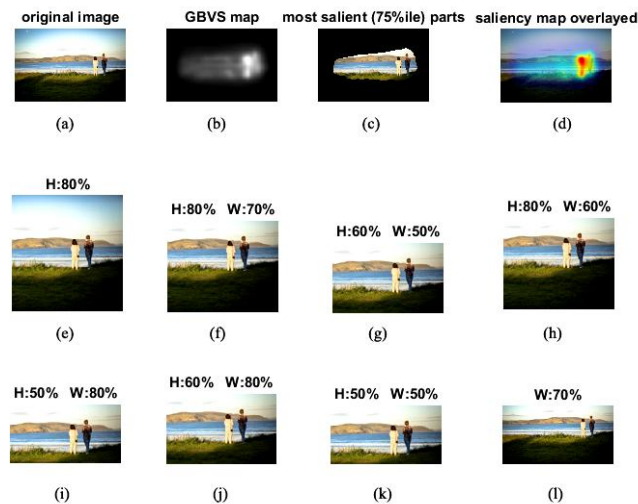
Fig. 7. Photo cropping. (a)Original image, with (b) Saliency Map, (c) Threshold and (d) Overlay from GBVS Function. (e)(l) Computed Cropping Regions based on user's Sx and Sy values. (e) Sx: 0.8, Sy: 1 (f) Sx: 0.8, Sy: 0.7 (g) Sx: 0.6, Sy: 0.5 (h) Sx: 0.8, Sy: 0.6 (i) Sx: 0.5, Sy: 0.8 (j) Sx: 0.6, Sy: 0.8 (k) Sx: 0.5, Sy: 0.5 (l) Sx: 1, Sy: 0.7.

The original image in [10] (a) is cropped and is shown in sub-figures (e) to (l). In figure (e), when height is 80%, the important elements i.e. the couple are secured along with surrounding necessary points very effectively. In figure (f), when height is 80% and width is 70%, the change in width causes the zooming effect and hence causes an effect on the height. Hence the height gets reduced. In this case the cropping is effective as the landscape is still maintained even if the sky is lost. In figure (g), when height is 60% and width is 50%, the unnecessary bottom some portion of the grass that is blacked out is cropped away efficiently. The landscape is maintained and the couple preserved and centered to the image. Figure (h), where height is 80% and width is 60%, produces the same result as [4] (a) but reduces the width a bit as a result of the zooming effect when both sides are cropped. Couple is a bit enlarged and more visible now. In figure (i), when height is 50% and width is 80%, the side portions of the image are effectively cropped but the landscape is maintained along with the couple. Only necessary information is shown from the land's bottom portion. In figure (j), when height is 60% and width is 80%, a similar result is produced but the height is 10% greater than figure (i). Figure (k), even has the same effect as figure (i). However, the width is reduced as the width is less by 30%. Lastly, in figure (l), when weight is 70% the sides of the images are carefully cropped but the zooming effect now causes the height to be shirked to fit the new 70% width dimension.
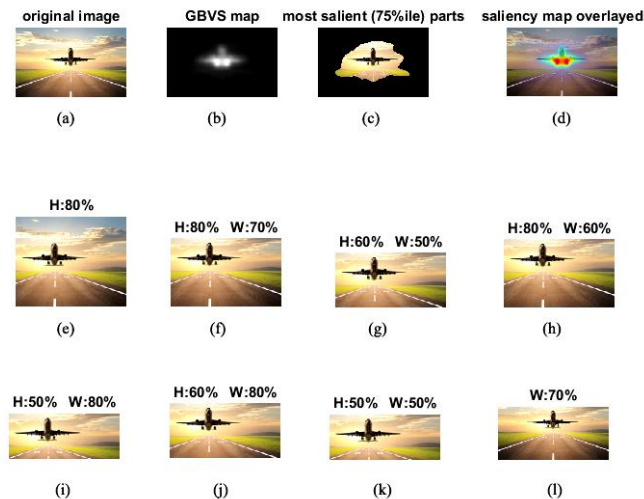
Fig. 8. Photo cropping. (a)Original image, with (b) Saliency Map, (c) Threshold and (d) Overlay from GBVS Function. (e)(l) Computed Cropping Regions based on user's Sx and Sy values. (e) Sx: 0.8, Sy: 1 (f) Sx: 0.8, Sy: 0.7 (g) Sx: 0.6, Sy: 0.5 (h) Sx: 0.8, Sy: 0.6 (i) Sx: 0.5, Sy: 0.8 (j) Sx: 0.6, Sy: 0.8 (k) Sx: 0.5, Sy: 0.5 (l) Sx: 1, Sy: 0.7.

The original image in [11] (a) is cropped and is shown in sub-figures (e) to (l). In figure (e), when height is 80%, the important elements i.e. the airplane is secured along with surrounding necessary points very effectively. In figure (f), when height is 80% and width is 70%, the change in width causes the zooming effect and hence causes an effect on the height. Hence the height gets reduced. In this case the cropping is effective as the surrounding landscape is still maintained even if the left sides are lost. In figure (g), when height is 60% and width is 50%, unnecessary bottom extended portions of the runway is cropped away efficiently. The landscape is maintained and the airplane. Figure (h), where height is 80% and width is 60%, produces the same result as [4] (a) but reduces the width a bit as a result of the zooming effect when both sides are cropped. The airplane is a bit enlarged and more visible now. In figure (i), when height is 50% and width is 80%, the side portions of the image are effectively cropped but the landscape is maintained along with the airplane. Only necessary information is shown from the land's bottom portion. In figure (j), when height is 60% and width is 80%, a similar result is produced but the height is 10% greater than figure (i). Figure (k), even has the same effect as figure (i). However, the width is reduced as the width is less by 30%. Lastly, in figure (l), when weight is 70% the sides of the images are carefully cropped but the zooming effect now causes the height to be shirked to fit the new 70% width dimension. This in turn causes everything to be preserved except the sky.
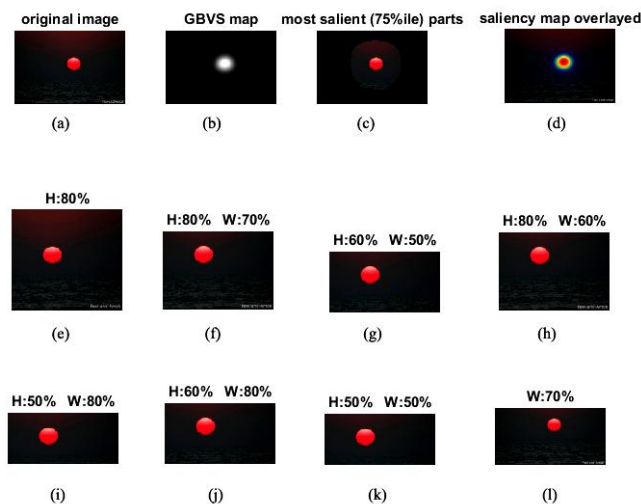
Fig. 9. Photo cropping. (a)Original image, with (b) Saliency Map, (c) Threshold and (d) Overlay from GBVS Function. (e)(l) Computed Cropping Regions based on user's Sx and Sy values. (e) Sx: 0.8, Sy: 1 (f) Sx: 0.8, Sy: 0.7 (g) Sx: 0.6, Sy: 0.5 (h) Sx: 0.8, Sy: 0.6 (i) Sx: 0.5, Sy: 0.8 (j) Sx: 0.6, Sy: 0.8 (k) Sx: 0.5, Sy: 0.5 (l) Sx: 1, Sy: 0.7.

The original image in [12] (a) is cropped and is shown in sub-figures (e) to (l). In figure (e), when height is 80%, the important elements i.e. the moon is secured along with surrounding necessary points very effectively. Only the light reflected is important with the moon in this photo. No other important points are found on far boundaries. In figure (f), when height is 80% and width is 70%, the change in width causes the zooming effect and hence causes an effect on the height. Hence the height gets reduced. In this case the cropping is effective as there are no surrounding elements that are lost. In figure (g), when height is 50% and width is 80%, important points are preserved. Figure (h), where height is 80% and width is 60%, produces the same result as [4] (e) but reduces the width a bit as a result of the zooming effect when both sides are cropped. The moon is slightly enlarged and a bit more visible now. In figure (i), when height is 50% and width is 80%, the moon is centered along the y axes and the width doesn't affect the image or doesn't cause any important data loss. In figure (j), when height is 60% and width is 80%, a similar result is produced but the height is 10% greater than figure (i). Figure (k), even has the same effect as figure (i). However, the width is reduced as the width is less by 30%. Lastly, in figure (l), when weight is 70% the sides of the images are carefully cropped but the zooming effect now causes the height to be shirked to fit the new 70% width dimension. This in turn causes the moon to be a bit smaller in perspective.

Random images are tested and results produced are clear to justify that salient objects are detected accurately and the cropping window is placed accordingly in the image. However after analyzing all the above figures, it can easily be said that if the size of the salient object is covering the entire frame of the image then during cropping, some important portions of that salient object will be lost. If not, in the least possible case, the nearest most important points and objects in the threshold will be lost. Furthermore, the change in width and height causes light but vivid changes in proportion.

To further analyze and prove that the algorithms applied work accurately in well sized and spread images rather than images where objects cover the whole real estate, some random images are tested on extreme cropping sizes while comparison is made to those images where the entire real estate is covered by the salient object(s) and threshold environment.
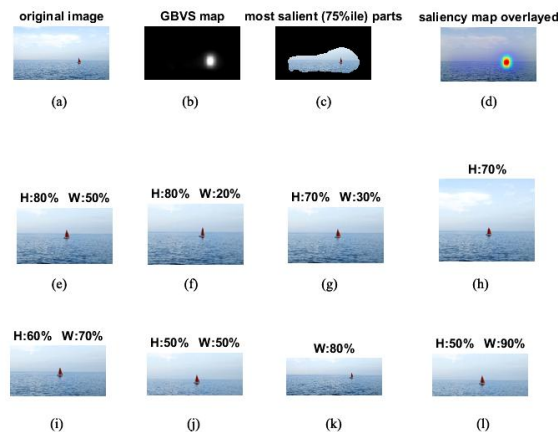
E. Extreeme Cases and Limitations

Fig. 10. Photo cropping. (a)Original image, with (b) Saliency Map, (c) Threshold and (d) Overlay from GBVS Function. (e)(l) Computed Cropping Regions based on user's Sx and Sy values. (e) Sx: 0.8, Sy: 0.5 (f) Sx: 0.8, Sy: 0.2 (g) Sx: 0.7, Sy: 0.3 (h) Sx: 0.7, Sy: 1 (i) Sx: 0.6, Sy: 0.7 (j) Sx: 0.5, Sy: 0.5 (k) Sx: 1, Sy: 0.8 (l) Sx: 0.5, Sy: 0.9.

The original image in [13] (a) is cropped and is shown in sub-figures (e) to (l). In figure (f), when height is 80% and width is 20%, the boat is secured along with surrounding necessary points very effectively even though the width is extremely reduced to only 20%. However, the boat is squeezed and is not quite clearly visible. Increasing the width in figure (g), produces a better result as the boat is now more visible.
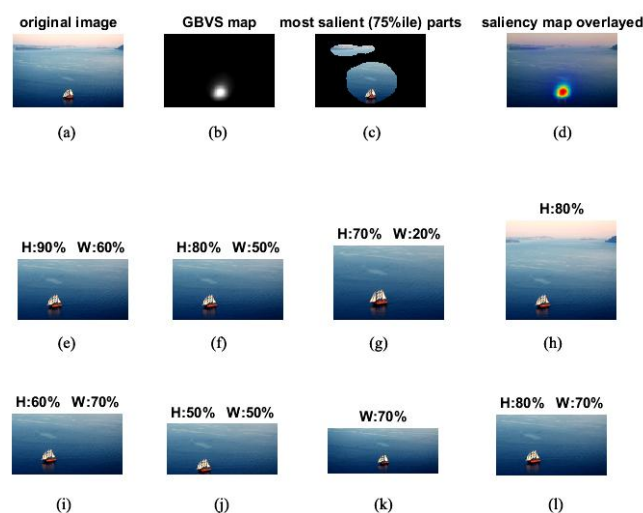

Fig. 11. Photo cropping. (a)Original image, with (b) Saliency Map, (c) Threshold and (d) Overlay from GBVS Function. (e)(l) Computed Cropping Regions based on user's Sx and Sy values. (e) Sx: 0.9, Sy: 0.6 (f) Sx: 0.8, Sy: 0.5 (g) Sx: 0.7, Sy: 0.2 (h) Sx: 0.8, Sy: 1 (i) Sx: 0.6, Sy: 0.7 (j) Sx: 0.5, Sy: 0.5 (k) Sx: 1, Sy: 0.7 (l) Sx: 0.8, Sy: 0.7.

The original image in [14] (a) is cropped and is shown in sub-figures (e) to (l). In figure (g), when height is 70% and width is 20%, the boat is secured along with surrounding necessary points very effectively even though the width is extremely reduced to only 20%. Further. The boat is enlarged and is more clearly visible now.
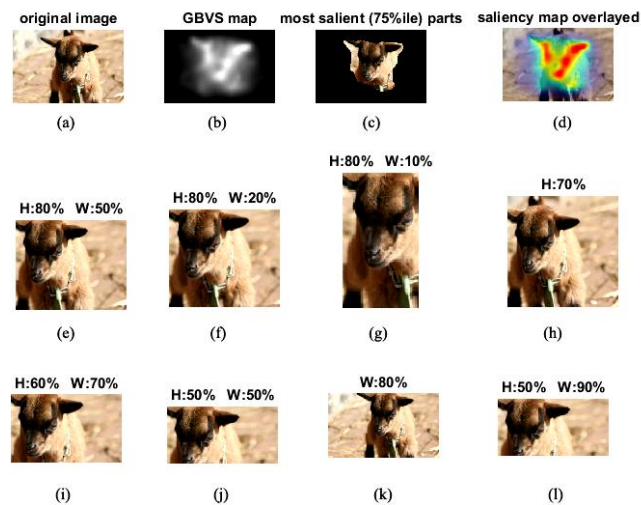

Fig. 12. Full-Scale image – Maximum Coverage Photo cropping. (a)Original image, with (b) Saliency Map, (c) Threshold and (d) Overlay from GBVS Function. (e)(l) Computed Cropping Regions based on user's Sx and Sy values. (e) Sx: 0.8, Sy: 0.5 (f) Sx: 0.8, Sy: 0.2 (g) Sx: 0.8, Sy: 0.1 (h) Sx: 0.7, Sy: 1 (i) Sx: 0.6, Sy: 0.7 (j) Sx: 0.5, Sy: 0.5 (k) Sx: 1, Sy: 0.8 (l) Sx: 0.5, Sy: 0.9.

The original image in [17] (a) is cropped and is shown in sub-figures (e) to (l). This figure is another example of a limitation in this study, it is not appropriate to be used in this case by the user. Again the most salient object i.e. the goat covers the whole image and hinders the aims and objectives of the system. For images of these sort, different approaches have to be implemented and studied. It requires further segmentation and reduction of foreground object in the overall region of the image. For example in figure (k), even when width is just reduced 20%, the goat is cut from the top. This results in nothing but data loss.

To conclude the analysis, Images are studied, algorithms are applied and images are cropped. Efficient results are produced using the method specified by the developers. However, careful attention has to be placed on the selection of images. If the salient object covers most of the area in the image then cropping these kind of images will definitely lead to loss of important data in these salient objects. Further, this requires for other methods to be researched to segment the salient objects effectively from the image to separate the foreground and background images. This foreground produced can then be scaled down separately before cropping to save the salient object from being chopped in the final result.

VII. CONCLUSION

Image processing is an aspect of image enhancement for different purpose and used in different areas. One important use of image processing is in applications to use in different devices is image resizing. Different devices have different size and image adjustment becomes an important exception. With the use of scalar cropping, important object of an image can be lost or distorted. There were various methods which were purposed by researchers to tackle with this limitation are not effective to use. But the context-aware cropping is

one of most significant to preserve the salient object. An automatic cropping of an image was able to resolve the issue. By using this technique, the maximum pixel value target its neighbours and determines the salient map around the object but this technique has some limitations as well.

However, no cropping technique introduces artefacts but it may fails to preserve or retain the all-important salient objects when image has scattered objects or regions. Moreover, image composition is also damaged because pixels are discarded. Another limitation which was analysed during the test plan is, if the important object is bigger in size, it may damage its important part and image will be damaged.

VIII. ACKNOWLEDGMENT

References

[1] Bovik, A. (2009). The essential guide to image processing. London: Academic Press.

[2] Avidan , S. & Shamir , A., 2013. Seam Carving for Content Aware Image Resizing, s.l.: s.n.

[3] Jin , Y., Wu , Q. & Liu , L., 2012. Aesthetic Photo Composition by Optimal Crop-and-Warp. ELSEVIER , p. 1.

[4] Kumar , S. S. & Babu, A. S., 2010. Automatic Image Retargeting Using salinecy Based Mesh Parameterization. 1(4), p. 3.

[5] Lin, Y., Varodayan , D. & Girod , B., 2015. Authenticating Cropped and Resized Images Using Distributed Source Coding and Expectation Maximization, Standford: s.n.

[6] Vaquero, D. et al., 2014. A Survey of Image Retargeting Techniques, Palo Alto: Nokia Research Center.

[7] Yan, J., Lin, S., Kang , S. B. & Tang , X., 2009. Learning the Change for Automatic Image Cropping, s.l.: Mi-crosoft Research Asia